Abstract

Recently, Xiao (2009) develops a novel estimation technique for quantile cointegrated time series in a static regression by extending the semiparametric approach by Phillips and Hansen (1990) and the parametrically augmented approach by Saikkonen (1991). This paper aims to extend the autoregressive distributed-lag approach of Pesaran and Shin (1998) into the quantile regression framework. This QARDL extension enables us to jointly analyse the short-run dynamics and the long-run cointegrating relationship across a range of quantiles. We derive the asymptotic distribution of QARDL estimators, and provide a general package in which the model can be estimated and tested within and across quantiles. Monte Carlo simulation results provide strong support for theoretical predictions. The main utilities of QARDL are demonstrated through the empirical application to the dividend policy in the U.S.

Keywords: QARDL, Quantile Regression, Long-run Cointegrating Relationship, Dividend Smoothing, Time-varying Rolling Estimation.

JEL classifications: C22, G35.

Acknowledgements: We are mostly grateful to the Co-editor, Han Hong, and two anonymous referees for their helpful comments. We are also grateful to Youngsoo Bae, Stephane Bonhomme, Yongsung Chang, Kausik Chaudhuri, Seonghoon Cho, In Choi, Kyungwook Choi, Viet Anh Dang, Ana-Maria Fuertes, Matthew Greenwood-Nimmo, Hwankoo Kang, Junhan Kim, Gary Koop, Sangsoo Park, Yangsoo Park, Peter Phillips, Kevin Reily, Xin Shen, Kyulee Shin, Peter Smith, Seungjoo Song, Peter Spencer, Till van Treeck, Mike Wickens, Ralf Wilke, seminar participants at the Bank of Korea, the Universities of Brunel, Korea, Leeds, Seoul, Sogang, Yonsei and York, Cass Business School, and the Institut für Makroökonomie und Konjunkturforschung (IMK, Dusseldorf), and conference delegates at the Conference in Honour of Professor PCB Phillips (University of York, 12-13 July 2012) for their helpful comments. Shin acknowledges partial support from the ESRC (Grant No. RES-000-22-3161). Kim is grateful for financial support from the Yonsei University Research Fund (2011-1-0244). The simulation codes written in GAUSS are downloadable at http://web.yonsei.ac.kr/jinseocho/research.htm. The usual disclaimer applies.
1 Introduction

The statistical properties of cointegrated time series have been extensively investigated over the last three decades. The fully modified OLS estimator of Phillips and Hansen (1990) has been the most widely used method in linear time series modelling. However, this semiparametric approach is static and shows only the long-run relationship between integrated series, leaving the associated short-run dynamics to be recovered through a potentially inefficient two-step approach, as in Engle and Granger (1987). In response, Pesaran and Shin (1998) propose the autoregressive distributed lag error correction model (ARDL-ECM) that allows us to simultaneously investigate the long-run relationship and the short-run dynamics in a fully parametric manner. In particular, they generalise the ARDL approach for cointegration and develop the asymptotic theory for estimation and inference. Furthermore, Pesaran, Shin, and Smith (2001) develop a pragmatic bounds-testing procedure for the existence of a stable long-run relationship, which is valid irrespective of whether the underlying regressors are either $I(1)$, mutually cointegrated, or $I(0)$.

Recently, the literature on quantile time series regression has been rapidly growing, e.g., Koenker and Xiao (2004, 2006). More importantly, Xiao (2009) advances a novel quantile cointegration approach in a static regression and develops the semiparametric fully modified and the parametrically augmented quantile estimators, which can be regarded as the quantile counterparts of the estimators proposed respectively by Phillips and Hansen (1990) and Saikkonen (1991). It is now well established that the quantile estimator is consistent and asymptotically normal when the stationarity condition is satisfied together with other regularity conditions, e.g., Koenker and Basset (1978), Phillips (1991a) and Koenker and Zhao (1996). In this regard, Xiao’s approach is a natural but challenging contribution to the context of quantile regression with nonstationary variables. As a pioneer of cointegration analysis, Granger (2010) provides further insightful discussions on the analysis of possibly cointegrated quantile time series. An increasing number of studies have also adopted Xiao’s quantile cointegration approach and documented evidence that the conventional cointegration analysis may not be sufficiently informative because it focuses only on the central area of the whole distribution, e.g., Lee and Zeng (2011), Burdekin and Siklos (2012), Wang (2012) and Tsong and Lee (2013).

In this paper, we aim to contribute to this growing literature by proposing the dynamic quantile ARDL-ECM (QARDL-ECM), in which we can simultaneously address both the long-run (cointegrating) relationship and the associated short-run dynamics across a range of quantiles in a fully parametric setting. Since this is a straightforward extension of the ARDL-ECM developed by Pesaran and Shin (1998) into the quantile regression context, it is expected that all of the optimal estimation properties can be obtained in a similar...

We provide an asymptotic theory for estimating and testing the QARDL model with nonstationary regressors. The QARDL estimators of the short-run dynamic parameters and the long-run cointegrating parameters are shown to asymptotically follow the (mixture) normal distribution. We also show that the null distribution of the Wald statistics for testing the restrictions on the short-run and the long-run parameters within and across quantiles weakly converges to a chi-squared distribution. Next, via Monte Carlo simulation studies, we find that the overall simulation results, focusing on the empirical size and power of the Wald test statistics, provide strong support for our theoretical predictions both in the case with fixed QARDL orders and in the case where the (unknown) QARDL orders are consistently selected based on the Bayesian information criterion (BIC).

In a seminal study on dividend policy, Lintner (1956) observes that firms gradually adjust dividends in response to changes in earnings toward the long-run target payout ratio. Empirical research at both the firm and aggregate levels generally supports Lintner’s partial adjustment framework, e.g., Fama and Babiak (1968) and Marsh and Merton (1987). Recently, Brav, Graham, Harvey and Michaely (2005) surveyed 384 financial executives to determine the factors that drive dividend and share repurchase decisions. Although the new survey evidence is mostly consistent with Lintner’s observations, the link between dividends and earnings was substantially weak. Using the firm-level data in the US, Leary and Michaely (2011) document that dividend smoothing has steadily increased over the past century, even before firms began using share repurchases in the mid-1980s. Chen, Da, and Priestley (2012) also demonstrate that aggregate dividends are dramatically more smoothed in the postwar period than in the prewar period.

However, all of these studies examine dividend behavior only at the conditional mean and do not investigate an important possibility that the dividend policy may be fundamentally heterogeneous across different quantiles of the conditional distribution of dividends. In this study, we aim to contribute to the existing literature on dividend policy by incorporating location (quantile) asymmetries in the long-run target payout ratios and the dynamic dividend adjustment at the aggregate level. In order to construct a simple but flexible model of the dividend process that captures the stylised facts on management behavior, we apply the QARDL model to the dataset over the period from1871Q3 to 2010Q2 extracted from Robert Shiller’s web page (http://www.econ.yale.edu/shiller).

The full sample estimation results demonstrate that there is strong evidence of location asymmetries between lower and medium-to-higher quantiles of dividends. In particular, we find that the long-run payout ratio is higher and the dividend smoothing is stronger at higher quantiles than lower quantiles of dividends.
Overall, our findings are consistent with the cross-sectional evidence by Leary and Michaely (2011) and the aggregate time series evidence in a global setting by Rangvid, Schmeling, and Schrimp (2012) that dividend smoothing is most common among large and mature firms with stable cash flows and that are not financially constrained, face low levels of asymmetric information, and are readily susceptible to agency conflicts.

We also allow for time-varying patterns of dividend policy by employing a rolling estimation technique with a window length of 320 quarters. A thorough examination of the time-varying QARDL estimation results provides a number of insightful findings on dividend policy in the US over the past century. First, dividend smoothing has become monotonically stronger over time. Similar monotonic downward trends have been observed for the impact coefficient with respect to contemporaneous changes in earnings. Both factors contribute to the extremely strong dividend smoothing reported in recent periods. Second, payout ratios have been monotonically decreasing over time, supporting the survey evidence by Brav et al. (2005) that the target payout ratio may no longer be the preeminent variable affecting payout decisions. More importantly, we find that the location asymmetries across different quantiles of the conditional distribution of dividends, which were clearly visible and pervasive in earlier periods, are less frequent in recent periods. These phenomena may indicate the establishment of financial deepening as a long-term process in the US, which serves to promote the stability of the whole financial system.

The paper is organized as follows. Section 2 introduces the QARDL model and derives the asymptotic distribution of both the short-run and the long-run estimators. Section 3 extends the QARDL model and its inference across multiple quantiles. Section 4 evaluates the finite-sample performance via Monte Carlo simulations. Section 5 presents the empirical application to an analysis of dividend smoothing and the long-run target payout ratio in the US. Section 6 presents concluding remarks and further research extensions. All of the proofs are relegated to the Appendix.

Before proceeding, we discuss some notational details. A function is denoted using an empty argument. Parameters without a subscript are used to indicate generic notations constituting the parameter space. Parameters attached with the subscript “∗” are those characterizing the data generating process. “⇒” and “[a]” denote “weakly converges” and “the smallest integer greater than a” respectively. We let $[a_{ij}]_{i=1,...,m,j=1,...,n}$ be an $m \times n$ matrix with a typical element $a_{ij}$ and denote $\mathbf{1}_\ell$ as an $\ell \times 1$ vector of ones. $\mathbf{A} \odot \mathbf{B}$ denotes the Hadamard product of $\mathbf{A}$ and $\mathbf{B}$. Other notations are standard.

---

1Denis and Osobov (2008) document that dividends are concentrated among the largest, most profitable payers in the US, Canada, UK, Germany, France, and Japan. In particular, they document that dividend payers account for more than 90% of the aggregate market capitalization except in the US and Canada, and the top 20% of dividend payers account for virtually all of the market capitalization of dividend payers. The concentration of dividends among the largest, most profitable firms is consistent with the life-cycle theory’s central prediction that the distribution of free cash flow is the primary determinant of dividend policy.
2 Quantile Autoregressive Distributed Lag Model

We consider the following quantile autoregressive distributed lag (QARDL) process:

\[ Y_t = \alpha^*(\tau) + \sum_{j=1}^{p} \phi_{j*}(\tau)Y_{t-j} + \sum_{j=0}^{q} \theta_{j*}(\tau)'X_{t-j} + U_t(\tau), \]  

(1)

where \( \tau \in (0, 1) \) is a quantile index, and \( p \) and \( q \) are lag orders. We suppose that \( U_t(\tau) \in \mathbb{R} \) is an independent and identically distributed (IID) process with a finite second moment, and \( X_t \in \mathbb{R}^k \) is assumed to be an integrated process of a stationary and ergodic process with population mean zero. It is further assumed that the \( k \) variables in \( X_t \) are not cointegrated among themselves.

Pesaran and Shin (1998) introduce the ARDL approach to cointegrated series and develop an asymptotic theory for estimating and establishing an inference on cointegrated series. Furthermore, Pesaran et al. (2001) extend the ARDL approach and develop a pragmatic testing procedure for the existence of a stable long-run relationship, showing that it is a valid procedure irrespective of whether the underlying regressors are either \( I(1) \), mutually cointegrated, or \( I(0) \) processes.\(^2\)

We analyse the QARDL process by extending the ARDL approach. To this end, we first reformulate (1) as follows:

\[ Y_t = \alpha^*(\tau) + \sum_{j=0}^{q-1} W_{t-j}'\delta_{j*}(\tau) + X_t'\gamma^*(\tau) + \sum_{j=1}^{p} \phi_{j*}(\tau)Y_{t-j} + U_t(\tau), \]  

(2)

where \( \gamma^*(\tau) := \sum_{j=0}^{q} \theta_{j*}(\tau) \), \( W_t := \Delta X_t \), and \( \delta_{j*}(\tau) := -\sum_{i=j+1}^{q} \theta_{i*}(\tau) \). All of the parameters in (2) measure the short-run dynamics, but the long-run relationship between \( Y_t \) and \( X_t \) is captured simply by reformulating (2) into the following long-run quantile regression equation:

\[ Y_t = \mu^*(\tau) + X_t'\beta^*(\tau) + R_t(\tau) \quad \text{and} \quad R_t(\tau) := \sum_{j=0}^{\infty} W_{t-j}'\xi_{0,j*}(\tau) + \sum_{j=0}^{\infty} \rho_{j*}(\tau)U_{t-j}(\tau), \]  

(3)

where

\[ \mu^*(\tau) := \left(1 - \sum_{i=1}^{p} \phi_{i*}(\tau)\right)^{-1} \alpha^*(\tau); \quad \beta^*(\tau) := \left(1 - \sum_{i=1}^{p} \phi_{i*}(\tau)\right)^{-1} \gamma^*(\tau); \quad \xi_{0,j*}(\tau) := -\sum_{\ell=j+1}^{\infty} \pi_{\ell*}(\tau) \]  

(4)

\(^2\)The flexibility and utility of the ARDL technique are reflected in the vast literature that adopts its applications for the analysis of a wide range of economic variables (e.g., the video, available at http://www.youtube.com/watch?v=d9E8BKsocis, which demonstrates its applications in Microfit and Eviews).
such that \( \{ \rho_{0*}(\tau), \rho_{1*}(\tau), \ldots \} \) and \( \{ \pi_{0*}(\tau), \pi_{1*}(\tau), \ldots \} \) respectively satisfy

\[
\frac{1}{1 - \sum_{j=1}^{p} \phi_{j*}(\tau)L^j} = \sum_{j=0}^{\infty} \rho_{j*}(\tau)L^j \quad \text{and} \quad (1 - L)^{-1} \left( \frac{\sum_{j=0}^{q} \theta_{j*}(\tau)L^j}{1 - \sum_{j=1}^{p} \phi_{j*}(\tau)L^j} \right) = \sum_{j=0}^{\infty} \pi_{j*}(\tau)L^j.
\]

Here, the static equation (3) is obtained by solving for \( Y_t \) from (2). The residual term \( R_t(\tau) \) represents the collection of serially correlated stationary variables irrelevant to the long-run relationship. We capture this by \( \beta_*(\tau) \) and call it the \textit{long-run parameter}. As is clear from its definition, \( \beta_*(\tau) \) can be estimated by first estimating \( \gamma_*(\tau) \) and \( \phi_*(\tau) := (\phi_{1*}(\tau), \ldots, \phi_{p*}(\tau))^t \).

Our main interests lie in developing the estimation theory for the long-run parameter \( \beta_*(\tau) \) by first estimating the short-run parameters. To this end we reformulate the QARDL process in (2) as

\[
Y_t = G'_t\lambda_*(\tau) + \tilde{Y}_t\phi_*(\tau) + U_t(\tau) = Z'_t\alpha_*(\tau) + U_t(\tau),
\]

where \( Z_t := (G'_t, \tilde{Y}_t)' := (G'_t, Y_{t-1}, \ldots, Y_{t-p})' := (1, W'_t, \ldots, W'_{t-q+1}, X'_t, Y_{t-1}, \ldots, Y_{t-p})' \), and \( \alpha_*(\tau) := [\lambda_*(\tau)', \phi_*(\tau)']' := [\alpha_*(\tau), \delta_*(\tau)', \gamma_*(\tau), \phi_*(\tau)']' := [\alpha_*(\tau), \delta_1*(\tau)', \ldots, \delta_{(q-1)*}(\tau)', \gamma_*(\tau)', \phi_*(\tau)']' \). Here, each element of \( \tilde{Y}_t \) in (5) has the following specific form:

\[
Y_{t-i} = \mu_*(\tau) + X'_t\beta_*(\tau) + \sum_{j=0}^{q-1} W'_{t-j}\xi_{i,j*}(\tau) + K_{t,i}(\tau), \quad i = 1, 2, \ldots, p,
\]

where we let

\[
\xi_{i,j*}(\tau) := \begin{cases} -\beta_*(\tau), & \text{if } i > j; \\ -\sum_{\ell=j-i}^{\infty} \pi_{\ell*}(\tau), & \text{if } i \leq j, \end{cases}
\]

and

\[
K_{t,i}(\tau) := \begin{cases} -\sum_{q=1}^{\infty} W'_{t-i-j}\xi_{0,j*}(\tau) + \sum_{j=0}^{\infty} \rho_{j*}(\tau)U_{t-i-j}(\tau), & \text{if } i \leq q; \\ -\sum_{j=0}^{q-1} W'_{t-q-j}\beta_*(\tau) + \sum_{j=0}^{\infty} W'_{t-i-j}\pi_{j*}(\tau) + \sum_{j=0}^{\infty} \rho_{j*}(\tau)U_{t-i-j}(\tau), & \text{if } i > q. \end{cases}
\]

Note that (6) is the lagged version of (3). Pesaran and Shin (1998) also provide a detailed derivation of (6) in their context. As (6) is iteratively used below, we rewrite it more compactly as

\[
\tilde{Y}_t = \Gamma_*(\tau)'G_t + \tilde{K}_t(\tau),
\]

\[5\]
where
\[
\Gamma_s(\tau) := \begin{bmatrix}
\mu_s(\tau) & \mu_s(\tau) & \cdots & \mu_s(\tau) \\
\xi_{1,0s}(\tau) & \xi_{2,0s}(\tau) & \cdots & \xi_{p,0s}(\tau) \\
\vdots & \vdots & \ddots & \vdots \\
\xi_{1,q-1s}(\tau) & \xi_{2,q-1s}(\tau) & \cdots & \xi_{p,q-1s}(\tau) \\
\beta_s(\tau) & \beta_s(\tau) & \cdots & \beta_s(\tau)
\end{bmatrix}
\quad \text{and} \quad
K_t(\tau) := \begin{bmatrix}
K_{t,1}(\tau) \\
K_{t,2}(\tau) \\
\vdots \\
K_{t,p}(\tau)
\end{bmatrix}.
\]

Note that the short-run parameters are presented as the coefficients of \(Z_t\) in (5), and that the long-run parameter \(\beta_s(\tau)\) can then be estimated using the definition in (4).

We provide remarks about the quantile dependency of the QARDL parameters. First, we allow the short-run and the long-run parameters to be quantile-dependent, implying that the QARDL parameters can be affected by the innovation \(U_t\) received at each period and thus can be different across quantiles. Therefore, the (dynamic) conditioning variables not only shift the location, but also alter the scale and shape of the conditional distribution of \(Y_t\). Furthermore, Xiao (2009) shows that the zero correlation between regressors and errors is a key condition for consistently estimating the quantile-dependent cointegrating vector. Notice that this is the main reason why Xiao allows the quantile-dependent cointegrating parameters only under the dynamic OLS approach using leads and lags, not under the semiparametric fully-modified framework, and develops the relevant asymptotic theory. In the ARDL context, this condition is equivalent to \(\mathbb{E}[\Delta X_t U_t] = 0\) and \(\mathbb{E}[Y_{t-j} U_t] = 0\) \((j = 1, \ldots, p)\), both of which are satisfied given sufficiently large lag orders of \(p\) and \(q\) (see Assumption 1 below). Second, it is straightforward to rewrite the QARDL process in (1) in the following error correction model (ECM):

\[
\Delta Y_t = \alpha_s(\tau) + \varsigma_s(\tau) Y_{t-1} + \gamma_s(\tau) X_{t-1} + \sum_{j=1}^{p-1} \phi_{js}(\tau) \Delta Y_{t-j} + \sum_{j=0}^{q-1} \tilde{\theta}_{js}(\tau) \Delta X_{t-j} + U_t(\tau) \quad (7)
\]

where \(\varsigma_s(\tau) := \sum_{j=1}^{p} \phi_{js}(\tau) - 1, \gamma_s(\tau) := \sum_{j=0}^{p} \theta_{js}(\tau), \tilde{\theta}_{0s}(\tau) = \theta_{0s}(\tau), \) and for \(j = 1, \ldots, p - 1, \phi_{js}(\tau) := -\sum_{h=j+1}^{p} \phi_{hs}(\tau) \) and \(\tilde{\theta}_{js}(\tau) = -\sum_{h=j+1}^{p} \theta_{hs}(\tau) \). We refer to (7) as the QARDL-ECM representation. Furthermore, the special case of the quantile-invariant homogeneous cointegration with
\( \beta_{*}(\tau) = \beta_{*} \) for all \( \tau \)'s, is given by

\[
\Delta Y_t = \alpha_{*}(\tau) + \zeta_{*}(\tau) (Y_{t-1} - \beta'_{*} X_{t-1}) + \sum_{j=1}^{p-1} \tilde{\phi}_{j,n}(\tau) \Delta Y_{t-j} + \sum_{j=0}^{q-1} \tilde{\theta}_{j,n}(\tau)' \Delta X_{t-j} + U_t(\tau).
\]

Even in this simpler case, we might still be interested in testing whether the speed of adjustment \( \zeta_{*}(\tau) \) is quantile dependent.

We now define our estimators. First, the reformulation in (5) can be used to estimate the unknown short-run parameters by applying the standard quantile regression approach, e.g., Koenker and Hallock (2001). For a given \( \tau \in (0, 1) \), we obtain the QARDL estimator as

\[
\tilde{\alpha}_n(\tau) := \arg \min_{\alpha(\tau)} \sum_{t=1}^{n} \varrho_{\tau} \{ Y_t - Z_t' \alpha(\tau) \},
\]

where \( \varrho_{\tau}(u) := u \psi_{\tau}(u) \) is the so-called check function with \( \psi_{\tau}(u) := \tau - I(u \leq 0) \). Conformably with \( \alpha_{*}(\tau) \), we partition \( \tilde{\alpha}_n(\tau) \) into

\[
\tilde{\alpha}_n(\tau) := \left( \tilde{\lambda}_n(\tau)', \tilde{\phi}_n(\tau)' \right)' := \left( \tilde{\alpha}_n(\tau), \tilde{\delta}_n(\tau)' \tilde{\gamma}_n(\tau)', \tilde{\phi}_n(\tau)' \right)'
\]

for future references. Next, we estimate the long-run parameter using the plug-in principle:

\[
\tilde{\beta}_n(\tau) := \left( 1 - \sum_{j=1}^{p} \tilde{\phi}_{j,n}(\tau) \right)^{-1} \tilde{\gamma}_n(\tau).
\]

It is clear that the behavior of \( \tilde{\beta}_n(\tau) \) governing the quantile long-run equilibrium is intrinsically related to the short-run estimators \( \tilde{\phi}_{j,n}(\tau) \) and \( \tilde{\gamma}_n(\tau) \).

The asymptotic properties of the quantile regression estimators are well-established under standard conditions. They are consistent and asymptotically normal under the stationarity condition along with other regularity conditions, e.g., Bloomfield and Steiger (1983), Pollard (1991), Phillips (1991a), and Koenker and Zhao (1996). In particular, Phillips (1991a) examines the dynamic LAD estimator, and Koenker and Zhao (1996) focus on the quantile regression model with conditional heteroskedasticity. Furthermore, Kim and White (2003) examine a misspecified quantile regression model with conditional heteroskedastic disturbances. Quantile regression models have been widely used in a number of fields, notably the analysis of stock market returns (e.g., Barnes and Huhes, 2002) and in labour economics (e.g., Martins and Pereira, 2004). Quantile regression has also become an important tool in risk management; see Engle and Man-
ganelli (2004) for the value at risk (VaR) evaluation; Adrian and Brunnermeier (2010) for conditional VaR (CoVaR); and Acharya et al. (2011) for marginal expected shortfall (MES).

The literature on quantile regressions with nonstationary variables has emerged very recently. In particular, Xiao (2009) advances a quantile cointegration approach in a static regression context and develops both the semiparametric fully modified estimator and the parametrically augmented estimator using leads and lags of the first-differenced regressors. These correspond to the quantile extensions of the estimators proposed by Phillips and Hansen (1990) and by Saikkonen (1991), respectively. These estimators are shown to follow the limiting mixture normal distributions even in the presence of endogenous regressors and/or serially correlated residuals such that the standard inference theory can be applied.

Xiao’s (2009) quantile cointegration modelling approach has been adopted by a number of studies, documenting evidence that the conventional cointegration analysis focusing on the (common) mean behaviors can be misleading. Lee and Zeng (2011) find that the response of spot oil prices to shocks in one-month futures is much greater with higher spot prices than with lower prices and argue that this finding is consistent with the prospect theory in that the value function is generally steeper for losses than for gains. Burdekin and Siklos (2012) utilise a post-Asian financial crisis sample of 1999–2010 and find substantial evidence of integration of the Shanghai stock market with the US and many regional stock markets. In particular, they find that cointegration is prevalent at the higher end of the distribution and argue that this clearly reflects the tendency for emerging market stock returns to be more concentrated at the higher end of the distribution. Tsong and Lee (2013) examine the empirical validity of the Fisher hypothesis for six OECD countries and find that the cointegrating coefficients between the nominal interest rate and inflation display the asymmetric location patterns: the cointegrating coefficient between the nominal rate and the inflation is smaller in the lower quantiles, than higher quantiles, leading to the so-called Fisher effect puzzle. In contrast, in the upper quantiles, the former adjusts on a one-to-one basis to changes in the latter, supporting the Fisher hypothesis. Furthermore, Wang (2012) studies inference on multiple structural breaks within quantile cointegrating regressions and proposes a fully modified estimator.

We note in passing that our proposed QARDL-ECM approach is different from those in the aforementioned studies. previous methods are typically built upon Engle and Granger’s (1987) static cointegration framework, whereas the QARDL estimator simultaneously addresses both the long-run relationship and the associated short-run dynamics across a wide range of quantiles.

We impose the following assumptions to examine the regular properties of the QARDL estimator.

**Assumption 1.** (i) The disturbance \( U_t \sim IID (0, \sigma_u^2) \) with \( \sigma_u^2 < \infty \). \( U_t \) has a continuous probability
density function (PDF) \( f(\cdot) \) and cumulative density function (CDF) \( F(\cdot) \) such that \( f(\cdot) > 0 \) and \( f_\tau := f \left[ F^{-1}(\tau) \right] < \infty \). In addition, \( f_{t|Z_t}(\cdot|Z_t) = f(\cdot); \)

(ii) \( X_t \) is a \( k \times 1 \) vector of integrated regressors such that \( W_t := \Delta X_t \) is a general linear multivariate stationary and ergodic process with \( \mathbb{E} [W_{ij}] = 0 \) and \( \mathbb{E} \left[ |W_{ij}|^2 \right] < \infty \) for \( j = 1, 2, \ldots, k; \)

(iii) \( U_t \) and \( W_s \) are independent for all \( t, s = 1, 2, \ldots; \)

(iv) The cointegration order of \( X_t \) is zero;

(v) For each \( \tau \in (0, 1), \) the roots of \( \left( 1 - \sum_1^P \phi_j(\tau) L^j \right) \) lie outside the unit circle, and for all \( i = 1, \ldots, k, \) \( \sum_{j=0}^\infty |\xi_{0,i}(\tau)| < \infty \) and \( \sum_{j=0}^\infty |\pi_{j,i}(\tau)| < \infty, \) where \( \xi_{0,i}(\tau) \) and \( \pi_{j,i}(\tau) \) are the \( i \)-th elements of \( \xi_{0,i}(\tau) \) and \( \pi_{j,i}(\tau), \) respectively; and

(vi) For each \( r, \tau \in (0, 1), \) \( B_{n}(\cdot, \tau) \Rightarrow B(\cdot, \tau), \) where

\[
B_{n}(r, \tau) := n^{-1/2} \sum_{i=1}^{[nr]} \begin{bmatrix} \bar{W}_t, \bar{K}_t(\tau)', \psi_t[U_t(\tau)], \psi_t[U_t(\tau)] \bar{W}_t', \psi_t[U_t(\tau)] K_t(\tau)' \end{bmatrix}',
\]

\[
\bar{K}_t(\tau) := K_t(\tau) - \mathbb{E}[K_t(\tau)], \; \bar{W}_t := \begin{bmatrix} W_t, W_{t-1}, \ldots, W_{t-q+1} \end{bmatrix}', \; \mathcal{B}(\cdot, \tau) := \begin{bmatrix} B_{W}(\cdot, \tau)' \; B_{K}(\cdot, \tau)' \; B_{\psi}(\cdot, \tau)' \end{bmatrix}', \; \Omega(r, \tau) := \lim_{n \to \infty} \mathbb{E}[B_{n}(r, \tau) B_{n}(\bar{r}, \tau)']
\]

A number of remarks are in order. First, the infinite density case is eliminated by Assumption 1(i).\(^3\) Second, \( \{U_t(\tau)\} \) is assumed to be an IID process for analytical convenience. Allowing \( \{U_t(\tau)\} \) to be a martingale difference array (MDA) does not alter the main results, although more tedious derivations are required. Third, Assumption 1(v) is imposed for the existence of a stable long-run relationship between \( Y_t \) and \( X_t. \) Fourth, the covariance structure of \( \Omega \) in Assumption 1(vi) is obtained as \( \lim_{t \to \infty} \mathbb{E}[B_{n}(r, \tau) B_{n}(\bar{r}, \tau)']. \) In particular, \( B_{W}(\cdot, \tau), B_{\psi}(\cdot, \tau) \) and \( B_{\psi,W}(\cdot, \tau) \) are independent according to Assumption 1(iii) and the fact that \( \mathbb{E}[\psi[U_t(\tau)]] = 0, \mathbb{E}[\bar{W}_s] = 0, \mathbb{E}[\bar{W}_s\psi[U_t(\tau)]] = 0, \mathbb{E}[\bar{W}_s\bar{W}_s'\psi[U_t(\tau)]] = 0, \) and \( \mathbb{E}[\bar{W}_s\psi[U_t(\tau)]]^2 = 0. \) This renders the covariances of \( B_{W}(\cdot, \tau), B_{\psi}(\cdot, \tau) \) and \( B_{\psi,W}(\cdot, \tau) \) to be zero. On the other hand, \( K_s(\tau) \) is not necessarily independent of \( U_t(\tau), \) although \( K_s(\tau) \) is a vector of stationary and ergodic processes based on the definition of \( K_{i,s}(\tau) \) and Assumptions 1(ii, iii). This aspect makes the covariance between \( B_{K}(\cdot)' \) and \( B_{\psi,K}(\cdot, \tau)' \) indeterministic. We impose the positive-definite matrix condition to

\(^3\)Cho, Han, and Phillips (2010) and Han, Cho, and Phillips (2011) examine the asymptotic behavior of the LAD estimator when the density function has infinite value at the median using the methodology of Knight (1998).
\( \Omega (r, \bar{r}, \tau) \) as its minimal condition. Fifth, by the definition of \( \mathbf{X}_t \), it is straightforward to derive that

\[
n^{-1} \sum_{t=1}^{n} \psi_\tau[U_t(\tau)]\mathbf{X}_t \Rightarrow \int_{0}^{1} \mathcal{B}_W(r) d\mathbb{B}_\psi(r, \tau) \quad \text{and} \quad n^{-2} \sum_{t=1}^{n} \mathbf{X}_t \mathbf{X}_t' \Rightarrow \int_{0}^{1} \mathcal{B}_W(r) \mathcal{B}_W(r)' \, dr,
\]

where \( \mathcal{B}_W(\cdot) \) is the first \( k \) elements of \( \mathcal{B}_W(\cdot) \). Since \( \psi_\tau[U_t(\tau)] \) is independent of \( \mathbf{W}_t \) from the definition of \( U_t(\tau) \), the one-sided long-run covariance is equal to zero. Sixth, the independence condition in Assumption 1(iii) is not as strong as it appears. In particular, the empirical application in Section 5 illustrates how one can control for endogeneity of regressors by projecting the regression errors on \( \Delta \mathbf{X}_t \). See also Pesaran and Shin (1998) and Pesaran et al. (2001). We, therefore, stick to this condition without loss of generality of our analysis.\(^4\) Seventh, we impose the positive-definite condition to \( \Omega (r, \bar{r}, \tau) \) to prevent identical or constant sample paths in \( \mathcal{B}(\cdot, \tau) \). Eighth, the high level condition in Assumption 1(vi) can be attained in a couple of ways. Given the other conditions, if \( \{(\mathbf{W}_t', \mathbf{K}_t(\tau)')'\} \) is an adapted mixingale of size \(-1\) with finite global variance, Theorem 3 in Scott (1973) provides the desired result. Alternatively, Assumption 1(vi) follows if the summability condition on Assumption 1(v) is further strengthened and their fourth moment is finite. See for example Phillips and Solo (1992) for this derivation. As one of these two assumptions does not necessarily imply the other, we opt to impose the high level assumption. Finally, Assumptions 1(i and iii) imply that \( E[\psi_\tau(U_t(\tau))|Y_{t-1}, Y_{t-2}, ..., \mathbf{X}_t, \mathbf{X}_{t-1}, ...] = 0 \), so that the \( \tau^{th} \)-quantile of the conditional distribution of \( U_t(\tau) \) is zero. From this zero-condition, the coefficients of the QARDL model in (1) are identified.

We provide the asymptotic distributions of the short-run estimators \( \tilde{\phi}_{j,n}(\tau) \) and \( \tilde{\gamma}_{n}(\tau) \) below.

**Theorem 1 (Short-Run Estimators).** \( \text{Under Assumption 1,} \)

(i) For each \( \tau \in (0, 1), \sqrt{n}(\tilde{\phi}_{n}(\tau) - \phi_*(\tau)) \Rightarrow \mathcal{Z}_\phi(\tau) \sim N[0, \Pi(\tau)], \text{where} \Pi(\tau) := \tau(1 - \tau) f_\tau^{-2}E\left[ \tilde{H}_t(\tau)\tilde{H}_t(\tau)' \right]^{-1}, \tilde{H}_t(\tau) := \mathbf{K}_t(\tau) - E\left[ \mathbf{K}_t(\tau)\mathbf{W}_t' \right] E\left[ \mathbf{W}_t\mathbf{W}_t' \right]^{-1} \mathbf{W}_t, \text{and} \mathbf{W}_t := [1, \mathbf{W}_t']; \text{and}

(ii) For each \( \tau \in (0, 1), \sqrt{n}(\tilde{\gamma}_{n}(\tau) - \gamma_*(\tau)) \Rightarrow \mathcal{Z}_\gamma(\tau) \sim N[0, \beta_*(\tau) \tau'\Pi(\tau) \tau_\beta \beta_*(\tau)']. \)

Therefore, \( (\tilde{\phi}_{n}(\tau)', \tilde{\gamma}_{n}(\tau)')' \) is consistent for \( (\phi_*(\tau)', \gamma_*(\tau)')' \), and its convergence rate is \( \sqrt{n} \). We outline the key steps of the proof and provide its details in the Appendix. Using Assumption 1, we show:

\[
\sqrt{n}(\tilde{\phi}_{n}(\tau) - \phi_*(\tau)) = f_\tau^{-1}E\left[ \tilde{H}_t(\tau)\tilde{H}_t(\tau)' \right]^{-1}\left( n^{-1/2}\tilde{H}(\tau)\Psi_\tau(U) \right) + o_\mathbb{P}(1),
\]

(8)

where \( \tilde{H}(\tau) := [\tilde{H}_1(\tau), \ldots, \tilde{H}_n(\tau)]' \), and \( \Psi_\tau(U) := [\psi_\tau[U_1(\tau)], \ldots, \psi_\tau[U_n(\tau)]]' \). It is also straightfor-

---

\(^4\)We also note that the independence assumption can be relaxed into the zero correlation condition, as in OLS regression, in which case the variance-covariance matrix will be expressed in the usual sandwich form.
ward to show that
\[ n^{-1/2} \tilde{H}(\tau)' \Psi_\tau(U) \overset{d}{\sim} N \left[ 0, \tau(1-\tau)E [\tilde{H}_t(\tau)\tilde{H}_t(\tau)'] \right] \]  
(9)
by applying the central limit theorem (CLT) for MDA. Theorem 1(i) follows by combining (8) and (9).

Pesaran and Shin (1998) derive the asymptotic covariance matrix in their context by implicitly imposing that \( E[\tilde{K}_t\tilde{W}_t] \) is zero. However, this is too strong unless \( \{W_t\} \) is an IID process. Hence, we relax this restriction in what follows. Second, we can prove Theorem 1(ii) by first showing that
\[ \sqrt{n} \left\{ (\tilde{\gamma}_n(\tau) - \gamma_*(\tau)) + \beta_*(\tau) \iota' p (\tilde{\phi}_n(\tau) - \phi_*(\tau)) \right\} = o_P(1), \]  
(10)
and then using the asymptotic distribution of \( \sqrt{n} \left( \tilde{\phi}_n(\tau) - \phi_*(\tau) \right) \) given in Theorem 1(i). It is clear from (10) that the asymptotic distribution of \( \tilde{\gamma}_n(\tau) \) depends upon that of \( \tilde{\phi}_n(\tau) \).

We examine the asymptotic behaviors of the long-run parameter estimator in the following theorem:

**Theorem 2 (Long-Run Estimator).** Under Assumption 1, and for each \( \tau \in (0,1) \),

(i) \( n(\tilde{\beta}_n(\tau) - \beta_*(\tau)) \Rightarrow f_\tau \left( 1 - \sum_{j=1}^{p} \phi_{j*}(\tau) \right) \left[ \int_0^1 \tilde{B}_W(r) \tilde{B}_W(r)'dr \right]^{-1} \int_0^1 \tilde{B}_W(r)d\psi(r,\tau), \)

where \( \tilde{B}_W(r) := B_W(r) - \int_0^1 B_W(r)dr \); and

(ii) \( nM^{1/2}(\tilde{\beta}_n(\tau) - \beta_*(\tau)) \Rightarrow Z_\beta(\tau) \sim N \left[ 0, \tau(1-\tau) \left\{ f_\tau(1 - \sum_{j=1}^{p} \phi_{j*}(\tau)) \right\}^{-2} I_k \right], \)

where \( M := n^{-2}X' \left[ I - \tilde{W} (\tilde{W}'\tilde{W})^{-1} \tilde{W}' \right] X \) and \( \tilde{W} := [\tilde{W}_1, \tilde{W}_2, \ldots, \tilde{W}_n]' \).

This implies that \( \tilde{\beta}_n(\tau) \) is consistent for \( \beta_*(\tau) \) at the convergence rate of \( n \). We highlight the key steps of our proof. First, we note the following identity:

\[ (\tilde{\beta}_n(\tau) - \beta_*(\tau)) \equiv \left[ 1 - \sum_{j=1}^{p} \tilde{\phi}_{j,n}(\tau) \right]^{-1} \left( \tilde{\gamma}_n(\tau) - \gamma_*(\tau) + \beta_*(\tau) \sum_{j=1}^{p} (\tilde{\phi}_{j,n}(\tau) - \phi_{j*}(\tau)) \right) . \]  
(11)
Using this identity, we show in the Appendix (see the derivation of (34)) that

\[
\begin{align*}
n \left\{ (\gamma_n(\tau) - \gamma_*(\tau)) + \beta_*(\tau) \sum_{j=1}^{p} (\phi_{j,n}(\tau) - \phi_j(\tau)) \right\} \\
= f^{-1}_\tau M^{-1} \left\{ n^{-1} X' \left[ I - \tilde{W} (\tilde{W}' \tilde{W})^{-1} \tilde{W}' \right] \Psi_\tau(U) \right\} + o_p(1).
\end{align*}
\]

We obtain this by showing that

\[
n \left\{ (\gamma_n(\tau) - \gamma_*(\tau)) + \beta_*(\tau) \sum_{j=1}^{p} (\phi_{j,n}(\tau) - \phi_j(\tau)) \right\}
\]

is equivalent to the last \(k\) elements of \(D_G \left\{ (\lambda_n(\tau) - \lambda_*(\tau)) + \Gamma_*(\tau) (\phi_n(\tau) - \phi_*(\tau)) \right\}\) in probability, where \(D_G := \text{diag} ([\sqrt{n} \varepsilon_{1+qk}, n \varepsilon_k])\). We derive that the latter is equal to \(f^{-1}_\tau M^{-1} \left\{ n^{-1} X' \left[ I - \tilde{W} (\tilde{W}' \tilde{W})^{-1} \tilde{W}' \right] \Psi_\tau(U) \right\}\) in probability (see the Appendix). As the right-hand side (RHS) of this equation is bounded in probability, we can apply Theorem 1(i) and obtain

\[
\begin{align*}
n \left( \tilde{\beta}_n(\tau) - \beta_*(\tau) \right) \\
= \left[ 1 - \sum_{j=1}^{p} \phi_{j,n}(\tau) \right]^{-1} \left( f_{\tau} M \right)^{-1} \left\{ n^{-1} X' \left[ I - \tilde{W} (\tilde{W}' \tilde{W})^{-1} \tilde{W}' \right] \Psi_\tau(U) \right\} + o_p(1).
\end{align*}
\]

From this, we establish the asymptotic distribution in Theorem 2(i). Next, it is easily seen that Theorem 2(ii) holds as a corollary of Theorem 2(i). By Assumption 1(iii), \(\tilde{W}\) and \(\Psi_\tau(U)\) are independent, and thus the weak limit in Theorem 2(i) can be viewed as a mixture of normal random variables. Theorem 2(ii) captures this implication such that its standardized version asymptotically follows a multivariate normal random variable.

The mixture normal distribution property of the long-run parameter estimator implies that the Wald statistic asymptotically follows a chi-squared distribution under a suitable null hypothesis of \(\beta_*(\tau)\). To investigate this property, we consider the following null and alternative hypotheses:

\[
\mathbb{H}_0^{(1)} : R_0 \beta_*(\tau) = r \quad \text{versus} \quad \mathbb{H}_1^{(1)} : R_1 \beta_*(\tau) \neq r,
\]

where \(R\) is an \(r \times k\) matrix and \(r\) is an \(r \times 1\) vector. The Wald statistic is defined as

\[
W_n(\beta) := \frac{n^2 f_{\tau}^2}{\tau(1-\tau)} \left( 1 - \sum_{j=1}^{p} \phi_{j,n}(\tau) \right)^2 \left( R_1 \tilde{\beta}_n(\tau) - r \right)' \left( RM^{-1} R' \right)^{-1} \left( R_1 \tilde{\beta}_n(\tau) - r \right),
\]
where \( \hat{f}_r \) is a consistent estimator of \( f_r \). Using the definition of \( \tilde{\beta}_n(\tau) \), the Wald statistic can be rewritten as

\[
\frac{n^2 \hat{f}_r^2}{\tau(1 - \tau)} \left[ R\tilde{\beta}_n(\tau) - \left( 1 - \sum_{j=1}^p \phi_{j,n}(\tau) \right) r \right]' (RM^{-1}R')^{-1} \left[ R\tilde{\beta}_n(\tau) - \left( 1 - \sum_{j=1}^p \phi_{j,n}(\tau) \right) r \right].
\]

Note that this Wald test statistic is a generalized version of the \( t \)-test statistic. More specifically, if we let

\[
\mathcal{T}_{m,n} := n \left( 1 - \sum_{j=1}^p \phi_{j,n}(\tau) \right) \frac{\hat{f}_r(R\hat{\beta}_n(\tau) - r_m)}{\{\tau(1 - \tau)R_mM^{-1}R_m\}^{1/2}}
\]

be the \( t \)-statistic, where \( R_m := [0'_{(m-1)\times 1}, 1, 0'_{(k-m)\times 1}] \) and \( r_m \) is a constant, then \( \mathcal{T}_{m,n}^2 \) is identical to the Wald statistic testing \( \mathbb{H}_0^{(0)} : \beta_m(\tau) = r_m \) against \( \mathbb{H}_1^{(0)} : \beta_m(\tau) \neq r_m \), where \( \beta_m(\tau) \) is the \( m \)-th element of \( \beta(\tau) \).

The following corollary provides the asymptotic behaviors of \( \mathcal{W}(\beta) \) and \( \mathcal{T}_{m,n} \) statistics.

**Corollary 1.** Suppose that Assumption 1 holds.

(i) For each \( m = 1, ..., k \),

(i.a) \( \mathcal{T}_{m,n} \overset{A}{\sim} N(0, 1) \) under \( \mathbb{H}_0^{(0)} \);

(i.b) \( \mathbb{P} (|\mathcal{T}_{m,n}| \geq c_n) \rightarrow 1 \) for any sequence \( \{c_n\} \) s.t. \( c_n = o(n) \) under \( \mathbb{H}_1^{(0)} \); and

(ii) If \( \text{rank}(R) = r \) and \( \hat{f}_r \rightarrow f_r \) in probability,

(ii.a) \( \mathcal{W}_n(\beta) \overset{A}{\sim} \chi_r^2 \) under \( \mathbb{H}_0^{(1)} \); and

(ii.b) \( \mathbb{P} (\mathcal{W}_n(\beta) \geq c'_n) \rightarrow 1 \) for any sequence \( \{c'_n\} \) s.t. \( c'_n = o(n^2) \) under \( \mathbb{H}_1^{(1)} \).

Corollary 1 is a straightforward consequence of Theorem 2(ii). In particular, \( \sum_{j=1}^p \phi_{j,n}(\tau) \) consistently estimates \( \sum_{j=1}^p \phi_{j*}(\tau) \) by Theorem 1 under Assumption 1. Thus, employing \( \tilde{\phi}_n(\tau) \) to define the test statistics does not alter their asymptotic null distributions.

We construct statistics for testing linear restrictions on the short-run parameters and consider the following null and alternative hypotheses:

\[
\mathbb{H}_0^{(2)} : Q\phi_* = q \quad \text{versus} \quad \mathbb{H}_1^{(2)} : Q\phi_* \neq q, \quad \text{and} \quad (12)
\]

\[
\mathbb{H}_0^{(3)} : R\gamma_* = r \quad \text{versus} \quad \mathbb{H}_1^{(3)} : R\gamma_* \neq r, \quad \text{and} \quad (13)
\]

where \( Q \) is an \( r \times p \) matrix and \( q \) is an \( r \times 1 \) vector. The asymptotic null distributions of the Wald statistics testing the hypotheses in (12) and (13) are regular by Theorem 1. Formally, we define the following Wald
test statistic:

\[ W_n(\phi) := n \left( Q\hat{\phi}_n(\tau) - q \right)' \left( Q\hat{\Pi}_n(\tau)Q' \right)^{-1} \left( Q\hat{\phi}_n(\tau) - q \right). \]

Here, \( \hat{\Pi}_n(\tau) \) is any consistent estimator for \( \Pi(\tau) := \tau(1 - \tau) f_{\tau}^{-2} E \left[ H_i H_i' \right]^{-1} \). By definition, \( \Pi(\tau) \) can be estimated in a number of ways. For example, given a consistent estimator \( \hat{f}_\tau \) for \( f_\tau \), we let:

\[ \hat{\Pi}_n(\tau) := \hat{f}_\tau^{-1}(1 - \tau) \left\{ \left( n^{-1} \hat{K}(\tau)'\hat{K}(\tau) \right) - \left( n^{-1} \hat{K}(\tau)'\hat{W} \right) \left( n^{-1} \hat{W}'\hat{W} \right)^{-1} \left( n^{-1} \hat{W}'\hat{K}(\tau) \right) \right\}, \quad (14) \]

where \( \hat{K}(\tau) := [\hat{K}_1, \hat{K}_2, \ldots, \hat{K}_n]' \), \( \hat{K}_i(\tau) := \hat{K}_{t,1}(\tau), \hat{K}_{t,2}(\tau), \ldots, \hat{K}_{t,p}(\tau) \)', and for each \( i \), \( \hat{K}_{t,i}(\tau) \) is the quantile regression error obtained from the regression of \( \hat{W}_t \) on \( Y_{t,i} - X_i'\hat{\beta}_n(\tau) \). By (6), for each \( t \), \( \hat{K}_i(\tau) \) is a consistent quantile regression error for \( K_i(\tau) \), and this leads to \( \hat{\Pi}_n(\tau) \rightarrow \Pi(\tau) \). In addition, many other estimators can be similarly defined.

The following corollary provides the asymptotic behaviors of the Wald statistic \( W(\phi) \).

**Corollary 2.** Suppose that Assumption 1 holds and \( \hat{\Pi}_n(\tau) \rightarrow \Pi(\tau) \) in probability. If \( \text{rank}(Q) = r \),

(i) \( W_n(\phi) \stackrel{d}{\sim} \chi_r^2 \) under \( \mathbb{H}_0^{(2)} \); and

(ii) \( \mathbb{P} \left( W_n(\phi) \geq c''_n \right) \rightarrow 1 \) for any sequence \( \{c''_n\} \) s.t. \( c''_n = o(n) \) under \( \mathbb{H}_1^{(2)} \).

Given the conditions in Corollary 2, it is straightforward to prove these consequences using Theorem 1. We, therefore, omit the proof.

We now examine the asymptotic behaviors of the Wald statistic testing \( \mathbb{H}_0^{(3)} \) versus \( \mathbb{H}_1^{(3)} \) under the null and alternative hypotheses in (13). The asymptotic distribution of \( \sqrt{n}(\hat{\gamma}_n(\tau) - \gamma_*(\tau)) \) is degenerate because its asymptotic covariance \( \beta_* \tau \mu_p' \Pi(\tau) \mu_p / \beta_* (\tau)' \) has a rank equal to unity by Theorem 1(ii). This implies that, if \( R \) has a rank greater than unity, the corresponding Wald test has a degenerate asymptotic null distribution. To prevent this feature, we let \( \text{rank}(R) = 1 \) and define the following Wald test statistic:

\[ W_n(\gamma) := n \left( R\tilde{\gamma}_n(\tau) - r \right)' \left( R\tilde{\beta}_n(\tau)\mu_p' \tilde{\Pi}_n(\tau)\mu_p \tilde{\beta}_n(\tau)' R' \right)^{-1} \left( R\tilde{\gamma}_n(\tau) - r \right). \]

We estimate the asymptotic covariance matrix by estimating \( \beta_*(\tau) \) and \( \Pi(\tau) \) through \( \tilde{\beta}_n(\tau) \) and \( \tilde{\Pi}_n(\tau) \), respectively. The regular asymptotic behaviors of this Wald test statistic can be summarised as follows:

**Corollary 3.** Suppose that Assumption 1 holds and \( \tilde{\Pi}_n(\tau) \rightarrow \Pi(\tau) \) in probability. If \( \text{rank}(R) = 1 \),

(i) \( W_n(\gamma) \stackrel{d}{\sim} \chi_1^2 \) under \( \mathbb{H}_0^{(3)} \); and

(ii) \( \mathbb{P} \left( W_n(\gamma) \geq c'''_n \right) \rightarrow 1 \) for any sequence \( \{c'''_n\} \) s.t. \( c'''_n = o(n) \) under \( \mathbb{H}_1^{(3)} \).
Corollary 3 is trivially implied by Theorem 1(ii). Due to the distributional degeneracy of \( \sqrt{n}(\gamma_n(\tau) - \gamma_s(\tau)) \), it is difficult to apply \( W_n(\gamma) \) if the rank of \( R \) is greater than 1. Nevertheless, this can be easily resolved if \( R \) and \( r \) are designed to test the relative ratio property between the elements in \( \gamma_s(\tau) \). As \( \gamma_s(\tau) \equiv \{1 - \sum_{j=1}^p \phi_j^s(\tau)\} \beta_s(\tau) \) by definition, a relative ratio property can be easily translated into that between the elements in \( \beta_s(\tau) \). Hence, we can test the same hypothesis using \( W_n(\beta) \).

3 Inference with Multiple Quantiles

In the context of quantile regression, we typically estimate the model using multiple quantile indices. In such a case, one may wish to test the null hypothesis that the short-run or the long-run parameters at a low quantile (say, \( \tau = 0.1 \)) are the same as those at an upper quantile (\( \tau = 0.9 \)). Any evidence of disparity can be regarded as an indication of an asymmetric behavior in the short-run or the long-run parameters associated with the distributional location of the dependent variable.

We develop a systematic inference procedure for testing hypotheses constructed by multiple quantile indices. Consider an \( s \) number of quantile indices, say \( \tau_1 < \ldots < \tau_s \), and let their corresponding multi-quantile short-run and long-run parameters be denoted as \( [\Phi_s(\tau)', \Gamma_s(\tau)']' := [\phi_s(\tau_1)', \ldots, \phi_s(\tau_s)'; \gamma_s(\tau_1)', \ldots, \gamma_s(\tau_s)']' \) and \( B_s(\tau) := [\beta_s(\tau_1)', \ldots, \beta_s(\tau_s)']' \), respectively. We wish to test the validity of linear restrictions on \( \Phi_s(\tau), \Gamma_s(\tau), \) and \( B_s(\tau) \) and express such hypotheses as

\[
\mathbb{H}_0' : F\Phi_s(\tau) = f \quad \text{versus} \quad \mathbb{H}_1' : F\Phi_s(\tau) \neq f, \tag{15}
\]

\[
\mathbb{H}_0'' : St\Gamma_s(\tau) = s \quad \text{versus} \quad \mathbb{H}_1'' : St\Gamma_s(\tau) \neq s, \tag{16}
\]

\[
\mathbb{H}_0''' : SB_s(\tau) = s \quad \text{versus} \quad \mathbb{H}_1''' : SB_s(\tau) \neq s, \tag{17}
\]

where \( F \) and \( f \) are \( h \times ps \) and \( h \times 1 \) pre-specified matrices, respectively, and \( S \) and \( s \) are \( h \times ks \) and \( h \times 1 \) pre-specified matrices with \( h \) being the number of restrictions. As a benchmark case, we test whether the short-run and long-run parameters are equal at two quantile indices. For example, if we intend to test \( \beta_s(\tau_1) = \beta_s(\tau_2) \), we let \( S = (I_k, -I_k), s = 0_{k \times 1} \) and \( B_s(\tau) := [\beta_s(\tau_1)', \beta_s(\tau_2)']' \).

We impose the following assumptions in addition to Assumption 1 such that the test statistics defined below behave appropriately for moderately large sample sizes.

**Assumption 2.** (i) \( \Xi(\tau) := [(f_{\tau}, f_{\tau})]^{-1} \{(\min[\tau_1, \tau_2] - \tau_i\tau_j) L(\tau_1, \tau_1) L(\tau_1, \tau_2) L(\tau_2, \tau_j) L(\tau_2, \tau_2)^{-1}\} s, j = 1, 2, \ldots, s \) is a \( ps \times ps \) positive-definite matrix, where \( s \in \mathbb{N} \) and \( L(\tau_1, \tau_j) := \mathbb{E} \left[ H_{\ell}(\tau_1) \overline{H}_{\ell}(\tau_j) \right] \).
(ii) For each \( r, \tau \in (0, 1) \), we let \( G_n (r, \tau) := n^{-1/2} \sum_{t=1}^{[nr]} \psi_T [U_t(\tau)] [\tilde{W}_t', \tilde{K}_t(\tau)']' \) and suppose that \( G_n (\cdot, \cdot) \Rightarrow G (\cdot, \cdot) \), where \( G (\cdot, \cdot) := [B_\psi (\cdot, \cdot), B_{\psi, \cdot} (\cdot, \cdot)', B_{\psi, \cdot} (\cdot, \cdot)', B_{\psi, \cdot} (\cdot, \cdot)']' \) is a multivariate Gaussian process with \( \mathbb{E} [G (\cdot, \cdot)] = 0 \), and for each \( r, \tau, \tilde{\tau} \in (0, 1) \), \( \mathbb{E} [G (r, \tau)G (\tilde{r}, \tilde{\tau})'] = (\min[r, \tilde{r}] - \tau \tilde{\tau}) \Pi (r, \tilde{r}, \tau, \tilde{\tau}) \) with \( \Pi (r, \tilde{r}, \tau, \tilde{\tau}) := \lim_{n \to \infty} \mathbb{E} [C_n (r; \tau)C_n (\tilde{r}; \tilde{\tau})] \) and \( C_n (r; \tau) := n^{-1/2} \sum_{t=1}^{[nr]} [\tilde{W}_t', \tilde{K}_t(\tau)']' \); and

(iii) \( \Sigma (\tau) := T (\tau) \circ P (\tau) \) is positive definite, where we let \( T (\tau) := [\min [\tau_i, \tau_j] - \tau_i \tau_j]_{i,j=1,2,\ldots,s} \) and \( P (\tau) := \left[ (f_{\tau_i} (1 - \sum_{t=1}^{P} \phi_{\ell_t} (i_t)))^{-1} (f_{\tau_j} (1 - \sum_{t=1}^{P} \phi_{\ell_t} (j_t)))^{-1} \right]_{i,j=1,2,\ldots,s} \).

There are several remarks. First, \( L (\cdot, \cdot) \) is introduced for notational simplicity. Second, if \( K_t (\cdot) \) is invariant to the level of \( \tau \), it is written as \( K_t \), and \( \Xi (\tau) \) is simplified to \( \Xi (\tau) = \left[ (f_{\tau_i} f_{\tau_j})^{-1} (\min [\tau_i, \tau_j] - \tau_i \tau_j) \right]_{i,j=1,2,\ldots,s} \otimes \mathbb{E} [K_t K_t'] - \mathbb{E} [K_t \tilde{W}_t] \mathbb{E} [\tilde{W}_t \tilde{W}_t']^{-1} \mathbb{E} [\tilde{W}_t K_t']^{-1} \). Assumption 1(vi) implies that \( \Xi (\tau) \) is positive definite in which case Assumption 2(i) is redundant. Third, it is possible to derive Assumption 2(ii) from Assumption 1 by imposing other primitive conditions, but we still introduce it for brevity. For \( \tau \in (0, 1) \), \( G (\cdot, \cdot) \overset{d}{=} [B_\psi (\cdot, \cdot), B_{\psi, \cdot} (\cdot, \cdot)', B_{\psi, \cdot} (\cdot, \cdot)', B_{\psi, \cdot} (\cdot, \cdot)']' \) has a positive-definite covariance matrix by Assumption 1(vi). It is clear that the multivariate Gaussian process \( [G (\cdot, \tau_1)', G (\cdot, \tau_2)', \ldots, G (\cdot, \tau_s)']' \) is well defined under Assumption 2. Finally, Assumption 2(iii) is imposed to derive a non-degenerate asymptotic distribution of the multi-quantile long-run estimator.

By Theorem 1, \( \hat{\Phi}_n (\tau) \) and \( \hat{\Gamma}_n (\tau) \) consistently estimate \( \Phi_* (\tau) \) and \( \Gamma_* (\tau) \), respectively. Hence, we first establish the asymptotic joint distribution of these short-run estimators by extending Theorem 1 to the multi-quantile framework and develop statistics for testing the hypotheses in (15) and (16).

**Theorem 3 (Multi-Quantile Short-Run Estimators).** Under Assumptions 1 and 2,

(i) \( \sqrt{n} \left( \hat{\Phi}_n (\tau) - \Phi_* (\tau) \right) \Rightarrow \mathcal{N} \left( 0_{ps \times 1}, \Xi (\tau) \right); \) and

(ii) \( \sqrt{n} \left( \hat{\Gamma}_n (\tau) - \Gamma_* (\tau) \right) \Rightarrow \mathcal{N} \left( 0_{ks \times 1}, \Lambda (\tau) \Xi (\tau) \Lambda (\tau)' \right), \) where \( \Lambda (\tau) \) is a \( ks \times ps \) block diagonal matrix with \( s \) diagonal blocks \( \beta_{\tau_i} (\tau_i) \ell_{\tau_i}^p \) for \( i = 1, 2, \ldots, s. \)

The key outlines for proving Theorem 3 are summarised as follows. We establish Theorem 3(i) by noting that

\[
\sqrt{n} \left( \hat{\phi}_n (\tau_j) - \phi_* (\tau_j) \right) = f_{\tau_j}^{-1} L (\tau_j, \tau_j)^{-1} \left( n^{-1/2} \sum_{t=1}^{n} \tilde{H}_t (\tau_j) \psi_T [U_t(\tau_j)] \right) + o_p (1). \tag{18}
\]

This equality follows from (8). As \( \Xi (\tau) \) is positive definite by Assumption 2(i), it is straightforward to
apply the multivariate CLT and conclude Theorem 3(i). Next, we prove Theorem 3(ii) by noting that

\[ \sqrt{n} \left( \tilde{\Gamma}_n(\tau) - \Gamma_*(\tau) \right) = -\sqrt{n} \Lambda(\tau) \left( \tilde{\Phi}_n(\tau) - \Phi_*(\tau) \right) + o_p(1), \]

which is obtained by extending (10) to the multi-quantile version. Using Theorem 3(i), we easily derive the limit distribution of \( \sqrt{n} \left( \tilde{\Gamma}_n(\tau) - \Gamma_*(\tau) \right) \). Here, the covariance matrix of \( Z_\Gamma(\tau) \) is not necessarily positive definite. Although \( \Xi(\tau) \) is positive definite, rank \( \Lambda(\tau) \Xi(\tau) \Lambda(\tau)' \) is at most \( s \). Note that

\[ \text{rank}[\Lambda(\tau) \Xi(\tau) \Lambda(\tau)'] \leq \min \{ \text{rank}(\Lambda(\tau)), \text{rank}(\Xi(\tau)) \}, \]

where \( \text{rank}(\Lambda(\tau)) = s \) and \( \text{rank}(\Xi(\tau)) = ps \), so that the highest rank of \( \Lambda(\tau) \Xi(\tau) \Lambda(\tau)' \) is \( s \). Theorem 3 is a multi-quantile generalisation of Theorem 1 and establishes that the multi-quantile short-run parameter estimators asymptotically follow the multivariate normal distribution. In the special case in which \( s = 1 \), the consequence of Theorem 3 is identical to that of Theorem 1.

We provide the asymptotic distribution of the multi-quantile long-run estimator in the following theorem, which generalises Theorem 2 to the case in which \( s > 1 \).

**Theorem 4** (Multi-Quantile Long-Run Estimator). **Under Assumptions 1 and 2,**

(i) \( n \left( \tilde{\mathbf{B}}_n(\tau) - \mathbf{B}_*(\tau) \right) \Rightarrow \left[ \mathbf{I}_s \otimes \left( \int_0^1 \tilde{\mathbf{B}}_W(r) \tilde{\mathbf{B}}_W(r)' dr \right)^{-1} \right] \mathbf{J}_\beta(\tau) \), where

\[ \mathbf{J}_\beta(\tau) := \left[ f_i^{-1} \left( 1 - \sum_{j=1}^p \phi_{js}(\tau_i) \right)^{-1} \int_0^1 \tilde{\mathbf{B}}_W(r)' d\mathbf{B}_W(r, \tau_i) \right]_{i=1,...,s} \]; and

(ii) \( n \left( \mathbf{I}_s \otimes \mathbf{M}^{1/2} \right) \left( \tilde{\mathbf{B}}_n(\tau) - \mathbf{B}_*(\tau) \right) \overset{d}{\sim} N \left[ \mathbf{0}, \Sigma(\tau) \otimes \mathbf{I}_{k \times k} \right]. \)

Theorem 4 trivially holds by Assumption 2. Here, the asymptotic covariance matrix \( \Sigma(\tau) \otimes \mathbf{I}_{k \times k} \) is positive definite, given that \( \Sigma(\tau) \) is positive definite by Assumption 2(iii). The asymptotic distribution of \( n \left( \mathbf{I}_s \otimes \mathbf{M}^{1/2} \right) \left( \tilde{\mathbf{B}}_n(\tau) - \mathbf{B}_*(\tau) \right) \) is not necessarily degenerate, although \( \sqrt{n} \left( \tilde{\Gamma}_n(\tau) - \Gamma_*(\tau) \right) \) may converge to a degenerate distribution for \( k < p \). This non-degeneracy is explicitly imposed by Assumption 2(iii).

It is straightforward to extend the single-quantile Wald statistics into the multi-quantile counterparts by using the asymptotic mixed normal distribution in Theorems 3 and 4. We consider the following test statistics:

\[ W_n(\Phi) := n \left( \mathbf{F} \bar{\Phi}_n(\tau) - f \right)' \left( \mathbf{F} \bar{\Xi}_n(\tau) \mathbf{F}' \right)^{-1} \left( \mathbf{F} \bar{\Phi}_n(\tau) - f \right), \]
to test the joint hypotheses in (15), (16), and (17), respectively, where for each \(i,j\), \(\Lambda\) is a consistent estimator for \(f\), and an additional condition, rank

\[
\Sigma(n) := \left[
\begin{array}{ccc}
\Sigma_{i,j}(\tau_i) & \cdots & \Sigma_{i,s}(\tau_i) \\
\vdots & \ddots & \vdots \\
\Sigma_{j,i}(\tau_j) & \cdots & \Sigma_{j,s}(\tau_j)
\end{array}
\right]
\]

Suppose that Assumptions 1 and 2 hold, and for Corollary 4.

We summarise the null and alternative asymptotic behaviors of the multi-quantile Wald test statistics \(\mathcal{W}_n(\Phi)\), \(\mathcal{W}_n(\Gamma)\), and \(\mathcal{W}_n(B)\) in the following corollary.

**Corollary 4.** Suppose that Assumptions 1 and 2 hold, and for \(j = 1, 2, \ldots, s\), \(f_{\tau_j} \to f_{\tau_j}\) in probability.

(i) If rank\((\Phi) = h\), and for each \(\tau_i\) and \(\tau_j\), \(\hat{\Lambda}(\tau_i, \tau_j) \to \Lambda(\tau_i, \tau_j)\) in probability, then

\[
\mathcal{W}_n(\Phi) \overset{d}{\sim} \chi^2_h \text{ under } H_0; \text{ and}
\]

\[
(i.a) \text{ } P(\mathcal{W}_n(\Phi) \geq d''_n) \to 1 \text{ for any sequence } \{d'_n\} \text{ s.t. } d''_n = o(n) \text{ under } H_1;
\]

(ii) If rank\((S\Lambda(\tau) \Xi(\tau)\Lambda(\tau)' S') = h \leq s\), and for each \(\tau_i\) and \(\tau_j\), \(\hat{\Lambda}(\tau_i, \tau_j) \to \Lambda(\tau_i, \tau_j)\) in probability, then

\[
\mathcal{W}_n(\Gamma) \overset{d}{\sim} \chi^2_h \text{ under } H_0; \text{ and}
\]

\[
(ii.a) \text{ } P(\mathcal{W}_n(\Gamma) \geq d''_n) \to 1 \text{ for any sequence } \{d'_n\} \text{ s.t. } d''_n = o(n) \text{ under } H_1; \text{ and}
\]

(iii) If rank\((S) = h\), then

\[
\mathcal{W}_n(B) \overset{d}{\sim} \chi^2_h \text{ under } H_0; \text{ and}
\]

\[
(iii.a) \text{ } P(\mathcal{W}_n(B) \geq d''_n) \to 1 \text{ for any sequence } \{d'_n\} \text{ s.t. } d''_n = o(n^2) \text{ under } H_1.
\]

As Corollary 4 trivially follows Theorems 3 and 4, its proof is omitted. Notice that Corollary 4 imposes an additional condition, rank\([S\Lambda(\tau) \Xi(\tau)\Lambda(\tau)' S'] \leq s\), as rank\([\Lambda(\tau) \Xi(\tau)\Lambda(\tau)'\] is at most \(s\).
4 Monte Carlo Simulations

In this section, we conduct Monte Carlo experiments to investigate the finite sample performances of our proposed estimators and test statistics. We separate our simulations into two parts. First, we aim to verify the theoretical claims developed in Section 3 by assuming that the QARDL orders are known and that the regularity conditions are satisfied. Second, we also consider the case in which the QARDL orders are unknown, and they are estimated by the information criterion.

4.1 Simulations with Known QARDL Orders

In this subsection, we focus on the validity of the joint testing procedure under the multiple quantiles framework as derived in Corollary 4. To this end, we consider the following QARDL(1,1) process:

\[
Y_t = \alpha_s(\tau) + \phi_s(\tau)Y_{t-1} + \theta_{0s}(\tau)X_t + \theta_{1s}(\tau)X_{t-1} + U_t(\tau) \quad \text{and} \quad X_t = X_{t-1} + W_t,
\]

(19)

where \(U_t(\tau) := U_t - F^{-1}(\tau), U_t \sim \text{IID } N(0,1), W_t := \rho_s R_{t-1} + (1 - \rho_s^2) R_t \) with \(R_t \sim \text{IID } N(0,1)\), and \(U_t\) is generated independently of \(R_t\). We consider three different quantile indices, \(\tau_1 = 0.25, \tau_2 = 0.50,\) and \(\tau_3 = 0.75\). By construction, the true values of \(\phi_s(\tau), \theta_{0s}(\tau),\) and \(\theta_{1s}(\tau)\) are set to be the same for each \(\tau\) although \(\alpha_s(\tau)\) can vary with \(\tau\).

We examine the finite sample performance of the Wald statistics using the QARDL(1,1) model. We consider the three Wald test statistics, denoted \(W_n(B), W_n(\Phi),\) and \(W_n(\Gamma)\). First, we test the following four hypotheses on the long-run parameters \(B_s(\tau) = [\beta_s(0.25), \beta_s(0.5), \beta_s(0.75)]'\): for \(j = 1, 2, 3,\) and \(4,\)

\[
H_0^{(j)}(B) : S_j B_s(\tau) = s_j \quad \text{versus} \quad H_1^{(j)}(B) : S_j B_s(\tau) \neq s_j,
\]

(20)

where \(S_1 = (1, -1, 0), S_2 = (0, 1, -1), S_3 = (1, 0, -1)'\) and \(S_4 = (S_1', S_2')'\) with \(s_1 = s_2 = s_3 = 0,\) and \(s_4 = [0, 0]'\). We denote these Wald statistics respectively as \(W_n^{(1)}(B), W_n^{(2)}(B), W_n^{(3)}(B),\) and \(W_n^{(4)}(B).\)

Second, we test the four restrictions on the short-run parameters, \(\Phi_s(\tau) = [\phi_s(0.25), \phi_s(0.5), \phi_s(0.75)]'\) and \(\Gamma_s(\tau) = [\gamma_s(0.25), \gamma_s(0.5), \gamma_s(0.75)]'\): for \(j = 1, 2, 3,\) and \(4,\)

\[
H_0^{(j)}(\Phi) : S_j \Phi_s(\tau) = s_j \quad \text{versus} \quad H_1^{(j)}(\Phi) : S_j \Phi_s(\tau) \neq s_j, \quad \text{and}
\]

(21)

\[
H_0^{(j)}(\Gamma) : S_j \Gamma_s(\tau) = s_j \quad \text{versus} \quad H_1^{(j)}(\Gamma) : S_j \Gamma_s(\tau) \neq s_j.
\]

(22)

For each \(j = 1, 2, 3,\) and \(4,\) these Wald statistics are denoted respectively as \(W_n^{(j)}(\Phi)\) and \(W_n^{(j)}(\Gamma).\)
To estimate the relevant statistics for testing the hypotheses in (20), (21), and (22), we need to consistently estimate the density function. For this purpose, we employ the following kernel density estimator:

$$
\hat{f}_{\tau_\ell} := \frac{1}{nh_B(\tau_\ell)} \sum_{t=1}^{n} \phi \left( \frac{\tilde{U}_t(\tau_\ell)}{h_B(\tau_\ell)} \right), \quad \ell = 1, 2, 3,
$$

where $h_B$ is the bandwidth proposed by Bofinger (1975).

$$
h_B(\tau_\ell) := n^{-1/5} \left[ 4.5\phi^{-1}(\Phi^{-1}(\tau_\ell))^{4/3} \left( (2\Phi^{-1}(\tau_\ell))^2 + 1 \right) \right]^{1/5},
$$

and $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal PDF and CDF, respectively. It is well established that $\hat{f}_{\tau_\ell}$ is a consistent estimator of $f_{\tau_\ell}$, e.g., Koenker and Xiao (2002).5 Next, we need to consistently estimate $\Pi(\tau)$ for evaluating $W_n^{(j)}(\Phi)$ and $W_n^{(j)}(\Gamma)$. We estimate (6) by the quantile regression, and compute $\{\hat{K}_{t,i}(\tau)\}$ as the quantile regression error. Using the sequence of $\{W_t\}$, we estimate $\Pi(\tau)$ by (14).

We summarise the null simulation setup as follows: First, we set the parameter values at $\alpha_* = 1$, $\phi_* = 0.25$, $\theta_{0*} = 2$, $\theta_{1*} = 3$, and $\rho_* = 0.5$, and examine the empirical rejection rates of the Wald tests under the null hypotheses. By construction, $\gamma_* = 5$ and $\beta_* = 20/3$. We set the number of replications at 5,000 and report the empirical rejection rates of $W_n^{(j)}(B)$, $W_n^{(j)}(\Phi)$, and $W_n^{(j)}(\Gamma)$ for six different sample sizes, $n = 50, 200, 400, 600, 800, 1,000, and 2,000$.

We now consider the finite sample testing performance for the long-run parameter. Table 1 presents the empirical rejection rates of $W_n^{(j)}(B)$ under the null hypothesis $H_n^{(j)}(B)$ at three different nominal levels, 1%, 5%, and 10%. In all four cases, the Wald statistics tend to moderately over-reject the null when the sample size is relatively small. Especially for $n = 50$, the size distortions seem to be non-negligible, but this is not a surprising finding given the common empirical findings in the stationary case that requires a sufficiently large sample size for an accurate estimation of the quantile regression. As the sample size increases, as expected, empirical levels of all four Wald statistics tend to converge to the nominal levels.

These findings generally support the theoretical predictions made in Corollary 4(iii.a) that the Wald statistics $W_n^{(j)}(B)$ and $W_n^{(4)}(B)$ converge to $\chi_1^2$ and $\chi_2^2$ asymptotically under $\mathbb{H}_0^{(j)}(B)$ and $\mathbb{H}_0^{(4)}(B)$, respectively ($j = 1, 2, 3$). Furthermore, we draw the empirical PDFs (EPDFs) and empirical CDFs (ECDFs) of the four

5 Alternatively, the bandwidth advocated by Hall and Sheather (1988) can be used to consistently estimate the density. However, this method requires the selection of a significance level of the test such that the simulations will depend upon this particular choice. This implies that we should conduct identical Monte Carlo experiments at different significance levels of the test. Hence, we use the simpler Bofinger’s (1975) approach.
Wald statistics $\mathcal{W}_n^{(j)}(B)$ ($j = 1, 2, 3, 4$) using the sample size $n = 5,000$. The first panels of Figures 1 and 2 show the empirical probability functions of $\mathcal{W}_n^{(j)}(B)$ for $j = 1, 2, 3$ and those of $\mathcal{W}_n^{(4)}(B)$, respectively. It is clear from these two figures that the ECDFs and EPDFs are almost identical to the asymptotic CDF and PDF of $\mathcal{X}_1^2$ and $\mathcal{X}_2^2$ variables, respectively. This confirms the theoretical predictions of Corollary 4(iii.a).

Insert Figures 1 and 2 around here. >>>>>>>>>>>

Second, we turn to testing the performance of the short-run parameters under the null. Tables 2 and 3 present the empirical rejection rates of $\mathcal{W}_n^{(j)}(\Phi)$ and $\mathcal{W}_n^{(j)}(\Gamma)$ under the null hypotheses $H_0^{(j)}(\Phi)$ and $H_0^{(j)}(\Gamma)$ at 1%, 5%, and 10%, respectively. These statistics tend to moderately over-reject the null when the sample size is relatively small. Nevertheless, as expected, empirical levels of all four Wald statistics tend to converge to the nominal levels as the sample size increases.

Insert Tables 2 and 3 around here. >>>>>>>>>>>

These aspects are as claimed in Corollary 4(i.a, ii.a): (i) $\mathcal{W}_n^{(j)}(\Phi)$ and $\mathcal{W}_n^{(4)}(\Phi)$ converge to $\mathcal{X}_1^2$ and $\mathcal{X}_2^2$ asymptotically under $H_0^{(j)}(\Phi)$ and $H_0^{(4)}(\Phi)$, respectively; (ii) $\mathcal{W}_n^{(j)}(\Gamma)$ and $\mathcal{W}_n^{(4)}(\Gamma)$ converge to $\mathcal{X}_1^2$ and $\mathcal{X}_2^2$ asymptotically under $H_0^{(j)}(\Gamma)$ and $H_0^{(4)}(\Gamma)$, $(j = 1, 2, 3)$. From the second and third panels of Figures 1 and 2, the ECDFs and EPDFs of $\mathcal{W}_n^{(j)}(\Phi)$ and $\mathcal{W}_n^{(j)}(\Gamma)$ for $j = 1, 2, 3$ and those of $\mathcal{W}_n^{(4)}(\Phi)$ and $\mathcal{W}_n^{(4)}(\Gamma)$ are almost identical to the asymptotic counterparts, affirming Corollaries 4(i.a, ii.a).

Next, we report the power performance of the Wald test statistics under the alternative. For this purpose, we let $s_j = 0.1$ ($j = 1, 2, 3$) and $s_4 = [0.1, 0.1]'$ and consider the sample sizes $\{50, 100, 200, 300, 400, 500\}$ for $\mathcal{W}_n^{(j)}(B)$ and $\mathcal{W}_n^{(j)}(\Phi)$ and $\{200, 400, 600, 800, 1200, 1600, 2000\}$ for $\mathcal{W}_n(\Gamma)$.

First, Table 4 reports the empirical rejection rates of the $\mathcal{W}_n^{(j)}(B)$ tests for the long-run parameter under the four alternative hypotheses $H_1^{(j)}(B)$ ($j = 1, 2, 3$) and $H_1^{(4)}(B)$, from which we find that the powers increase quickly with the sample size. This is a satisfactory finding given that the alternative hypotheses we consider are not much different from the null. Overall, this evidence provides strong support for the theoretical predictions in Corollary 4(iii.b) that $\mathcal{W}_n^{(j)}(B)$ is consistent.

Insert Table 4 around here. >>>>>>>>>>>

Second, we present the empirical rejection rates of $\mathcal{W}_n^{(j)}(\Phi)$ and $\mathcal{W}_n^{(j)}(\Gamma)$ for the short-run parameters under the alternative hypotheses, respectively, in Tables 5 and 6. The power of $\mathcal{W}_n^{(j)}(\Phi)$ approaches unity very quickly, faster than that of $\mathcal{W}_n^{(j)}(B)$. Surprisingly, however, we observe that the power of $\mathcal{W}_n^{(j)}(\Gamma)$ approaches unity at a much slower rate than those of $\mathcal{W}_n^{(j)}(B)$ or $\mathcal{W}_n^{(j)}(\Phi)$. 

21
Two remarks are in order. First, the finite sample performance of the Wald tests depends on the distance between quantiles. Remind that the distance between quantile indices for $\mathbb{H}_0^{(1)}(\cdot)$ and $\mathbb{H}_0^{(2)}(\cdot)$ is 0.25 (i.e., $|\tau_1 - \tau_2| = |\tau_2 - \tau_3| = 0.25$), whereas the distance is 0.5 under $\mathbb{H}_0^{(3)}(\cdot)$ ($|\tau_1 - \tau_3| = 0.5$). The level distortions of $W_n^{(j)}(B)$, $W_n^{(j)}(\Phi)$, and $W_n^{(j)}(\Gamma)$ are smaller for $\mathbb{H}_0^{(1)}(\cdot)$ and $\mathbb{H}_0^{(2)}(\cdot)$ than for $\mathbb{H}_0^{(3)}(\cdot)$. On the other hand, the empirical powers of $W_n^{(j)}(B)$, $W_n^{(j)}(\Phi)$, and $W_n^{(j)}(\Gamma)$ are higher against $\mathbb{H}_0^{(1)}(\cdot)$ and $\mathbb{H}_0^{(2)}(\cdot)$ than against $\mathbb{H}_0^{(3)}(\cdot)$. Second, it is easily seen from Tables 4, 5, and 6 that the empirical powers of $W_n^{(4)}(B)$, $W_n^{(4)}(\Phi)$, and $W_n^{(4)}(\Gamma)$ are higher than those of $W_n^{(j)}(B)$, $W_n^{(j)}(\Phi)$, and $W_n^{(j)}(\Gamma)$ ($j = 1, 2, 3$). Nevertheless, this result should be somewhat discounted because the level distortions of $W_n^{(4)}(B)$, $W_n^{(4)}(\Phi)$, and $W_n^{(4)}(\Gamma)$ are also higher than those of $W_n^{(j)}(B)$, $W_n^{(j)}(\Phi)$, and $W_n^{(j)}(\Gamma)$ ($j = 1, 2, 3$), as we can see from Tables 1, 2, and 3, respectively. This implies that testing hypotheses with a smaller quantile interval tends to deliver more precise inference results, especially when the sample size is relatively small. Alternatively, sufficiently large samples are required for reliable inference when the quantile interval is relatively wide and/or multiple restrictions are tested.

4.2 Simulations with Unknown QARDL Orders

We extend the simulation set-up and consider the case with unknown QARDL orders. To this end, we construct the GDP as follows:

$$Y_t = \alpha_s(\tau) + \phi_s(\tau)Y_{t-1} + \theta_0(\tau)X_t + \theta_1(\tau)X_{t-1} + U_t(\tau) \quad \text{and} \quad X_t = X_{t-1} + W_t, \quad (23)$$

where $U_t(\tau) := U_t - F^{-1}(\tau)$, $U_t := \sigma_uR_{t-1}^{(1)} + (1 - \sigma_u^2)R_t^{(1)}$, and $W_t := \rho_uR_{t-1}^{(2)} + (1 - \rho_u^2)R_t^{(2)}$ with $(R_{t-1}^{(1)}, R_t^{(2)})' \sim \text{IID } N(0, I_2)$. Notice that the only difference of the DGP condition in (23) from that of (19) is that we now allow $\{U_t\}$ to be an MA(1) process with $\sigma_u \in (0, 1)$. If $\sigma_u = 0$, then the DGP condition is identical to that in (19). Otherwise, the regularity conditions in Assumptions 1 and 2 may not hold in general. In particular, we are interested in examining how the Wald test statistics perform in the presence of the MA error terms.\(^6\)

In this case, we should estimate the unknown QARDL orders. For this purpose, we apply the Bayesian information criteria (BIC). Specifically, we consider 49 models by letting $(p, q) = \{(1, 1), (1, 2), \ldots, (7, 6), (7, 7)\}$ and select the best fitting model with the smallest number of BIC. Note that the conditional mean equation is the weighted average of the quantile equations and BIC can consistently estimate the ARDL

\(^6\)We appreciate an anonymous referee’s suggestion for this additional simulation.
orders, as also confirmed by Pesaran and Shin (1998). This implies that the QARDL orders selected by BIC
may be consistently higher for some quantiles than for other quantiles.

Our simulations are conducted across three quantile levels, \( \tau_1 = 0.25, \tau_2 = 0.50 \) and \( \tau_3 = 0.75 \)
as follows: For the long-run parameters, \( \mathbf{B}_s(\tau) = [\beta_s(0.25), \beta_s(0.5), \beta_s(0.75)]' \) and the short-run parameters, \( \mathbf{\Gamma}_s(\tau) := [\gamma_s(0.25), \gamma_s(0.5), \gamma_s(0.75)]' \), we use \( \mathbf{W}_{n}^{(j)}(\mathbf{B}) \) and \( \mathbf{W}_{n}^{(j)}(\mathbf{\Gamma}) \) to test the hypotheses:

\[
\mathbb{H}_0^{(j)}(\mathbf{B}) : \mathbf{S}_j \mathbf{B}_s(\tau) = s_j \text{ versus } \mathbb{H}_1^{(j)}(\mathbf{B}) : \mathbf{S}_j \mathbf{B}_s(\tau) \neq s_j \quad \text{and} \quad \mathbb{H}_0^{(j)}(\mathbf{\Gamma}) : \mathbf{S}_j \mathbf{\Gamma}_s(\tau) = s_j \text{ versus } \mathbb{H}_1^{(j)}(\mathbf{\Gamma}) : \mathbf{S}_j \mathbf{\Gamma}_s(\tau) \neq s_j \quad \text{for } j = 1, 2, 3, 4.
\]

Next, for \( \Phi_s(\tau) := [\phi_{1s}(0.25), \ldots, \phi_{ps}(0.25), \phi_{1s}(0.50), \ldots, \phi_{ps}(0.50), \phi_{1s}(0.75), \ldots, \phi_{ps}(0.75)]' \), we use \( \mathbf{W}_{n}^{(j)}(\Phi) \) to test the hypotheses:

\[
\mathbb{H}_0^{(j)}(\Phi) : \mathbf{S}_j \Phi_s(\tau) = s_j \text{ versus } \mathbb{H}_1^{(j)}(\Phi) : \mathbf{S}_j \Phi_s(\tau) \neq s_j, \quad \text{for } j = 1, 2, 3, 4, \text{ where } \mathbf{S}_1 = (1, 0'_{p-1}, -1, 0'_{p-1}, 1), \mathbf{S}_2 = (0, 0'_{p-1}, 1, 0'_{p-1}, -1, 0'_{p-1}), \mathbf{S}_3 = (1, 0'_{p-1}, 0, 0'_{p-1}, -1, 0'_{p-1}), \text{ and } \mathbf{S}_4 = (\mathbf{S}_1', \mathbf{S}_2').
\]

In principle, if the different lag orders are selected based on BIC, the null hypotheses can be different for the short-run parameters (of course, the same for the long-run parameter). Nevertheless, the essence of inference is intact when testing the coefficients on the first lagged dependent variable. For example, \( \mathbf{W}_{n}^{(1)}(\Phi) \) tests \( \phi_{1s}(0.25) = \phi_{1s}(0.50) \), and \( \mathbf{W}_{n}^{(4)}(\Phi) \) tests \( \phi_{1s}(0.25) = \phi_{1s}(0.50) = \phi_{1s}(0.75) \). Under the DGP condition in (23), we expect the empirical rejection rates of the Wald test statistics to be close to the nominal level. For \( j = 1, 2, \) and \( 3, \mathbf{W}_{n}^{(j)}(\mathbf{B}) \), \( \mathbf{W}_{n}^{(j)}(\Phi) \), and \( \mathbf{W}_{n}^{(j)}(\mathbf{\Gamma}) \) follow \( \chi^2_1 \); and \( \mathbf{W}_{n}^{(4)}(\mathbf{B}) \), \( \mathbf{W}_{n}^{(4)}(\Phi) \), and \( \mathbf{W}_{n}^{(4)}(\mathbf{\Gamma}) \) follow \( \chi^2_2 \). We consider sample sizes of 50, 200, 400, 600, 800, 1,000, and 2,000, and we let \( \sigma_* = 0.00, 0.10, 0.20, 0.30, \) and 0.40. Notice that, as \( \sigma_* \) increases, \( \{U_t\} \) becomes more serially correlated; and thus, the larger size distortion is expected for \( \sigma_* = 0.4 \). We compute the Wald test statistics using the same methodology as in the previous subsection and present our simulation results in Tables 7, 8, and 9. For brevity, we examine the empirical rejection rates at 5% significance only.

\[\text{Insert Tables 7, 8, and 9 around here.}\]

The overall simulation results can be summarised as follows. First, the sizes of the Wald test statistics all tend to the nominal level as the sample size increases, implying that the Wald tests are asymptotically well approximated by the claimed null distributions even in the presence of MA(1) errors. Second, in order to examine the asymptotic validity at other nominal levels, we also draw the EPDFs and ECDFs of the Wald tests for \( \sigma_* = 0.4 \) and \( n = 2,000 \). Figures 3 and 4 show that the claimed asymptotic null distributions well approximate the empirical distribution functions. Third, as expected, we note that there are somewhat non-negligible size distortions especially in the small sample (say, \( n = 50 \)) as \( \sigma_* \) increases.

\[\text{Insert Figures 3 and 4 around here.}\]
5 Empirical Application

In a classic study on dividend policy, based on interviews with 28 managers, Lintner (1956) observes that firms gradually adjust dividends in response to changes in earnings. Lintner also observes that firms are reluctant to make dividend changes that have to be reversed in the near future. An important implication of this finding is that managers make dividend adjustments in response to unanticipated and non-transitory changes in their firms’ earnings in order to attain a long-run target payout ratio. Empirical research generally supports Lintner’s partial adjustment framework at both the firm and aggregate levels, e.g., Fama and Babiak (1968), Marsh and Merton (1987), Garrett and Priestley (2000), von Eije and Megginson (2008) and Andres et al. (2009). In the literature, however, the estimates of adjustment speed are widely different. Fama and Babiak (1968) and Aivazian et al. (2006) estimate speeds 0.37 and 0.24, respectively. Using the US real estate investment trusts data from 1992-2003, Hanyunga and Stephens (2009) report widely different estimates of the speed of adjustment: 0.028 and 0.04 at a quarterly frequency and 0.379 and 0.371 at an annual frequency by the panel OLS and Tobit methods, respectively. These large differences are mainly due to the well-known small-T bias in dynamic AR regression (e.g., Hurwicz, 1950; Nickell, 1981).

Recently, Brav et al. (2005) surveyed 384 financial executives to determine the factors that drive dividend and share repurchase decisions. The most important findings are as follows: executives try to avoid reducing dividends per share (93.8% agreed) and aims to maintain a smooth dividend stream (89.6%); they are reluctant to make changes that will be reversed (77.9%) because there are negative consequences to cutting dividends (88.1%). These responses are mostly consistent with Lintner’s observations, although the link between dividends and earnings has weakened in the subsequent 50 years. With the new survey evidence in mind, Leary and Michaely (2011) address an important issue of why firms smooth dividends by studying the cross-sectional differences in the behavior of dividend smoothing and providing empirical evidence against alternative explanations. Their main findings suggest that dividend smoothing is most common in the U.S. in large and mature firms that are not financially constrained, possess ce low levels of asymmetric information, and are readily susceptible to agency conflicts. Furthermore, they document that dividend smoothing has steadily increased over the past century, even before firms began to repurchase shares in the mid-1980s. Chen et al. (2012) also show that aggregate dividends are dramatically more smoothed in the post-world war II period (1946-2006) than in the prewar period (1871-1945). In particular, they report that, during the post-world war II period, dividends adjust to the earnings target at a speed about one-fourth that of the pre-war period. This implies that dividend growth is unpredictable for the firms that have most smoothed their dividends but predictable for the firms that have least smoothed their dividends. By a simulation analysis,
they draw two important conclusions. First, even if dividends are supposed to be strongly predictable without smoothing, dividend smoothing can negate this predictability. Second, dividend smoothing leads to a persistent dividend yield. Motivated by these conclusions, Rangvid et al. (2012) investigate the relationship between dividend predictability and dividend smoothing in a global framework constructed by aggregated data of dividend yields, prices, and total returns in 50 countries at a quarterly frequency. They find that dividend predictability is driven by cross-country differences in firm characteristics, dividend smoothing, and institutions. In particular, aggregate dividend growth is highly predictable by dividend yields in medium-sized and smaller countries, although this predictability disappears for large countries.

However, all of these empirical findings are obtained from examining dividend behaviors only at the conditional mean, and the literature has not yet investigated an important possibility that the degree of dividend smoothing can be fundamentally heterogeneous across different quantiles of dividend distribution. The relationships between the dependent variable and covariates may differ depending on the location of the dependent variable in its own distribution. We contribute to the existing literature on dividend policy by incorporating location asymmetries into dividend adjustment and target payout ratio at the aggregate level.

Lintner (1956) suggests that firms partially adjust toward their target dividend level as follows:

$$\Delta D_t = a_\ast - \zeta_\ast (D_t^\ast - D_{t-1}) + \epsilon_t,$$

where $D_t$ and $D_t^\ast$ are the current (observed) and target levels of dividends at time $t$, respectively, and $|\zeta_\ast|$ is the speed of adjustment or smoothing coefficient (expected to lie between 0 and 1), which represents how quickly firms adjust toward the target dividend. It is widely suggested that the target dividend has a long-run relationship with the current earning $E_t$ as follows:

$$D_t^\ast = \beta_\ast E_t,$$

where $\beta_\ast$ is the target payout ratio. Combining (24) and (25) yields the partial adjustment process:

$$\Delta D_t = a_\ast + \zeta_\ast D_{t-1} + \theta_\ast E_t + \epsilon_t,$$

where $\theta_\ast := -\zeta_\ast \beta_\ast$.

---

7The literature provides a number of ways to measure permanent earnings. Marsh and Merton (1987) suggest using the one-period lagged real stock price. Garrett and Priestley (2000) apply the Kalman filtering method and separate permanent earnings from a transitory component in the reported earnings. Alternatively, Andres et al. (2009) suggest using cash flow as an alternative measure of (permanent) earnings. In this study, we consider aggregated earnings in its simplest form.
We find that $D_t$ and $E_t$ are non-stationary, as in Garrett and Priestley (2000) and Mougoue and Rao (2003). Hence, the target dividend relationship in (25) is indeed a cointegrating relationship between $D_t$ and $E_t$ (these are confirmed by the unreported unit root and the cointegration test results in order to save space). More importantly, the errors $\epsilon_t$ in (26) are highly likely to be serially correlated. Therefore, it is more general to consider the following ARDL extension of the simple partial adjustment model in (26):

$$\Delta D_t = \alpha \ast + \zeta s D_{t-1} + \gamma s E_{t-1} + \sum_{j=1}^{p-1} \lambda j s \Delta D_{t-j} + \sum_{j=0}^{q-1} d j s \Delta E_{t-j} + \varepsilon_t,$$  

(27)

where $\varepsilon_t$ is assumed to be no longer serially correlated given sufficiently large lag orders of $p$ and $q$. Notice, however, that $\Delta E_t$ may still be contemporaneously correlated with the error term $\varepsilon_t$, in which case we can control for such correlation by employing a projection of $\varepsilon_t$ on $\Delta E_t$ as follows:

$$\varepsilon_t = \omega \ast \Delta E_t + U_t$$  

(28)

where $U_t$ is not correlated with $\Delta E_t$ by construction. Substituting (28) in (27), we obtain the final ARDL($p$, $q$) specification as

$$\Delta D_t = \alpha \ast + \zeta s (D_{t-1} - \beta s E_{t-1}) + \sum_{j=1}^{p-1} \lambda j s \Delta D_{t-j} + \sum_{j=0}^{q-1} d j s \Delta E_{t-j} + U_t,$$  

(29)

where $\beta_s := -\frac{\gamma}{\zeta}$, $\delta s := d0 s + \omega s$ and $\delta j s := dj s$ for $j = 1, \ldots, q - 1$. Unless $\omega s = 0$ in (28), the ARDL($p$, $q$) model in (29) is unable to identify a contemporaneous causal relationship between $\Delta D_t$ and $\Delta E_t$ due to the use of the projection in (28). However, under the additional and acceptable assumption that earning changes can immediately cause dividend changes, but not vice versa, we can interpret the new coefficient on $\Delta E_t$, $\delta s$, as an impact reaction parameter. Furthermore, all the other parameters, including $\beta_s$ and $\zeta s$ in (29), can be estimated free of endogeneity since $\Delta E_t$ is included in the regression.

In particular, we are interested in the following four parameters: (i) the ECM parameter $\zeta s$ measuring the degree of dividend smoothing, (ii) the long-run (cointegrating) target payout ratio $\beta s$, (iii) the momentum effect of the dividend growth captured by $\lambda s := \sum_{j=1}^{p-1} \lambda j s$, which measures the cumulative impact of past dividend growth on the current dividend growth, i.e. $\sum_{j=1}^{p-1} \partial \Delta D_t / \partial \Delta D_{t-j}$, and (iv) the momentum effect of the earnings growth, i.e. $\sum_{j=1}^{p-1} \partial \Delta E_t / \partial \Delta E_{t-j}$ (there is pervasive evidence that

\footnote{In general, the sign of $\omega s$ may dictate whether the impact effect captured by $\delta s$ in (29) is overreacting ($\omega s > 0$) or underreacting ($\omega s < 0$). See Nguyen and Shin (2012) for details.}
stock returns with positive momentum in the short-run are followed by reversals in the long-run, e.g., Koijen et al. (2009)), and (iv) the impact reaction of dividend growth to earnings growth, \( \delta_{0*} \), which measures the effect of contemporaneous change in earnings on the current dividend growth, i.e., \( \partial \Delta D_t / \partial \Delta E_t \). Based on existing empirical and theoretical studies (e.g., Koijen et al., 2009; Chen et al., 2012; and Nguyen and Shin, 2012), our priors about these four parameters are given as follows:

\[
\zeta_* < 0, \quad 0 < \beta_* < 1, \quad \lambda_* \geq 0, \quad \text{and} \quad \delta_{0*} \geq 0.
\]

We employ the dataset collected by Shiller (2005) for Irrational Exuberance. We construct a quarterly dataset on real price, real dividend, and real earnings for the Standard and Poor’s 500 stocks over the period 1871Q3 - 2010Q2 with 558 quarterly observations. As a baseline case, we first estimate the ARDL(3,1) model in (29) by OLS. Employing the BIC lag selection procedure, we find that the appropriate lag orders are \( p = 3 \) and \( q = 1 \), respectively. The conditional mean estimation results for the four key parameters are reported in Table 10(a). The ECM coefficient is \(-0.04\), which implies that the adjustment speed is only about 4%, whereas the long-run payout coefficient is 0.36. Further, the momentum and the impact reaction parameters are estimated at 0.48 and 0.01, respectively. Combined together, we conclude that the changes in dividend appear to be driven more strongly by history than the changes in earnings. The magnitude and the sign of all four key coefficients are generally in line with our priors.

Next, in order to apply the QARDL approach, we consider the quantile counterpart of the ARDL(3,1) model as follows:

\[
\Delta D_t = \alpha_* (\tau) + \zeta_* (\tau) D_{t-1} + \gamma_* (\tau) E_{t-1} + \sum_{j=1}^{2} \lambda_{j*} (\tau) \Delta D_{t-j} + \delta_* (\tau) \Delta E_t + U_t (\tau) \quad \text{for } \tau \in (0, 1).
\]

To the best of our knowledge, our dynamic quantile regression application to an investigation of the long-run target payout policy and the associated dividend smoothing is the first attempt in the literature. Based on the existing literature discussed above, we aim to address the following issues that may challenge many studies relying upon the conditional mean models:

- **Issue 1:** Locational asymmetry is associated with the notion that the key four parameters may depend

---

\(^9\)When estimating the QARDL model, we acknowledge that the lag orders, \( p \) and \( q \), should be selected according to quantile in a data-coherent manner. However, extending the analysis to allow for lag order selection or approximability of infinite order models could lead to an overly long and complicated paper and is beyond the scope of the current study. We have briefly addressed this issue in the Monte Carlo section using BIC. Making \( p \) and \( q \) dependent on the quantile index \( \tau \) not only increases the computation complexity, but also reduces the comparability across different quantile estimation results. Thus, we find it reasonable to employ the common lag orders selected via the BIC applied to the conditional mean model across quantiles in order to make quantile estimation results comparable to the baseline conditional mean model, as similarly implemented in Covas et al. (2012).
on the current location of the dividend within its conditional distribution. In particular, we are keen to
determine whether the long-run target payout ratio and dividend smoothing are heterogeneous across
different quantiles. This allows for a natural distinction among firms with low- and high-dividend
paying policies.

- **Issue 2**: Through employing the robust rolling estimation technique, we wish to investigate the time-
varying patterns of dividend policy. Given the empirical evidence that the link between dividends and
earnings has recently weakened (Brav et al., 2005) and that aggregate dividends are dramatically more
smoothed in the post-world war II period than in the pre-world war II period (Chen et al. 2012), we
are particularly interested in elucidating whether the location asymmetries are monotonic or pervasive
over the whole periods.

The full sample quantile estimation results are reported in Table 10(b), showing that most of the coef-
ficients are statistically significant across all of the quantiles that we consider. The only exceptions are the
estimates of \( \delta^*_\tau \) at higher quantiles.

We also plot the estimation results in Figure 5, which displays the quantile estimates of the four key
parameters \( \zeta^*_\tau, \beta^*_\tau, \lambda^*_\tau = \sum_{j=1}^{2} \lambda^*_j(\tau), \) and \( \delta^*_\tau \) with 90% confidence intervals against quantile
indices ranging from 0.05 to 0.95.\(^{10}\)

The estimation results are summarised as follows. First, the quantile estimates of the ECM parameter
\( |\hat{\zeta}(\tau)| \) start with a 6% adjustment speed at the low quantiles (\( \tau = 0.05 \) and 0.1) and decrease monotonically
as the quantile increases. The values reach a minimum of 3% at \( \tau = 0.4 \) and stay between 3% and 4% at
higher quantiles. This result suggests location asymmetry. In particular, we find that dividend smoothing
is stronger in medium-to-higher quantiles than in lower quantiles of dividends. That is, the dividend policy
is more conservative at higher quantiles. This evidence is consistent with the hypotheses of the free cash
flow problem and/or agency conflicts resulting from market frictions. For example, Easterbrook (1984)
and DeAngelo and DeAngelo (2007) predict a positive relationship between the level of dividends and
smoothing. Second, the quantile estimates of the long-run target payout ratio \( \hat{\beta}(\tau) \) increase monotonically

\(^{10}\)We employed the wild bootstrap method proposed by Feng et al. (2011) to produce confidence intervals because bootstrapping
can provide a better approximation of the underlying sampling distribution than the asymptotic theory.
with increasing quantiles, reaching a peak of 0.40 at $\tau = 0.3$ and remaining at similar levels at higher quantiles, except the highest quantile $\tau = 0.95$ where the long-run payout ratio drops to 0.33. Such stable and high levels contrast with the low long-run payout ratio observed at low quantiles (e.g., $\hat{\beta}(\tau) = 0.27$ at $\tau = 0.05$), providing evidence of location asymmetries. Third, we find that the quantile estimates of the momentum parameter $\hat{\lambda}(\tau)$ generally decrease as the quantile increases, decreasing from 0.58 at $\tau = 0.05$ to 0.30 at $\tau = 0.9$. Thus, the momentum effects are strongly observed at the lower quantiles. Finally, the impact reaction coefficient $\hat{\delta}(\tau)$ tends to decrease with increase in quantile, although the magnitudes are negligibly small across all quantiles, ranging between 0.02 and −0.01. This suggests that a change in current earnings flow that is viewed by management as essentially transitory would be unlikely to give rise to a noticeable and immediate change in dividends.

The full sample estimation results clearly confirm that there is strong evidence of location asymmetries between lower and medium-to-higher quantiles for all four key parameters. Overall, these findings are consistent with the cross-sectional evidence by Leary and Michaely (2011) and the aggregate time series evidence shown in a global setting by Rangvid et al. (2012) that large and mature firms with stable cash-flow and return processes have a greater tendency to smooth dividends. Our findings of the quantile-dependent cointegrating relationship between dividends and earnings provide further support for the number of recent studies that report similar results; namely, the quantile-varying cointegrating relationship between stock price and dividend in Xiao (2009), between spot and future oil prices in Lee and Zeng (2011), and between the nominal interest rate and inflation in Tsong and Lee (2013). As discussed by Xiao (2009), a plausible explanation for quantile-varying cointegration is that the underlying relationship between integrated time series may vary over time due to (heterogeneous) shocks arising at each point of time. If so, the quantile cointegrating framework is naturally fitted for such a situation because quantile coefficients can be viewed as random coefficients, as explained in Koenker and Xiao (2006), in which such randomness is driven by a common shock arriving at each time period.

Considering that the sample period is quite long at more than 140 years, we find it more prudent to allow for time-varying patterns of dividend policy. Any model estimated over a long span of time that does not incorporate structures of time-varying mechanisms can result in only the average tendency of the dividend policy as examined using a range of regime-switching models. In this paper, we instead employ a robust rolling estimation technique with a window length of 320 quarters, a figure that should balance the data requirement of the QARDL model with our desire to examine the richest possible range of regimes. To this end, we re-estimate the QARDL(3,1) model by successively moving the estimation window forward by one quarter until we reach the end of the sample. The time-series plots of the rolling quantile estimates
of the four parameters are displayed in Figure 6. For clarity, we provide only the point estimates across three quantile indices, \( \tau = (0.25, 0.5, 0.75) \), because superimposing confidence intervals would make them incomprehensible.\(^{11}\)

\[ \text{Insert Figure 6 around here.} \]

The rolling quantile estimates of \( |\zeta^* (\tau)| \) display quite strong time-varying patterns. In general, we observe that dividend smoothing has been stronger in recent periods, a finding consistent with Leary and Michaely (2011) and Chen et al. (2012). In the earlier periods, the location asymmetry was stronger with an order \( |\hat{\zeta}(0.25)| > |\hat{\zeta}(0.5)| > |\hat{\zeta}(0.75)| \), where the average adjustment speeds were about 8.50% and 4.00% at \( \tau = 0.25 \) and \( \tau = 0.75 \), respectively. This ordering is generally consistent with the full sample results. On the other hand, in the later periods, the location asymmetry was much weaker. Interestingly, we observe a reversed order with \( |\hat{\zeta}(0.75)| > |\hat{\zeta}(0.5)| > |\hat{\zeta}(0.25)| \) at some later periods, although their differences are negligible and insignificant. In particular, we find that the speed of adjustment at lower quantiles begins to fall with a steep downward trend. For example, the adjustment speed at \( \tau = 0.25 \) decreases dramatically from about 9% in the earlier period to below 1% in the recent period (the late 2000s). In contrast, the adjustment speed seems quite stable over the whole period at the higher quantile \( \tau = 0.75 \).

Next, the rolling quantile estimates of \( \beta^* (\tau) \) display a strong downward time-varying pattern, showing that payout ratios have become significantly lower in recent periods. This finding generally supports the study by Fama and French (2001) who report a substantial decline in the proportion of U.S. firms paying dividends. A significant decrease in the residual propensity to pay dividends was observed even after controlling for firm characteristics. Chen et al. (2012) also document that the aggregate payout ratio has declined in the postwar period. Notice, however, that this (monotonic) downward trend is observed only at \( \tau = 0.5 \) and 0.75. The payout ratio at \( \tau = 0.25 \) tends to increase from the middle of the sample period (approximately after the 1960s) and starts to decrease from the 1970s. The location asymmetry was stronger in the earlier periods with an order \( \hat{\beta} (0.75) > \hat{\beta} (0.5) > \hat{\beta} (0.25) \), where the average payout ratios are about 35% and 60% at \( \tau = 0.25 \) and \( \tau = 0.75 \), respectively. On the other hand, we observe weaker location asymmetry in the later periods but with a reversed order \( \hat{\beta} (0.25) > \hat{\beta} (0.5) > \hat{\beta} (0.75) \). Especially since 1980s, the location asymmetry was almost negligible, and the payout ratio remains generally at a low level, except at the noted spikes.\(^{12}\) This finding can be explained by the evidence provided by Skinner (2008).

\(^{11}\)The complete figures with 90% confidence intervals are available upon request.

\(^{12}\)Three conspicuous spikes are observed in the payout ratio coefficient at \( \tau = 0.25 \) during the late 2000s. These coincide with the corresponding three spikes observed in the ECM parameter, where estimates of \( \hat{\zeta} (0.25) \) are very close to zero. As the long-run coefficient is evaluated by \( \hat{\beta} (0.25) = -\hat{\gamma} (0.25) / \hat{\zeta} (0.25) \), we should bear in mind that these three extreme estimates may be unreliably inflated.
who observes that repurchases are increasingly responsive to earnings by substituting themselves for dividends, especially over the last two decades, since stock repurchases emerged as significant in the early 1980s.\footnote{One of the main reasons behind this structural shift is Rule 10b–18 introduced in 1982 which provided a non-exclusive safe harbor for issuer repurchase. Furthermore, the Securities Exchange Commission proposed amendments of the rule in 2010 to clarify and modernize the safe harbor provision in light of developments in automated trading systems and technology.} Hence, the most recent ongoing downward trend can be attributed to the popularity of repurchase as a means of disbursement of temporal increases of cash flow.

Turning to the rolling quantile estimates of $\lambda_*(\tau)$, we observe volatile time-varying patterns. The location asymmetry is stronger in the earlier periods with an order $\lambda(0.25) > \lambda(0.5) > \lambda(0.75)$, in which the average momentum coefficients are about 0.45 and 0.32 at $\tau = 0.25$ and $\tau = 0.75$, respectively. The location asymmetry is less significant in the later periods, although they tend to increase substantially at $\tau = 0.75$ (from 0.30 to 0.60) such that significant location asymmetry is observed very recently with $\lambda(0.75) > \lambda(0.5) \simeq \lambda(0.25)$. From this evidence, we conclude that dividend policy has evolved in such a way that its own lags have become the more important predictor in the postwar period. Using the univariate regression of dividend change on its own lags, Chen et al. (2012) also document a similar finding that the autoregressive coefficient is 0.061 for the prewar period (statistically insignificant), whereas it is 0.687 in the postwar period (significant).

Finally, the rolling quantile estimates of $\delta_*(\tau)$ show that there is a strong downward trend from around 12–14% in the earliest of our sample periods to almost zero in very recent period. These time-varying patterns are somewhat similar to those observed from $|\hat{\zeta}(\tau)|$, and both contribute to the extremely strong dividend smoothing reported in more recent periods. The location asymmetry looks insignificant over the whole rolling period, with no clear pattern or order at different quantiles. In the earlier periods (mostly in the prewar period), managers tend to make a noticeable change in dividends immediately with respect to current earnings changes, although such impact reactions will decrease monotonically over time. These downward trends in conjunction with a zero bound in the most recent periods clearly confirm that managers now favor repurchase out of current earnings changes because they are more flexible than dividends, as documented in the survey evidence by Brav et al. (2005).

As our last step in data examination, we provide formal testing results for the location asymmetries for all four parameters over three selected quantiles $\tau = 0.25, 0.5$ and 0.75; i.e. we wish to test if each parameter is constant across quantiles. For example, if $\beta_*(\tau)$ is the parameter of interest, then we consider the following four null hypotheses: $H_{01}^\beta : \beta_*(0.25) = \beta_*(0.5)$, $H_{02}^\beta : \beta_*(0.5) = \beta_*(0.75)$, $H_{03}^\beta : \beta_*(0.25) = \beta_*(0.75)$, and $H_{04}^\beta : \beta_*(0.25) = \beta_*(0.5) = \beta_*(0.75)$. To this end, we employ the Wald test statistics proposed in Section 3, denoted as $W_{n}^{(1)}(\beta)$, $W_{n}^{(2)}(\beta)$, $W_{n}^{(3)}(\beta)$, and $W_{n}^{(4)}(\beta)$. Figure 7 plots the $p$-values of these Wald
statistics using rolling estimation. In the earlier periods, the null hypotheses are strongly rejected across all three quantiles. In later periods, such location asymmetries become statistically significant only between $\beta_*(0.5)$ and $\beta_*(0.75)$ and between $\beta_*(0.25)$ and $\beta_*(0.75)$. Figure 8 displays the testing result for $\zeta_*(\tau)$. We find that the locational asymmetries are significantly observed only in the earlier periods, especially between $\zeta_*(0.25)$ and $\zeta_*(0.75)$. On the other hand, in the later periods, such location asymmetries become statistically insignificant, except in a few recent periods. These results confirm our earlier findings that strong locational asymmetries are observed in both target payout ratio and smoothing pattern between the low and the high quantiles only in the earlier periods.

Figures 9 and 10 provide the test results for $\lambda_*(\tau)$ and $\delta_*(\tau)$, respectively. The $p$-values for $\lambda_*(\tau)$ indicate that there is weak evidence in favor of the location asymmetry between $\lambda_*(0.25)$ and $\lambda_*(0.75)$ only in the very early periods. With regard to $\delta_*(\tau)$, we find that there exist some evidence of location asymmetry across quantiles around the 1990s.

In summary, we conclude that a thorough examination using the proposed QARDL method sheds further light on the evolution of dividend policy in the U.S. over the past century. First, we document that dividend smoothing has become monotonically stronger over time. Similar monotonic downward trending patterns have been observed in the impact coefficient with respect to changes in current earnings, and they reach almost zero in the most recent periods. Both factors contribute to the extremely strong dividend smoothing reported in very recent periods. Second, our results clearly display that payout ratios have been monotonically decreasing over time and have recently stayed below 30%, providing support for the survey evidence by Brav et al. (2005) that the target payout ratio may no longer be the preeminent decision variable affecting payout decisions. Furthermore, we find that the location asymmetries across different quantiles of the conditional distribution of dividends, which are clearly visible in the earlier periods, are mostly negligible in the most recent periods. These findings may indicate the establishment of financial deepening as a consequence of the long-term process to promote the stability of the whole financial system in the U.S. Rangvid et al. (2012) also provide international evidence that developed countries with more stable returns and dividend processes and with a higher quality of legal systems and corporate governance, such as the U.S., the UK, and Japan, tend to smooth dividends more than other countries.
6 Concluding Remarks

Recently, the literature on quantile time series regression, especially with nonstationary variables, has been rapidly increasing (e.g., Granger, 2010). In particular, Xiao (2009) advances a novel quantile cointegration estimation technique in a static regression. In this paper, we propose the dynamic QARDL modelling approach to simultaneously address the long-run relationship between integrated time series as well as the associated short-run dynamics across a range of quantiles of the conditional distribution of the dependent variable.

We have derived the asymptotic theory for the QARDL model with nonstationary regressors. The QARDL estimators of both the short-run dynamic and the long-run (cointegrating) parameters are shown to asymptotically follow a (mixture) normal distribution. Hence, the null distribution of the Wald statistics for testing the restrictions on both the short-run and the long-run parameters weakly converge to a chi-squared distribution. Given that models are usually investigated in terms of multiple quantiles, we also provided a general package in which the model can be estimated across multiple quantiles, and any linear restrictions on the parameters involving multiple quantiles can be tested using the standard Wald statistics. Overall, Monte Carlo simulation results provide strong support for theoretical predictions.

The key strengths of the QARDL framework have been demonstrated in the empirical analysis of dividend policy in the U.S. using quarterly data on real dividends and earnings for the S&P 500 stocks over the period 1871Q3 through 2010Q2. Following the seminal study by Lintner (1956), the huge number of empirical studies at both the firm and aggregate levels generally support Lintner’s partial adjustment framework, although none of these studies has investigated the important possibility that the dividend policy may be fundamentally heterogeneous across a range of dividend quantiles. Through the use of the time-varying rolling QARDL estimation, in particular, we produced a number of insightful findings. First, dividend smoothing has become monotonically stronger over time, mainly contributing to the extremely strong smoothing observed in recent periods. Second, payout ratios have monotonically declines over time, suggesting that the target payout ratio may no longer be the preeminent decision variable. Importantly, location (quantile) asymmetries were clearly visible in the earlier only, but have dramatically weakened recently.

Finally, we note several avenues for further researches following the current study. First, our framework could easily be adapted to address a range of further issues. Foremost among these must be the development of a formal testing mechanism that could address the issues of whether the long-run cointegration relationship exists at each of quantiles. As we can now analyse the quantile-dependent cointegration relationship, we find it more prudent to consider an important possibility of cointegration holding at some quantiles while
a spurious relationship at other quantiles. In a similar context, Koenker and Xiao (2004) and Galvao (2009) suggest that the stationary property of individual time series of interest may change across different quantiles. Considering that the distribution of nonstationary variables is changing over time, it is a challenging issue to analyse such stochastic trends of nonstationary data within the scope of the quantile regression. In this regard, we also find it a challenging task to develop a formal testing procedure for the presence of the quantile-dependent cointegration relationships. We conjecture that a quantile-dependent cointegration testing framework can be possibly developed by extending the findings of Pesaran et al. (2001), although this approach significantly differs from the existing approaches adopted by Koenker and Xiao (2004, 2006) for the quantile unit root test and by Xiao (2009) for the quantile cointegration test. Second, we find it quite useful to develop a quantile regression extension of the asymmetric ARDL framework advanced by Shin, Yu, and Greenwood-Nimmo (2013) should be considered. This combined approach is expected to provide us the flexible econometric framework, which can help us to identify several forms of distinct asymmetry that can be applied to a wide variety of fundamentally asymmetric processes in Economics and Finance. Finally, given that conventional estimation procedures can be significantly affected by the presence of conditional heteroskedasticity (e.g., Seo, 2012), it would be desirable to explicitly control for time-varying volatilities in the QARDL framework.

Appendix: Proofs

We first provide a number of preliminary lemmas and a corollary that will be used in proving the main theorems.

**Lemma A1.** Under Assumption 1,

(i) $\sum_{t=1}^{n} W_{t-i} = O_P(\sqrt{n})$ for $i = 0, 1, \ldots$; and

(ii) for each $\tau$, $\sum_{t=1}^{n} \bar{K}_t(\tau) = O_P(\sqrt{n})$ and $\sum_{t=1}^{n} K_t(\tau) = O_P(n)$. □

**Proof of Lemma A1:** By letting $r = 1$, Assumption 1(vi) implies that $n^{-1/2} \sum_{t=1}^{n} \bar{W}_t \Rightarrow \mathcal{B}_W(1)$ and $n^{-1/2} \sum_{t=1}^{n} \bar{K}_t(\tau) \Rightarrow \mathcal{B}_K(1, \tau)$. Furthermore, we can apply the ergodic theorem to $n^{-1} \sum_{t=1}^{n} K_t(\tau) \Rightarrow \mathbb{E}[K_t(\tau)]$ in probability, which completes the proof. □

**Lemma A2.** Under Assumption 1,

(i) $n^{-1} \sum_{t=1}^{n} W_{t-i} W_{t-j}' \rightarrow \mathbb{E}[W_{t-i} W_{t-j}']$ almost surely (a.s.) for $i, j = 0, 1, \ldots, q - 1$;

(ii) $\sum_{t=1}^{n} W_{t-i} X_t' = O_P(n)$ and $\sum_{t=1}^{n} U_{t-i}(\tau) X_t = O_P(n^{3/2})$, where $i = 0, 1, \ldots$;

(iii) For each $\tau$, $n^{-1} \sum_{t=1}^{n} K_t(\tau) K_t(\tau)' \rightarrow \mathbb{E}[K_t(\tau) K_t(\tau)']$ a.s.;
(iv) For each \( \tau \), \( n^{-1} \sum_{t=1}^{n} K_t(\tau) W_{t-i} \rightarrow \mathbb{E}[K_t(\tau)W_{t-i}] \) a.s., where \( i = 0, 1, \ldots, q - 1 \); and

(v) For each \( \tau \), \( n^{-3/2} \sum_{t=1}^{n} K_t(\tau) X_t' \Rightarrow \mathbb{E}[K_t(\tau)] \int_{0}^{\tau} B_W(\tau)^{'} d\tau. \)

\[ \square \]

**Proof of Lemma A2:** (i) By applying the ergodic theorem to Assumption 1(ii), we obtain the result in Lemma A2(i).

(ii) We follow the proof of Proposition 18.1(d) in Hamilton (1994, pp. 562–563) by letting his \( u_{t-s} \) and \( \xi_{t-1} \) be our \( W_{t-i} \) and \( X_t \), respectively. We then apply Assumption 1(vi), and derive \( \sum_{t=1}^{n} W_{t-i}X_t' = O_p(n) \) by induction. By Assumption 1(vi), \( n^{-1/2} \sum_{t=1}^{n} W_t = B_W(\cdot) \). Then, it is straightforward to show that \( \sum_{t=1}^{n} U_{t-i}(\tau)X_t = O_p(n) \) by the continuous mapping theorem and Assumptions 1(i) and 1(ii). Furthermore, it is elementary to show that \( \sum_{t=1}^{n} X_t = O_p(n^{3/2}) \). From these, it follows that \( \sum_{t=1}^{n} U_{t-i}(\tau)X_t = O_p(n^{3/2}). \)

(iii) We first that \( \{K_{t,i}(\tau)\} \) is a stationary and ergodic process. First, consider when \( i \leq q \). Then,

\[ K_{t,i}(\tau) = -\sum_{j=q-1}^{\infty} \xi_{0,j*}(\tau)'W_{t-i-j} + \sum_{j=0}^{\infty} \rho_{j*}(\tau)U_{t-i-j}(\tau). \]

From the definition of \( \rho_{j*}(\tau) \) and Assumption 1(v), \( \{\sum_{j=0}^{\infty} \rho_{j*}(\tau)U_{t-i-j}(\tau)\} \) is stationary and ergodic. Assumption 1(v) and Theorem 4.4.1 of Brockwell and Davis (1991, p. 122) also imply that \( \{\sum_{j=q-1}^{\infty} \xi_{0,j*}(\tau)'W_{t-i-j}\} \) is stationary, which implies that \( \{\sum_{j=q-1}^{\infty} \xi_{0,j*}(\tau)'W_{t-i-j}\} \) is ergodic by Theorem 3.35 of White (2001). Thus, \( \{K_{t,i}(\tau)\} \) is stationary and ergodic when \( i \leq q \). We next consider when \( i > q \). Then,

\[ K_{t,i}(\tau) = -\beta_s(\tau)' \sum_{j=0}^{i-q-1} W_{t-q-j} + \sum_{j=0}^{\infty} \pi_{j*}(\tau)'W_{t-i-j} + \sum_{j=0}^{\infty} \rho_{j*}(\tau)U_{t-i-j}(\tau), \]

and \( \{\sum_{j=0}^{\infty} \pi_{j*}(\tau)' \sum_{t=1}^{n} W_{t-i-j} + \sum_{j=0}^{\infty} \rho_{j*}(\tau) \sum_{t=1}^{n} U_{t-i-j}(\tau)\} \) is stationary and ergodic by the same logic as above. Furthermore, \( \{\sum_{j=0}^{i-q-1} W_{t-q-j}\} \) is a sequence of finite sums of the stationary and ergodic processes, so that \( \{K_{t,i}(\tau)\} \) is stationary and ergodic even when \( i > q \).

Given this, Assumption 1(vi) implies that \( \mathbb{E}[K_{t,i}(\tau)^2] < \infty \). By applying the ergodic theorem to the definition of \( K_t(\tau) \), it follows that \( n^{-1} \sum_{t=1}^{n} K_{t,i}(\tau)K_{t,j}(\tau) \rightarrow \mathbb{E}[K_{t,i}(\tau)K_{t,j}(\tau)] \) a.s. for \( i, j = 1, 2, \ldots, p \). This verifies Lemma A2(iii).

(iv) Since \( K_t(\tau) \) and \( W_{t-i} \) are stationary and ergodic processes by Assumption 1(ii), \( \{K_t(\tau)W_{t-i}\} \) is a stationary and ergodic process by Theorem 3.35 of White (2001). Moreover, by Assumption 1(vi), \( \mathbb{E}[K_{t,i}(\tau)^2] < \infty \) and \( \mathbb{E}[W_{t-j,\ell}^2] < \infty \) for \( j = 1, \ldots, p \) and \( \ell = 1, \ldots, k \). Hence, the result in Lemma A2(iv) follows from the ergodic theorem.
(v) First, we consider the case with \( i \leq q \). Then,

\[
\sum_{t=1}^{n} K_{t,i}(\tau)X'_t = -\sum_{j=q+1}^{\infty} \xi_{0,j*}(\tau)' \sum_{t=1}^{n} W_{t-i,j}X'_t + \sum_{j=0}^{\infty} \rho_{j*}(\tau) \sum_{t=1}^{n} U_{t-i-j}(\tau)X'_t.
\]

We now have that \( \sum_{t=1}^{n} W_{t-i,j}X_t = O_p(n) \) and \( \sum_{t=1}^{n} U_{t-i-j}(\tau)X_t = O_p(n^{3/2}) \) by Lemma A2(ii). Therefore, Assumption 1(v) implies \( \sum_{t=1}^{n} K_{t,i}(\tau)X'_t = O_p(n^{3/2}) \). Furthermore, \( U_{t-i-j}(\tau) \) and \( X_t \) are independent by the definition of \( U_{t-i-j}(\tau) \). This implies that \( \sum_{t=1}^{n} (U_{t-i-j}(\tau) - \mathbb{E}[U_{t-i-j}(\tau)])X_t = O_p(n) \) by theorem 17. 3 of Hamilton (1994, pp. 505–506). This implies that \( n^{-3/2} \sum_{t=1}^{n} K_{t,i}(\tau)X'_t = \sum_{j=0}^{\infty} \rho_{j*}(\tau)\mathbb{E}[U_{t-i-j}(\tau)]n^{-3/2} \sum_{t=1}^{n} X'_t = \mathbb{E}[K_{t,i}(\tau)] \int_{0}^{1} \mathcal{B}_W(r)dr. \)

Next, consider the case with \( i > q \). Then, Lemma A2(ii) allows that

\[
\sum_{t=1}^{n} K_{t,i}(\tau)X'_t = -\beta_{*}(\tau)' \sum_{j=0}^{i-q-1} \sum_{t=1}^{n} W_{t-q-j}X'_t + \sum_{j=0}^{\infty} \pi_{j*}(\tau)' \sum_{t=1}^{n} W_{t-i-j}X'_t + \sum_{j=0}^{\infty} \rho_{j*}(\tau) \sum_{t=1}^{n} U_{t-i-j}(\tau)X'_t.
\]

By the same reason as above, \( \sum_{t=1}^{n} W_{t-q-j}X'_t, \sum_{t=1}^{n} W_{t-i-j}X'_t, \) and \( \sum_{t=1}^{n} U_{t-i-j}(\tau)X'_t \) are \( O_p(n^{3/2}) \). Thus, Assumption 1(v) implies that \( \sum_{t=1}^{n} K_{t,i}(\tau)X'_t \) is \( O_p(n^{3/2}) \). Therefore, \( n^{-3/2} \sum_{t=1}^{n} K_{t,i}(\tau)X'_t = n^{-3/2} \sum_{t=1}^{n} \sum_{j=0}^{\infty} \rho_{j*}(\tau) \sum_{t=1}^{n} U_{t-i-j}(\tau)X'_t + o_p(1) \). The rest of the proof is identical to the case when \( i \leq q \). This completes the proof. \( \square \)

**Lemma A3.** *Under Assumption 1,*

\[
\sum_{t=1}^{n} \begin{bmatrix} n^{-3/2}X_t \\ n^{-2}X_tX'_t \\ n^{-1}\psi_{r}[U_t(\tau)]X_t \end{bmatrix} \Rightarrow \begin{bmatrix} \int_{0}^{1} \mathcal{B}_W(r)dr \\ \int_{0}^{1} \mathcal{B}_W(r)\mathcal{B}_W(r)dr \\ \int_{0}^{1} \mathcal{B}_W(r)d\mathcal{B}_W(r,\tau) \end{bmatrix}.
\]

**Proof of Lemma A3:** By Assumption 1(vi), we have \( n^{-1/2} \sum_{t=1}^{\lfloor n(\cdot) \rfloor} W_t \Rightarrow \mathcal{B}_W(\cdot) \). Then, application of the continuous mapping theorem and Lemma 3.1(e) in Phillips and Durlauf (1986) delivers the desired result. \( \square \)

The following Corollary immediately follows from the previous Lemmas.

**Corollary A1.** *Under Assumption 1,*
(i) Let $D_G := \text{diag}([\sqrt{n}t_{1+qk}, nt_k'])$ and $G := [G_1, \ldots, G_n]'. Then,

$$D_G^{-1}G'GD_G^{-1} = \sum_{t=1}^{n} \begin{bmatrix}
    n^{-1} & n^{-1}W_t' & n^{-3/2}X_t' \\
    n^{-1}W_t & n^{-1}W_tW_t' & n^{-3/2}W_tX_t' \\
    n^{-3/2}X_t & n^{-3/2}X_tW_t & n^{-2}X_tX_t'
\end{bmatrix}
\Rightarrow \begin{bmatrix}
    1 & 0' & \int_0^1 B_W(r)dr \\
    0 & \mathbb{E}[W_tW_t'] & 0' \\
    \int_0^1 B_W(r)dr & 0 & \int_0^1 B_W(r)B_W(r)dr
\end{bmatrix};$$

(ii) Furthermore,

$$D_G^{-1}G'\Psi'(U) = \sum_{t=1}^{n} \begin{bmatrix}
    n^{-1/2}\psi_t[U_t(\tau)] \\
    n^{-1/2}\psi_t[U_t(\tau)]W_t \\
    n^{-1}\psi_t[U_t(\tau)]X_t
\end{bmatrix} \Rightarrow \begin{bmatrix}
    B_\psi(1, \tau) \\
    B_\psi W(1, \tau) \\
    \int_0^1 \bar{B}_W(r)dB_\psi(r, \tau)
\end{bmatrix};$$

(iii) $D_H^{-1}G'K(\tau) = O_p(1)$, where $D_H := \text{diag}([nt_{1+qk}, n^{3/2}t_k'])$;

(iv) $M := n^{-2}X'[I - \bar{W}(\bar{W}'\bar{W})^{-1}\bar{W}']X \Rightarrow \int_0^1 \bar{B}_W(r)\bar{B}_W(r)dr$; and

(v) $n^{-1}X'[I - \bar{W}(\bar{W}'\bar{W})^{-1}\bar{W}']\Psi'(U) \Rightarrow \int_0^1 \bar{B}_W(r)dB_\psi(r, \tau). \quad \Box$

Proof of Corollary A1: (i) Lemmas A1(i), A2(i) and A2(ii) imply that

$$\left\{ n^{-1}\sum_{t=1}^{n} W_t, n^{-1}\sum_{t=1}^{n} W_tW_t', n^{-3/2}\sum_{t=1}^{n} W_tX_t' \right\} \rightarrow_p \{ 0, \mathbb{E}[W_tW_t'], 0 \}.$$

Next, Lemma A3 implies that

$$\left\{ n^{-3/2}X_t, n^{-2}X_tX_t \right\} \Rightarrow \left\{ \int_0^1 \bar{B}_W(r)dr, \int_0^1 \bar{B}_W(r)\bar{B}_W(r)dr \right\}.$$

Combining these two results we obtain the desired result in Corollary A1(i).

(ii) Assumption 1(vi) implies that

$$\left\{ n^{-1/2}\sum_{t=1}^{n} \psi_t[U_t(\tau)], n^{-1/2}\sum_{t=1}^{n} \psi_t[U_t(\tau)]W_t \right\} \Rightarrow \{ B_\psi(1, \tau), B_\psi W(1, \tau) \}.$$

Moreover, Lemma A3 implies that $n^{-1}\sum_{t=1}^{n} \psi_t[U_t(\tau)]X_t \Rightarrow \int_0^1 B_W(r)dB_\psi(r, \tau)$. By combining these results, we show that the asymptotic limit of $D_G^{-1}G'\Psi'(U)$ is equal to Corollary A1(ii).
(iii) Notice that $D_{H}^1 G'K(\tau) = \left[ n^{-1} \sum_{t=1}^{n} K_t(\tau) \tilde{W}'_t, n^{-3/2} \sum_{t=1}^{n} K_t(\tau) X'_t \right]'$. Then, the desired result in Corollary A1(iii) follows from Lemmas A1(ii), A2(iv) and A2(v).

(iv) Note that

$$M = n^{-2}X'\left[I - \tilde{W}(\tilde{W}'\tilde{W})^{-1}\tilde{W}'\right]X = n^{-2}X'X - (n^{-3/2}X'\tilde{W})(n^{-1}\tilde{W}'\tilde{W})^{-1}(n^{-3/2}\tilde{W}'X).$$

By Lemma A3, we have $n^{-2}X'X = O_P(1)$; by Corollary A1(i) and Assumption 1(vi), $(n^{-1}\tilde{W}'\tilde{W})^{-1} \rightarrow \text{diag}[1, \mathbb{E}(\tilde{W}_t\tilde{W}'_t)]$ a.s.; and by Corollary A1(i) $n^{-3/2}\tilde{W}'X \Rightarrow \left[ \int_0^1 \tilde{B}_W(r) dr, 0 \right]$. Therefore, we have

$$M = n^{-2}X'X - (n^{-3/2}X'\ell_n)(n^{-3/2}X'\ell_n)' + o_P(1)$$

$$\Rightarrow \int_0^1 \tilde{B}_W(r)\tilde{B}_W(r)' dr - \int_0^1 \tilde{B}_W(r) dr \int \tilde{B}_W(r)' dr = \int_0^1 \tilde{B}_W(r)\tilde{B}_W(r)' dr.$$

(v) We note that $n^{-1}X'\Psi_\tau(U) = \int_0^1 \tilde{B}_W(r) dB_\psi(r, \tau) = O_P(1)$ by Corollary A1(ii), $n^{-3/2}\tilde{W}'X \Rightarrow \left[ \int_0^1 \tilde{B}_W(r) dr, 0 \right]$ by Corollary A1(i), $(n^{-1}\tilde{W}'\tilde{W})^{-1} \rightarrow \text{diag}[1, \mathbb{E}[\tilde{W}_t\tilde{W}_t']]$ a.s. by Corollary A1(i) and Assumption 1(vi), and $n^{-1/2}\tilde{W}'\Psi_\tau(U) = \int_0^1 dB_\psi(r, \tau) = O_P(1)$ by Corollary A1(ii). Therefore,

$$n^{-1}X'\Psi_\tau(U) - n^{-3/2}X'\tilde{W}(n^{-1}\tilde{W}'\tilde{W})^{-1}n^{-1/2}\tilde{W}'\Psi_\tau(U)$$

$$\Rightarrow \int_0^1 \tilde{B}_W(r) dB_\psi(r, \tau) - \int_0^1 \tilde{B}_W(r) dr \int_0^1 dB_\psi(r, \tau) = \int_0^1 \tilde{B}_W(r) dB_\psi(r, \tau).$$

This completes the proof.

\textbf{Lemma A4.} Under Assumptions 1 and 2,

(i)

$$n^{-1/2} \sum_{t=1}^{[n(\cdot)]} \left[ \psi_{\tau_1}[U_t(\tau_1)], \ldots, \psi_{\tau_n}[U_t(\tau_n)] \right]' \Rightarrow \left[ B_{\psi}(\cdot, \tau_1), \ldots, B_{\psi}(\cdot, \tau_n) \right]'; \quad \text{and}$$

(ii) $J(\tau) \Rightarrow J_\beta(\tau)$, where

$$J(\tau) := \left[ \begin{array}{c}
\left\{ f_{\tau_1} \left( 1 - \sum_{j=1}^{p} \phi_{j^*}(\tau_1) \right) \right\}^{-1} n^{-1}X' \left[ I - \tilde{W}(\tilde{W}'\tilde{W})^{-1}\tilde{W}' \right] \Psi_{\tau_1}(U) \\
\left\{ f_{\tau_2} \left( 1 - \sum_{j=1}^{p} \phi_{j^*}(\tau_2) \right) \right\}^{-1} n^{-1}X' \left[ I - \tilde{W}(\tilde{W}'\tilde{W})^{-1}\tilde{W}' \right] \Psi_{\tau_2}(U) \\
\vdots \\
\left\{ f_{\tau_n} \left( 1 - \sum_{j=1}^{p} \phi_{j^*}(\tau_n) \right) \right\}^{-1} n^{-1}X' \left[ I - \tilde{W}(\tilde{W}'\tilde{W})^{-1}\tilde{W}' \right] \Psi_{\tau_n}(U)
\end{array} \right].$$

\textit{Proof of Lemma A4:} (i) The result in Lemma A4(i) is obviously implied by Assumption 2.
(ii) We note from the proof of Corollary A1(v) that, for each \( j = 1, 2, \ldots, s \),
\[
\frac{1}{n} X' [I - \tilde{W} (\tilde{W}' \tilde{W})^{-1} \tilde{W}'] \Psi_{\tau_j} (U) = \frac{1}{n} X' \Psi_{\tau_j} (U) - \left[ \frac{1}{n^{3/2}} \sum_{t=1}^{n} X'_t \right] \left[ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \psi_{\tau_j} (U_t) \right] = o_{\mathbb{P}} (1),
\]
where we use that for each \( j \), \( n^{-1/2} \tilde{W} \Psi_{\tau_j} (U) = O_{\mathbb{P}} (1) \) by Assumption 2. Applying Lemma A4(i) and the continuous mapping theorem, we obtain the desired result in Lemma A4(ii). \( \blacksquare \)

Although Lemmas A3 and A4, and Corollary A1 state the weak convergency results as if they are independent, their weak limits are jointly achieved under Assumptions 1 and 2. This is mainly because the weak limits in Lemmas A3 and A4, and Corollary A1 are the variations of the weak limits jointly obtained by Assumptions 1 and 2, respectively.

Using Lemmas A1–A4, we now prove the main results: Theorems 1–4.

Proof of Theorem 1: (i) We first note that
\[
\varrho_{\tau} \{ Y_t - Z_t' \tilde{\alpha}_n (\tau) \} = \varrho_{\tau} \{ U_t (\tau) - D_n^{-1} \tilde{v}_n (\tau)' Z_t \},
\]
where we let \( \tilde{v}_n (\tau) := D_n \{ \tilde{\alpha}_n (\tau) - \alpha_*(\tau) \} \) and \( D_n := \text{diag} \left( [\sqrt{n} t'_{1+q,k}, n t'_{k+p}] \right) \). Thus, minimizing \( \sum_{t=1}^{n} \varrho_{\tau} \{ Y_t - Z_t' \alpha \} \) with respect to \( \alpha \) is equivalent to minimizing
\[
Q_{\tau,n}(v) := \sum_{t=1}^{n} \left[ \varrho_{\tau} \{ U_t (\tau) - (D_n^{-1} v)' Z_t \} - \varrho_{\tau} \{ U_t (\tau) \} \right]
\]
with respect to \( v \). Notice that this objective function can be rewritten as
\[
Q_{\tau,n}(v) = - \sum_{t=1}^{n} v' D_n^{-1} Z_t \psi_{\tau} [U_t (\tau)]
\]
\[
+ \sum_{t=1}^{n} \{ U_t (\tau) - v' D_n^{-1} Z_t \} I \left[ v' D_n^{-1} Z_t < U_t (\tau) < 0 \right]
\]
\[
- \sum_{t=1}^{n} \{ U_t (\tau) - v' D_n^{-1} Z_t \} I \left[ 0 < U_t (\tau) < v' D_n^{-1} Z_t \right]
\]
where \( \varrho_{\tau} (x - y) - \varrho_{\tau} (x) = -y \varrho_{\tau} (x) + (x - y) [I(0 > x > y) - I(0 < x < y)] \) for \( x \neq 0 \).

We derive the asymptotic behavior of each element in the RHS of (30) by combining the techniques in Pesaran and Shin (1998) and Xiao (2009). First, notice that Assumptions 1(i, ii, iii, and vi) imply Assump-
tions A, B, and C of Xiao (2009). Thus, we can use his results to prove Theorem 1 as follows:

\[
\sum_{t=1}^{n} \{ U_t(\tau) - \nu_d^{-1}Z_t \} I \left[ \nu_d^{-1}Z_t < U_t(\tau) < 0 \right] = \frac{1}{2} \psi \int_{\nu_d^{-1}(Z'd)\nu_d^{-1}v} I \left[ \nu_d^{-1}Z_t < 0 \right] + o_p(1)
\]

\[
\sum_{t=1}^{n} \{ \nu_d^{-1}Z_t - U_t(\tau) \} I \left[ 0 < U_t(\tau) < \nu_d^{-1}Z_t \right] = \frac{1}{2} \psi \int_{\nu_d^{-1}(Z'd)\nu_d^{-1}v} I \left[ \nu_d^{-1}Z_t > 0 \right] + o_p(1),
\]

where \( Z := [Z_1, Z_2, \ldots, Z_n]' \). Hence, \( Q(n)(v) = -\nu_d^{-1} \left( \sum_{t=1}^{n} Z_t \psi[U_t(\tau)] \right) + \frac{1}{2} \psi \int_{\nu_d^{-1}(Z'd)\nu_d^{-1}v} I \left[ \nu_d^{-1}Z_t > 0 \right] + o_p(1) \), implying that \( D_{\nu_d^{-1}v} = \tilde{\alpha}_n(\tau) - \alpha_n(\tau) = f^{-1}_\tau(Z'd)^{-1}Z'\bar{\Psi}_\tau(U) + o_p(1) \).

Next, partitioning \( Z \) into \([G, \bar{Y}]\), where \( G := [G_1, G_2, \ldots, G_n]' \) and \( \bar{Y} := [\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_n]' \), we can show that

\[
\tilde{\Phi}_n(\tau) - \Phi_\tau(\tau) = f^{-1}_\tau[\bar{Y}'P_G \bar{Y}]^{-1}\bar{Y}'P_G \Psi_\tau(U),
\]

where \( P_G := I - G(G'G)^{-1}G' \) and \( \tilde{\alpha}_n(\tau) - \alpha_n(\tau) = \left[ \left( \tilde{\lambda}_n(\tau) - \lambda_n(\tau) \right)' , \left( \tilde{\phi}_n(\tau) - \phi_n(\tau) \right)' \right]' \). We also note that \( \bar{Y} = G\Gamma(\tau) + K(\tau) \), where \( K(\tau) := K(\tau) := [K_1(\tau), \ldots, K_n(\tau)]' \). Since \( P_GG = 0 \), hence

\[
\tilde{Y}'P_G \bar{Y} = K(\tau)'K(\tau) - K(\tau)'G(G'G)^{-1}G'K(\tau),
\]

\[
\tilde{Y}'P_G \Psi_\tau(U) = K(\tau)'\Psi_\tau(U) - K(\tau)'G(G'G)^{-1}G'\Psi_\tau(U).
\]

Furthermore, \( n^{-1}K(\tau)'K(\tau) \to E[K_i(\tau)K_i(\tau)'] \) a.s. by Lemma A2(iii), and \( K(\tau)'GD_{G}^{-1}(D_{G}^{-1}G'G D_{G}^{-1})^{-1}D_{G}^{-1}G'K(\tau) = O_p(n) \) by Corollary A1(iii). More specifically, it follows that \( n^{-1/2}K(\tau)'GD_{G}^{-1} \to_p E[K_i(\tau), E[K_i(\tau)W_i]' , E[K_i(\tau)] W_i]' B_i(\tau) d\tau \). Furthermore, we note that \( n^{-1/2}D_{D} = D_H \) and from this, it follows that \( K(\tau)'GD_{H}^{-1}(D_{G}^{-1}G'G D_{G}^{-1})^{-1}D_{H}^{-1}G'K(\tau) \to_p E[K_i(\tau)\bar{W}_i]' \bar{W}_i]'^{-1} \bar{W}_iK_i(\tau)' \) by Corollary A1(i). This implies that

\[
n^{-1} \tilde{Y}'P_G \bar{Y} \to_p E[K_i(\tau)K_i(\tau)'] - E[K_i(\tau)\bar{W}_i]' \bar{W}_i]'^{-1} \bar{W}_iK_i(\tau)' = E[\bar{H}_i(\tau)\bar{K}_i(\tau)'].
\]

In a similar manner, \( K(\tau)'GD_{G}^{-1}(D_{G}^{-1}G'G D_{G}^{-1})^{-1}D_{G}^{-1}G'\Psi_\tau(U) = O_p(\sqrt{n}) \) by Corollaries A1(i,ii,iii). Hence,

\[
n^{-1/2}K(\tau)'GD_{G}^{-1}(D_{G}^{-1}G'G D_{G}^{-1})^{-1}D_{G}^{-1}G'\Psi_\tau(U) = E[K_i(\tau)\bar{W}_i]' \bar{W}_i]'^{-1} \left( n^{-1/2} \bar{W}_i\Psi_\tau(U) \right) + o_p(1),
\]

40
which implies that

\[ n^{-1/2} \hat{Y}'P_G \Psi_\tau(U) = n^{-1/2} K(\tau)' \Psi_\tau(U) - E \left[ K_\tau(\tau) \hat{W}_\tau \right] E \left[ \hat{W}_\tau \hat{W}_\tau' \right]^{-1} n^{-1/2} \tilde{W}' \Psi_\tau(U) + o_p(1) \]

\[ = n^{-1/2} \tilde{H}(\tau)' \Psi_\tau(U) + o_p(1) \sim N \left\{ 0, \tau(1 - \tau) E \left[ \tilde{H}_\tau(\tau) \tilde{H}_\tau(\tau)' \right]^{-1} \right\} \]

according to Assumption 1(vi). This in turn implies that

\[ \sqrt{n} \left( \tilde{\phi}_n(\tau) - \phi_*(\tau) \right) \sim N \left[ 0, \tau(1 - \tau) f^{-2}_\tau E \left[ \tilde{H}_\tau(\tau) \tilde{H}_\tau(\tau)' \right]^{-1} \right]. \]

We note that the asymptotic variance is identical to \( \Pi(\tau) \) by definition.

(ii) To prove Theorem 1(ii), we use the fact that \( \tilde{Y} = G \Gamma_*(\tau) + K(\tau) \). We then show that

\[ \tilde{\lambda}_n(\tau) - \lambda_*(\tau) = f^{-1}_\tau (G'G)^{-1} G' \Psi(U) - (G'G)^{-1} G' \tilde{Y} \left( \tilde{\phi}_n(\tau) - \phi_*(\tau) \right) \]

\[ = f^{-1}_\tau (G'G)^{-1} G' \Psi(U) - \Gamma_*(\tau) \left( \tilde{\phi}_n(\tau) - \phi_*(\tau) \right) - (G'G)^{-1} G' K(\tau) \left( \tilde{\phi}_n(\tau) - \phi_*(\tau) \right). \] (31)

If we let \( \tilde{\varsigma}_n(\tau) := \tilde{\lambda}_n(\tau) + \Gamma_*(\tau) \tilde{\phi}_n(\tau) \) and \( \varsigma_*(\tau) := \lambda_*(\tau) + \Gamma_*(\tau) \phi_*(\tau) \), it easily follows that \( \tilde{\varsigma}_n(\tau) - \varsigma_*(\tau) := (\tilde{\lambda}_n(\tau) - \lambda_*(\tau)) + \Gamma_*(\tau) \left( \tilde{\phi}_n(\tau) - \phi_*(\tau) \right) \), so we obtain from (31) that

\[ \sqrt{n} (\tilde{\varsigma}_n(\tau) - \varsigma_*(\tau)) = \sqrt{n} f^{-1}_\tau D^{-1}_G (D^{-1}_G G'G^{-1}_G)^{-1} D^{-1}_G G' \Psi_\tau(U) \]

\[ - \sqrt{n} D^{-1}_G (D^{-1}_G G'G^{-1}_G)^{-1} D^{-1}_G G' K(\tau) \left( \tilde{\phi}_n(\tau) - \phi_*(\tau) \right). \]

Furthermore, Corollaries A1(i, ii) imply that \( D^{-1}_G G'G^{-1}_G = O_p(1) \) and \( D^{-1}_G G' \Psi_\tau(U) = O_p(1) \). Also, \( D^{-1}_G G' K(\tau) = O_p(1) \) by Corollary A1(iii). We have already shown that Theorem 1(i) implies that \( \tilde{\phi}_n(\tau) - \phi_*(\tau) = o_p(1) \) and \( \sqrt{n} D^{-1}_G = O(1) \). By combining all of these results, we obtain that

\[ \sqrt{n} (\tilde{\varsigma}_n(\tau) - \varsigma_*(\tau)) = \sqrt{n} f^{-1}_\tau D^{-1}_G (D^{-1}_G G'G^{-1}_G)^{-1} D^{-1}_G G' \Psi_\tau(U) + o_p(1) \]

\[ = f^{-1}_\tau N \left( D^{-1}_G G'G^{-1}_G \right)^{-1} D^{-1}_G G' \Psi_\tau(U) + o_p(1), \] (32)

where \( N := \text{diag}(k'_{1+qk}, 0'_{k \times 1}) \). Thus, the last \( k \) elements of \( \sqrt{n} (\tilde{\varsigma}_n(\tau) - \varsigma_*(\tau)) \) are \( o_p(1) \), which implies that

\[ \sqrt{n} \left\{ \left( \tilde{\gamma}_n(\tau) - \gamma_*(\tau) \right) + \beta_*(\tau) t'_p \left( \tilde{\phi}_n(\tau) - \phi_*(\tau) \right) \right\} = o_p(1) \]
because

\[
(\bar{\zeta}_n(\tau) - \zeta_*(\tau)) = \begin{bmatrix}
\bar{\alpha}_n(\tau) - \alpha_*(\tau) \\
\bar{\delta}_{0,n}(\tau) - \delta_{0*}(\tau) \\
\vdots \\
\bar{\delta}_{q-1,n}(\tau) - \delta_{q-1*}(\tau) \\
\bar{\gamma}_n(\tau) - \gamma_*(\tau)
\end{bmatrix} + \begin{bmatrix}
\mu_*(\tau) & \mu_*(\tau) & \cdots & \mu_*(\tau) \\
\xi_{1,0*}(\tau) & \xi_{2,0*}(\tau) & \cdots & \xi_{p,0*}(\tau) \\
\vdots & \vdots & \ddots & \vdots \\
\xi_{1,q-1*}(\tau) & \xi_{2,q-1*}(\tau) & \cdots & \xi_{p,q-1*}(\tau) \\
\beta_*(\tau) & \beta_*(\tau) & \cdots & \beta_*(\tau)
\end{bmatrix} \begin{pmatrix}
\bar{\phi}_n(\tau) - \phi_*(\tau)
\end{pmatrix}.
\]

By the asymptotic result in Theorem 1(i), the desired result in Theorem 1(ii) follows. □

**Proof of Theorem 2:** We focus on the asymptotic behavior of

\[
n \left\{ (\bar{\gamma}_n(\tau) - \gamma_*(\tau)) + \beta_*(\tau) \sum_{j=1}^{p} (\bar{\phi}_{j,n}(\tau) - \phi_{j*}(\tau)) \right\},
\]

which is equal to the last \(k\) elements of \(D_G (\bar{\zeta}_n(\tau) - \zeta_*(\tau))\) by definition. Notice that

\[
D_G (\bar{\zeta}_n(\tau) - \zeta_*(\tau)) = (D_G^{-1} G' G D_G^{-1})^{-1} \left\{ f_\tau^{-1} D_G^{-1} G' \Psi_\tau(U) - D_G^{-1} G' K(\tau) (\bar{\phi}_n(\tau) - \phi_*(\tau)) \right\} \\
= f_\tau^{-1} (D_G^{-1} G' G D_G^{-1})^{-1} D_G^{-1} G' \Psi_\tau(U) + o_p(1)
\]

because \(D_G^{-1} G' G D_G^{-1} = O_p(1), D_G^{-1} G' K(\tau) = O_p(1), \) and \(\bar{\phi}_n(\tau) - \phi_*(\tau) = O_p(n^{-1/2})\) by Corollaries A1(i), A1(iii), and Theorem 1(i), respectively. The last \(k\) elements of \((D_G^{-1} G' G D_G^{-1})^{-1} D_G^{-1} G' \Psi_\tau(U)\) are also equal to \(M^{-1} n^{-1} X' [I - \bar{W} (\bar{W}' \bar{W})^{-1} \bar{W}'] \Psi_\tau(U)\). Therefore, (33) implies that

\[
n \left\{ (\bar{\gamma}_n(\tau) - \gamma_*(\tau)) + \beta_*(\tau) \sum_{j=1}^{p} (\bar{\phi}_{j,n}(\tau) - \phi_{j*}(\tau)) \right\} \\
= f_\tau^{-1} M^{-1} \left\{ n^{-1} X' [I - \bar{W} (\bar{W}' \bar{W})^{-1} \bar{W}'] \Psi_\tau(U) \right\} + o_p(1) = O_p(1).
\]

Second, we focus on the relationship between

\[
\left\{ (\bar{\gamma}_n(\tau) - \gamma_*(\tau)) + \beta_*(\tau) \sum_{j=1}^{p} (\bar{\phi}_{j,n}(\tau) - \phi_{j*}(\tau)) \right\}
\]
and \( \left( \tilde{\beta}_n(\tau) - \beta_*(\tau) \right) \). Using the identity in (11), we have

\[
n \left( \tilde{\beta}_n(\tau) - \beta_*(\tau) \right) = \left\{ 1 - \sum_{j=1}^{p} \phi_{js}(\tau) \right\}^{-1} n \left\{ (\bar{\gamma}_n(\tau) - \gamma_*(\tau)) + \beta_*(\tau) \sum_{j=1}^{p} \left( \bar{\phi}_{j,n}(\tau) - \phi_{js}(\tau) \right) \right\} + o_P(1) \tag{35}
\]

because Theorem 1(i) implies that \( \bar{\phi}_n(\tau) = \phi_*(\tau) + o_P(1) \), \( \left| \sum_{j=1}^{q} \phi_{js}(\tau) \right| < 1 \) by Assumption 1(v), and

\[
n \{ (\bar{\gamma}_n(\tau) - \gamma_*(\tau)) + \beta_*(\tau) \sum_{j=1}^{p} \left( \bar{\phi}_{j,n}(\tau) - \phi_{js}(\tau) \right) \} = O_P(1) \text{ by (34).} \]

Finally, combining the results in (34) and (35), we obtain:

\[
n \left( \tilde{\beta}_n(\tau) - \beta_*(\tau) \right) = \left\{ f_\tau \left( 1 - \sum_{j=1}^{p} \phi_{js}(\tau) \right) \right\}^{-1} M^{-1} \left\{ n^{-1} X'[I - \tilde{W}(\tilde{W}'\tilde{W})^{-1}\tilde{W}']\Psi_\tau(U) \right\} + o_P(1) \tag{36}
\]

the result of which, combined with Corollary A1(iv,v), and the continuous mapping theorem, delivers the desired result in Theorem 2(ii).

We turn to proving Theorem 2(ii). By the result in (36) and using the fact that \( M = O_P(1) \) by Corollary A1(iv), we obtain that

\[
n M^{1/2} \left( \tilde{\beta}_n(\tau) - \beta_*(\tau) \right) = \left\{ f_\tau \left( 1 - \sum_{j=1}^{p} \phi_{js}(\tau) \right) \right\}^{-1} M^{-1/2} \left\{ n^{-1} X'[I - \tilde{W}(\tilde{W}'\tilde{W})^{-1}\tilde{W}']\Psi_\tau(U) \right\} + o_P(1).
\]

Corollary A1(iv,v) provides the asymptotic limits of \( M \) and \( n^{-1} X'[I - \tilde{W}(\tilde{W}'\tilde{W})^{-1}\tilde{W}']\Psi_\tau(U) \). Noting that \( \tilde{W} \) and \( \Psi_\tau(U) \) are independent, the desired result in Theorem 2(ii) follows using the mixture normality of Phillips (1991b). ■

Proof of Theorem 3: (i) It follows from (8) that for each \( j = 1, 2, \ldots, s \),

\[
\sqrt{n} \left( \bar{\phi}_n(\tau_j) - \phi_*(\tau_j) \right) = f_{\tau_j}^{-1} \mathbb{E} \left[ \tilde{H}_t(\tau_j)\tilde{H}_t(\tau_j)' \right]^{-1} \left( n^{-1/2} \tilde{H}(\tau_j)'\Psi_{\tau_j}(U) \right) + o_P(1). \tag{37}
\]

Assumption 1(vi) implies that \( \mathbb{E} \left[ \tilde{H}_t(\tau_j)\tilde{H}_t(\tau_j)' \right] \) is positive definite. Using (37), we also note that the
asymptotic covariance matrix between $\sqrt{n} \left( \phi_n(\tau) - \phi_s(\tau) \right)$ and $\sqrt{n} \left( \tilde{\phi}_n(\tau) - \phi_s(\tau) \right)$ is obtained as

$$f_{\tau_i}^{-1} f_{\tau_j}^{-1} (\min[\tau_i, \tau_j] - \tau_i \tau_j) E \left[ \mathbf{H}_t(\tau_i) \mathbf{H}_t(\tau_j) \right]^{-1} E \left[ \tilde{\mathbf{H}}_t(\tau_i) \tilde{\mathbf{H}}_t(\tau_j) \right] E \left[ \tilde{\mathbf{H}}_t(\tau_j) \tilde{\mathbf{H}}_t(\tau_j)' \right]^{-1}$$

$$= f_{\tau_i}^{-1} f_{\tau_j}^{-1} (\min[\tau_i, \tau_j] - \tau_i \tau_j) I(\tau_i, \tau_j)^{-1} L(\tau_i, \tau_j) L(\tau_j, \tau_j)^{-1}$$

by the definition of $L(\tau_i, \tau_j)$. That is, the $i$-th row and $j$-th column block matrix of $\Xi(\tau)$ are obtained. By Assumption 2(i), $\Xi(\tau)$ is positive definite, and we can apply the multivariate CLT using this to obtain that

$$\sqrt{n} \left( \tilde{\Phi}_n(\tau) - \Phi_s(\tau) \right) \xrightarrow{p} N[0, \Xi(\tau)].$$

(ii) By (10), for each $j = 1, \ldots, s$, $\sqrt{n} \left\{ (\tilde{\gamma}_n(\tau_j) - \gamma_s(\tau_j)) + \beta_s(\tau_j) \phi_s(\tau_j) - \phi_s(\tau_j) \right\} = o_P(1)$, which also implies that $\sqrt{n} \left[ \tilde{\Gamma}_n(\tau) - \Gamma_s(\tau) \right] = -\sqrt{n} \left[ \Lambda(\tau) (\Phi_n(\tau) - \Phi_s(\tau)) \right] + o_P(1)$. Hence, Theorem 3(i) implies that $\sqrt{n} \left[ \tilde{\Gamma}_n(\tau) - \Gamma_s(\tau) \right] \xrightarrow{p} N(0, \Lambda(\tau) \Xi(\tau) \Lambda(\tau)')$. This completes the proof. ■

Proof of Theorem 4: (i) It is easily seen from the definition of $\tilde{B}_n(\tau)$ and the result in (36) that

$$n \left[ \tilde{B}_n(\tau) - B_s(\tau) \right] = [I_s \otimes M^{-1}] J(\tau) + o_P(1), \tag{38}$$

and $J(\tau) \Rightarrow J_\beta(\tau)$ by Lemma A4(ii). Thus, Corollary A1(iv) implies that

$$n \left[ \tilde{B}_n(\tau) - B_s(\tau) \right] \Rightarrow [I_s \otimes \left( \int_0^1 \tilde{B}_W(r) \tilde{B}_W(r)' dr \right)^{-1}] J_\beta(\tau).$$

(ii) Using (38), we note that $\text{rank}[I_s \otimes M^{1/2}] = ks$ implies that $n[I_s \otimes M^{1/2}] \left[ \tilde{B}_n(\tau) - B_s(\tau) \right] = [I_s \otimes M^{-1/2}] J(\tau) + o_P(1)$. Hence, applying the continuous mapping theorem, we obtain:

$$n \left[ I_s \otimes M^{1/2} \right] \left[ \tilde{B}_n(\tau) - B_s(\tau) \right] \Rightarrow [I_s \otimes \left( \int_0^1 \tilde{B}_W(r) \tilde{B}_W(r)' dr \right)^{-1/2}] J_\beta(\tau).$$

For each $i = 1, 2, \ldots, s$, $B_\psi(\cdot, \tau_i)$ is independent of $\tilde{B}_W(\cdot)$, so that we can obtain the desired result by applying the mixture interpretation of Phillips (1991b). ■

References


Table 1: Empirical Levels of $W_{n}^{(1)}(B)$, $W_{n}^{(2)}(B)$, $W_{n}^{(3)}(B)$, and $W_{n}^{(4)}(B)$. Notes: (i) Number of iterations: 5,000. (ii) DGP: $Y_{t} = \alpha_{s} + \phi_{s} Y_{t-1} + \theta_{0s} X_{t} + \theta_{1s} X_{t-1} + U_{t}$, $X_{t} := X_{t-1} + W_{t}$, $W_{t} := \rho_{s} R_{t-1} + (1 - \rho_{s}^{2}) R_{t}$, and $(U_{t}, R_{t})' \sim$ IID $N(0, I_{2})$. (iii) Model: $Y_{t} = \alpha_{s}(\tau) + W_{t}' \delta_{s}(\tau) + X_{t}' \gamma_{s}(\tau) + \phi_{s}(\tau) Y_{t-1} + U_{t}(\tau)$. (iv) The null hypothesis: $s_{1} = s_{2} = s_{3} = 0$, and $s_{4} = [0, 0]'$ for $W_{n}^{(1)}(B)$, $W_{n}^{(2)}(B)$, $W_{n}^{(3)}(B)$, and $W_{n}^{(4)}(B)$, respectively.
<table>
<thead>
<tr>
<th>Level \ Sample Size</th>
<th>50</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1,000</th>
<th>2,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>3.22</td>
<td>1.96</td>
<td>1.20</td>
<td>1.42</td>
<td>1.32</td>
<td>1.20</td>
<td>1.04</td>
</tr>
<tr>
<td>5%</td>
<td>9.80</td>
<td>6.68</td>
<td>5.88</td>
<td>6.08</td>
<td>4.98</td>
<td>4.82</td>
<td>5.40</td>
</tr>
<tr>
<td>10%</td>
<td>16.70</td>
<td>12.36</td>
<td>11.42</td>
<td>11.40</td>
<td>9.72</td>
<td>10.54</td>
<td>10.88</td>
</tr>
</tbody>
</table>

\( W_n^{(1)}(\Phi) \)

<table>
<thead>
<tr>
<th>Level \ Sample Size</th>
<th>50</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1,000</th>
<th>2,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>3.76</td>
<td>1.50</td>
<td>1.24</td>
<td>1.30</td>
<td>0.98</td>
<td>1.32</td>
<td>0.98</td>
</tr>
<tr>
<td>5%</td>
<td>10.66</td>
<td>6.44</td>
<td>5.74</td>
<td>5.88</td>
<td>5.26</td>
<td>5.44</td>
<td>5.14</td>
</tr>
<tr>
<td>10%</td>
<td>17.50</td>
<td>11.96</td>
<td>11.34</td>
<td>11.40</td>
<td>9.72</td>
<td>10.54</td>
<td>10.88</td>
</tr>
</tbody>
</table>

\( W_n^{(2)}(\Phi) \)

<table>
<thead>
<tr>
<th>Level \ Sample Size</th>
<th>50</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1,000</th>
<th>2,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>6.38</td>
<td>2.48</td>
<td>1.46</td>
<td>1.72</td>
<td>1.32</td>
<td>1.50</td>
<td>1.18</td>
</tr>
<tr>
<td>5%</td>
<td>15.04</td>
<td>8.24</td>
<td>6.08</td>
<td>6.74</td>
<td>5.38</td>
<td>5.92</td>
<td>5.84</td>
</tr>
<tr>
<td>10%</td>
<td>23.54</td>
<td>14.40</td>
<td>11.90</td>
<td>12.14</td>
<td>10.02</td>
<td>10.84</td>
<td>11.14</td>
</tr>
</tbody>
</table>

\( W_n^{(3)}(\Phi) \)

<table>
<thead>
<tr>
<th>Level \ Sample Size</th>
<th>50</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1,000</th>
<th>2,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>6.22</td>
<td>2.22</td>
<td>1.46</td>
<td>1.44</td>
<td>1.16</td>
<td>1.46</td>
<td>0.94</td>
</tr>
<tr>
<td>5%</td>
<td>14.90</td>
<td>8.40</td>
<td>6.28</td>
<td>6.96</td>
<td>5.54</td>
<td>5.98</td>
<td>5.76</td>
</tr>
<tr>
<td>10%</td>
<td>22.94</td>
<td>13.68</td>
<td>11.98</td>
<td>12.78</td>
<td>10.30</td>
<td>10.84</td>
<td>10.84</td>
</tr>
</tbody>
</table>

\( W_n^{(4)}(\Phi) \)

Table 2: **Empirical Levels of** \( W_n^{(1)}(\Phi) \), \( W_n^{(2)}(\Phi) \), \( W_n^{(3)}(\Phi) \), and \( W_n^{(4)}(\Phi) \). **Notes:** (i) Number of iterations: 5,000. (ii) DGP: \( Y_t = \alpha + \phi Y_{t-1} + \theta_0 x_t + \theta_1 x_{t-1} + U_t \), \( X_t := X_{t-1} + W_t \), \( W_t := \rho_s R_{t-1} + (1 - \rho_s^2) R_t \), and \( (U_t, R_t)' \sim \text{IID } N(0, I_2) \). (iii) Model: \( Y_t = \alpha(\tau) + W_t' \delta_s(\tau) + X_t' \gamma_s(\tau) + \phi(\tau) Y_{t-1} + U_t(\tau) \). (iv) The null hypothesis: \( s_1 = s_2 = s_3 = 0 \), and \( s_4 = [0, 0]' \) for \( W_n^{(1)}(\Phi) \), \( W_n^{(2)}(\Phi) \), \( W_n^{(3)}(\Phi) \), and \( W_n^{(4)}(\Phi) \), respectively.
<table>
<thead>
<tr>
<th>Level \ Sample Size</th>
<th>$W_n^{(1)}(\Gamma)$</th>
<th>$W_n^{(2)}(\Gamma)$</th>
<th>$W_n^{(3)}(\Gamma)$</th>
<th>$W_n^{(4)}(\Gamma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>1%</td>
<td>2.44</td>
<td>1.62</td>
<td>1.20</td>
<td>1.36</td>
</tr>
<tr>
<td>5%</td>
<td>8.08</td>
<td>6.64</td>
<td>5.90</td>
<td>6.00</td>
</tr>
<tr>
<td>10%</td>
<td>14.62</td>
<td>12.26</td>
<td>11.26</td>
<td>11.08</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>1%</td>
<td>2.60</td>
<td>1.40</td>
<td>1.18</td>
<td>1.22</td>
</tr>
<tr>
<td>5%</td>
<td>8.74</td>
<td>6.04</td>
<td>5.58</td>
<td>6.04</td>
</tr>
<tr>
<td>10%</td>
<td>14.94</td>
<td>11.56</td>
<td>11.14</td>
<td>10.82</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>1%</td>
<td>5.54</td>
<td>2.16</td>
<td>1.40</td>
<td>1.60</td>
</tr>
<tr>
<td>5%</td>
<td>13.54</td>
<td>7.84</td>
<td>5.84</td>
<td>6.60</td>
</tr>
<tr>
<td>10%</td>
<td>20.88</td>
<td>13.82</td>
<td>11.68</td>
<td>12.00</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>1%</td>
<td>4.60</td>
<td>1.92</td>
<td>1.40</td>
<td>1.44</td>
</tr>
<tr>
<td>5%</td>
<td>12.70</td>
<td>7.68</td>
<td>6.14</td>
<td>6.82</td>
</tr>
<tr>
<td>10%</td>
<td>19.06</td>
<td>13.24</td>
<td>11.48</td>
<td>12.56</td>
</tr>
</tbody>
</table>

Table 3: **Empirical Levels of $W_n^{(1)}(\Gamma)$, $W_n^{(2)}(\Gamma)$, $W_n^{(3)}(\Gamma)$, and $W_n^{(4)}(\Gamma)$**. Notes: (i) Number of iterations: 5,000. (ii) DGP: $Y_t = \alpha + \phi Y_{t-1} + \theta_0 X_t + \theta_1 X_{t-1} + U_t$, $X_t := X_{t-1} + W_t$, $W_t := \rho R_{t-1} + (1 - \rho^2) R_t$, and $(U_t, R_t)' \sim$ IID $N(0, I_2)$. (iii) Model: $Y_t = \alpha(\tau) + W_t^\prime \delta_s(\tau) + X_t^\prime \gamma_s(\tau) + \phi(\tau) Y_{t-1} + U_t(\tau)$. (iv) The null hypothesis: $s_1 = s_2 = s_3 = 0$, and $s_4 = [0, 0]'$ for $W_n^{(1)}(\Gamma)$, $W_n^{(2)}(\Gamma)$, $W_n^{(3)}(\Gamma)$, and $W_n^{(4)}(\Gamma)$, respectively.
<table>
<thead>
<tr>
<th>Level \ Sample Size</th>
<th>( W_n^{(1)}(B) )</th>
<th>( W_n^{(2)}(B) )</th>
<th>( W_n^{(3)}(B) )</th>
<th>( W_n^{(4)}(B) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1% )</td>
<td>25.05 61.60 94.15 99.45 99.95 100.0</td>
<td>27.30 60.85 94.95 99.35 99.90 100.0</td>
<td>23.20 49.90 86.45 97.20 99.95 100.0</td>
<td>50.15 83.45 99.30 100.0 100.0 100.0</td>
</tr>
<tr>
<td>( 5% )</td>
<td>41.75 74.90 97.80 99.85 99.95 100.0</td>
<td>42.50 75.00 98.00 99.75 99.90 100.0</td>
<td>37.00 64.15 93.30 98.95 99.85 100.0</td>
<td>64.20 90.20 99.70 100.0 100.0 100.0</td>
</tr>
<tr>
<td>( 10% )</td>
<td>51.25 80.80 98.65 99.90 100.0 100.0</td>
<td>51.50 81.25 98.70 99.90 100.0 100.0</td>
<td>46.70 71.30 95.65 99.30 99.85 100.0</td>
<td>70.65 93.00 99.80 100.0 100.0 100.0</td>
</tr>
</tbody>
</table>

Table 4: **Empirical Powers of** \( W_n^{(1)}(B), W_n^{(2)}(B), W_n^{(3)}(B), \) and \( W_n^{(4)}(B) \). Notes: (i) Number of iterations: 5,000. (ii) DGP: \( Y_t = \alpha + \phi Y_{t-1} + \theta_0 X_t + \theta_1 X_{t-1} + U_t, X_t := X_{t-1} + W_t, W_t := \rho \cdot R_{t-1} + (1 - \rho^2) R_t, \) and \((U_t, R_t)' \sim \text{IID } N(0, \mathbf{I}_2)\). (iii) Model: \( Y_t' = \alpha_s(\tau) + W_t' \delta_s(\tau) + X_t' \gamma_s(\tau) + \phi_s(\tau) Y_{t-1} + U_t(\tau)\). (iv) The alternative hypothesis: \( s_1 = s_2 = s_3 = 0.1, \) and \( s_4 = [0.1, 0.1]' \) for \( W_n^{(1)}(B), W_n^{(2)}(B), W_n^{(3)}(B), \) and \( W_n^{(4)}(B) \), respectively.
<table>
<thead>
<tr>
<th>Level \ Sample Size</th>
<th>$W_n^{(1)}(\Phi)$</th>
<th>$W_n^{(2)}(\Phi)$</th>
<th>$W_n^{(3)}(\Phi)$</th>
<th>$W_n^{(4)}(\Phi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>62.04 89.32 99.74 99.98 100.0 100.0</td>
<td>61.50 88.52 99.62 100.0 100.0 100.0</td>
<td>48.12 71.18 94.76 99.46 99.92 100.0</td>
<td>93.68 99.78 100.0 100.0 100.0</td>
</tr>
<tr>
<td>5%</td>
<td>78.66 96.42 99.98 100.0 100.0 100.0</td>
<td>79.18 96.12 100.0 100.0 100.0 100.0</td>
<td>65.22 85.50 98.52 99.88 100.0 100.0</td>
<td>97.56 99.92 100.0 100.0</td>
</tr>
<tr>
<td>10%</td>
<td>85.46 98.06 100.0 100.0 100.0 100.0</td>
<td>85.68 97.90 100.0 100.0 100.0 100.0</td>
<td>74.20 90.78 99.83 99.92 100.0 100.0</td>
<td>98.62 99.98 100.0 100.0</td>
</tr>
</tbody>
</table>

Table 5: **Empirical Powers of** $W_n^{(1)}(\Phi)$, $W_n^{(2)}(\Phi)$, $W_n^{(3)}(\Phi)$, and $W_n^{(4)}(\Phi)$. **Notes:** (i) Number of iterations: 5,000. (ii) DGP: $Y_t = \alpha + \phi_* Y_{t-1} + \theta_0 X_t + \theta_1 X_{t-1} + U_t$, $X_t := X_{t-1} + W_t$, $W_t := \rho_* R_{t-1} + (1 - \rho_*^2) R_t$, and $(U_t, R_t)' \sim$ IID $N(0, I_2)$. (iii) Model: $Y'_t = \alpha_*(\tau) + W'_t \delta_*(\tau) + X'_t \gamma_*(\tau) + \phi_*(\tau) Y_{t-1} + U_t(\tau)$. (iv) The alternative hypothesis: $s_1 = s_2 = s_3 = 0.1$, and $s_4 = [0.1, 0.1]'$ for $W_n^{(1)}(\Phi)$, $W_n^{(2)}(\Phi)$, $W_n^{(3)}(\Phi)$, and $W_n^{(4)}(\Phi)$, respectively.
<table>
<thead>
<tr>
<th>Level \ Sample Size</th>
<th>( W_n^{(1)}(\Gamma) )</th>
<th>( W_n^{(2)}(\Gamma) )</th>
<th>( W_n^{(3)}(\Gamma) )</th>
<th>( W_n^{(4)}(\Gamma) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1% )</td>
<td>5.36 8.88 12.44 17.80 27.26 37.20 50.62</td>
<td>5.52 8.42 12.88 17.06 27.42 38.28 48.58</td>
<td>5.00 6.48 7.94 9.70 15.42 20.62 28.62</td>
<td>10.66 18.10 27.10 37.10 56.60 73.10 85.00</td>
</tr>
<tr>
<td>( 5% )</td>
<td>15.70 22.10 29.82 37.54 49.78 62.20 72.94</td>
<td>15.12 21.80 29.72 36.28 49.96 61.80 71.04</td>
<td>13.68 16.64 19.60 24.84 32.80 41.86 50.22</td>
<td>24.22 36.48 49.72 61.30 77.96 89.12 94.84</td>
</tr>
<tr>
<td>( 10% )</td>
<td>23.00 31.80 41.32 48.88 62.36 73.78 81.78</td>
<td>22.62 31.82 40.46 48.12 62.18 73.92 80.44</td>
<td>21.28 25.26 29.56 35.66 44.60 54.28 62.16</td>
<td>34.34 49.36 61.72 72.30 85.58 94.16 97.32</td>
</tr>
</tbody>
</table>

Table 6: Empirical Powers of \( W_n^{(1)}(\Gamma) \), \( W_n^{(2)}(\Gamma) \), \( W_n^{(3)}(\Gamma) \), and \( W_n^{(4)}(\Gamma) \). Notes: (i) Number of iterations: 5,000. (ii) DGP: \( Y_t = \alpha_0 + \phi_y Y_{t-1} + \theta_0 x_t + \theta_1 x_{t-1} + U_t \), \( X_t := X_{t-1} + W_t \), \( W_t := \rho_x R_{t-1} + (1 - \rho_x^2) R_t \), and \( (U_t, R_t)' \sim \text{IID } N(0, I_2) \). (iii) Model: \( Y_t = \alpha_\tau + W_t^\prime \delta_\tau + X_t^\prime \gamma_\tau + \phi_\tau Y_{t-1} + U_t(\tau) \). (iv) The alternative hypothesis: \( s_1 = s_2 = s_3 = 0.1 \), and \( s_4 = [0.1, 0.1]' \) for \( W_n^{(1)}(\Gamma) \), \( W_n^{(2)}(\Gamma) \), \( W_n^{(3)}(\Gamma) \), and \( W_n^{(4)}(\Gamma) \), respectively.
<table>
<thead>
<tr>
<th>$\sigma_\ast$ \ Sample Size</th>
<th>50</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1,000</th>
<th>2,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>13.60</td>
<td>7.68</td>
<td>6.68</td>
<td>5.68</td>
<td>6.26</td>
<td>5.92</td>
<td>5.28</td>
</tr>
<tr>
<td>0.10</td>
<td>13.46</td>
<td>7.66</td>
<td>7.10</td>
<td>6.48</td>
<td>5.90</td>
<td>5.74</td>
<td>5.12</td>
</tr>
<tr>
<td>0.20</td>
<td>12.98</td>
<td>7.70</td>
<td>6.76</td>
<td>5.74</td>
<td>5.78</td>
<td>5.92</td>
<td>5.28</td>
</tr>
<tr>
<td>0.30</td>
<td>14.32</td>
<td>7.52</td>
<td>7.32</td>
<td>6.60</td>
<td>6.22</td>
<td>5.80</td>
<td>5.40</td>
</tr>
<tr>
<td>0.40</td>
<td>13.08</td>
<td>8.28</td>
<td>7.24</td>
<td>6.00</td>
<td>6.48</td>
<td>6.06</td>
<td>5.90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\sigma_\ast$ \ Sample Size</th>
<th>50</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1,000</th>
<th>2,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>13.84</td>
<td>7.16</td>
<td>6.44</td>
<td>5.44</td>
<td>6.06</td>
<td>5.44</td>
<td>5.92</td>
</tr>
<tr>
<td>0.10</td>
<td>13.48</td>
<td>8.80</td>
<td>7.04</td>
<td>6.30</td>
<td>5.96</td>
<td>5.34</td>
<td>4.72</td>
</tr>
<tr>
<td>0.20</td>
<td>12.32</td>
<td>7.88</td>
<td>6.54</td>
<td>6.24</td>
<td>5.92</td>
<td>5.82</td>
<td>5.20</td>
</tr>
<tr>
<td>0.30</td>
<td>13.64</td>
<td>7.48</td>
<td>7.04</td>
<td>6.40</td>
<td>5.32</td>
<td>5.62</td>
<td>5.46</td>
</tr>
<tr>
<td>0.40</td>
<td>12.14</td>
<td>6.98</td>
<td>6.64</td>
<td>6.74</td>
<td>5.18</td>
<td>5.80</td>
<td>5.58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\sigma_\ast$ \ Sample Size</th>
<th>50</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1,000</th>
<th>2,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>21.44</td>
<td>9.00</td>
<td>7.56</td>
<td>6.34</td>
<td>6.92</td>
<td>6.12</td>
<td>5.62</td>
</tr>
<tr>
<td>0.10</td>
<td>20.38</td>
<td>9.36</td>
<td>8.32</td>
<td>7.18</td>
<td>6.58</td>
<td>6.60</td>
<td>4.98</td>
</tr>
<tr>
<td>0.20</td>
<td>19.96</td>
<td>9.90</td>
<td>7.52</td>
<td>6.68</td>
<td>5.88</td>
<td>6.76</td>
<td>5.22</td>
</tr>
<tr>
<td>0.30</td>
<td>21.50</td>
<td>9.86</td>
<td>8.28</td>
<td>7.24</td>
<td>6.24</td>
<td>6.60</td>
<td>5.80</td>
</tr>
<tr>
<td>0.40</td>
<td>20.36</td>
<td>10.04</td>
<td>8.52</td>
<td>7.44</td>
<td>6.82</td>
<td>6.52</td>
<td>6.58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\sigma_\ast$ \ Sample Size</th>
<th>50</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1,000</th>
<th>2,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>21.64</td>
<td>9.42</td>
<td>7.68</td>
<td>6.10</td>
<td>7.10</td>
<td>6.00</td>
<td>5.98</td>
</tr>
<tr>
<td>0.10</td>
<td>20.66</td>
<td>10.00</td>
<td>7.86</td>
<td>7.20</td>
<td>6.36</td>
<td>6.28</td>
<td>5.32</td>
</tr>
<tr>
<td>0.20</td>
<td>20.06</td>
<td>9.80</td>
<td>7.28</td>
<td>6.80</td>
<td>6.64</td>
<td>6.74</td>
<td>4.96</td>
</tr>
<tr>
<td>0.30</td>
<td>21.84</td>
<td>9.36</td>
<td>8.66</td>
<td>7.24</td>
<td>6.56</td>
<td>6.24</td>
<td>5.28</td>
</tr>
<tr>
<td>0.40</td>
<td>19.60</td>
<td>9.84</td>
<td>7.92</td>
<td>7.42</td>
<td>6.54</td>
<td>6.26</td>
<td>6.00</td>
</tr>
</tbody>
</table>

Table 7: Empirical Levels of $\tilde{W}_n^{(1)}(B)$, $\tilde{W}_n^{(2)}(B)$, $\tilde{W}_n^{(3)}(B)$, and $\tilde{W}_n^{(4)}(B)$ (level of significance: 5%). Notes: (i) Number of iterations: 5,000. (ii) DGP: $Y_t = \alpha_\ast + \phi_\ast Y_{t-1} + \theta_0 \ast X_t + \rho_\ast R_{t-1} + U_t$, $X_t := X_{t-1} + W_t$, $U_t := \sigma_\ast R_{t-1}^{(1)} + (1 - \sigma_\ast^2)R_t^{(2)}$, $W_t := \rho_\ast R_{t-1}^{(2)} + (1 - \rho_\ast^2)R_t^{(2)}$, and $(R_t^{(1)}, R_t^{(2)}) \sim$ IID $N(0, I_2)$. (iii) Model: $Y_t = \alpha_\ast(\tau) + \sum_{j=0}^{q-1} W_{t-j}^\prime \delta_{j+}(\tau) + X_t^\prime \gamma_\ast(\tau) + \sum_{j=1}^{p} \phi_{j+}(\tau) Y_{t-j} + U_t(\tau)$. (iv) The null hypothesis: $s_1 = s_2 = s_3 = 0$, and $s_4 = [0, 0]'$ for $\tilde{W}_n^{(1)}(B)$, $\tilde{W}_n^{(2)}(B)$, $\tilde{W}_n^{(3)}(B)$, and $\tilde{W}_n^{(4)}(B)$, respectively. (v) BIC is applied to determine the QARDL orders $p$ and $q$. 
Table 8: Empirical Levels of $\hat{\gamma}_n^{(1)}(\Phi)$, $\hat{\gamma}_n^{(2)}(\Phi)$, $\hat{\gamma}_n^{(3)}(\Phi)$, and $\hat{\gamma}_n^{(4)}(\Phi)$ (Level of Significance: 5%). Notes: (i) Number of iterations: 5,000. (ii) DGP: $Y_t = \alpha_\ast + \phi_\ast Y_{t-1} + \theta_0 X_t + \theta_1 X_{t-1} + U_t$, $X_t := X_{t-1} + W_t$, $U_t := \sigma_\ast R^{(2)}_t$, $W_t := \rho_\ast R^{(2)}_{t-1} + (1 - \rho_\ast^2) R^{(2)}_t$, and $(R^{(1)}_t, R^{(2)}_t) \sim \text{IID } \mathcal{N}(0, \mathbf{I}_2)$. (iii) Model: $Y_t = \alpha_\ast(\tau) + \sum_{j=0}^{q-1} W_{t-j} \delta_\ast(\tau) + X_t \gamma_\ast(\tau) + \sum_{j=1}^{p} \phi_\ast(\tau) Y_{t-j} + U_t(\tau)$. (iv) The null hypothesis: $s_1 = s_2 = s_3 = 0$, and $s_4 = [0, 0]'$ for $\hat{\gamma}_n^{(1)}(\Phi)$, $\hat{\gamma}_n^{(2)}(\Phi)$, $\hat{\gamma}_n^{(3)}(\Phi)$, and $\hat{\gamma}_n^{(4)}(\Phi)$, respectively. (v) BIC is applied to determine the QARDL orders $p$ and $q$. 

<table>
<thead>
<tr>
<th>$\sigma_\ast$ \ Sample Size</th>
<th>$\hat{\gamma}_n^{(1)}(\Phi)$</th>
<th>$\hat{\gamma}_n^{(2)}(\Phi)$</th>
<th>$\hat{\gamma}_n^{(3)}(\Phi)$</th>
<th>$\hat{\gamma}_n^{(4)}(\Phi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>0.00</td>
<td>12.58</td>
<td>7.48</td>
<td>6.70</td>
<td>5.54</td>
</tr>
<tr>
<td>0.10</td>
<td>12.66</td>
<td>8.14</td>
<td>6.84</td>
<td>6.22</td>
</tr>
<tr>
<td>0.20</td>
<td>12.00</td>
<td>7.26</td>
<td>6.74</td>
<td>6.32</td>
</tr>
<tr>
<td>0.30</td>
<td>10.88</td>
<td>7.32</td>
<td>6.62</td>
<td>6.24</td>
</tr>
<tr>
<td>0.40</td>
<td>10.00</td>
<td>7.04</td>
<td>6.22</td>
<td>5.94</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>0.00</td>
<td>13.66</td>
<td>7.56</td>
<td>6.20</td>
<td>5.78</td>
</tr>
<tr>
<td>0.10</td>
<td>11.42</td>
<td>7.62</td>
<td>6.74</td>
<td>6.06</td>
</tr>
<tr>
<td>0.20</td>
<td>12.00</td>
<td>7.66</td>
<td>6.00</td>
<td>5.92</td>
</tr>
<tr>
<td>0.30</td>
<td>11.20</td>
<td>7.20</td>
<td>6.84</td>
<td>6.12</td>
</tr>
<tr>
<td>0.40</td>
<td>10.38</td>
<td>7.08</td>
<td>6.32</td>
<td>5.68</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>0.00</td>
<td>20.38</td>
<td>9.46</td>
<td>7.32</td>
<td>6.46</td>
</tr>
<tr>
<td>0.10</td>
<td>19.68</td>
<td>10.24</td>
<td>7.18</td>
<td>7.00</td>
</tr>
<tr>
<td>0.20</td>
<td>19.46</td>
<td>9.56</td>
<td>7.16</td>
<td>6.44</td>
</tr>
<tr>
<td>0.30</td>
<td>16.40</td>
<td>8.78</td>
<td>7.08</td>
<td>7.26</td>
</tr>
<tr>
<td>0.40</td>
<td>16.90</td>
<td>8.94</td>
<td>7.14</td>
<td>6.34</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>0.00</td>
<td>20.78</td>
<td>9.20</td>
<td>7.58</td>
<td>5.96</td>
</tr>
<tr>
<td>0.10</td>
<td>19.16</td>
<td>9.72</td>
<td>7.42</td>
<td>6.64</td>
</tr>
<tr>
<td>0.20</td>
<td>18.78</td>
<td>9.40</td>
<td>7.24</td>
<td>6.84</td>
</tr>
<tr>
<td>0.30</td>
<td>16.68</td>
<td>8.84</td>
<td>7.32</td>
<td>7.06</td>
</tr>
<tr>
<td>0.40</td>
<td>15.72</td>
<td>8.96</td>
<td>7.20</td>
<td>6.36</td>
</tr>
<tr>
<td>$\sigma_s \backslash$ Sample Size</td>
<td>50</td>
<td>200</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>0.00</td>
<td>12.26</td>
<td>7.32</td>
<td>6.70</td>
<td>5.44</td>
</tr>
<tr>
<td>0.10</td>
<td>11.78</td>
<td>8.02</td>
<td>6.70</td>
<td>6.16</td>
</tr>
<tr>
<td>0.20</td>
<td>11.46</td>
<td>7.04</td>
<td>6.74</td>
<td>6.38</td>
</tr>
<tr>
<td>0.30</td>
<td>10.36</td>
<td>7.18</td>
<td>6.58</td>
<td>6.10</td>
</tr>
<tr>
<td>0.40</td>
<td>9.74</td>
<td>6.68</td>
<td>6.10</td>
<td>5.92</td>
</tr>
</tbody>
</table>

Table 9: Empirical Levels of $\widetilde{W}_n^{(1)} (\Gamma)$, $\widetilde{W}_n^{(2)} (\Gamma)$, $\widetilde{W}_n^{(3)} (\Gamma)$, and $\widetilde{W}_n^{(4)} (\Gamma)$ (Level of Significance: 5%). Notes: (i) Number of iterations: 5,000. (ii) DGP: $Y_t = \alpha^* + \phi^* Y_{t-1} + \theta_{0^*} X_t + \theta_{1^*} X_{t-1} + U_t$, $X_t := X_{t-1} + W_t$, $U_t := \sigma^* \epsilon_{t-1} + (1 - \sigma^2) \epsilon_{t-2}$, and $(R_t^{(1)}, R_t^{(2)})' \sim$ IID $N(0, I_2)$. (iii) Model: $Y_t = \alpha^*(\tau) + \sum_{j=0}^{q-1} W_{t-j} \delta_j^*(\tau) + X_t^T \gamma^*(\tau) + \sum_{j=1}^{p} \phi_j^*(\tau) Y_{t-j} + U_t(\tau)$. (iv) The null hypothesis: $s_1 = s_2 = s_3 = 0$, and $s_4 = [0, 0]'$ for $\widetilde{W}_n^{(1)} (\Gamma)$, $\widetilde{W}_n^{(2)} (\Gamma)$, $\widetilde{W}_n^{(3)} (\Gamma)$, and $\widetilde{W}_n^{(4)} (\Gamma)$, respectively. (v) BIC is applied to determine the QARDL orders $p$ and $q$. 58
### OLS Estimation Results

<table>
<thead>
<tr>
<th>ζ*</th>
<th>β*</th>
<th>λ*</th>
<th>δ*</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.04</td>
<td>0.36</td>
<td>0.48</td>
<td>0.01</td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.013)</td>
<td>(0.039)</td>
<td>0.007</td>
</tr>
</tbody>
</table>

### Quantile Estimation Results

<table>
<thead>
<tr>
<th>Quantile index (τ)</th>
<th>ζ∗(τ)</th>
<th>β∗(τ)</th>
<th>λ∗(τ)</th>
<th>δ∗(τ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>-0.06</td>
<td>0.27</td>
<td>0.58</td>
<td>0.02</td>
</tr>
<tr>
<td>(0.014)</td>
<td>(0.030)</td>
<td>(0.043)</td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>0.10</td>
<td>-0.06</td>
<td>0.33</td>
<td>0.55</td>
<td>0.02</td>
</tr>
<tr>
<td>(0.006)</td>
<td>(0.020)</td>
<td>(0.051)</td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>-0.04</td>
<td>0.37</td>
<td>0.55</td>
<td>0.02</td>
</tr>
<tr>
<td>(0.009)</td>
<td>(0.050)</td>
<td>(0.057)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>-0.04</td>
<td>0.40</td>
<td>0.53</td>
<td>0.02</td>
</tr>
<tr>
<td>(0.007)</td>
<td>(0.027)</td>
<td>(0.053)</td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>0.40</td>
<td>-0.03</td>
<td>0.36</td>
<td>0.48</td>
<td>0.02</td>
</tr>
<tr>
<td>(0.008)</td>
<td>(0.041)</td>
<td>(0.059)</td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>-0.03</td>
<td>0.35</td>
<td>0.47</td>
<td>0.03</td>
</tr>
<tr>
<td>(0.008)</td>
<td>(0.043)</td>
<td>(0.066)</td>
<td>(0.013)</td>
<td></td>
</tr>
<tr>
<td>0.60</td>
<td>-0.03</td>
<td>0.38</td>
<td>0.50</td>
<td>0.02</td>
</tr>
<tr>
<td>(0.006)</td>
<td>(0.023)</td>
<td>(0.067)</td>
<td>(0.013)</td>
<td></td>
</tr>
<tr>
<td>0.70</td>
<td>-0.04</td>
<td>0.36</td>
<td>0.45</td>
<td>0.02</td>
</tr>
<tr>
<td>(0.006)</td>
<td>(0.022)</td>
<td>(0.065)</td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>0.80</td>
<td>-0.04</td>
<td>0.39</td>
<td>0.37</td>
<td>0.01</td>
</tr>
<tr>
<td>(0.008)</td>
<td>(0.024)</td>
<td>(0.053)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>0.90</td>
<td>-0.04</td>
<td>0.39</td>
<td>0.30</td>
<td>-0.00</td>
</tr>
<tr>
<td>(0.009)</td>
<td>(0.033)</td>
<td>(0.063)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>-0.04</td>
<td>0.33</td>
<td>0.44</td>
<td>-0.01</td>
</tr>
<tr>
<td>(0.008)</td>
<td>(0.039)</td>
<td>(0.050)</td>
<td>(0.009)</td>
<td></td>
</tr>
</tbody>
</table>

Table 10: OLS and Quantile Estimation Results Based on the Whole Sample. Notes: (i) Standard errors are in parentheses, and those for the long-run coefficient are calculated via the delta method. (ii) OLS estimation results are based on the following model: \( \Delta D_t = \alpha_\tau + \zeta_\tau D_{t-1} + \gamma_\tau E_{t-1} + \lambda_\tau \Delta D_{t-1} + \delta_\tau \Delta E_t + U_t = \alpha_s + \zeta_s (D_{t-1} - \beta_s E_{t-1}) + \lambda_s \Delta D_{t-1} + \delta_s \Delta E_t + U_t \). (iii) Quantile estimation results are based on the following model: \( \Delta D_t = \alpha_s (\tau) + \zeta_s (\tau) D_{t-1} + \gamma_s (\tau) E_{t-1} + \lambda_s (\tau) \Delta D_{t-1} + \delta_s (\tau) \Delta E_t + U_t (\tau) = \alpha_s (\tau) + \zeta_s (\tau) (D_{t-1} - \beta_s (\tau) E_{t-1}) + \lambda_s (\tau) \Delta D_{t-1} + \delta_s (\tau) \Delta E_t + U_t (\tau) \).
ECDFs of $W_n^{(1)}(B)$, $W_n^{(2)}(B)$, and $W_n^{(3)}(B)$

EPDFs of $W_n^{(1)}(B)$, $W_n^{(2)}(B)$, and $W_n^{(3)}(B)$

ECDFs of $W_n^{(1)}(\Phi)$, $W_n^{(2)}(\Phi)$, and $W_n^{(3)}(\Phi)$

EPDFs of $W_n^{(1)}(\Phi)$, $W_n^{(2)}(\Phi)$, and $W_n^{(3)}(\Phi)$

ECDFs of $W_n^{(1)}(\Gamma)$, $W_n^{(2)}(\Gamma)$, and $W_n^{(3)}(\Gamma)$

EPDFs of $W_n^{(1)}(\Gamma)$, $W_n^{(2)}(\Gamma)$, and $W_n^{(3)}(\Gamma)$

Figure 1: Cumulative Distribution and Density Functions of $W_n(B)$, $W_n(\Phi)$, and $W_n(\Gamma)$. Number of Iterations: 10,000. Number of Observations: 5,000. DGP: $Y_t = \alpha_s + \phi_s Y_{t-1} + \theta_0 X_t + \theta_1 X_{t-1} + U_t$, $X_t := X_{t-1} + W_t$, $W_t := \rho_s R_{t-1} + (1 - \rho_s^2) R_t$, and $(U_t, R_t)' \sim$ IID $N(0, I_2)$. Here, $W_n^{(1)}(B)$, $W_n^{(2)}(B)$, and $W_n^{(3)}(B)$ test the null hypotheses when $s_1 = s_2 = s_3 = 0$, respectively; $W_n^{(1)}(\Phi)$, $W_n^{(2)}(\Phi)$, and $W_n^{(3)}(\Phi)$ test the null hypotheses when $s_1 = s_2 = s_3 = 0$, respectively; and $W_n^{(1)}(\Gamma)$, $W_n^{(2)}(\Gamma)$, and $W_n^{(3)}(\Gamma)$ test the null hypotheses when $s_1 = s_2 = s_3 = 0$, respectively.
Figure 2: **Cumulative Distribution and Density Functions of** $W_n(B)$, $W_n(\Phi)$, and $W_n(\Gamma)$.
Number of Iterations: 10,000. Number of Observations: 5,000. DGP: $Y_t = \alpha_* + \phi_* Y_{t-1} + \theta_0 X_t + \theta_1 X_{t-1} + U_t$, $X_t := X_{t-1} + W_t$, $W_t := \rho_* R_{t-1} + (1 - \rho_*^2) R_t$, and $(U_t, R_t)' \sim$ IID $N(0, I_2)$. Here, $W_n^{(4)}(B)$ tests the null hypothesis when $s_4 = [0, 0]'$; $W_n^{(4)}(\Phi)$ tests the null hypothesis when $s_4 = [0, 0]'$; and $W_n^{(4)}(\Gamma)$ tests the null hypothesis when $s_4 = [0, 0]'$. 
ECDFs of $\tilde{W}_n^{(1)}(B)$, $\tilde{W}_n^{(2)}(B)$, and $\tilde{W}_n^{(3)}(B)$.

EPDFs of $\tilde{W}_n^{(1)}(B)$, $\tilde{W}_n^{(2)}(B)$, and $\tilde{W}_n^{(3)}(B)$.

ECDFs of $\tilde{W}_n^{(1)}(\Phi)$, $\tilde{W}_n^{(2)}(\Phi)$, and $\tilde{W}_n^{(3)}(\Phi)$.

EPDFs of $\tilde{W}_n^{(1)}(\Phi)$, $\tilde{W}_n^{(2)}(\Phi)$, and $\tilde{W}_n^{(3)}(\Phi)$.

ECDFs of $\tilde{W}_n^{(1)}(\Gamma)$, $\tilde{W}_n^{(2)}(\Gamma)$, and $\tilde{W}_n^{(3)}(\Gamma)$.

EPDFs of $\tilde{W}_n^{(1)}(\Gamma)$, $\tilde{W}_n^{(2)}(\Gamma)$, and $\tilde{W}_n^{(3)}(\Gamma)$.

Figure 3: Cumulative Distribution and Density Functions of $\tilde{W}_n(B)$, $\tilde{W}_n(\Phi)$, and $\tilde{W}_n(\Gamma)$.

Number of Iterations: 10,000. Number of Observations: 2,000. DGP: $Y_t = \alpha_s + \phi_s Y_{t-1} + \theta_0 + \theta_1 X_{t-1} + U_t$, $U_t := \sigma_u R_{t-1}^{(1)} + (1 - \sigma_u^2) R_{t-1}$, $X_t := X_{t-1} + W_t$, $W_t := \rho_w R_{t-1}^{(2)} + (1 - \rho_w^2) R_{t-1}$, and $(R_{t-1}^{(1)}, R_{t-1}^{(2)})' \sim \text{IID } N(0, I_2)$. Here, $\tilde{W}_n^{(1)}(B)$, $\tilde{W}_n^{(2)}(B)$, and $\tilde{W}_n^{(3)}(B)$ test the null hypotheses when $s_1 = s_2 = s_3 = 0$, respectively; $\tilde{W}_n^{(1)}(\Phi)$, $\tilde{W}_n^{(2)}(\Phi)$, and $\tilde{W}_n^{(3)}(\Phi)$ test the null hypotheses when $s_1 = s_2 = s_3 = 0$, respectively; and $\tilde{W}_n^{(1)}(\Gamma)$, $\tilde{W}_n^{(2)}(\Gamma)$, and $\tilde{W}_n^{(3)}(\Gamma)$ test the null hypotheses when $s_1 = s_2 = s_3 = 0$, respectively.
Figure 4: Cumulative Distribution and Density Functions of $\tilde{W}_n(B), \tilde{W}_n(\Phi),$ and $\tilde{W}_n(\Gamma)$. Number of Iterations: 10,000. Number of Observations: 2,000. DGP: $Y_t = \alpha_0 + \phi_0 Y_{t-1} + \theta_0 X_t + \theta_1 X_{t-1} + U_t$, $U_t := \sigma_0 R_{t-1}^{(1)} + (1 - \sigma_2^2) R_t$, $X_t := X_{t-1} + W_t$, $W_t := \rho_0 R_{t-1}^{(2)} + (1 - \rho_2^2) R_t^{(2)}$, and $(R_t^{(1)}, R_t^{(2)})' \sim$ IID $N(0, I_2)$. Here, $\tilde{W}_n^{(4)}(B)$ tests the null hypothesis when $s_4 = [0, 0]'$; $\tilde{W}_n^{(4)}(\Phi)$ tests the null hypothesis when $s_4 = [0, 0]'$; and $\tilde{W}_n^{(4)}(\Gamma)$ tests the null hypothesis when $s_4 = [0, 0]'$. 
Figure 5: Parameter Estimates Using the Whole Sample. These are estimated parameters (the middle solid line) using all available observations for different quantile levels: 0.05, 0.10, ... 0.95, with 90% confidence intervals (the outer dotted lines).
Figure 6: Parameter Estimates $\zeta^*(\tau)$ Using the Rolling Window Method. These are estimated parameters using the rolling window method, and each window has 320 observations. Three different quantile levels are considered for this computation: 0.25, 0.5, 0.75. The horizontal axis indicates the last date for the corresponding estimation window. For example, the first estimation window uses 320 quarterly observations from 1871Q1 to 1950Q4 so that the first date on the horizontal axis is 1950Q4. The total number of out-of-sample observations is 239.
Figure 7: $p$-VALUES OF $W_n^1(\beta)$ TEST STATISTICS. The figures show the estimated $p$-values of $W_n^1(\beta)$, $W_n^2(\beta)$, $R_n^3(\beta)$, and $W_n^4(\beta)$ test statistics, and the statistics test the equality of $\beta_*(0.25)$, $\beta_*(0.5)$, and $\beta_*(0.75)$. That is, $W_n^1(\beta)$ tests $\beta_*(0.25) = \beta_*(0.5)$; $W_n^2(\beta)$ tests $\beta_*(0.5) = \beta_*(0.75)$; $W_n^3(\beta)$ tests $\beta_*(0.25) = \beta_*(0.75)$; and $W_n^4(\beta)$ tests $\beta_*(0.25) = \beta_*(0.5) = \beta_*(0.75)$. The horizontal axis indicates the last date for the corresponding estimation window. For example, the first estimation window uses 320 quarterly observations from 1871Q1 to 1950Q4 so that the first date on the horizontal axis is 1950Q4. The total number of out-of-sample observations is 239.
Figure 8: \( p \)-VALUES OF \( W_n(\zeta) \) TEST STATISTICS. The figures show the estimated \( p \)-values of \( W_n^{(1)}(\zeta) \), \( W_n^{(2)}(\zeta) \), \( W_n^{(3)}(\zeta) \), and \( W_n^{(4)}(\zeta) \) test statistics, and the statistics test the equality of \( \zeta^*(0.25) \), \( \zeta^*(0.5) \), and \( \zeta^*(0.75) \). That is, \( W_n^{(1)}(\zeta) \) tests \( \zeta^*(0.25) = \zeta^*(0.5) \); \( W_n^{(2)}(\zeta) \) tests \( \zeta^*(0.5) = \zeta^*(0.75) \); \( W_n^{(3)}(\zeta) \) tests \( \zeta^*(0.25) = \zeta^*(0.75) \); and \( W_n^{(4)}(\zeta) \) tests \( \zeta^*(0.25) = \zeta^*(0.5) = \zeta^*(0.75) \). The horizontal axis indicates the last date for the corresponding estimation window. For example, the first estimation window uses 320 quarterly observations from 1871Q1 to 1950Q4 so that the first date on the horizontal axis is 1950Q4. The total number of out-of-sample observations is 239.