Econometric Theory, **29**, 2013, 659–672.
doi:10.1017/S0266466612000710

SEMIPARAMETRIC FUNCTIONAL COEFFICIENT MODELS WITH INTEGRATED COVARIATES

YIGUO SUN
University of Guelph

ZONGWU CAI
University of North Carolina at Charlotte

QI LI
Texas A&M University
and
Capital University of Economics and Business

Cai, Li, and Park (*Journal of Econometrics*, 2009) and Xiao (*Journal of Econometrics*, 2009) developed asymptotic theories for estimators of semiparametric varying coefficient models when regressors are integrated processes but the smooth coefficients are functionals of stationary processes. Using a recent result from Phillips (*Econometric Theory*, 2009), we extend this line of research by allowing for both the regressors and the covariates entering the smooth functionals to be integrated variables. We derive the asymptotic distribution for the proposed semiparametric estimator. An empirical application is presented to examine the purchasing power parity hypothesis between U.S. and Canadian dollars.

1. INTRODUCTION

Econometric analysis involving integrated time series is quite popular in macroeconometrics and financial econometrics. With the availability of useful technical tools developed for integrated time series analysis by Jeganathan (2004), Wang and Phillips (2009a, 2009b), Phillips (2009), and Kasparis and Phillips (2012), researchers can apply flexible nonparametric/semiparametric techniques to analyze integrated time series data; see Juhl (2005), Wang and Phillips (2009a, 2009b), Phillips (2009), Cai, Li, and Park (2009), Xiao (2009), and Kasparis and Phillips, among many others.

We thank two referees and Giuseppe Cavaliere (co-editor) for their insightful comments that greatly improved our paper. Sun acknowledges financial support from the Social Sciences and Humanities Foundation of Canada (grant #410-2009-0109). Cai thanks the National Science Foundation (NSF) of China (projects #71131008 and #70971113) for partial financial support, as well as the "1,000 Talents Project" of the Chinese central government, Fujian Province and Xiamen city governments, the Cheung Kong Scholarship Project from the Chinese Ministry of Education, the Minjiang Scholarship Project from Fujian Province, and Xiamen University. Li thanks the NSF of China (projects #70773005 and #71131008) and research funds provided by the Capital University of Economics and Business (China) for partial financial support. Address correspondence to Qi Li, Dept. of Economics, Texas A&M University, College Station, TX 77843-4228, USA; e-mail: qi@econmail.tamu.edu.

In this note we consider the semiparametric functional coefficient model

$$Y_t = X_t^T \beta(Z_t) + u_t, \tag{1}$$

where both X_t (a $d \times 1$ vector) and Z_t (a scalar) are integrated variables (i.e., I(1) variables), u_t is a stationary process, and $\beta(\cdot) = (\beta_1(\cdot), \ldots, \beta_d(\cdot))^T$ is a $d \times 1$ vector of unknown functional coefficients. Asymptotic theories of local linear estimator of $\beta(\cdot)$ have been developed in the literature under three cases: (i) both X_t and Z_t are stationary by Cai, Fan, and Yao (2000); (ii) Z_t is I(0) and X_t contains both I(0) and I(1) variables by Cai et al. (2009) and Xiao (2009); and (iii) Z_t is I(1) and X_t is I(0) by Cai et al. (2009). However, to the best of our knowledge, asymptotic analysis for the case that both X_t and Z_t are I(1) variables is not available in the existing literature. In this note we close this gap by providing asymptotic theories of the local linear estimator of $\beta(\cdot)$ for the case that both X_t and Z_t are I(1) variables.

2. THE LOCAL LINEAR ESTIMATOR AND ASYMPTOTIC RESULTS

Denote the j^{th} derivative of $\beta(z)$ by $\beta^{(j)}(z) = \partial^j \beta(z) / \partial z^j$, where j is a positive integer. For each $l = 1, \ldots, d$, we have the Taylor expansion of $\beta_l(Z_t)$ at $z_0 \in R$,

$$\beta_l(Z_t) = \beta_l(z_0) + \beta_l^{(1)}(z_0)(Z_t - z_0) + \beta_l^{(2)}(z_0)(Z_t - z_0)^2/2 + r_l(Z_t, z_0), \tag{2}$$

where z_0 is either a fixed constant or $z_0 = z_{n,0} = c_0\sqrt{n}$ for some nonzero constant c_0 (so that $z_0 = O(\sqrt{n})$). Our estimator solves the minimization problem

$$\begin{pmatrix} \widehat{a} \\ \widehat{b} \end{pmatrix} = \arg\min_{(a,b)} \sum_{t=1}^{n} \left[Y_t - a^T X_t - b^T X_t (Z_t - z_0) \right]^2 K\left(\frac{Z_t - z_0}{h} \right), \tag{3}$$

where $\widehat{a} = \widehat{\beta}(z_0)$ estimates $\beta(z_0)$ and $\widehat{b} = \widehat{\beta}^{(1)}(z_0)$ estimates $\beta^{(1)}(z_0)$. Let $\widehat{\theta}(z_0) = \left(\widehat{\beta}(z_0)^T, \widehat{\beta}^{(1)}(z_0)^T\right)^T$ and $\theta(z_0) = \left(\beta(z_0)^T, \beta^{(1)}(z_0)^T\right)^T$. Then it is easy to see that

$$\widehat{\theta}(z_0) = \theta(z_0) + A_{n1}(z_0)^{-1}\left[A_{n2}(z_0) + A_{n3}(z_0) + A_{n4}(z_0)\right], \tag{4}$$

where $A_{n1}(z_0) = \sum_{t=1}^n K\left(\frac{Z_t - z_0}{h}\right) Q_t Q_t^T$, $A_{n2}(z_0) = \frac{1}{2}\sum_{t=1}^n K\left(\frac{Z_t - z_0}{h}\right)(Z_t - z_0)^2 Q_t X_t^T \beta^{(2)}(z_0)$, $A_{n3}(z_0) = \sum_{t=1}^n K\left(\frac{Z_t - z_0}{h}\right) Q_t X_t^T r(Z_t, z_0)$, $A_{n4}(z_0) = \sum_{t=1}^n K\left(\frac{Z_t - z_0}{h}\right) Q_t u_t$, $Q_t = (X_t^T, X_t^T(Z_t - z_0))^T$, $r(\cdot,\cdot) = [r_1(\cdot,\cdot), \ldots, r_d(\cdot,\cdot)]^T$, and $r_l(\cdot,\cdot)$ is defined in (2) for $l = 1, \ldots, d$.

Define a $(2d) \times (2d)$ diagonal matrix $H_n = \text{diag}(I_d, hI_d)$, where I_d denotes a $d \times d$ identity matrix. We will show that

$$n^{3/4} h^{1/2} H_n \left(\widehat{\theta}(z_0) - \theta(z_0) - h^2 C(z_0) \right) \xrightarrow{d} MN,$$

where MN is a mixed normal random vector, and $C(z_0)$ is a nonstochastic constant depending on z_0 and is defined in Theorem 1 below.

The result of this paper heavily relies on a recent covariance asymptotic result provided by Phillips (2009). In particular, Phillips (2009) develops a limit theory for the sample covariance of the form

$$\frac{1}{\sqrt{n}h}\sum_{t=1}^{n} g\left(\frac{X_t}{\sqrt{n}}\right) T\left(\frac{Z_t-z_0}{h}\right) \xrightarrow{d} \int_{-\infty}^{\infty} T(r)dr \int_0^1 g(B_X(s))dL_{B_Z}(s,c_0), \quad \textbf{(5)}$$

where $T(\cdot)$ is an integrable function, (B_X, B_Z) is a vector of Brownian motion, L_{B_Z} is the local time of B_Z, and $c_0 = \lim_{n\to\infty}(z_0/\sqrt{n})$ is a finite constant ($c_0 = 0$ if z_0 is a finite constant). For the Brownian motion $B_Z(\cdot)$, its local time is defined by $L_{B_Z}(t,s) = \lim_{\epsilon\to 0}(2\epsilon)^{-1}\int_0^t I\{|B_Z(r)-s| < \epsilon\}dr$.

We now turn to the derivation of the asymptotic theory of $\widehat{\theta}(z_0) = \left(\widehat{\beta}(z_0)^T, \widehat{\beta}^{(1)}(z_0)^T\right)^T$. We first list some regularity conditions. Let $\psi_t = \left(X_{t+1}^T, Z_{t+1}, V_t\right)^T$ be a column vector of partial sums of dimension $m = d+2$; i.e., $\psi_t = \psi_0 + \sum_{i=1}^t \eta_t$ for $1 \le t \le n$, $n \ge 1$, where $\eta_t = \left(\eta_{x,t+1}^T, \eta_{z,t+1}, u_t\right)^T$ with $\eta_{x,t+1} = X_{t+1} - X_t$, $\eta_{z,t+1} = Z_{t+1} - Z_t$, and $u_t = V_t - V_{t-1}$ for all t. We will use Theorem 1 of Phillips (2009) to study the asymptotic behaviors of $A_{n,j}(z_0)$ for $j = 1, 2, 3$. Below we list sufficient conditions to ensure the use of Theorem 1 of Phillips (2009).

Assumption A1. Assume $\{\eta_t\}$ is a zero-mean, strictly stationary, α-mixing sequence of size $-p/(p-2)$ with $p = 2+\delta$ and $\|\eta_t\|_{2+\tilde{\delta}} < C < \infty$ for some small $0 < \delta < \tilde{\delta} \le 1$. In addition, $\lim_{n\to\infty} n^{-1}\mathbb{E}\left(\psi_n\psi_n^T\right) = \Sigma_\eta < \infty$ and $\psi_0 = O_p(1)$.

Under Assumption A1 and by Corollary 4.2 of Wooldridge and White (1988), we apply the multivariate functional central limit theorem to the vector of triangular arrays $\psi_{t,n} = \psi_t/\sqrt{n}$ for $1 \le t \le n$ such that

$$\frac{\psi_{[nr],n}}{\sqrt{n}} = \frac{\psi_0}{\sqrt{n}} + \frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} \eta_t \Rightarrow B(r) \quad \text{for } r \in [0,1], \quad \textbf{(6)}$$

where $[a]$ denotes the integer part of a, $B(\cdot)$ is an m-dimensional multivariate Brownian motion with zero mean and covariance matrix Σ_η, and "$\Rightarrow$ " denotes the weak convergence on the Skorohod space $\mathcal{D}[0,1]^m$. Partitioning $B(r)$ conformably with ψ_t gives $B(r) = \left(B_X(r)^T, B_Z(r), B_u(r)\right)^T$. Throughout this note, we use C to denote a genetic finite positive constant, which may take different values at different places.

Assumption A2. The kernel function $K(v)$ is a symmetric (around zero) bounded probability density function with a bounded support [-1,1].

Assumption A3. (i) Assume $\beta(z)$ is three-time continuously differentiable, and $\beta^{(3)}(z)$ satisfies a Hölder condition: For all $z, z' \in \mathcal{R}$, $\left\|\beta^{(3)}(z) - \beta^{(3)}(z')\right\| \le C\left|z-z'\right|^\alpha$ for some $\alpha > 0$, where $\|\cdot\|$ is the Euclidean norm. (ii) When $z_0 = z_{n,0} = \sqrt{n}c_0$ with $c_0 \neq 0$, $\lim_{n\to\infty}\beta^{(j)}\left(z_{n,0}\right)$ is finite for $j = 2, 3$.

We denote by $f_t^x(x)$ the probability density function of $X_t/\sqrt{t}$ and $f_{s,t}^z(z|x)$ the conditional probability density function of $(Z_t - Z_s)/\sqrt{t-s}$ conditional on $(n^{-1/2}X_s, \mathcal{F}_{n,s})$ for $0 \le s < t \le n$, where we set $Z_0 = 0$ for simplicity, and $\mathcal{F}_{n,s} = \sigma\{\eta_i : 1 \le i \le s\}$ is the smallest sigma field containing the past history of $\{\eta_i : 1 \le i \le s \le n\}$.

Assumption A4. (i) Assume $\max_{0\le s<t\le n}\sup_{z\in R}\sup_{x\in R^d} f_{s,t}^z(z|x) \le C < \infty$. Also, for all $x \in R^d$ and $0 \le s < t \le n$ and for some $l > 0$, we have (a local Hölder condition) $\sup_x \left| f_{s,t}^z\left(\frac{hu+z_0}{\sqrt{t-s}}|x\right) - f_{s,t}^z\left(\frac{z_0}{\sqrt{t-s}}|x\right)\right| \le C\left|\frac{hu}{\sqrt{t-s}}\right|^l$. (ii) There is a nonnegative function $f(x)$ satisfying $f_t^x(x) \le f(x)$ for all t and $\int \|x\|^q f(x)\,dx < \infty$ for some $q > 2$.

Assumption A5. Assume $nh^2 \to \infty$ and $h \to 0$ as $n \to \infty$.

Note that if z_0 is a fixed constant, Assumption A3(i) is a mild assumption, as it only requires that $\beta(z)$ is three-time continuously differentiable at z_0. However, when $z_0 = \sqrt{n}c_0$ with $c_0 \ne 0$, Assumption A3(ii) also requires that $\beta^{(j)}(\cdot)$ is a bounded function for $j = 2, 3$. Assumption A4 is a modification of Assumption 2.3 in Phillips (2009). Assumption A5 ensures that the smoothing parameter h converges to zero at a proper rate for the application of Theorem 1 of Phillips (2009).

In addition, although $\sup_{1\le t\le n}|Z_t| = O_p(n)$, $\sup_{1\le t\le n}|Z_t| = O\left(\sqrt{n\ln\ln n}\right)$ holds almost surely (see Thm. 2 of Rio, 1995). Therefore, there is a small probability that Z_t may take values at far tails beyond $O\left(\sqrt{n}\right)$. To avoid data sparsity, we estimate $\beta(z)$ at $z_0 \in \mathcal{M}_n$, where $\mathcal{M}_n = \left\{z \in \mathcal{R} : \ z/\sqrt{n} \text{ is bounded as } n \to \infty\right\}$. Also, $A \otimes B$ refers to the Kronecker product. The direct application of Theorem 1 of Phillips (2009) gives the following results, and the proof is delayed to the Appendix.

LEMMA 1.

(i) Under Assumptions A1, A2, A4, and A5, for $z_0 \in \mathcal{M}_n$, we have

$$H_n^{-1}\frac{A_{n1}(z_0)}{n^{3/2}h}H_n^{-1} \xrightarrow{d} S_\mu \otimes \varkappa(c_0),$$

where $H_n = \operatorname{diag}\{I_d, hI_d\}$, $\varkappa(c_0) = \int_0^1 B_X(r)B_X(r)^T\,dL_{B_Z}(r, c_0)$, $S_\mu = \operatorname{diag}\{1, \mu_2\}$ *and* $\mu_j = \int_{-\infty}^{\infty} v^j K(v)\,dv$ *(j is a nonnegative integer).*

(ii) If Assumption A3 also holds, we have

$$H_n^{-1}\frac{A_{n2}(z_0)}{n^{3/2}h^3} \xrightarrow{d} \begin{bmatrix} \mu_2\varkappa(c_0)\beta^{(2)}(z_0)/2 \\ \mathbf{0}_{d\times 1} \end{bmatrix} \quad \text{and}$$

$$\frac{H_n^{-1}A_{n3}(z_0)}{n^{3/2}h^4} \xrightarrow{d} \begin{bmatrix} \mathbf{0}_{d\times 1} \\ \mu_4\varkappa(c_0)\beta^{(3)}(z_0)/6 \end{bmatrix},$$

where $\mathbf{0}_{d\times 1}$ is a $d \times 1$ vector of zeros.

The results in Lemma 1 can be strengthened to convergence in the probability result as shown in Phillips (2009) if the weak convergence result (6) is strengthened to the uniform convergence in the probability result. We will do so in the Appendix when we prove the main result of this paper.

To obtain the limit result for $A_{n1}(z_0)^{-1}A_{n4}(z_0)$, we make the following additional assumption.

Assumption B. Assume $\{u_t\}$ is a stationary process satisfying $\mathbb{E}\left(u_t|\mathcal{F}_{n,t-1}\right) = 0$, $\mathbb{E}\left(u_t^2|\mathcal{F}_{n,t-1}\right) \overset{a.s.}{\rightarrow} \sigma_u^2 < \infty$, and $\sup_{1\le t\le n}\mathbb{E}\left(|u_t|^q\,|\mathcal{F}_{n,t-1}\right) < C < \infty$ for some $q > 2$, where $\overset{a.s.}{\rightarrow}$ denotes the convergence with probability one.

Assumption B implies that X_t and Z_t are exogenous regressors. Given the recent work of Wang and Phillips (2009b), who show that estimation consistency can be achieved under endogeneity, we conjecture that Assumption B may be relaxed to allow for endogenous regressors. We leave this possible extension as a future research topic. The main result of this note is given below, and the proof is given in the Appendix.

THEOREM 1. *Under Assumptions A1–A5 and B and* $n^{3/2}h^5 = O(1)$, *for* $z_0 \in \mathcal{M}_n$, *we have*

$$\left(H_n^{-1}A_{n1}H_n^{-1}\right)^{1/2}\left[H_n\left(\widehat{\theta}(z_0)-\theta(z_0)\right)-h^2C(z_0)\right] \overset{d}{\rightarrow} N\left(0,\Sigma^*\right), \tag{7}$$

where $C(z_0) = \left(\frac{1}{2}\mu_2\beta^{(2)}(z_0)^T,\ \mathbf{0}_{1\times d}\right)^T$, $\Sigma^* = \sigma_u^2\mathrm{diag}\left(\nu_0,\frac{\nu_2}{\mu_2}\right)\otimes I_{2d}$, $\nu_j = \int K^2(u)u^j du$, *and* $\mathbf{0}_{1\times d}$ *is a* $1\times d$ *vector of zeros.*

By Lemma 1, we know that $H_n^{-1}A_{n1}H_n^{-1} = O_p\left(n^{3/2}h\right)$. Hence, Theorem 1 implies that the asymptotic variance of $\widehat{\beta}(z_0)$ is of order $O\left(n^{-1/2}(nh)^{-1}\right)$. In contrast, Cai et al. (2009) show that the asymptotic variance of $\widehat{\beta}(z_0)$ is of order $O\left(n^{1/2}(nh)^{-1}\right)$ when X_t is I(0) and Z_t is I(1), and it is of order $O\left(n^{-1}(nh)^{-1}\right)$ when X_t is I(1) and Z_t is I(0). Also, it is well known that the asymptotic variance of $\widehat{\beta}(z_0)$ is of order $O\left((nh)^{-1}\right)$ when both X_t and Z_t are I(0). Therefore, combining the existing results with our Theorem 1, we learn that a nonstationary Z_t inflates the asymptotic variance by an order of $n^{1/2}$ (for a given stochastic process X_t), while a nonstationary X_t deflates the asymptotic variance by an order of n^{-1} (for a given stochastic process Z_t). Moreover, from the asymptotic result of Theorem 1 that $\widehat{\beta}(z_0)-\beta(z_0) = h^2\beta^{(2)}(z_0)\mu_2/2 + O_p\left(n^{-3/4}h^{-1/2}\right)$, we know that $\widehat{\beta}(z_0)-\beta(z_0) = O_p\left(n^{-0.6}\right)$ if one selects the bandwidth $h = O\left(n^{-3/10}\right)$.

3. AN ILLUSTRATIVE EMPIRICAL APPLICATION

We reinvestigate the purchasing power parity (PPP) hypothesis using Canadian and U.S. price index and exchange rate data. The PPP theory is typically tested under the setup

$$s_t = \beta_0 + \beta_1 p_t + \beta_2 p_t^* + u_t, \tag{8}$$

where s_t, p_t, and p_t^* are the logarithms of the nominal exchange rate expressed as Canadian dollars per unit of the U.S. dollar, the Canadian and U.S. aggregate price levels, respectively. The PPP theory predicts that $\{u_t\}$ is an $I(0)$ process. One may further test that $\beta_0 = 0$ and $\beta_1 = -\beta_2 = 1$.

We use monthly data for the period from January 1974 to December 2009. The aggregate price index is measured by the producer price index (PPI) base-weighted to the year 2000. The augmented Dickey-Fuller (or ADF) test, the generalized least squares (GLS) modified ADF (or ADF-GLS) test of Elliott, Rothenberg, and Stock (1996), and Phillips and Perron's (1988, or PP) Z_α test are used to test for unit roots. None of the three unit root tests reject that s_t, p_t, and p_t^* are I(1) variables in line with the conventional wisdom. We then apply three cointegration tests to test for a potential cointegrating relation among the three integrated variables. First, when we apply the Engle-Granger (EG) residual-based cointegrating test to model (8), the EG test statistic equals -1.697 with a p value of 0.97, which indicates a failure to reject a spurious regression at any conventional significance level. Second, when we apply Xiao and Phillips' (2002) test, the t-statistic for the coefficient of the linear time trend is -2.98 with a p value of 0.003, which means that this coefficient is significantly different from zero even at the 1% level. Hence, Xiao and Phillips' test strongly rejects the null of a cointegrating relation. In addition, Johansen's cointegrating rank test statistic for no cointegrating relation against at least one cointegrating relation yields a value of 28.37, smaller than the 5% critical value of 31.59. Therefore, linear model-based testing results fail to support the long-run PPP theory.

Next, we use a flexible semiparametric varying coefficient model to reexamine the PPP hypothesis. The linear model is restrictive, as it implies that the bilateral exchange rate is solely determined by the price levels from the two countries. However, based on the sticky-price theory of exchange rate determination, exchange rate movements also respond to monetary shocks. Due to sticky prices, the goods markets adjust to the monetary shocks more slowly than do asset markets. Hence, in addition to the aggregate price levels, some other economic variables, such as interest rate differentials between two nations, also affect exchange rate formation and adjust more quickly to monetary shocks than the aggregate price indexes do. Therefore, we will allow the price coefficients to depend on the interest rate differential between the United States and Canada. Specifically, we denote by $z_t = T_{US,t} - T_{CA,t}$ the difference between the two countries' 10-year Treasury bond rates, where the rates are scaled up by 100. Allowing the coefficients in model (8) to vary with respect to the long-term bond rate differential, we estimate the semiparametric model[1]

$$\tilde{s}_t = \beta_1(z_t)\,\tilde{p}_t + \beta_2(z_t)\,\tilde{p}_t^* + u_t. \tag{9}$$

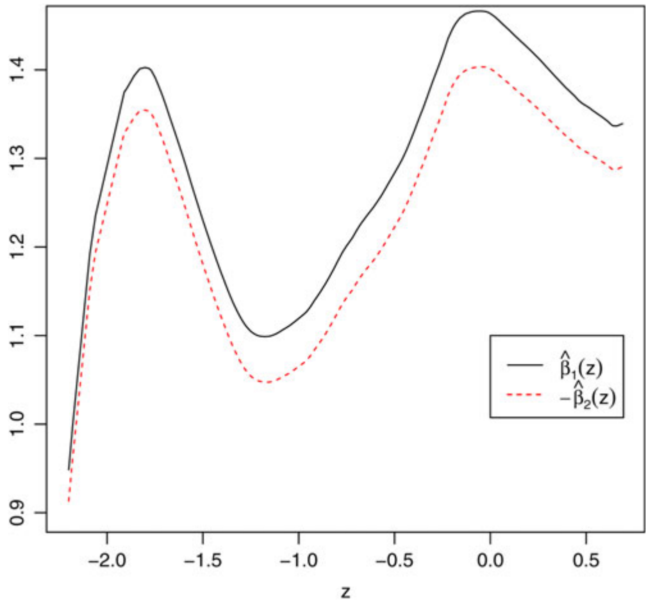

FIGURE 1. Estimated coefficient curves from the semiparametric PPP model (9).

Applying the ADF-GLS test statistic to the interest rate differential, we cannot reject the null hypothesis that z_t follows a unit root process at the 5% significance level. Therefore, we treat z_t as an integrated series. We use the Epanechinikov kernel, and the smoothing parameter h is selected by the least squares cross-validation method. Figure 1 plots the semiparametric estimates of $\beta_1(\cdot)$ and $-\beta_2(\cdot)$. Examining the graph, we find that the estimated coefficient curves $\widehat{\beta}_1(z)$ and $-\widehat{\beta}_2(z)$ have similar shapes.

We use Xiao's (2009) t-statistic to test whether u_t is an I(0) process; i.e., we regress $\widehat{u}_t^2$ (the squared estimated residuals) on an intercept term and a linear time trend. The t-statistic for the time trend coefficient equals 1.26 with a p-value of 0.787, where Newey-West standard errors are calculated with 8 lags. Thus, we cannot reject a zero slope coefficient at the 5% significance level. This implies that the semiparametric PPP model (9) leads to a stable (cointegrating) relationship among the stochastic trends of s_t, p_t, and p_t^*. Therefore, in contrast to the result obtained from the linear model (8), we find that the PPP hypothesis holds for the U.S. and Canadian markets when allowing for the coefficients of the aggregate price indexes to vary with respect to a relevant macroeconomic variable: the 10-year Treasury bond rate differential between the two nations.

Finally, we make some comments on the use of Xiao's (2009) test in our context. As a co-editor correctly points out, Xiao's model differs from our case because Xiao only considers the case that $\{Z_t\}$ is a stationary process while we consider the case that $\{Z_t\}$ is a unit root nonstationary process. We conjecture that, under some proper regularity conditions, it may be possible to show that Xiao's test is valid for the case of an $I(1)$ Z_t process. We have conducted some simulations, and the results seem to support our conjecture. These simulation results are available from the authors upon request. We leave a rigorous proof of extending Xiao's (2009) test to our case as a future research topic.

NOTE

1. Here, $\tilde{y}_t = y_t - \widehat{\alpha} t$, where $\widehat{\alpha}$ is the ordinary least squares (OLS) estimate of the slope parameter in the regression of y_t on an intercept term and a time trend t, $t = 1, 2, \ldots, n$, where $y = s$ or p or p^*. For a relatively large n, this detrending procedure removes asymptotically the nonzero drift of an I(1) process. Then, the detrended data is in line with our theory derived in the paper. Therefore, equation (9) is aimed at capturing the relation among stochastic trends embedded within exchange rates and aggregate price indexes from U.S. and Canadian markets.

REFERENCES

Cai, Z., J. Fan, & Q. Yao (2000) Functional coefficient regression models for nonlinear time series. *Journal of the American Statistical Association* 95, 941–956.

Cai, Z., Q. Li, & J.Y. Park (2009) Functional-coefficient models for nonstationary time series data. *Journal of Econometrics* 148, 101–113.

Elliott, G., T.J. Rothenberg, & J.H. Stock (1996) Efficient tests for an autoregressive unit root. *Econometrica* 64, 813–836.

Jacod, J. & A.N. Shiryaev (2003) *Limit Theorems for Stochastic Processes*. 2nd ed. Springer-Verlag.

Jeganathan, P. (2004) Convergence of functionals of sums of random variables to local times of fractional stable motions. *The Annals of Probability* 32, 1771–1795.

Juhl, T. (2005) Functional coefficient models under unit root behavior. *Econometrics Journal* 8, 197–213.

Kasparis, I. & P.C.B. Phillips (2012) Dynamic misspecification in nonparametric cointegration. *Journal of Econometrics* 168, 270–284.

Kuelbs, J. & W. Philipp (1980) Almost sure invariance principles for partial sums of mixing B-valued random variables. *The Annals of Probability* 8, 1003–1036.

Park, J.Y. & P.C.B. Phillips (2001) Nonlinear regressions with integrated time series. *Econometrica* 69, 117–161.

Phillips, P.C.B. (1991) Spectral regression for cointegrated time series. In W.A. Barnett, J. Powell, & G.E. Tauchen (eds.), *Nonparametric Semiparametric Methods in Econometrics and Statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, pp. 413–435. Cambridge University Press.

Phillips, P.C.B. (2009) Local limit theory and spurious nonparametric regression. *Econometric Theory* 25, 1466–1497.

Phillips, P.C.B. & P. Perron (1988) Testing for unit roots in time series regression. *Biometrika* 75, 335–346.

Revuz, D. & M. Yor (2005) *Continuous Martingales and Brownian Motion*, 3rd ed. Fundamental Principles of Mathematical Sciences 293. Springer-Verlag.

Rio, E. (1995) The functional law of the iterated logarithm for stationary strongly mixing sequences. *The Annals of Probability* 23, 1188–1203.

Wang, Q. & P.C.B. Phillips (2009a) Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory* 25, 710–738.

Wang, Q. & P.C.B. Phillips (2009b) Structural nonparametric cointegrating regression. *Econometrica* 77, 1901–1948.

Wang, Q. & P.C.B. Phillips (2011) Specification testing for nonlinear cointegrating regression. Cowles Foundation Discussion Paper no. 1779.
Wooldridge, J.M. & H. White (1988) Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory* 4, 210–230.
Xiao, Z. (2009). Functional coefficient co-integration models. *Journal of Econometrics* 152, 81–92.
Xiao, Z. & P.C.B. Phillips (2002) A CUSUM test for cointegration using regression residuals. *Journal of Econometrics* 108, 43–61.

APPENDIX: Mathematical Proofs

Throughout this section we denote $d_n = \sqrt{n}/h$, $T_j(v) = v^j K(v)$, and $\mu_j = \int T_j(v)\,dv$ for some nonnegative integer j.

Proof of Lemma 1. Letting $X_{t,n} = X_t/\sqrt{n}$ and $Z_{t,n} = Z_t/\sqrt{n}$, we have

$$S_{n,j} \stackrel{def}{=} \left(n^{3/2}h\right)^{-1} \sum_{t=1}^{n} X_t X_t^T T_j\left(\frac{Z_t - z_0}{h}\right) = \frac{d_n}{n} \sum_{t=1}^{n} X_{t,n} X_{t,n}^T T_j\left(d_n\left(Z_{t,n} - \frac{z_0}{\sqrt{n}}\right)\right).$$

Under Assumptions A1, A2, A4, and A5, following step by step the proof of Theorem 1 of Phillips (2009) (or by using (5)), it can be shown that

$$S_{n,j} \stackrel{d}{\to} \mu_j \varkappa(c_0). \tag{A.1}$$

For the term $A_{n,3}(z_0)$, we apply the Taylor expansion and obtain $r(Z_t, z_0) = \beta(Z_t) - \beta(z_0) - \beta^{(1)}(z_0)(Z_t - z_0) - \beta^{(2)}(z_0)(Z_t - z_0)^2/2 = \beta^{(3)}\left(\tilde{Z}_t\right)(Z_t - z_0)^3/6$, where $\tilde{Z}_t = \lambda_t Z_t + (1-\lambda_t) z_0 = \lambda_t (Z_t - z_0) + z_0$ for some $\lambda_t \in [0,1]$. By Assumption A3, we have $\left\|\beta^{(3)}\left(\tilde{Z}_t\right) - \beta^{(3)}(z_0)\right\| \le C\left|\tilde{Z}_t - z_0\right|^{\alpha} = C(h\lambda_t)^{\alpha} |(Z_t - z_0)/h|^{\alpha}$, which follows $\sup_{|Z_t - z_0| \le h} \left\|\beta^{(3)}\left(\tilde{Z}_t\right) - \beta^{(3)}(z_0)\right\| \le Ch^{\alpha}$. As $K(\cdot)$ has a compact support $[-1,1]$ by Assumption A2, we obtain

$$\left(n^{3/2}h^4\right)^{-1} H_n^{-1} A_{n3}(z_0) = \begin{bmatrix} S_{n,3} \\ S_{n,4} \end{bmatrix} \frac{\beta^{(3)}(z_0)}{6} + O\left(h^{\alpha}\right) \frac{d_n}{n} \sum_{t=1}^{n} K\left(\frac{Z_t - z_0}{h}\right) \times \begin{bmatrix} \left\|X_{t,n}X_{t,n}^T\right\| \left|\frac{Z_t - z_0}{h}\right|^3 \\ \left\|X_{t,n}X_{t,n}^T\right\| \left|\frac{Z_t - z_0}{h}\right|^4 \end{bmatrix}. \tag{A.2}$$

Applying (A.1) into (A.2) proves Lemma 1, where the second term is of order $O_p\left(h^{\alpha}\right)$ by (5). Also, $S_{n,3} = o_p(1)$ because $\mu_3 = 0$.

Proof of Theorem 1. The proof below uses similar arguments as in the proof of Theorem 3.2 in Wang and Phillips (2009a) and Wang and Phillips (2011). We rewrite equation (4) as

$$H_n\left[\widehat{\theta}(z_0) - \theta(z_0)\right] = \left[\frac{H_n^{-1} A_{n1}(z_0) H_n^{-1}}{n^{3/2}h}\right]^{-1} \times \left[h^2 \frac{H_n^{-1} A_{n2}(z_0)}{n^{3/2}h^3} + h^3 \frac{H_n^{-1} A_{n3}(z_0)}{n^{3/2}h^4} + \frac{H_n^{-1} A_{n4}(z_0)}{n^{3/2}h}\right]. \tag{A.3}$$

When equation (6) holds true, by Theorem 4 of Kuelbs and Phillipp (1980) and Assumption A1, there exists a random sequence $\psi_{t,n}^0 \overset{d}{=} \psi_{t,n}$ in a suitable probability space $[\Omega,\ \mathcal{F}, P]$ such that

$$\sup_{0\le r\le 1}\left\|\psi_{[nr],n}^0 - B(r)\right\| = o_p(1), \tag{A.4}$$

where $D_{t,n} \overset{d}{=} G_{t,n}$ means that the two stochastic processes $D_{t,n}$ and $G_{t,n}$ have the same distribution. This kind of technique has been used many times in time series econometrics in the last 20 years (e.g., Phillips, 1991).

For $r \in [0,1]$, we write $\psi_{[nr],n}^0 = \left(\psi_{[nr],n,x}^0{}^T, \psi_{[nr],n,z}^0, \psi_{[nr],n,u}^0\right)^T \equiv \left(\psi_{n,x}^0(r)^T,\right.$ $\left.\psi_{n,z}^0(r), \psi_{n,u}^0(r)\right)^T$. By Lemma 2.1 in Park and Phillips (2001), $\psi_{n,u}^0(\cdot)$ can be constructed such that $\psi_{n,u}^0(r) \overset{d}{=} B_u(\tau_{nt}/n)$ for $1 \le t \le n$ (for all $n \ge 1$), where $\tau_{n,t}$ is an increasing sequence of stopping times with respect to $\mathcal{F}_{n,t}^0$ in $[\Omega, \mathcal{F}, P]$ with

$$\mathcal{F}_{n,t}^0 = \sigma\left\{B_u(r), r \le \tau_{n,t}/n;\ \psi_{n,x}^0\left(\frac{s-1}{n}\right), \psi_{n,z}^0\left(\frac{s-1}{n}\right), s = 1,\ldots,t\right\} \tag{A.5}$$

satisfying $\tau_{n,0} = 0$ and

$$\sup_{1\le t\le n}\left|\frac{\tau_{n,t} - t}{n^\zeta}\right| \overset{a.s}{\to} 0 \quad \text{as } n \to \infty \tag{A.6}$$

for any $\zeta > \max(1/2, 2/q)$. Therefore, we have, for all $1 \le t \le n$,

$$\frac{X_t}{\sqrt{n}} \overset{d}{=} \psi_{n,x}^0\left(\frac{t-1}{n}\right), \qquad \frac{Z_t}{\sqrt{n}} \overset{d}{=} \psi_{n,z}^0\left(\frac{t-1}{n}\right), \quad \text{and}$$
$$\frac{u_t}{\sqrt{n}} \overset{d}{=} B_u\left(\frac{\tau_{n,t}}{n}\right) - B_u\left(\frac{\tau_{n,t-1}}{n}\right). \tag{A.7}$$

Combining (A.3) and (A.7), we obtain

$$H_n\left[\widehat{\theta}(z_0) - \theta(z_0)\right] \overset{d}{=} \Delta_{n1}^{-1}\left(h^2\Delta_{n2} + h^3\Delta_{n3} + \Delta_{n4}\right), \tag{A.8}$$

where $\Delta_{n1} = (d_n/n)\sum_{t=1}^n K\left(\omega_n^z\left(\frac{t-1}{n}\right)\right) Q_n\left(\frac{t-1}{n}\right) Q_n\left(\frac{t-1}{n}\right)^T$, $\Delta_{n2} = (d_n/n)\sum_{t=1}^n T_2\left(\omega_n^z\left(\frac{t-1}{n}\right)\right) Q_n\left(\frac{t-1}{n}\right)\psi_{n,x}^0\left(\frac{t-1}{n}\right)^T \beta^{(2)}(z_0)$, $\Delta_{n3} = (d_n/(6n))\sum_{t=1}^n T_3\left(\omega_n^z\left(\frac{t-1}{n}\right)\right) Q_n\left(\frac{t-1}{n}\right)\psi_{nx}^0\left(\frac{t-1}{n}\right)^T \beta^{(3)}\left(\lambda_t h\omega_n^z\left(\frac{t-1}{n}\right) + z_0\right)$ for some $\lambda_t \in [0,1]$, $\Delta_{n4} = \frac{d_n}{n}\sum_{t=1}^n K\left(\omega_n^z\left(\frac{t-1}{n}\right)\right) Q_n\left(\frac{t-1}{n}\right)\left[B_u\left(\frac{\tau_{n,t}}{n}\right) - B_u\left(\frac{\tau_{n,t-1}}{n}\right)\right]$, $\omega_n^z(t/n) = d_n\left(\psi_{nz}^0(t/n) - z_0/\sqrt{n}\right)$, and $Q_n(t/n) = \left[\psi_{nx}^0(t/n)^T, \omega_n^z(t/n)\psi_{nx}^0(t/n)^T\right]^T$.

Under (A.4), equation (14) of Theorem 1 of Phillips (2009, p. 1475) states that the convergence in the distribution result of (5) can be strengthened to the uniform convergence in probability result. Consequently, the convergence in distribution results in Lemma 1 can be strengthened to the convergence in probability results,

$$\Delta_{n1} \overset{p}{\to} \begin{bmatrix} 1 & 0 \\ 0 & \mu_2 \end{bmatrix} \otimes \varkappa(c_0), \qquad \Delta_{n2} \overset{p}{\to} \begin{bmatrix} \frac{1}{2}\mu_2\varkappa(c_0)\beta^{(2)}(z_0) \\ 0_{d\times 1} \end{bmatrix} \quad \text{and} \quad \Delta_{n3} = O_p(1). \tag{A.9}$$

Using (A.9) and noting that $n^{3/2}h^5 = O(1)$, we immediately have

$$\sqrt{n^{3/2}h}\,\Delta_{n1}^{-1}\left[h^2\Delta_{n2} - h^2 C(z_0)\right] = o_p(1) \quad \text{and}$$
$$\sqrt{n^{3/2}h}h^3\Delta_{n1}^{-1}\Delta_{n3} = O_p\left(\sqrt{n^{3/2}h^7}\right) = o_p(1)\,. \qquad \textbf{(A.10)}$$

To derive the asymptotic result for the last term in (A.8) (i.e., Δ_{n4}), we construct the continuous martingale process

$$M_n(r) = \sqrt{d_n}\sum_{t=1}^{j-1} K\left(\omega_n^z\left(\frac{t-1}{n}\right)\right) Q_n\left(\frac{t-1}{n}\right)\left[B_u\left(\frac{\tau_{n,t}}{n}\right) - B_u\left(\frac{\tau_{n,t-1}}{n}\right)\right]$$
$$+ \sqrt{d_n}K\left(\omega_n^z\left(\frac{j-1}{n}\right)\right) Q_n\left(\frac{j-1}{n}\right)\left[B_u(r) - B_u\left(\frac{\tau_{n,j-1}}{n}\right)\right]$$

for $\tau_{n,j-1}/n < r \leq \tau_{n,j}/n$, $j = 1,2,\ldots,n$.

Hence, we have

$$\sqrt{n^{3/2}h}\,\Delta_{n4} = M_n\left(\frac{\tau_{n,n}}{n}\right) = M_n(1)\left[1 + o_p(1)\right], \qquad \textbf{(A.11)}$$

where the second equality uses (A.6). Assumption B ensures that $M_n(\cdot)$ is a continuous martingale vanishing at 0 and has a quadratic variation process $[M_n]$ given by

$$[M_n]_r = \sigma_u^2 d_n \sum_{t=1}^{j-1} K^2\left(\omega_n^z\left(\frac{t-1}{n}\right)\right) Q_n\left(\frac{t-1}{n}\right) Q_n\left(\frac{t-1}{n}\right)^T\left(\frac{\tau_{n,t}}{n} - \frac{\tau_{n,t-1}}{n}\right)$$
$$+ \sigma_u^2 d_n K^2\left(\omega_n^z\left(\frac{j-1}{n}\right)\right) Q_n\left(\frac{j-1}{n}\right) Q_n\left(\frac{j-1}{n}\right)^T\left(r - \frac{\tau_{n,j-1}}{n}\right) \qquad \textbf{(A.12)}$$

for $\tau_{n,j-1}/n < r \leq \tau_{n,j}/n$, $j = 1,2,\ldots,n$.

Applying the proof method used in the proof of Lemma 1, we obtain that

$$[M_n]_r \xrightarrow{p} \sigma_u^2 \begin{bmatrix}\nu_0 & 0\\ 0 & \nu_2\end{bmatrix} \otimes \varkappa_r(c_0) \stackrel{def}{=} \Lambda(r) \qquad \textbf{(A.13)}$$

holds uniformly over $r \in [0,1]$, where $\nu_j = \int u^j K^2(u)\,du$ and $\varkappa_r(c_0) = \int_0^r B_X(s) B_X(s)^T\, dL_{B_Z}(s, c_0)$.

Now, for any $\lambda \in R^{2d}$, we define $G_{n,\lambda}(r) = \lambda^T M_n(r)$. Now $G_{n,\lambda}(\cdot)$ is a continuous martingale and has a quadratic variation process given by $[G_{n,\lambda}] = \lambda^T [M_n]\lambda$. For any $a \in R^d$ and $b \in R$, the covariance process of $\left(G_{n,\lambda}, a^T B_X + bB_Z\right)$ is given by, for $\tau_{n,j-1}/n < r \leq \tau_{n,j}/n$, $j = 1,2,\ldots,n$,

$$\left[G_{n,\lambda}, a^T B_X + bB_Z\right]_r$$
$$= \sqrt{d_n}\sum_{t=1}^{j-1} K\left(\omega_n^z\left(\frac{t-1}{n}\right)\right)\lambda^T Q_n\left(\frac{t-1}{n}\right)\left(\frac{\tau_{n,t}}{n} - \frac{\tau_{n,t-1}}{n}\right)\left(a^T\Sigma_{xu} + b\Sigma_{zu}\right)$$
$$+ \sqrt{d_n}K\left(\omega_n^z\left(\frac{j-1}{n}\right)\right)\lambda^T Q_n\left(\frac{j-1}{n}\right)\left(r - \frac{\tau_{n,j-1}}{n}\right)\left(a^T\Sigma_{xu} + b\Sigma_{zu}\right)$$
$$= \left(a^T\Sigma_{Xu} + b\Sigma_{Zu}\right)\frac{\sqrt{d_n}}{n}\sum_{t=1}^{[nr]} K\left(\omega_n^z\left(\frac{t-1}{n}\right)\right)\lambda^T Q_n\left(\frac{t-1}{n}\right)\left[1 + o_p(1)\right]$$
$$= O_p\left(d_n^{-1/2}\right) = o_p(1) \quad \text{as } n \to \infty, \qquad \textbf{(A.14)}$$

where $\Sigma_{Xu} = cov(B_X, B_u)$ and $\Sigma_{Zu} = cov(B_Z, B_u)$, and Theorem 1 of Phillips (2009) (i.e., (5)) is used to derive (A.14).

Define a sequence of time changes $\rho_{n,\lambda}(t) = \inf\{s \in [0,1] : [G_{n,\lambda}]_s > t\}$. Equation (A.14) implies

$$\left[G_{n,\lambda}, a^T B_X + b B_Z\right]_{\rho_{n,\lambda}(t)} \xrightarrow{p} 0 \quad \text{as } n \to \infty \tag{A.15}$$

for each $t \in [0,1]$. Let $B_\lambda^n(t) = G_{n,\lambda}(\rho_{n,\lambda}(t))$. By Theorem 1.6 of Revuz and Yor (2005, p. 181) B_λ^n is the DDS (Dambis, Dubins-Schwarz) Brownian motion of the continuous martingale $G_{n,\lambda}$. We then apply the asymptotic version of Knight theorem (Thm. 2.3 of Revuz and Yor, 2005, p. 524) to obtain $B_\lambda^n \Longrightarrow W_G$, where W_G is a standard Brownian motion process independent of $a^T B_X + b B_Z$. Since the above results hold for all $a \in R^d$ and $b \in R$, it implies that W_G is independent of both B_X and B_Z. Therefore, W_G is independent of $\lambda^T \Lambda(1)\lambda$, where $[G_{n,\lambda}]_1 \xrightarrow{p} \lambda^T \Lambda(1)\lambda$. Combining with $G_{n,\lambda}(r) = B_\lambda^n([G_{n,\lambda}]_r)$ for $r \in [0,1]$, we have $G_{n,\lambda}(1) = B_\lambda^n([G_{n,\lambda}]_1) \xrightarrow{d} \lambda^T \Lambda(1)^{1/2} N$, where $N \sim N(0, I_{2d})$ and $\Lambda(1)^{1/2}$ is the square root of the matrix $\Lambda(1)$; i.e., $\Lambda(1) = \Lambda(1)^{1/2}\Lambda(1)^{1/2}$.

For the q defined in Assumption B and any sequence $\{\alpha_t\}$, we have $\mathbb{E}\left(\max_{1\le t\le n}|\alpha_t|\right) = \mathbb{E}\left(\max_{1\le t\le n}|\alpha_t|^q\right)^{1/q} \le \mathbb{E}\left(\sum_{t=1}^n |\alpha_t|^q\right)^{1/q} \le \left(\sum_{t=1}^n \mathbb{E}|\alpha_t|^q\right)^{1/q}$, where the last inequality follows from Jensen's inequality. Applying this result to $\alpha_t \overset{def}{=} \sqrt{d_n}\, T_j\left(\frac{Z_t - z_0}{h}\right)\lambda^T X_t u_t / n$, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\max_{1\le t\le n}\left|\sqrt{d_n}\, T_j\left(\frac{Z_t - z_0}{h}\right)\right|\left|\frac{\lambda^T X_t}{\sqrt{n}}\right|\left|\frac{u_t}{\sqrt{n}}\right|\right] \\
&\quad \le \frac{\sqrt{d_n}}{n}\left[\sum_{t=1}^n \mathbb{E}\left(\left|T_j\left(\frac{Z_t - z_0}{h}\right)\right|\left|\lambda^T X_t\right||u_t|\right)^q\right]^{1/q} \\
&\quad \le \frac{\sqrt{d_n}}{n}\left[\max_{1\le t\le n}\mathbb{E}\left(|u_t|^q\,|\mathcal{F}_{n,t-1}\right)\sum_{t=1}^n \mathbb{E}\left(\left|T_j\left(\frac{Z_t - z_0}{h}\right)\right|\left|\lambda^T X_t\right|\right)^q\right]^{1/q} \\
&\quad \le C\frac{\sqrt{d_n}}{n}\left[\sum_{t=1}^n \mathbb{E}\left(\left|T_j\left(\frac{Z_t - z_0}{h}\right)\right|\left|\lambda^T X_t\right|\right)^q\right]^{1/q} \\
&\quad \le C\frac{\sqrt{d_n}}{n}\left[\sum_{t=1}^n h\left(\sqrt{t}\right)^{q-1}\right]^{1/q} \\
&\quad = \left(\sqrt{n}h\right)^{-1/2+1/q} = o(1) \quad \text{since } \sqrt{n}h \to \infty \quad \text{as } n \to \infty \quad \text{and } q > 2,
\end{aligned} \tag{A.16}$$

where the last inequality in (A.16) follows from the change of the variable argument ($\xi_{z,t} = Z_t/\sqrt{t}$, $\xi_{x,t} = X_t/\sqrt{t}$ and $v = \left(\sqrt{t}\xi_{z,t} - z_0\right)/h$ below)

$$\begin{aligned}
&\mathbb{E}\left(\left|T_j\left(\frac{Z_t - z_0}{h}\right)\right|\left|\lambda^T X_t\right|\right)^q \\
&\quad = \left(\sqrt{t}\right)^q \mathbb{E}\left[\left|\lambda^T \xi_{x,t}\right|^q \int \left|T_j\left(\frac{\sqrt{t}\xi_{z,t} - z_0}{h}\right)\right|^q f_{0,t}^z\left(\xi_{z,t}|\xi_{x,t}\right) d\xi_{z,t}\right] \\
&\quad = h\left(\sqrt{t}\right)^{q-1}\mathbb{E}\left[\left|\lambda^T \xi_{x,t}\right|^q \int \left|T_j(v)\right|^q f_{0,t}^z\left(\frac{hv + z_0}{\sqrt{t}}|\xi_{x,t}\right) dv\right]
\end{aligned}$$

$$
\begin{aligned}
&\leq Ch\left(\sqrt{t}\right)^{q-1}\int\left|T_j(v)\right|^q dv\,\mathbb{E}\left(\left|\lambda^T\xi_{x,t}\right|^q\right) \quad \text{by Assumption A4(i)}\\
&\leq Ch\left(\sqrt{t}\right)^{q-1}\int\left|\lambda^T x\right|^q f(x)\,dx \quad \text{by Assumption A4(ii)}\\
&\leq Ch\left(\sqrt{t}\right)^{q-1} \quad \text{by Assumption A4(ii).}
\end{aligned}
$$

By Corollary 6.30 of Jacod and Shiryaev (2003, p. 385), we obtain

$$
\left(G_{n,\lambda}(1),\left[G_{n,\lambda}\right]_1\right)\stackrel{d}{\to}\left(\left(\lambda^T\Lambda(1)\lambda\right)^{1/2}N,\lambda^T\Lambda(1)\lambda\right). \tag{A.17}
$$

Applying the Cramér-Wold device, we therefore know that $(M_n(1),[M_n]_1)\stackrel{d}{\to}\left(\Lambda(1)^{1/2}N,\Lambda(1)\right)$. By the continuous mapping theorem, it follows that

$$
[M_n]_1^{-1/2}M_n(1)\stackrel{d}{\to}N(0,I_{2d}). \tag{A.18}
$$

From

$$
[M_n]_1=\sigma_u^2 d_n/n\sum_{t=1}^n K^2\left(\omega_n^z\left(\frac{t-1}{n}\right)\right)Q_n\left(\frac{t-1}{n}\right)Q_n\left(\frac{t-1}{n}\right)^T\left[1+o_p(1)\right], \tag{A.19}
$$

$$
\Delta_{n1}=(d_n/n)\sum_{t=1}^n K\left(\omega_n^z\left(\frac{t-1}{n}\right)\right)Q_n\left(\frac{t-1}{n}\right)Q_n\left(\frac{t-1}{n}\right)^T, \tag{A.20}
$$

and using the short-hand notation $\omega_{n,t}^z=\omega_n^z\left(\frac{t-1}{n}\right)$, we have

$$
\begin{aligned}
&\begin{bmatrix}1 & 0\\ 0 & \mu_2^{-1}\end{bmatrix}\otimes I_d\,\Delta_{n1}-\sigma_u^{-2}\begin{bmatrix}\upsilon_0^{-1} & 0\\ 0 & \upsilon_2^{-1}\end{bmatrix}\otimes I_d\,[M_n]_1\\
&=\frac{d_n}{n}\sum_{t=1}^n\begin{bmatrix}K\left(\omega_{n,t}^z\right)-\upsilon_0^{-1}K^2\left(\omega_{n,t}^z\right) & \left[K\left(\omega_{n,t}^z\right)-\upsilon_0^{-1}K^2\left(\omega_{n,t}^z\right)\right]\omega_{n,t}^z\\ \left[\mu_2^{-1}K\left(\omega_{n,t}^z\right)-\upsilon_2^{-1}K^2\left(\omega_{n,t}^z\right)\right]\omega_{n,t}^z & \left[\mu_2^{-1}K\left(\omega_{n,t}^z\right)-\upsilon_2^{-1}K^2\left(\omega_{n,t}^z\right)\right]\left(\omega_{n,t}^z\right)^2\end{bmatrix}\\
&\quad\otimes\left[\psi_{nx}^0(t/n)\psi_{nx}^0(t/n)^T\right]+o_p(1)\\
&\stackrel{p}{\to}\int_{-\infty}^{\infty}\begin{bmatrix}K(r)-\upsilon_0^{-1}K^2(r) & \left[K(r)-\upsilon_0^{-1}K^2(r)\right]r\\ \left[\mu_2^{-1}K(r)-\upsilon_2^{-1}K^2(r)\right]r & \left[\mu_2^{-1}K(r)-\upsilon_2^{-1}K^2(r)\right]r^2\end{bmatrix}dr\\
&\quad\otimes\int_0^1 B_X(s)B_X(s)^T\,dL_{B_Z}(s,c_0)\\
&=0.
\end{aligned}
\tag{A.21}
$$

Combining (A.9), (A.13), and (A.21), we obtain

$$
\Delta_{n1}^{-1/2}[M_n]_1^{1/2}\stackrel{p}{\to}\sqrt{\sigma_u^2}\begin{bmatrix}\sqrt{\nu_0} & 0\\ 0 & \sqrt{\frac{\upsilon_2}{\mu_2}}\end{bmatrix}\otimes I_{2d}\stackrel{def}{=}\Sigma^{*1/2}. \tag{A.22}
$$

Since Σ^* is a constant matrix, combining (A.18) and (A.22) and applying the Slustky lemma, we get

$$
\Delta_{n1}^{-1/2}M_n(1)=\Delta_{n1}^{-1/2}[M_n]_1^{1/2}[M_n]_1^{-1/2}M_n(1)\stackrel{d}{\to}N\left(0,\Sigma^*\right). \tag{A.23}
$$

By (A.8), (A.9), (A.10), (A.11), and (A.23) we have shown that

$$\begin{aligned}
&\left(H_n^{-1} A_{n1} H_n^{-1}\right)^{1/2} \left\{H_n \left[\widehat{\theta}(z_0) - \theta(z_0)\right] - h^2 C(z_0)\right\} \\
&\quad \stackrel{d}{=} \left(n^{3/2} h\, \Delta_{n1}\right)^{1/2} \left[\Delta_{n1}^{-1} \left(h^2 \Delta_{n2} + h^3 \Delta_{n3} + \Delta_{n4}\right) - h^2 C(z_0)\right] \\
&\quad = \Delta_{n1}^{-1/2} \sqrt{n^{3/2} h}\, \Delta_{n4} + o_p(1) \\
&\quad = \Delta_{n1}^{-1/2} M_n(1) + o_p(1) \\
&\quad \stackrel{d}{\to} N(0, \Sigma^*).
\end{aligned}$$

This completes the proof of Theorem 1.