Journal of Econometrics 148 (2009) 101–113

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Functional-coefficient models for nonstationary time series data[☆]

Zongwu Cai[a,b,c,∗], Qi Li[d,e], Joon Y. Park[d]

[a] Department of Mathematics & Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
[b] Department of Economics, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
[c] Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen, Fujian 361005, China
[d] Department of Economics, Texas A&M University, College Station, TX 77843, USA
[e] Department of Economics, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Available online 15 October 2008

Keywords:
Nonstationary
Nonlinearity
Semiparametric estimation

ABSTRACT

This paper studies functional coefficient regression models with nonstationary time series data, allowing also for stationary covariates. A local linear fitting scheme is developed to estimate the coefficient functions. The asymptotic distributions of the estimators are obtained, showing different convergence rates for the stationary and nonstationary covariates. A two-stage approach is proposed to achieve estimation optimality in the sense of minimizing the asymptotic mean squared error. When the coefficient function is a function of a nonstationary variable, the new findings are that the asymptotic bias of its nonparametric estimator is the same as the stationary covariate case but convergence rate differs, and further, the asymptotic distribution is a mixed normal, associated with the local time of a standard Brownian motion. The asymptotic behavior at boundaries is also investigated.

Published by Elsevier B.V.

1. Introduction

Nonparametric estimation techniques offer numerous advantages relative to parametric techniques, due mainly to their flexibility and robustness to functional form misspecification, and have been embraced by applied researchers in social, behavioral and economic sciences. Asymptotic theory underlying nonparametric estimators and test statistics for many commonly used models has been well established for independent and identically distributed (iid) data as well as for weakly dependence data. However, little is known about the behavior with nonstationary (in particular, integrated with order one, denoted by I(1)) data, which have predominately been modeled linearly. The early nonparametric asymptotic analyses with nonstationary data include Phillips and Park (1998), Park and Hahn (1999), Chang and Martinez-Chombo (2003) and Juhl (2005). Phillips and Park (1998) and Juhl (2005) considered nonparametric estimation of regression models when the true data generating process is a linear unit root process, while the others considered the models linearized in the nonstationary variables. More recently,[1] Wang and Phillips (forthcoming, 2008) considered nonparametric estimation of a regression model with an I(1) regressor and Xiao (forthcoming) considered a varying coefficient model with I(1) regressors appearing in the parametric component of the model. Finally, Karlsen et al. (2007) considered nonparametric estimation of a regression model for a different (a more general) type of nonstationary processes, a subclass of the class of null recurrent Markov chains.

In this paper, we tackle a more general set-up for a class of semiparametric models with non-stationary covariates. Specifically, we focus on the popular varying coefficient regression model with some nonstationary covariates

$$Y_t = \beta(Z_t)^{\mathrm{T}} X_t + \varepsilon_t, \quad 1 \le t \le n, \tag{1.1}$$

where Y_t, Z_t and ε_t are scalar, $X_t = (X_{t1}, \dots, X_{td})^{\mathrm{T}}$ is a vector of covariates with dimension d, $\beta(\cdot)$ is a $d \times 1$ column vector function, and the superscript T denotes transpose of a matrix. For ease notation, we assume that Z_t is univariate case. Extension to multivariate Z_t involves fundamentally no new ideas but

☆ We thank Peter Robinson for his constructive discussions and comments on an early version of this paper. We also thank an associate editor and two anonymous referees for their helpful comments that greatly improved the paper. Cai's research was supported, in part, by the National Science Foundation grant DMS-0404954 and funds provided by the University of North Carolina at Charlotte. Li's research was partially supported by the Private Enterprize Research Center, Texas A&M University. Park's research is partially supported by the National Science Foundation.

∗ Corresponding author at: Department of Mathematics & Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223, USA.

E-mail addresses: zcai@uncc.edu (Z. Cai), qi@econmail.tamu.edu (Q. Li), jpark@tamu.edu (J.Y. Park).

[1] The first version of this paper was written independently of these recent works on nonparametric estimation of regression models with non-stationary covariates.

0304-4076/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.jeconom.2008.10.003

complicated notations. We observe (Y_t, X_t, Z_t) for $t = 1, \ldots, n$. When $\{(X_t, Z_t, \varepsilon_t)\}$ is stationary (denoted by I(0)) or iid, various versions of (1.1) have been considered by many authors, including but not limited to, for example, Chen and Tsay (1993), Hastie and Tibshirani (1993), Cai et al. (2000), Li et al. (2002), and among others. When ε_t is stationary and $Z_t = t$, Eq. (1.1) has been tackled by Robinson (1989, 1991), Cai (2007) and Chen and Hong (2007) for stationary X_t, by Park and Hahn (1999) and Chang and Martinez-Chombo (2003) for nonstationary X_t, and by Cai and Wang (2008) for nearly integrated X_t. When $X_t = 1$ and Z_t is I(1), Eq. (1.1) becomes a standard univariate nonparametric regression model as considered by Wang and Phillips (forthcoming, 2008) and Karlsen et al. (2007). Finally, when Z_t is I(0) and X_t is I(1), model (1.1) reduces to the case considered by Xiao (forthcoming).

The advantage of a varying coefficient model specification, compared with an unrestricted nonparametric regression, is that it attenuates the "curse of dimensionality" problem. It also includes many popular semiparametric models as special cases. For example, when X_t contains a constant, say the first component $X_{t1} = 1$, we can write $X_t^{\mathrm{T}} = (1, \tilde{X}_t^{\mathrm{T}})$. Further, if the coefficient vector associated with $\tilde{X}_t$ is a vector of constants, say γ, then the varying coefficient model reduces to a partially linear model $E(Y_t|X_t, Z_t) = \beta_1(Z_t) + \tilde{X}_t^{\mathrm{T}}\gamma$; see, e.g., Robinson (1988).

The remainder of the paper is organized as follows: Section 2 discusses the case when Z_t is stationary. Here, local linear estimators of coefficient functions are developed, and their asymptotic properties are established. A two-step estimation procedure is also proposed when some covariates are nonstationary and the rest are stationary. Section 3 considers the case when Z_t is nonstationary. Nonparametric kernel smoothing of the coefficient functions is developed and its asymptotic behavior is investigated. Concluding remarks are presented in Section 4. Proofs of the main results of the paper are given in two Appendices.

2. Models with stationary Z_t

We consider first the case when some or all components of X_t are I(1) and Z_t is strictly stationary. For expositional simplicity, we re-express (1.1) as the following varying coefficient model

$$Y_t = \beta(Z_t)^{\mathrm{T}} X_t + \varepsilon_t = \beta_1(Z_t)^{\mathrm{T}} X_{t1} + \beta_2(Z_t)^{\mathrm{T}} X_{t2} + \varepsilon_t, \quad 1 \le t \le n, \tag{2.1}$$

where X_{t1}, Z_t, and ε_t are stationary, X_{t2} is an I(1) vector, $\beta(Z_t) = (\beta_1(Z_t)^{\mathrm{T}}, \beta_2(Z_t)^{\mathrm{T}})^{\mathrm{T}}$, and $X_t = (X_{t1}^{\mathrm{T}}, X_{t2}^{\mathrm{T}})^{\mathrm{T}}$, where X_{ti} is a $d_i \times 1$ vector, $i = 1, 2$, $d_1 + d_2 = d$, and the first component of X_{t1} is identically one. In what follows, we assume that $E(\varepsilon_t \mid X_t, Z_t) = 0$ which implies that X_t and Z_t are uncorrelated with ε_t. Note that Y_t is allowed to be stationary or nonstationary. For example, model (2.1) can be applied to the analysis of purchasing power of parity, in which $X_{t2}^{\mathrm{T}} = (P_t, P_t^*, E_t)$ (and no X_{t1}), where P_t and P_t^* are the price levels of the domestic and a foreign country, E_t is the exchange rate between the domestic and the foreign currencies, and $Z_t = I_t - I_t^*$ is the difference between the domestic interest rate I_t and the foreign interest rate I_t^*. Then if Y_t is an I(0) variable, we say that P_t, P_t^* and E_t are co-integrated with a varying coefficient co-integration vector $\beta(Z_t)$ which is a vector of smooth functions of Z_t. This setting is more general than the usual assumption that β is a vector of constant parameters in the usual purchasing power of parity analysis.

2.1. Local linear estimation

It is well known in the literature; see, e.g., Fan and Gijbels (1996), that a local linear fitting has several nice properties, over the classical Nadaraya–Watson (local constant) method, such as high statistical efficiency in an asymptotic minimax sense, design-adaptation, and automatic edge correction. We estimate $\beta(\cdot)$ using a local linear fitting from observations $\{(X_t, Z_t, Y_t)\}_{t=1}^n$. We assume throughout the paper that $\beta(\cdot)$ is twice continuously differentiable, so that for any given grid point z, we use a local approximation as $\beta(z) + \beta^{(1)}(z)(Z_t - z)$ to approximate $\beta(Z_t)$, where $\beta^{(s)}(z) = \mathrm{d}^s\beta(z)/\mathrm{d}z^s$. Define

$$\begin{pmatrix} \widehat{\theta}_0 \\ \widehat{\theta}_1 \end{pmatrix} = \operatorname{argmin}_{\theta_0, \theta_1} \sum_{t=1}^n \left[Y_t - \theta_0^{\mathrm{T}} X_t - (Z_t - z)\,\theta_1^{\mathrm{T}} X_t\right]^2 \times K_h(Z_t - z), \tag{2.2}$$

where $K_h(u) = h^{-1}K(u/h)$, $K(\cdot)$ is a kernel function satisfying Assumption A3 below, $\widehat{\theta}_0 = \widehat{\beta}(z)$ estimates $\beta(z)$, and $\widehat{\theta}_1 = \widehat{\beta}^{(1)}(z)$ estimates $\beta^{(1)}(z)$. Then, $\widehat{\beta}(z)$ and $\widehat{\beta}^{(1)}(z)$ can be expressed as

$$\begin{pmatrix} \widehat{\beta}(z) \\ \widehat{\beta}^{(1)}(z) \end{pmatrix} = \left[\sum_{t=1}^n \begin{pmatrix} X_t \\ (Z_t - z)\,X_t \end{pmatrix}^{\otimes 2} K_h(Z_t - z)\right]^{-1} \times \sum_{t=1}^n \begin{pmatrix} X_t \\ (Z_t - z)X_t \end{pmatrix} Y_t\, K_h(Z_t - z), \tag{2.3}$$

where $A^{\otimes 2} = A\,A^{\mathrm{T}}$ ($A^{\otimes 1} = A$) for a vector or matrix A.

2.2. Notations and assumptions

Since X_{t2} is a vector of I(1) processes, it can be re-expressed as $X_{t2} = X_{t-1,2} + \eta_t = X_{02} + \sum_{s=1}^t \eta_s$ $(t \ge 1)$, where $\{\eta_s\}$ is an I(0) process with mean zero and variance Ω_η. Then,

$$\frac{X_{[nr]2}}{\sqrt{n}} \equiv \frac{X_{t2}}{\sqrt{n}} = \frac{X_{0,2}}{\sqrt{n}} + \frac{1}{\sqrt{n}} \sum_{s=1}^t \eta_s = \frac{X_{0,2}}{\sqrt{n}} + \frac{1}{\sqrt{n}} \sum_{s=1}^{[nr]} \eta_s,$$

where $r = t/n$ and $[x]$ denotes the integer part of x. Under some regularity conditions, Donsker's theorem; see, for example, Theorems 14.1 and 19.2 in Billingsley (1999) for iid η_t and ρ-mixing η_t, respectively, generalizes in an obvious way to the multivariate cases and leads to

$$X_{[nr]2}/\sqrt{n} \Longrightarrow W_{\eta,2}(r) \quad \text{as } n \to \infty, \tag{2.4}$$

where $W_{\eta,2}(\cdot)$ is a d_2-dimensional Brownian motion on $[0,\ 1]$ with covariance matrix Σ_η and "$\Longrightarrow$" represents weak convergence. In particular, it follows from Merlevéde et al. (2006) that (2.4) holds if $\{\eta_t\}$ is a stationary strong (α-)mixing sequence satisfying, for some $\delta_0 > 0$,

$$E|\eta_t|^{2+\delta_0} < \infty, \quad \text{and} \quad \sum_{k=1}^{\infty} k^{(2+\delta_0)/\delta_0}\,\alpha(k) < \infty, \tag{2.5}$$

where $\alpha(\cdot)$ is the mixing coefficient; see, e.g., Hall and Heyde (1980) for more discussion on α-mixing process. Also, for any Borel measurable and totally Lebesgue integrable function $\Gamma(\cdot)$, one has

$$\frac{1}{n}\sum_{t=1}^n \Gamma(X_{[nr]2}/\sqrt{n}) \xrightarrow{d} \int_0^1 \Gamma(W_{\eta,2}(s))\mathrm{d}s \quad \text{as } n \to \infty,$$

where $\xrightarrow{d}$ denotes the convergence in distribution, so that, for $l = 1, 2$,

$$\frac{1}{n}\sum_{t=1}^n \left(X_{t2}/\sqrt{n}\right)^{\otimes l} \xrightarrow{d} \int_0^1 [W_{\eta,2}(r)]^{\otimes l}\mathrm{d}r \equiv W_{\eta,2}^{(l)} \quad \text{as } n \to \infty; \tag{2.6}$$

see Theorem 1.2 in Berkes and Horváth (2006) for details. Under stronger regularity conditions, (2.4) can be strengthened to the following strong approximation result

$$\sup_{0 \le r \le 1} \|X_{[nr]2}/\sqrt{n} - W_{\eta,2}(r)\| = O(n^{-\theta_*}\log^{\lambda_*}(n)) \tag{2.7}$$