Journal of the Korean Statistical Society (2005), **34**: *4, pp 345–366*

EFFICIENT ESTIMATION OF THE COINTEGRATING VECTOR IN ERROR CORRECTION MODELS WITH STATIONARY COVARIATES †

BYEONGSEON SEO[1]

ABSTRACT

This paper considers the cointegrating vector estimator in the error correction model with stationary covariates, which combines the stationary vector autoregressive model and the nonstationary error correction model. The cointegrating vector estimator is shown to follow the locally asymptotically mixed normal distribution. The variance of the estimator depends on the covariate effect of stationary regressors, and the asymptotic efficiency improves as the magnitude of the covariate effect increases. An economic application of the money demand equation is provided.

AMS 2000 subject classifications. Primary 62M10; Secondary 91B84.
Keywords. Cointegrating relationship, efficient estimation, stationary covariates.

1. INTRODUCTION

This paper aims to explore the asymptotic distribution of the cointegrating vector estimator in the vector error correction model (ECM) with stationary covariates. The stationary covariates have been used in many economic studies to consider the effect of stationary policy variables or to improve the model fitness on empirical grounds. However, the stationary covariates have been disregarded or treated less importantly in the cointegrated system, and statistical inference on the cointegrating vector has been based on the distribution theory, which does not allow for stationary variables. In this paper, we allow stationary variables in the cointegrated system and develop the distribution theory for the cointegrating vector estimator.

Received August 2005; accepted December 2005.
†This work was supported by Korea Research Foundation Grant (KRF-2002-041-B00054).
[1]Department of Economics, College Station, TX 77843-4228, USA (e-mail : seo@econmail.tamu.edu)

Suppose we have a nonstationary time series x_t, which is generated by an error correction model (ECM). We also have a stationary variable z_t generated by a stationary vector autoregressive (VAR) model as follows:

$$\Delta x_t = \Pi x_{t-1} + \sum_{i=1}^{l} C_{1i}\Delta x_{t-i} + \sum_{i=1}^{m} C_{2i} z_{t-i} + \nu_t$$
$$z_t = \sum_{i=1}^{l} \Phi_{1i}\Delta x_{t-i} + \sum_{i=1}^{m} \Phi_{2i} z_{t-i} + e_t$$

where

$$\begin{pmatrix} \nu_t \\ e_t \end{pmatrix} \sim i.i.d.(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{\nu\nu} & \Sigma_{\nu e} \\ \Sigma_{e\nu} & \Sigma_{ee} \end{pmatrix}).$$

The ECM error is correlated with the VAR error, so the likelihood function can be optimized by joint estimation. The joint model can be decomposed into the conditional and the marginal models. The cointegrating vector β and the adjustment vector α can be estimated without loss of efficiency, when there are no restrictions on the parameters across equations, from the following conditional error correction model:

$$\Delta x_t = \Pi x_{t-1} + \sum_{i=1}^{l} \Gamma_i \Delta x_{t-i} + \sum_{i=0}^{m} \Psi_i z_{t-i} + u_t,$$

where $\Pi = \alpha\beta'$, $u_t = \nu_t - \Psi_0 e_t$, $\Psi_0 = \Sigma_{\nu e}\Sigma_{ee}^{-1}$, $\Gamma_i = C_{1i} - \Psi_0\Phi_{1i}$, and $\Psi_j = C_{2j} - \Psi_0\Phi_{2j}$ for $i = 1, 2, \ldots, l$ and $j = 1, 2, \ldots, m$.

The conditional ECM includes the stationary variable z_t while the marginal ECM does not use the variable. If the ECM error is uncorrelated with the VAR error, the conditional ECM is equivalent to the marginal ECM. Thus, the role of the stationary covariate appears when the VAR error is correlated with the ECM error.

It is natural to ask whether we can rely on the existing distribution theory even after we include stationary variables. First of all, the proof for mixed normality of the cointegrating vector estimator does not seem trivial as we assume. In addition, it is necessary to examine the efficiency of the estimator because the joint estimation in general improves the efficiency of estimators in time series models with stationary variables. It has been assumed that the asymptotic distribution of the cointegrating vector estimator is invariant to the stationary regressors. In this study, we consider stationary covariates explicitly and provide the distribution theory for the cointegrating vector estimator.

The distribution theory of the cointegrating vector estimator has been developed by Johansen (1988, 1991), Phillips and Hansen (1990), Phillips (1991), Saikkonen (1991, 2001), and Stock and Watson (1993). Although there exists vast literature on the cointegrating vector, the literature on theoretical aspects of stationary covariates in the model with nonstationary variables is still sparse. Saikkonen (1991) considered the stationary covariates in the nonstationary regression model. The unit root tests with stationary covariates are explored by Hansen (1995) and Elliot and Jansson (2003), and inference on the cointegration rank in the ECM with stationary covariates has been developed by Seo (1998) and Rahbek and Mosconi (1999). This study is to explore the distribution of the cointegrating vector estimator in the vector ECM with stationary covariates.

We also discuss the consequences on the cointegrating vector estimator when we neglect the stationary covariates. The stationary covariates have been overlooked in time series models with nonstationary variables, and the effect of the omitted stationary covariates on the distribution of the cointegrating vector has not been explored. Here, we develop the associated distribution theory for the cointegrating vector estimator in the misspecified model, and thereby contribute to the literature.

In this study, we show that the cointegrating vector estimator follows the locally asymptotically mixed normal distribution. The variance of the cointegrating vector estimator depends on the effect of stationary covariates, and asymptotic efficiency improves as the magnitude of the covariate effect increases. The Monte Carlo simulation study indicates that the covariate effect generates a significant amount of efficiency gain.

Our model is related to the partial system or the structural error correction model because the partial system is specified by conditioning some variables of interest on the other remaining nonstationary variables. In this respect, the literature on the structural error correction model such as Boswijk (1995), Ericsson (1995), Johansen (1992), and Harbo *et al.* (1998) is worth mentioning. However, our model considers the conditional stationary variables explicitly. Because the stationary variables do not involve the cointegrating relationship, our model is different from the partial system.

The next section deals with the model. Section 3 provides the asymptotic theory for the cointegrating vector estimator. The omission of the stationary covariates and its consequences on the cointegrating vector estimator are discussed in Section 4. Section 5 extends the analysis to the models with deterministic trends. Section 6 provides simulation evidence. An economic application is provided in Section 7.

2. THE MODEL

Consider the p-dimensional nonstationary time series x_t and the k-dimensional stationary variable z_t, $t = 1, \ldots, n$, which follow an error correction model (ECM) and a stationary vector autoregressive (VAR) model, respectively, as follows:

$$\Delta x_t = \Pi x_{t-1} + \Gamma(L)\Delta x_t + \Psi(L) z_t + u_t \tag{2.1}$$

$$z_t = \Phi_1(L)\Delta x_t + \Phi_2(L) z_t + e_t, \tag{2.2}$$

where $\Gamma(L) = \sum_{i=1}^{l} \Gamma_i L^i$, $\Psi(L) = \sum_{i=0}^{m} \Psi_i L^i$, $\Phi_1(L) = \sum_{i=1}^{l} \Phi_{1i} L^i$, and $\Phi_2(L) = \sum_{i=1}^{m} \Phi_{2i} L^i$.

We assume that the error u_t is a vector-valued independent sequence with $\Sigma = E(u_t u_t') < \infty$. Note that the ECM error u_t is uncorrelated with the VAR error of the stationary covariates.

Our model is the standard ECM with the stationary covariates. In empirical studies, the estimated errors often do not satisfy the regularity conditions, and they are correlated with stationary economic variables. Also, the stationary policy variables may be considered in macroeconomic models. Thus, the stationary covariates have been used in many studies such as Johansen and Juselius (1992), Baba *et al.* (1992), and Juselius (1995).

The ECM (2.1) can be written as

$$\Pi(L) x_t = v_t, \tag{2.3}$$

where $\Pi(L) = (1-L)I - \Pi L - \Gamma(L)(1-L) - \Psi(L)(I - \Phi_2(L))^{-1}\Phi_1(L)(1-L)$ and $v_t = u_t + \Psi(L)(I - \Phi_2(L))^{-1} e_t$.

ASSUMPTION 2.1. *(a) All roots of* $\det(\Pi(L)) = 0$ *lie outside or on the unit circle.*

(b) All roots of $\det(I - \Phi_2(L)) = 0$ *lie outside the unit circle.*

If the cointegration rank is known and equals r, then there exist $p \times r$ full column rank matrices α and β satisfying $\Pi = \alpha\beta'$. Let $\alpha_\perp$ and $\beta_\perp$ be $p \times (p-r)$ full column rank matrices such that $\alpha_\perp' \alpha = 0$ and $\beta_\perp' \beta = 0$. From the representation theorem by Engle and Granger (1987), the error correction model (2.1) has the following representation:

$$\Delta x_t = C(L)v_t \tag{2.4}$$

$$x_t = C(1)\sum_{i=1}^{t} v_i + C^*(L)v_t \tag{2.5}$$

$$w_t = \beta' x_t = \beta' C^*(L)v_t, \tag{2.6}$$

where $C(1) = \beta_\perp(\alpha_\perp' \Pi^*(1)\beta_\perp)^{-1}\alpha_\perp'$, $\Pi^*(L) = (\Pi(L)-\Pi(1))/(1-L)$, and $C^*(L) = (C(L) - C(1))/(1 - L)$.

The data generating process x_t has stochastic trends and a stationary component. If $\Pi = \alpha\beta'$, the null space of $C(1)$ is spanned by the cointegration space. Hence, $\beta' C(1) = 0$ and $C(1)\alpha = 0$. The stochastic trends in x_t are eliminated if we multiply the cointegrating vector. Thus, the cointegrating relationship $w_t = \beta' x_t$ is stationary.

Also, from the representation theorem, the stationary covariate z_t has the following representation:

$$z_t = D_1(L)u_t + D_2(L)e_t,$$

where $D_1(L) = (I-\Phi_2(L))^{-1}\Phi_1(L)C(L)$ and $D_2(L) = (I-\Phi_2(L))^{-1}[I+\Phi_1(L)C(L)\Psi(L)(I\ \Phi_2(L))^{-1}]$.

We define $\mathcal{F}_{t-1}$ as the σ-field generated by $\{x_{t-i}, z_t, z_{t-i}, i = 1, 2, \cdots\}$. We denote $\mathcal{G}_{t-1}$ as the sub-σ-field generated by $\{x_{t-i}, z_{t-i}, i = 1, 2, \cdots\}$, which satisfies $\mathcal{G}_{t-1} \subset \mathcal{F}_{t-1}$. We assume the following conditions:

ASSUMPTION 2.2. *(a)* $E(u_t|\mathcal{F}_{t-1}) = 0$ *and* $E(e_t|\mathcal{G}_{t-1}) = 0$.

(b) $\sup_t E|\eta_t|^q < \infty$ *for some* $q > 2$, *where* $\eta_t = (u_t', e_t')'$.

(c) $\sum_{k=1}^{\infty} k|B_k| < \infty$, *where* $v_t = u_t + B(L)e_t = u_t + \sum_{k=0}^{\infty} B_k L^k e_t$ *and* $B(L) = \Psi(L)(I - \Phi_2(L))^{-1}$.

(d) $\sum_{k=1}^{\infty} k^2|C_k| < \infty$, *where* $\Delta x_t = C(L)v_t = \sum_{k=0}^{\infty} C_k v_{t-k}$.

(e) $\sum_{k=1}^{\infty} k|D_{1k}| < \infty$ *and* $\sum_{k=1}^{\infty} k|D_{2k}| < \infty$, *where* $D_j(L) = \sum_{k=0}^{\infty} D_{jk}L^k$ *for* $j = 1, 2$.

Assumption 2.2-(a) implies that the error process $\eta_t = (u_t', e_t')'$ allows for conditional heteroskedasticity. Assumptions 2.2-(b) and (d) imply that the process $\{\Delta x_t, w_t\}$ is uniformly 2^+ bounded. In the same way, the process $\{z_t\}$ is uniformly 2^+ bounded under Assumptions 2.2-(b) and (e).

LEMMA 2.1. *Under Assumptions 2.1-2.2,*

$$\begin{pmatrix} n^{-1/2}\sum_{t=1}^{[nr]} u_t \\ n^{-1/2}\sum_{t=1}^{[nr]} v_t \end{pmatrix} \Rightarrow \begin{pmatrix} U(r) \\ V(r) \end{pmatrix} = BM(\begin{pmatrix} \Sigma & \Sigma \\ \Sigma & \Omega_{vv} \end{pmatrix}),$$

where $\Omega_{vv} = \Sigma + \Psi(1)(I - \Phi_2(1))^{-1}\Sigma_{ee}(I - \Phi_2(1))^{-1'}\Psi'(1)$.

3. MAIN RESULTS

We assume that the matrix Π in (2.1) is of rank r $(0 < r < p)$. Thus, there exist $p \times r$ full-column rank matrices α and β which satisfy the following:

$$\Pi = -\Pi(1) = \alpha\beta'.$$

The cointegrating relationship $\beta' x_t$ is stationary as defined in Engle and Granger (1987). Our model allows for stationary covariates z_t, and it is assumed that the stationary covariates are weakly exogenous, as defined by Engle *et al.* (1983), to the cointegrating relationship. As discussed in Johansen (1992) and Boswijk (1995), the cointegrating vector β can be estimated efficiently in the model (2.1) under weak exogeneity.

We use the following normalization of the cointegrating vector:

$$w_t(\boldsymbol{\beta}) = x_{1t} + \boldsymbol{\beta}' x_{2t}, \tag{3.1}$$

where x_{1t} is r-dimensional, x_{2t} is $(p-r)$-dimensional, and $\boldsymbol{\beta}$ is a $(p-r)\times r$ matrix.

The cointegrating vector can be identified from the normalization. Our representation of the cointegrating relationship has been used in many studies such as Phillips (1991).

Let $s_t = (s_{1t}', s_{2t}')'$, where $s_{1t} = (\Delta x_{t-1}', \Delta x_{t-2}', \ldots, \Delta x_{t-l}')'$ and $s_{2t} = (z_t', z_{t-1}', \ldots, z_{t-m}')'$. The ECM (2.1) can be written as follows:

$$\Delta x_t = \alpha w_{t-1}(\boldsymbol{\beta}) + \Gamma s_{1t} + \Psi s_{2t} + u_t,$$

where $\Gamma = (\Gamma_1, \Gamma_2, \ldots, \Gamma_l)$ and $\Psi = (\Psi_0, \Psi_1, \ldots, \Psi_m)$.

We define the parameter vector $\theta = (\text{vec}(\boldsymbol{\beta})', \theta_2')' \in \Theta$, where $\theta_2 = \text{vec}(\alpha, \Gamma, \Psi, \Sigma)$ We denote θ_0 as the true parameter value.

Under the auxiliary condition $u_t \sim N(0, \Sigma)$, the likelihood function can be defined as follows:

$$\mathcal{L}_n(\theta) = -\frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{t=1}^{n} u_t'(\theta)\Sigma^{-1}u_t(\theta), \tag{3.2}$$

where $u_t(\theta)$ is defined as u_t in (2.1).

We denote $\hat{\theta}$ as the MLE of θ. That is,

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} \mathcal{L}_n(\theta).$$

The maximum likelihood estimator $\hat{\theta}$ maximizes the likelihood function, and the first order condition is given by

$$g_n(\hat{\theta}) = \frac{\partial \mathcal{L}_n(\hat{\theta})}{\partial \theta} = 0,$$

where

$$\begin{aligned}
\frac{\partial \mathcal{L}_n(\theta)}{\partial \boldsymbol{\beta}} &= \sum_{t=1}^{n} x_{2t-1} u_t' \Sigma^{-1} \alpha, \\
\frac{\partial \mathcal{L}_n(\theta)}{\partial \alpha'} &= \sum_{t=1}^{n} w_{t-1} u_t' \Sigma^{-1}, \\
\frac{\partial \mathcal{L}_n(\theta)}{\partial \Gamma'} &= \sum_{t=1}^{n} s_{1t} u_t' \Sigma^{-1}, \\
\frac{\partial \mathcal{L}_n(\theta)}{\partial \Psi'} &= \sum_{t=1}^{n} s_{2t} u_t' \Sigma^{-1}, \text{ and} \\
\frac{\partial \mathcal{L}_n(\theta)}{\partial \Sigma^{-1}} &= \frac{n}{2}\Sigma - \frac{1}{2}\sum_{t=1}^{n} u_t u_t'.
\end{aligned}$$

The MLE of $\boldsymbol{\beta}$ is denoted as $\hat{\boldsymbol{\beta}}$, which can be calculated by reduced rank regression (Ahn and Reinsel, 1988) or canonical analysis (Box and Tiao, 1977). Other slope parameters can be estimated by least squares once the cointegrating vector is estimated.

LEMMA 3.1. *Under Assumptions 2.1-2.2 and* $\Pi = \alpha\beta'$,

$$\frac{1}{\sqrt{n}} x_{[nr]} \Rightarrow C(1)V(r), \tag{3.3}$$

where $C(1) = \beta_\perp(\alpha_\perp' \Pi^*(1)\beta_\perp)^{-1}\alpha_\perp'$.

We define $(p-r) \times (p-r)$ full-rank matrix $\beta_{2\perp}$, which is the partitioned matrix of $\beta_\perp$. We denote $C_2 = \beta_{2\perp}(\alpha_\perp' \Pi^*(1)\beta_\perp)^{-1}\alpha_\perp'$, which corresponds to x_{2t}, so that $\frac{1}{\sqrt{n}} x_{2[nr]} \Rightarrow C_2 V(r)$.

We define the standard Brownian motions $B_1(r)$ and $B_2(r)$ as follows:

$$\begin{pmatrix} (\alpha'\Sigma^{-1}\alpha)^{-1/2}\alpha'\Sigma^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} u_t \\ (\alpha_{\perp}'\Omega_{vv}\alpha_{\perp})^{-1/2}\alpha_{\perp}'\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} v_t \end{pmatrix} \Rightarrow \begin{pmatrix} B_1(r) \\ B_2(r) \end{pmatrix} = BM(\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}). \tag{3.4}$$

We note that $B_1(r)$ and $B_2(r)$ are r and $(p-r)$-dimensional, respectively, and these two Brownian motions are mutually independent.

THEOREM 3.1. *Under Assumptions 2.1-2.2 and* $\Pi = \alpha\beta'$,

$$n(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \;\Rightarrow\; A\,(\int B_2 B_2')^{-1} \int B_2 dB_1'\;(\alpha'\Sigma^{-1}\alpha)^{-1/2'} \tag{3.5}$$

where $A = [(\alpha_{\perp}'\Omega_{vv}\alpha_{\perp})^{1/2'}(\alpha_{\perp}'\Pi^*(1)\beta_{\perp})^{-1'}\beta_{2\perp}']^{-1}$.

The asymptotic distribution of the cointegrating vector estimator $n(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ depends on two independent Brownian motions. Thus, the cointegrating vector estimator follows the locally asymptotically mixed normal (LAMN) distribution:

$$n(\mathrm{vec}(\hat{\boldsymbol{\beta}}) - \mathrm{vec}(\boldsymbol{\beta})) \sim N(0, (\alpha'\Sigma^{-1}\alpha)^{-1} \otimes A(\int B_2 B_2')^{-1}A').$$

It should be noted that the variance of the cointegrating vector estimator depends on the covariate effect. Because $\Omega_{vv} = \Sigma + \Psi(1)(I - \Phi_2(1))^{-1}\Sigma_{ee}(I - \Phi_2(1))^{-1'}\Psi'(1)$, the covariate effect decreases A, which reduces the variance of the cointegrating vector estimator. This result corresponds to the asymptotic efficiency of the cointegrating vector estimator in the nonstationary regression model with stationary covariates, which is shown by Saikkonen (1991).

If there is no covariate effect, the distribution of the cointegrating vector is the same as that found in Johansen (1991). Therefore, the stationary covariates provide information, which causes the asymptotic efficiency of the cointegrating vector estimator. Asymptotic efficiency improves as the magnitude of the covariate effect increases.

In addition, the dependence of the stationary covariates amplifies the covariate effect. Thus, the variance of the cointegrating vector decreases as the matrix $I - \Phi_2(1)$ approaches zero. Also, the variance depends on α and $\alpha_{\perp}$. As α approaches 0, the cointegrating relationship gets weaker, and the variance of the estimator increases to infinity. On the other hand, the cointegrating relationship becomes evident and its variance decreases as $\alpha_{\perp}$ approaches 0.

4. DISCUSSION: OMITTED STATIONARY REGRESSORS

Suppose the cointegrating vector is estimated in the following model:

$$\Delta x_t = \alpha w_{t-1}(\boldsymbol{\beta}) + \Gamma(L)\Delta x_t + \nu_t. \tag{4.1}$$

From the representation theorem, the error ν_t in the misspecified model can be written as follows:

$$\nu_t = u_t + \Psi(L)z_t = H(L)v_t,$$

where $H(L) = I + \Psi(L)(I - \Phi_2(L))^{-1}\Phi_1(L)C(L)$.

The error ν_t contains the current and lagged values of stationary covariates. Thus, the error in the misspecified model is serially correlated and correlated with the regressor. In principle, serial correlation can be handled by taking a large VAR lag-length. However, the innovation of the covariate model is independent of that of the ECM, and thus perfect whitening may not be attained with a penalized information criterion function in a practical sense. Furthermore, if stationary covariates are omitted, the error ν_t is correlated with the regressor, which violates the regularity condition. Here, we discuss the plausible pitfalls in the estimation of the cointegrating relationship when the stationary covariates are excluded.

We denote $\tilde{\boldsymbol{\beta}}$ as the cointegrating vector estimator of the misspecified model, which is defined as the solution to the objective function as follows:

$$\tilde{\mathcal{L}}_n(\boldsymbol{\beta}, \alpha, \Gamma, \Sigma_{\nu\nu}) = -\frac{n}{2}\log|\Sigma_{\nu\nu}| - \frac{1}{2}\sum_{t=1}^{n} \nu_t' \Sigma_{\nu\nu}^{-1} \nu_t, \tag{4.2}$$

where $\Sigma_{\nu\nu} = E(\nu_t \nu_t')$.

ASSUMPTION 4.1. $\sum_{k=1}^{\infty} k|H_k| < \infty$, *where* $\nu_t = H(L)v_t = \sum_{k=0}^{\infty} H_k v_{t-k}$.

In this section, we use the following asymptotic results:

LEMMA 4.1. *Under Assumptions 2.1, 2.2, 4.1 and* $\Pi = \alpha\beta'$,

$$n^{-1}\sum_{t=1}^{n} x_{t-1}v_t' \Rightarrow C(1)\int_0^1 VdV' + M$$

$$n^{-1}\sum_{t=1}^{n} x_{t-1}\nu_t' \Rightarrow C(1)\int_0^1 VdV'H'(1) + \Upsilon,$$

where $\Upsilon = MH'(1) + \sum_{k=0}^{\infty} J_k \sum_{i=k}^{\infty} H_i'$, $J_k = E(\Delta x_t v_{t+k}')$, $M = C(1)\Lambda + E((C^*(L)v_{t-1})v_t')$, *and* $\Lambda = \sum_{k=1}^{\infty} E(v_t v_{t+k}')$.

We denote $\Omega_{\nu\nu}$ as the long-run variance of ν_t. To derive the asymptotic distribution of the cointegrating vector estimator of the misspecified model, we define the Brownian motions $\tilde{B}_2(r)$ and $B_2(r)$ as follows:

$$\begin{pmatrix} M_1^{-1/2}\alpha'\Sigma_{\nu\nu}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}\nu_t \\ (\alpha_\perp'\Omega_{vv}\alpha_\perp)^{-1/2}\alpha_\perp'\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}v_t \end{pmatrix} \Rightarrow \begin{pmatrix} \tilde{B}_2(r) \\ B_2(r) \end{pmatrix} = BM(\begin{pmatrix} I & G \\ G' & I \end{pmatrix}) \tag{4.3}$$

where $G = M_1^{-1/2}\alpha'\Sigma_{\nu\nu}^{-1}H(1)\Omega_{vv}\alpha_\perp(\alpha_\perp'\Omega_{vv}\alpha_\perp)^{-1/2'}$ and $M_1 = \alpha'\Sigma_{\nu\nu}^{-1}\Omega_{\nu\nu}\Sigma_{\nu\nu}^{-1}\alpha$.

If there is no covariate effect, $H(L) = I$ and $\Sigma_{\nu\nu} = \Omega_{vv} = \Sigma$. Thus, G becomes zero, and two Brownian motions $\tilde{B}_2(r)$ and $B_2(r)$ are mutually independent. However, these two Brownian motions are correlated if the covariate effect is nonzero. We define the standard Brownian motion $B_{1.2}(r)$, which is independent of $B_2(r)$ and satisfies $\tilde{B}_2(r) = GB_2(r) + (I - GG')^{1/2}B_{1.2}(r)$.

THEOREM 4.1. *Under Assumptions 2.1-2.2, 4.1 and* $\Pi = \alpha\beta'$,

$$n(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \Rightarrow A(\int B_2B_2')^{-1}(\int B_2d\tilde{B}_2' + \Delta)M_1^{1/2'}M_2^{-1}, \tag{4.4}$$

where $\Delta = \Omega_{vv}^{-1/2}C_2^{-1}\Upsilon\Sigma_{\nu\nu}^{-1}\alpha M_1^{-1/2'}$ *and* $M_2 = \alpha'\Sigma_{\nu\nu}^{-1}\alpha$.

Also, we have

$$n(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \Rightarrow A(\int B_2B_2')^{-1}(f(G) + \Delta)M_1^{1/2'}M_2^{-1}, \tag{4.5}$$

where $f(G) = \int B_2dB_2'G' + \int B_2d\tilde{B}_{1.2}'(I - GG')^{1/2'}$.

First, the distribution of the cointegrating vector estimator $n(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})$ depends on the functional $f(G)$, which is a mixed combination of the mixture normal and the nonstandard distributions. If there is no covariate effect, then the correlation matrix G becomes zero, and the cointegrating vector follows a mixed normal distribution. As the covariate effect increases, the correlation increases, and the limiting distribution tends to depart from normality. Because the nonstandard distribution is skewed and leptokurtic, standard inference generates invalid results.

Second, the error of the misspecified model is serially correlated, and thus the cointegrating vector estimator entails the asymptotic bias Δ. If there is no covariate effect, the asymptotic bias disappears as the error ν_t becomes independent. However, the covariate effect generates the asymptotic bias, which leads to inferential difficulty when the standard distribution theory is applied to the cointegrating vector estimator.

The simulation evidence shows that the distribution of the t-statistic based on $\tilde{\beta}$ is close to the normal distribution if the covariate effect is not strong. However, as the covariate effect increases, the distribution of the t-statistic shows wide variance, asymmetry, and fat-tailed behavior even when the VAR-lag length is selected large by the Akaike information criterion. Therefore, the asymptotic distribution of the cointegrating vector estimator depends on the covariate effect if the covariates are omitted. The limiting distribution is nonstandard, and thus the misspecified model leads to the departure from the standard distribution theory.

5. MODELS WITH DETERMINISTIC TRENDS

We consider the vector error correction model with deterministic trends as follows:

$$\Delta x_t = \mu k_t + \Pi x_{t-1} + \sum_{i=1}^{l} \Gamma_i \Delta x_{t-i} + \sum_{i=0}^{m} \Psi_i z_{t-i} + u_t.$$

If $k_t = 1$, then the model allows an intercept. If $k_t = (1,t)'$, then the model allows an intercept and a linear trend. For the general $(p \times q)$ coefficient matrix μ, we assume that $\alpha_{\perp}' \mu_q = 0$, where μ_q is the q-th column of μ, which corresponds to the trend of the highest order. Thus, we preclude the possibility that the demeaned or detrended variables contain the deterministic trend.

If the information on the deterministic trend is given, efficiency can be improved by using the model with the restriction on the deterministic trend. However, this restriction may cost a specification error, and thus we do not impose the restriction on the deterministic trend. Because our model uses the detrended data, our results are robust to the deterministic trend.

We define the detrended variable x_t^* as follows.

$$x_t^* = x_t - (\sum_{t=1}^{n} x_t s_t^{*'})(\sum_{t=1}^{n} s_t^* s_t^{*'})^{-1} s_t^*,$$

where $s_t^* = (k_t', s_{1t}', s_{2t}')'$.

Because x_t^* removes deterministic trends, the asymptotic distribution of the cointegrating vector estimator based on the detrended variable is invariant to the deterministic trends. In the model with detrended variables, we can extend the previous analysis without difficulty.

$$n^{-1/2}x^*_{[nr]} \Rightarrow C(1)V(r) - \int_0^1 C(1)V(r)K'(r)dr[\int_0^1 K(r)K'(r)dr]^{-1}K(r) \equiv C(1)V^*(r),$$

where $K(r) = 1$ for the model with an intercept, and $K(r) = (1,r)'$ for the model with an intercept and a linear trend.

COROLLARY 5.1. *Suppose* $\alpha'_{\perp}\mu_q = 0$. *Under Assumptions 1-2 and* $\Pi = \alpha\beta'$,

$$n(\hat{\beta} - \beta) \Rightarrow A(\int B^*_2 B^{*'}_2)^{-1} \int B^*_2 dB'_1(\alpha'\Sigma^{-1}\alpha)^{-1/2'} \tag{5.1}$$

where $B^*_2(r) = B_2(r) - \int_0^1 B_2(r)K'(r)dr[\int_0^1 K(r)K'(r)dr]^{-1}K(r)$.

In the models with deterministic trends, the asymptotic distribution of the cointegrating vector estimator is a mixed normal with a variance of $(\alpha'\Sigma^{-1}\alpha)^{-1} \otimes A(\int B^*_2 B^{*'}_2)^{-1}A'$. Thus, the asymptotic distribution of the cointegrating vector estimator is invariant to the deterministic trend.

6. SIMULATION EVIDENCE

This section examines the small sample performances of the cointegrating vector estimator by using the Monte Carlo simulation. The simulations are based on a bivariate error correction model with the stationary covariates z_t, which follow a bivariate VAR process.

$$\Delta x_t = \Pi x_{t-1} + C_1\Delta x_{t-1} + C_2 z_{t-1} + \eta_t \tag{6.1}$$

$$z_t = \Phi_1\Delta x_{t-1} + \Phi_2 z_{t-1} + e_t, \tag{6.2}$$

where $x_t = (x_{1t}, x_{2t})'$, $z_t = (z_{1t}, z_{2t})'$, $(\eta'_t, e'_t)' \sim i.i.d.\ (0, \Omega)$, and $\Omega = \begin{pmatrix} I & \rho I \\ \rho I & I \end{pmatrix}$.

We design the experiments on the distribution of the cointegrating vector estimators in the model (6.1) with

$$\Pi = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \begin{pmatrix} 1 \\ \beta \end{pmatrix}'.$$

The error process (η_t, e_t) is generated by the Gauss random number generator. We assume that the errors follow the multivariate normal distribution $N(0, \Omega)$.

TABLE 6.1 *Distribution of cointegrating vector estimators*

Parameters (ϕ_1, ϕ_2, ψ)		Estimator RMSE	Estimator MAE	T-statistic Mean	T-statistic S.D.	T-statistic Skewness	T-statistic Kurtosis	Coverage Rates 5%	Coverage Rates 95%
(0.0, 0.00, 0.00)	$\hat{\beta}$	0.0138	0.0102	0.0261	1.0628	0.0208	3.1090	0.0566	0.0637
	$\tilde{\beta}$	0.0138	0.0102	0.0290	1.0697	0.0478	3.1172	0.0576	0.0648
(0.0, 0.25, 0.25)	$\hat{\beta}$	0.0134	0.0099	0.0255	1.0643	0.0441	3.1274	0.0576	0.0643
	$\tilde{\beta}$	0.0138	0.0102	0.0279	1.0709	0.0483	3.1519	0.0565	0.0661
(0.0, 0.25, 0.50)	$\hat{\beta}$	0.0118	0.0087	0.0221	1.0600	0.0394	3.1295	0.0574	0.0631
	$\tilde{\beta}$	0.0140	0.0101	0.0193	1.0723	0.0037	3.1736	0.0600	0.0644
(0.0, 0.25, 0.75)	$\hat{\beta}$	0.0090	0.0067	0.0196	1.0583	0.0486	3.1275	0.0581	0.0630
	$\tilde{\beta}$	0.0137	0.0101	0.0110	1.0761	0.0130	3.1438	0.0585	0.0635
(0.0, 0.50, 0.25)	$\hat{\beta}$	0.0134	0.0099	0.0258	1.0659	0.0429	3.1392	0.0567	0.0648
	$\tilde{\beta}$	0.0138	0.0102	0.0279	1.0709	0.0483	3.1519	0.0565	0.0661
(0.0, 0.50, 0.50)	$\hat{\beta}$	0.0119	0.0088	0.0221	1.0617	0.0375	3.1428	0.0570	0.0626
	$\tilde{\beta}$	0.0140	0.0101	0.0193	1.0723	0.0037	3.1736	0.0600	0.0644
(0.0, 0.50, 0.75)	$\hat{\beta}$	0.0091	0.0067	0.0186	1.0588	0.0471	3.1340	0.0587	0.0621
	$\tilde{\beta}$	0.0137	0.0101	0.0110	1.0761	0.0130	3.1438	0.0585	0.0635
(0.0, 0.75, 0.25)	$\hat{\beta}$	0.0136	0.0100	0.0263	1.0667	0.0411	3.1477	0.0566	0.0646
	$\tilde{\beta}$	0.0138	0.0102	0.0279	1.0709	0.0483	3.1519	0.0565	0.0661
(0.0, 0.75, 0.50)	$\hat{\beta}$	0.0121	0.0089	0.0218	1.0608	0.0338	3.1549	0.0563	0.0636
	$\tilde{\beta}$	0.0140	0.0101	0.0193	1.0723	0.0037	3.1736	0.0600	0.0644
(0.0, 0.75, 0.75)	$\hat{\beta}$	0.0093	0.0069	0.0160	1.0542	0.0437	3.1282	0.0569	0.0617
	$\tilde{\beta}$	0.0137	0.0101	0.0110	1.0761	0.0130	3.1438	0.0585	0.0635
(0.2, 0.00, 0.25)	$\hat{\beta}$	0.0095	0.0070	0.0195	1.0560	0.0189	3.1160	0.0555	0.0605
	$\tilde{\beta}$	0.0106	0.0077	-0.0076	1.0550	-0.0004	3.1260	0.0603	0.0584
(0.2, 0.00, 0.50)	$\hat{\beta}$	0.0081	0.0060	0.0169	1.0519	0.0224	3.1340	0.0553	0.0605
	$\tilde{\beta}$	0.0105	0.0077	-0.0043	1.0627	-0.0059	3.1504	0.0582	0.0604
(0.2, 0.00, 0.75)	$\hat{\beta}$	0.0059	0.0044	0.0149	1.0506	0.0298	3.1094	0.0574	0.0596
	$\tilde{\beta}$	0.0105	0.0077	-0.0005	1.0704	0.0017	3.1929	0.0591	0.0621
(0.2, 0.25, 0.25)	$\hat{\beta}$	0.0091	0.0067	0.0187	1.0581	0.0185	3.1429	0.0563	0.0612
	$\tilde{\beta}$	0.0104	0.0076	-0.0280	1.0622	-0.0100	3.1448	0.0632	0.0581
(0.2, 0.25, 0.50)	$\hat{\beta}$	0.0077	0.0057	0.0168	1.0565	0.0277	3.1618	0.0567	0.0611
	$\tilde{\beta}$	0.0104	0.0076	-0.0233	1.0726	-0.0092	3.1526	0.0608	0.0604
(0.2, 0.25, 0.75)	$\hat{\beta}$	0.0056	0.0041	0.0159	1.0576	0.0369	3.1373	0.0594	0.0604
	$\tilde{\beta}$	0.0103	0.0076	-0.0162	1.0803	0.0030	3.1867	0.0630	0.0616
(0.2, 0.50, 0.25)	$\hat{\beta}$	0.0083	0.0061	0.0178	1.0596	0.0173	3.1811	0.0570	0.0625
	$\tilde{\beta}$	0.0101	0.0074	-0.0723	1.0908	-0.0277	3.1908	0.0714	0.0571
(0.2, 0.50, 0.50)	$\hat{\beta}$	0.0069	0.0051	0.0172	1.0610	0.0354	3.1821	0.0580	0.0618
	$\tilde{\beta}$	0.0101	0.0074	-0.0635	1.1134	-0.0157	3.2051	0.0723	0.0606
(0.2, 0.50, 0.75)	$\hat{\beta}$	0.0050	0.0037	0.0178	1.0626	0.0396	3.1510	0.0580	0.0608
	$\tilde{\beta}$	0.0101	0.0074	-0.0555	1.1251	0.0135	3.2015	0.0736	0.0656
(0.2, 0.75, 0.25)	$\hat{\beta}$	0.0063	0.0046	0.0162	1.0627	0.0325	3.1911	0.0587	0.0618
	$\tilde{\beta}$	0.0096	0.0069	-0.1963	1.2217	-0.0215	3.1715	0.1154	0.0645
(0.2, 0.75, 0.50)	$\hat{\beta}$	0.0052	0.0038	0.0176	1.0654	0.0422	3.1755	0.0587	0.0629
	$\tilde{\beta}$	0.0096	0.0069	-0.1777	1.2531	-0.0222	3.2308	0.1182	0.0704
(0.2, 0.75, 0.75)	$\hat{\beta}$	0.0037	0.0027	0.0170	1.0636	0.0311	3.1138	0.0573	0.0599
	$\tilde{\beta}$	0.0095	0.0069	-0.1611	1.2630	-0.0173	3.2171	0.1146	0.0755

The experiments on the distribution of the estimators are based on a sample size of 250 and 10,000 simulation replications. In the simulation, we fix $\beta = -1$, $\alpha_1 = -1$, and $\alpha_2 = 0$. The correlation ρ varies among $(0, 0.25, 0.50, 0.75)$. We vary ϕ_1 among $(0, 0.2)$, ϕ_2 among $(0, 0.25, 0.5, 0.75)$ with $\phi_1 + \phi_2 < 1$, where $\Phi_1 = \phi_1 I$ and $\Phi_2 = \phi_2 I$. Also, we vary C_1 and C_2 by setting $C_1 = \phi_1 I$ and $C_2 = \phi_1 I$.

Table 6.1 shows the root mean squared error (RMSE) and the mean absolute error (MAE) of the cointegrating vector estimators. When there is no covariate effect, the cointegrating vector with stationary covariates is almost equivalent to the estimator without covariates. As the covariate effect increases, the RMSE and MAE of the MLE $\hat{\beta}$ decrease sharply, while those of the estimator $\tilde{\beta}$ remain stable. The estimator $\tilde{\beta}$ is calculated at the VAR lag-length that is chosen by the AIC for each simulated data. For example, at $(\phi_1, \phi_2) = (0.2, 0.25)$ the RMSE of the MLE $\hat{\beta}$ decreases about 28% as the covariate effect ρ increases from 0 to 0.5, and the MAE decreases about 25%. Thus, the efficiency gain of the MLE of the cointegrating vector improves in the covariate effect.

We examine the size performance of the t-statistics for the null hypothesis $\mathcal{H}_0 : \beta = -1$. Table 6.1 shows the descriptive statistics and the coverage rates of the t-statistics based on the cointegrating vector estimators $\hat{\beta}$ and $\tilde{\beta}$. The coverage rate is defined as $P(T < v_{0.05})$ for the lower 5% size and $P(T > v_{0.95})$ for the upper 5% size, where T is the t-statistic and v is the critical value.

The descriptive statistics of the t-statistics based on the estimator $\hat{\beta}$ are consistent with the properties of the normal distribution. The coverage rates are very close to the true size for most parameter values, and thus statistical inference on the cointegrating vector can be based on the standard theory. However, the t-statistic based on the estimator $\tilde{\beta}$ reveals a large amount of size distortion, large variance, asymmetry, and leptokurtic behavior as the covariate effect increases. For example, at $(\phi_1, \phi_2, \rho) = (0.2, 0.75, 0.5)$ the t-statistic based on the estimator $\tilde{\beta}$ shows that the standard deviation is larger than unity, and so the lower and upper 5% coverage rates are overly stated. The simulation study indicates that inference based on the standard distribution may bring invalid results.

The simulation is designed to show the gain in the marginal case of the covariate effect. The simulation results support the main results when we include non-zero off-diagonal coefficients, coefficients on the lagged values, and the intercept. As we allow for stronger covariate effect, the efficiency of the cointegrating vector estimator improves, and at the same time wide variation of the estimator of the misspecified model becomes more pronounced.

TABLE 7.1 *Money demand equation: 1960Q1-1999Q4*

	$m_t - p_t$		y_t		r_t	
	With Stationary Covariates					
$\hat{\beta}$	1		-1.3903	(0.2345)	0.0124	(0.0058)
$\hat{\alpha}$	-0.0593	(0.0124)	0.0115	(0.0166)	-0.0927	(1.5100)
R^2	0.6989		0.4178		0.4645	
Log-likelihood	590.77		553.56		-141.39	
$Q(1)$	0.0563	[0.812]	0.0677	[0.795]	0.1247	[0.724]
$Q(12)$	9.5024	[0.660]	7.3507	[0.834]	15.791	[0.201]
ARCH LM(1)	0.6818	[0.410]	0.0003	[0.986]	0.0320	[0.858]
ARCH LM(12)	0.3911	[0.965]	0.6055	[0.834]	2.1492	[0.018]
	Without Stationary Covariates					
$\tilde{\beta}$	1		-1.1188	(0.3166)	-0.0144	(0.0042)
$\tilde{\alpha}$	-0.0411	(0.0136)	0.0300	(0.0129)	2.7507	(1.4304)
R^2	0.6552		0.3725		0.3625	
Log-likelihood	580.36		547.79		-154.81	
$Q(1)$	0.0087	[0.926]	0.0083	[0.927]	0.0006	[0.981]
$Q(12)$	12.042	[0.442]	7.1259	[0.849]	20.272	[0.062]
ARCH LM(1)	0.7832	[0.378]	0.0177	[0.894]	4.6330	[0.033]
ARCH LM(12)	0.8926	[0.556]	0.7394	[0.711]	10.225	[0.000]

NOTE :*The standard errors are in the parentheses, and the p-values are in the square brackets.*

7. EMPIRICAL APPLICATION

This section provides an empirical analysis on the cointegrating relationship of the U.S. money demand equation. The money demand relation has been assessed in many empirical studies such as Baba *et al.* (1992). Here, we consider the estimation and cointegration inference of the money demand relation with the stationary covariates.

The data set is extracted from the Federal Reserve Economic Database (FRED) for the period 1960Q1-1999Q4: m_t is M2, p_t the price index, y_t real GDP, and r_t the TB 3-month interest rate. All variables are in logarithms except the short-run interest rate. We also consider the stationary covariates: z_{1t} is the oil price change, and z_{2t} is the risk factor, which is defined as the difference between the Moody's Baa and Aaa corporate bond rates. The monthly data are converted to the quarterly data by taking the three-month average.

Because the real money balances and real income contain growth terms, we include a constant and a linear trend in the error correction model. The VAR model of (z_{1t}, z_{2t}) allows a constant. The augmented Dickey-Fuller test rejects

TABLE 7.2 *Cointegration tests: 1960Q1-1999Q4*

	LR Statistic	Asymptotic p-value
With Covariates	39.515	0.010
Without Covariates	34.609	0.153

NOTE : *The canonical correlation R is estimated at* $(0.247, 0.973, 1.000)$.

the unit root hypothesis of the oil price change and the risk factor. The empirical results show that the covariates are weakly exogenous to the cointegrating relationship. The lag lengths of the ECM and the VAR model picked sufficiently large by Akaike information criterion (AIC) are $l = 5$ and $m = 1$, respectively.

The long-run money demand relationship and adjustment coefficients are estimated in Table 7.1. When the stationary covariates are used, the standard error of the long-run income elasticity decreases compared to the estimate without covariates. The income elasticity of the M2 demand is estimated as greater than unity. The estimated interest semi-elasticity is statistically significant and corresponds to the economic model. On the other hand, the interest elasticity is positive if the stationary covariates are not used. When we use the covariates, the adjustment coefficient of the money equation becomes larger, which implies that the adjustment process becomes faster. The R-squared coefficients and the log-likelihoods increase in each equation when we use the stationary covariates.

The Ljung-Box Q-statistics show that the estimated errors are not serially correlated at the lag length 1 in each equation. However, at the lag length 12, the serial correlation appears in the interest rate equation when the stationary covariates are not used. The residual of the interest rate equation shows irregular movement around the early 1980s, when the Fed operating system was changed. The ARCH LM test rejects the null hypothesis of no ARCH effect in the interest rate equation. Although the stationary covariates help reduce the ARCH effect, the interest rate equation still contains heteroskedastic errors because of the policy change.

Table 7.2 shows that the null hypothesis of no cointegration is maintained at the 5% size if we use the LR test statistic without stationary covariates. When the stationary covariates are used, the LR statistic rejects the null at the 5% size based on the asymptotic p-value. The asymptotic p-values are calculated by using the canonical correlation estimates, as suggested in Seo (1998). The magnitude of the covariate effect can be measured by the canonical correlation, which is estimated at $(0.247, 0.973, 1.000)$. Thus, inference on the cointegrating relationship with stationary covariates suggests an empirical evidence in favor of the presence of the long-run money demand relation.

8. CONCLUDING REMARKS

In this study, we find that the stationary covariates affect the asymptotic distribution of the cointegrating vector estimator. The distribution of the cointegrating vector estimator is a mixed normal, and the efficiency of the estimator improves as the covariate effect increases. This study also shows that the omission of the stationary covariates affects the distribution theory of the cointegrating vector estimator.

Although stationary covariates have been used in economic models, the estimation and inference in the cointegrated system have been based on the distribution theory, which does not allow stationary covariates. Furthermore, in the econometric model with nonstationary variables, the theoretical aspects of the stationary covariates have been disregarded or treated less importantly. Because it shows the distribution theory in the cointegrated system with stationary covariates, this study is useful and necessary.

APPENDIX: MATHEMATICAL PROOFS

PROOF OF LEMMA 2.1.
The invariance principle of Phillips and Durlauf (1986) implies

$$\begin{pmatrix} n^{-1/2}\sum_{t=1}^{[nr]} u_t \\ n^{-1/2}\sum_{t=1}^{[nr]} e_t \end{pmatrix} \Rightarrow \begin{pmatrix} U(r) \\ E(r) \end{pmatrix} = BM\left(\begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma_{ee} \end{pmatrix}\right).$$

We show $V_{[nr]} \Rightarrow V(r) = BM(\Omega_{vv})$, where $V_{[nr]} = n^{-1/2}\sum_{t=1}^{[nr]} v_t$ and $v_t = u_t + B(L)e_t$.

Define $B^*(L) = \frac{B(L)-B(1)}{1-L}$. Since

$$\begin{aligned} \sup_t \|B^*(L)e_t\|_q &\leq \sum_{j=0}^{\infty} |B_j^*| \ \sup_t \|e_t\|_q \\ &\leq \sum_{j=0}^{\infty}\sum_{k=j+1}^{\infty} |B_k| \ \sup_t \|e_t\|_q = \sum_{k=1}^{\infty} k|B_k| \ \sup_t \|e_t\|_q < \infty \end{aligned}$$

for some $q > 2$,

$$P(sup_{r\in[0,1]}n^{-1/2}|V_{[nr]} - \sum_{t=1}^{[nr]} u_t - B(1)\sum_{t=1}^{[nr]} e_t| > \epsilon)$$
$$\leq P(sup_{r\in[0,1]}n^{-1/2}|B^*(L)e_{[nr]}| > \epsilon) \to 0.$$

Thus, Assumptions 2.1-2.2 imply

$$n^{-1/2}\sum_{t=1}^{[nr]} v_t \Rightarrow V(r) = BM(\Omega_{vv}),$$

where $\Omega_{vv} = \Sigma + B(1)\Sigma_{ee}B(1)'$. □

PROOF OF LEMMA 3.1. We show $n^{-1/2}x_{[nr]} \Rightarrow C(1)V(r)$. We need to show

$$P(sup_{r\in[0,1]}n^{-1/2}|x_{[nr]} - C(1)\sum_{t=1}^{[nr]} v_t| > \epsilon)$$
$$\leq P(sup_{r\in[0,1]}n^{-1/2}|C^*(L)v_{[nr]}| > \epsilon) \to 0.$$

We can show that $\{C^*(L)v_t\}$ is uniformly square integrable because Assumptions 2.1-2.2, 4.1 imply

$$\sup_t \|C^*(L)v_t\|_q \leq \sum_{j=0}^{\infty} |C_j^*| \sup_t \|v_t\|_q$$
$$\leq \sum_{j=0}^{\infty}\sum_{k=j+1}^{\infty} |C_k| \sup_t \|v_t\|_q = \sum_{k=1}^{\infty} k|C_k| \sup_t \|v_t\|_q < \infty$$

for some $q > 2$.

Thus, we get the desired result. □

PROOF OF THEOREM 3.1. The Hessian matrix $H_n(\theta)$ can be defined as

$$H_n(\theta) = \begin{pmatrix} \frac{\partial^2 \mathcal{L}_n(\theta)}{\partial b \partial b'} & \frac{\partial^2 \mathcal{L}_n(\theta)}{\partial b \partial \theta_2'} \\ \frac{\partial^2 \mathcal{L}_n(\theta)}{\partial \theta_2 \partial b'} & \frac{\partial^2 \mathcal{L}_n(\theta)}{\partial \theta_2 \partial \theta_2'} \end{pmatrix},$$

where $b = \mathrm{vec}(\boldsymbol{\beta})$.

We denote $D_{1n} = \mathrm{diag}(n, \ldots, n)$ and $D_{2n} = \mathrm{diag}(\sqrt{n}, \ldots, \sqrt{n})$, which correspond to the parameter vectors $\mathrm{vec}(\boldsymbol{\beta})$ and θ_2, respectively. Define a diagonal matrix $D_n = \mathrm{diag}(D_{1n}, D_{2n})$.

From the representation theorem, $\Delta x_t = C(L)v_t$, $w_t = \beta' C^*(L)v_t$, and $z_t = D_1(L)u_t + D_2(L)e_t$. Assumptions 2.1-2.2 imply that $\sup_t \|\Delta x_t\|_q < \infty$, $\sup_t \|w_t\|_q < \infty$, and $\sup_t \|z_t\|_q < \infty$ for some $q > 2$.

First, we show that the normalized Hessian matrix $D_n^{-1} H_n(\theta) D_n^{-1}$ is asymptotically block-diagonal, which can be implied by $n^{-1}\sum_{t=1}^n x_{2t-1}s_t' = O_p(1)$. This result has been proved by Phillips (1988) for the linear process and by Hansen (1992) for the strong mixing process.

Thus, $n^{-1}\sum_{t=1}^n x_{2t-1}s_t' = O_p(1)$ and $\frac{1}{n^{3/2}}\frac{\partial^2 \mathcal{L}_n(\theta)}{\partial b \partial \theta_2'} \to^p 0$, and therefore the normalized Hessian matrix is block-diagonal.

Next, we get the following asymptotic results:

$$
\begin{aligned}
-\frac{1}{n^2}\frac{\partial^2 \mathcal{L}_n(\theta)}{\partial b \partial b'} &= \alpha'\Sigma^{-1}\alpha \otimes \frac{1}{n^2}\sum_{t=1}^n x_{2t-1}x_{2t-1}' \\
&\Rightarrow \alpha'\Sigma^{-1}\alpha \otimes C_2 \int VV'C_2'
\end{aligned}
$$

$$
\frac{1}{n}\sum_{t=1}^n x_{2t-1}u_t' \Rightarrow \int C_2 V dU'.
$$

Therefore, we can show that

$$
\begin{aligned}
n(\hat{\boldsymbol\beta} - \boldsymbol\beta) &= (\frac{1}{n^2}\sum_{t=1}^n x_{2t-1}x_{2t-1}')^{-1}\frac{1}{n}\sum_{t=1}^n x_{2t-1}u_t'\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1} + o_p(1) \\
&\Rightarrow (\int C_2 VV'C_2)^{-1}\int C_2 VdU'\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1} \\
&= A(\int B_2 B_2')^{-1}\int B_2 dB_1'(\alpha'\Sigma^{-1}\alpha)^{-1/2'},
\end{aligned}
$$

where $A = [(\alpha_\perp'\Omega_{vv}\alpha_\perp)^{1/2'}(\alpha_\perp'\Pi^*(1)\beta_\perp)^{-1'}\beta_{2\perp}']^{-1}$. □

PROOF OF LEMMA 4.1. Show that $n^{-1}\sum_{t=1}^n x_{t-1}v_t' \Rightarrow C(1)\int_0^1 VdV' + M$, where $M = C(1)\Lambda + E((C^*(L)v_{t-1})v_t')$ and $\Lambda = \sum_{k=1}^\infty E(v_t v_{t+k}')$.

$$
\begin{aligned}
n^{-1}\sum_{t=1}^n x_{t-1}v_t' &= n^{-1}\sum_{t=1}^n C(1)V_{t-1}v_t' + n^{-1}\sum_{t=1}^n C^*(L)v_{t-1}v_t' \\
&\Rightarrow C(1)(\int_0^1 VdV' + \Lambda) + E((C^*(L)v_{t-1})v_t'),
\end{aligned}
$$

where $V_t = \sum_{i=1}^t v_i$.

Set $\xi_{t,k} = \sum_{j=1}^{k} \Delta x_{t-j}$ and $J_k = E(\Delta x_t v_{t+k}')$.

$$\begin{aligned} n^{-1}\sum_{t=1}^{n} x_{t-1}\nu_t' &= n^{-1}\sum_{t=1}^{n} x_{t-1}(H(L)v_t)' \\ &= \sum_{k=0}^{\infty} n^{-1}\sum_{t=1}^{n} V_{t-k-1}v_{t-k}'H_k' + \sum_{k=1}^{\infty} n^{-1}\sum_{t=1}^{n}(\xi_{t,k}v_{t-k}'H_k') \\ &\Rightarrow C(1)\int_0^1 VdV'H'(1) + \Upsilon, \end{aligned}$$

where $\Upsilon = MH'(1) + \sum_{k=0}^{\infty} J_k \sum_{i=k}^{\infty} H_i'$.

PROOF OF THEOREM 4.1.

The estimator $\tilde{\theta} = (\mathrm{vec}(\tilde{\beta})', \mathrm{vec}(\tilde{\alpha})', \mathrm{vec}(\tilde{\Gamma})', \mathrm{vec}(\tilde{\Sigma}_{\nu\nu})')'$ satisfies the first order condition $\frac{\partial \tilde{\mathcal{L}}_n(\tilde{\theta})}{\partial \theta} = 0$, where

$$\begin{aligned} \frac{\partial \tilde{\mathcal{L}}_n(\theta)}{\partial \beta} &= \sum_{t=1}^{n} x_{2t-1}\nu_t'\Sigma_{\nu\nu}^{-1}\alpha, \\ \frac{\partial \tilde{\mathcal{L}}_n(\theta)}{\partial \alpha'} &= \sum_{t=1}^{n} w_{t-1}\nu_t'\Sigma_{\nu\nu}^{-1}, \\ \frac{\partial \tilde{\mathcal{L}}_n(\theta)}{\partial \Gamma'} &= \sum_{t=1}^{n} s_{1t}\nu_t'\Sigma_{\nu\nu}^{-1}, \text{ and} \\ \frac{\partial \tilde{\mathcal{L}}_n(\theta)}{\partial \Sigma_{\nu\nu}^{-1}} &= \frac{n}{2}\Sigma_{\nu\nu} - \frac{1}{2}\sum_{t=1}^{n} \nu_t\nu_t'. \end{aligned}$$

Because $n^{-1}\sum_{t=1}^{n} x_{2t-1}w_{t-1}' = O_p(1)$ and $n^{-1}\sum_{t=1}^{n} x_{2t-1}s_{1t}' = O_p(1)$, the normalized Hessian matrix of (4.2) is asymptotically block-diagonal, and thus under Assumptions 2.1-2.2, 4.1 and $\mathcal{H}_0 : \Pi = \alpha\beta'$, we get the following result.

$$\begin{aligned} n(\tilde{\beta} - \beta) &= (n^{-2}\sum_{t=1}^{n} x_{2t-1}x_{2t-1}')^{-1}n^{-1}\sum_{t=1}^{n} x_{2t-1}\nu_t'\Sigma_{\nu\nu}^{-1}\alpha(\alpha'\Sigma_{\nu\nu}^{-1}\alpha)^{-1} + o_p(1) \\ &\Rightarrow (C_2\int VV'C_2')^{-1}(\int C_2VdV'H(1)' + \Upsilon)\Sigma_{\nu\nu}^{-1}\alpha(\alpha'\Sigma_{\nu\nu}^{-1}\alpha)^{-1} \\ &= A(\int B_2B_2')^{-1}(\int B_2d\tilde{B}_2' + \Delta)M_1^{1/2'}M_2^{-1} \end{aligned}$$

where $\Delta = \Omega_{vv}^{-1/2}C_2^{-1}\Upsilon\Sigma_{\nu\nu}^{-1}\alpha M_1^{-1/2'}$, $M_1 = \alpha'\Sigma_{\nu\nu}^{-1}\Omega_{\nu\nu}\Sigma_{\nu\nu}^{-1}\alpha$, and $M_2 = \alpha'\Sigma_{\nu\nu}^{-1}\alpha$.

ACKNOWLEDGEMENTS

I would like to thank Ted Juhl, Inmoo Kim, Sahmyong Kim, Qi Li, Joon Park, and the seminar participants at Sungkyunkwan University and the Texas Camp Econometrics for useful comments and suggestions. Special thanks are owed to two anonymous referees, who provided invaluable comments and suggestions. All errors are my responsibility.

REFERENCES

AHN, S. K., AND G. C. REINSEL (1988). "Nested Reduced-Rank Autoregressive Models for Multiple Time Series", *Journal of the American Statistical Association*, **83**, 849–856.

BABA, Y., D. HENDRY, AND R. STARR (1992). "The Demand for M1 in the U. S. A.", 1960-1988, *Review of Economic Studies*, **59**, 25–61.

BOSWIJK, H. P. (1995). "Efficient Inference on Cointegration Parameters in Structural Error Correction Models", *Journal of Econometrics*, **69**, 133–158.

BOX, G. E. P., AND G. C. TIAO (1977). "A Canonical Analysis of Multiple Time Series", *Biometrika*, **64**, 355–365.

ELLIOT, G., AND M. JANSSON (2003). "Testing for Unit Root with Stationary Covariates", *Journal of Econometrics*, **115**, 75–89.

ENGLE, R. F., AND C. W. J. GRANGER (1987). "Cointegration and Error Correction Representation, Estimation, and Testing", *Econometrica*, **55**, 251–276.

ENGLE, R. F., D. F. HENDRY, AND J. F. RICHARD (1983). "Exogeneity", *Econometrica*, **51**, 277–304.

ERICSSON, N. R. (1995). "Conditional and Structural Error Correction Models", *Journal of Econometrics*, **69**, 159–171.

HANSEN, B. E. (1992). "Convergence to Stochastic Integrals for Dependent Heterogeneous Processes", *Econometric Theory*, **8**, 489–500.

HANSEN, B. E. (1995). "Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power", *Econometric Theory*, **11**, 1148–1171.

HARBO, I., S. JOHANSEN, B. NIELSON, AND A. RAHBEK (1998). "Asymptotic Inference on Cointegration Rank in Partial Systems", *Journal of Business and Economic Statistics*, **16**, 388–399.

JOHANSEN, S. (1988). "Statistical Analysis of Cointegrating Vectors", *Journal of Economic Dynamics and Control*, **12**, 231–254.

JOHANSEN, S. (1991). "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, **59**, 1551–1580.

JOHANSEN, S. (1992). "Cointegration in Partial Systems and the Efficiency of Single-Equation Analysis", *Journal of Econometrics*, **52**, 389–402.

JOHANSEN, S., AND K. JUSELIUS (1992). "Testing Structural Hypotheses in a Multivariate Cointgeration Analysis of the PPP and the UIP for UK", *Journal of Econometrics*, **53**, 211–244.

JUSELIUS, K. (1995). "Do Purchasing Power Parity and Uncovered Interest Parity Hold in the Long Run? An Example of Likelihood Inference in a Multivariate Time Series Model", *Journal of Econometrics*, **69**, 211–240.

PHILLIPS, P. C. B. (1988). "Weak Convergence of Sample Covariance Matrices to Stochastic Integrals via Martingale Approximations", *Econometric Theory*, **4**, 528–533.

PHILLIPS, P.C.B. (1991). "Optimal Inference in Cointegrated System", *Econometrica*, **59**, 283–306.

PHILLIPS, P.C.B., AND S.N. DURLAUF (1986). "Multiple Time Series with Integrated Variables", *Review of Economic Studies*, **53**, 473–495.

PHILLIPS, P.C.B., AND B.E. HANSEN (1990). "Statistical Inference in Instrumental Variables Regression with I(1) Process", *Review of Economic Studies*, **57**, 99–124.

RAHBEK, A. AND R. MOSCONI (1999). "Cointgeration rank inference with stationary regressors in VAR models, Econometrics Journal", **2**, 76–91.

SAIKKONEN, P. (1991). "Asymptotically Efficient Estimation of Cointegration Regressions", *Econometric Theory*, **7**, 1–21.

SAIKKONEN, P. (2001). "Consistent Estimation in Cointegrated Vector Autoregressive Processes with Nonlinear Time Trends in Cointegrating Relations", *Econometric Theory*, **17**, 296 - 326.

SEO, B. (1998). "Statistical Inference on Cointegration Rank in Error Correction Models with Stationary Covariates", *Journal of Econometrics*, **85**, 339–386.

STOCK, J. AND M. WATSON (1993). "A Simple Estimator of Cointegrating Vectors in Highly Order Integrated Systems", *Econometrica*, **61**, 783–820.