Regressions for Partially Identified, Cointegrated Simultaneous Equations*

In Choi†
Kookmin University, Korea
Peter C. B. Phillips
Cowles Foundation for Research in Economics, Yale University, USA
&
University of Auckland

August, 1997

Abstract

This paper studies regressions for partially identified equations in simultaneous equations models (SEMs) where all the variables are I(1) and cointegrating relations are present. Asymptotic properties of OLS and 2SLS estimators under partial identification are derived. The results show that the identifiability condition is important for consistency of estimates in nonstationary SEMs as it is for stationary SEMs. Also, OLS and 2SLS estimators are shown to have different rates of convergence and divergence under partial identification, though they have the same rates of convergence and divergence for the two polar cases of full identification and total lack of identifiability. Even in the case of full identification, however, the OLS and 2SLS estimators have different distributions in the limit. Fully modified OLS regression and leads-and-lags regression methods are also studied. The results show that these two estimators have nuisance parameters in the limit under general assumptions on the regression errors and are not suitable for structural inference. The paper proposes 2SLS versions of these two nonstationary regression estimators that have mixture normal distributions in the limit under general assumptions on the regression errors, that are more efficient than the unmodified estimators, and that are suited to statistical inference using asymptotic chi-squared distributions. Some simulation results are also reported.

*The first draft of this paper was presented at the Inaugural Meeting of the New Zealand Econometric Study Group in Auckland, February 1996. The paper was typed by the authors in SW2.5. Phillips thanks the NSF for support under Grant #SBR 94-22922. Choi gratefully acknowledges financial support from the Korea Research Foundation.

†**Address correspondence to:** In Choi, Department of Economics, Kookmin University, 861–1, Chongnung-dong, Songbuk-gu, Seoul, 136–702 KOREA. E-mail: ichoi@kmu.kookmin.ac.kr

1 Introduction

Through the 1950's to the 1970's, simultaneous equations models (SEMs) were one of the major research themes in econometrics. Important results in the area are summarized in Hausman (1983), Hsiao (1983) and Phillips (1983). However, research on the SEM has attracted relatively less attention these days partly due to a perceived failure in large macroeconomic models and partly due to criticisms of simultaneous equations modelling, following work by Lucas (1976) and Sims (1980).

Nonetheless, the SEM still seems to be in widespread emprical use for forecasting and policy analyses especially in such non-academic institutions as private banks and government economics departments [see Adams and Klein (1991) for discussions of various structural models in present use]. Some would ascribe this phenomenon to slow dissemination of knowledge. But nearly two decades have passed since the Lucas and Sims critiques, and so this does not seem to be an appropriate explanation. The phenomenon may also be explained because it is hard to find methods that work obviously better than the SEM for policy analysis and forecasting. As an alternative to the SEM, vector autoregressions (VARs) were introduced by Sims for econometric policy analyses and forecasting. Though this approach has proved useful for many research problems, VARs are reduced forms that are not directly linked to economic theory and estimation becomes difficult when there are a large number of variables. Computations of general equilibrium models are often used for policy analyses [cf., Shoven and Whalley (1984), Kydland and Prescott (1995)], but the viability of these methods depends on how such issues as robustness and model selection are dealt with, so that researchers can employ these methods with confidence. In the light of these and many other considerations, it seems likely that the SEM will continue to be used in practical econometric work.

The question of how to use the SEM for nonstationary time series has not attracted much attention from researchers. This is unfortunate because it is now well accepted that many economic time series have stochastic trends and can be represented in terms of I(1) or near-I(1) processes. In particular, the identification of coefficient parameters in the nonstationary SEM, which is one of the most basic practical problems in constructing SEMs, has been little studied. An important exception is Hsiao (1995), who gives identification conditions and derives the asymptotic distributions of the ordinary least squares (OLS) and three stage least squares estimators for the nonstationary, dynamic SEM. But the effects of identification failure (i.e. failure of the rank condition) on estimation were not studied there or elsewhere.

The first purpose of this paper is to study the effects of identification failure on the OLS and two stage least squares (2SLS) estimators in nonstationary SEMs. These effects on the instrumental variables estimator in stationary SEMs were analyzed in Choi and Phillips (1992). Identification failure in nonstationary SEM is especially interesting, because it is well known that OLS is consistent when the regressors are I(1) and the regression errors are I(0), so that it may seem that traditional identification conditions are not as important for regressions with I(1) regressors. But the results in Section 2 of this paper show that the identification condition is important for consistency of both OLS and 2SLS estimators in nonstationary SEMs. Hsiao (1995) shows that the OLS estimator is not consistent for the nonstationary, dynamic SEM even when the identification condition holds. His results are obtained for nonstationary SEMs with lagged dependent and exogenous variables for which standard $\sqrt{T}$ asymptotics apply [cf., Park and Phillips (1989) and Sims, Stock and Watson (1990)] and then serial correlation between the regressors and regression errors results in inconsistent OLS estimation. The model we use here does not involve lagged variables, and instead the short-run dynamics in our model are nonparametrically incorporated into the regression errors, as in many articles in the area of nonstationary time series [e.g., Phillips and Hansen (1990), Stock and Watson (1993)].

The second purpose of this paper is to study the Phillips - Hansen (1990) fully modified OLS estimator and the leads-and-lags estimator [cf. Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993)] applied to a fully identified single equation in a nonstationary SEM. The results from this study will show that these two estimators do not have mixture normal distributions and associated Wald tests do not have chi-square distributions in the limit. To remedy these problems, this paper proposes 2SLS versions of the two regression estimators. These estimators have mixture normal distributions in the limit under general assumptions on the regression errors and are more efficient than the two OLS estimators. Moreover, Wald tests using the 2SLS versions of the estimators do have chi-square distributions in the limit.

The plan of this paper is as follows. Section 2 introduces the model and assumption, and derives the asymptotic distributions of the OLS and 2SLS estimators for partially identified cointegrated simultaneous equations. Section 3 derives the asymptotic distributions of the Phillips - Hansen fully modified OLS estimator and the leads-and-lags estimator applied to a fully identified single equation in the nonstationary SEM. Moreover, 2SLS versions of the two regression estimators are proposed and their asymptotic distributions are derived in this section. Small-scale simulation results are also reported. Section 4 contains a summary of the paper and some further remarks.

A few words on our notation: all limits are taken as $T \to \infty$ unless otherwise specified; weak convergence and convergence in probability are denoted $\Rightarrow$ and $\xrightarrow{p}$, respectively; the t-th row of the matrix A is denoted A_t; $f_{aa}(\cdot)$ denotes the spectral density matrix of the time series $\{a_t\}$; $P_X = X(X'X)^{-1}X'$ and $Q_X = I - P_X$; and $r(A)$ is the rank of the matrix A.

2 Asymptotic Theory of Regressions for Partially Identified, Cointegrated Simultaneous Equations

2.1 Framework and Assumptions

Suppose that a true structural relation is represented by the regression equation

$$y_1 = Y_2\beta + Z_1\gamma + u, \tag{2.1}$$

where y_1 $(T \times 1)$ and Y_2 $(T \times n)$ denote $n+1$ endogenous variables and, Z_1 $(T \times k_1)$ is a matrix of k_1 exogenous variables, and u is a disturbance vector. The meaning of exogenous variables here is different from that in the standard SEM, and will be explained shortly.

The reduced form of equation (2.1) is written in partitioned form as

$$[y_1, Y_2] = [Z_1, Z_2] \begin{bmatrix} \pi_1 & \Pi_1 \\ \pi_2 & \Pi_2 \end{bmatrix} + [v_1, V_2] \tag{2.2}$$

or

$$Y = Z\Pi + V,$$

where Z_2 $(T \times k_2)$ is a matrix of exogenous variables excluded from equation (2.1).

We assume

Assumption 1 *$k_2 \geq n$.*

Assumption 2 *$Z_{1t} = [1, t, ..., t^p, Z'_{*t}]'$ and p is known.*

Assumption 3 *$w_t = [\Delta Z'_{*t}, \Delta Z'_{*2t}, v_{1t}, V'_{2t}]'$ $(\equiv [w'_{1t}, w'_{2t}, w_{3t}, w'_{4t}]')$ is a linear process such that*

$$T^{-1/2} \textstyle\sum_{t=1}^{[Tr]} w_t \Rightarrow B(r)$$

and

$$T^{-1} \textstyle\sum_{t=1}^{T} w_t w'_t \xrightarrow{p} \sum, \; T^{-1} \sum_{t=1}^{T} w_t w'_{t=1} \xrightarrow{p} \sum_{-},$$

where Δ is the usual backward difference operator and $B(r)$ is a vector Brownian motion with a positive definite covariance matrix Ω. The following partitions of $B(r)$, $\sum$ and Ω will be used repeatedly throughout the paper.

$$B(r) = [\overset{k_1-1-p}{B_1(r)'} \;\; \overset{k_2}{B_2(r)'} \;\; \overset{1}{B_3(r)} \;\; \overset{n}{B_4(r)'}]'$$

$$\sum = \left[\textstyle\sum_{ij}\right]_{i,j=1,\ldots,4}, \; \textstyle\sum_{ij} = E(w_{it}w'_{jt})$$

$$\Omega = [\Omega_{ij}]_{i,j=1,\ldots,4}, \; \Omega_{ij} = E[B_i(r)B_j(r)'] = \textstyle\sum_{ij} + \Lambda_{ij} + \Lambda'_{ij},$$

where $\Lambda_{ij} = \sum_{k=2}^{\infty} E(w_{it}w_{j(t+k)})$. Also, we let $\Gamma_{ij} = \sum_{ij} + \Lambda_{ij}$.

Several aspects of these assumptions merit comment. First, Assumption 1 implies that the order condition for the identification of equation (2.1) is satisfied.

Second, Assumption 2 allows y_{1t} and Y_{2t} to have nonstochastic trend components, as is usually assumed in the cointegration literature. In practice, it will suffice to assume either $p = 0$ or $p = 1$.

Third, Assumption 3 implies that y_{1t}, Y_{2t}, Z_{*t}, $Z_{2t} = I(1)$ and that the vector process $[y_{1t}, Y'_{2t}, Z'_{*t}]'$ is cointegrated in the sense of Engle and Granger (1987) with the cointegrating vector $[1, \beta', \gamma']'$. In practice, the observed variables Z_{*t} and Z_{2t} may also contain time polynomials. In such a case, we may rewrite equation (2.1) as

$$y_1 = Y_2\beta + \Lambda\gamma_\Lambda + Z_*\gamma_Z + u$$

where the t-th row of the matrix Λ is $[1, t, ...t^p]$. Then, when the order p is greater than or equal to the highest order of the time polynomials in Z_{*t} and Z_{2t}, the OLS and 2SLS estimates of β and γ_Z are invariant to the presence of time polynomials in Z_{*t} and Z_{2t}. Therefore, assuming that Z_{*t} and Z_{2t} are pure stochastic processes is not restrictive at all as long as our interest lies in estimating the coefficients of Y_2 and Z_*.

Fourth, the limit conditions in Assumption 3 are required for the development of asymptotic results in later sections. A variety of primitive conditions for Assumption 3 to hold are possible, are well known, and are discussed in Phillips and Solo (1992).

Fifth, Assumption 3 assumes that $\{V_t\}$ is a linear process of unknown form. This feature contrasts with the traditional SEM [cf., Hausman (1983), Hsiao (1983), and Phillips (1983)] where the disturbance terms for the reduced form model typically have no dynamic structure or follow some simple autoregressive form.

Last, assuming that the long-run variance matrix Ω is positive definite, as in Assumption 3, implies that the vector $[Z'_{*t}, Z'_{2t}]'$ is not cointegrated. As in Park (1990) and Saikkonen (1993), we call Z_{*t} and Z_{2t} exogenous variables in such a case. Therefore, the exogenous variables here can be interpreted as common trends that introduce nonstationarity into the SEM. In practice, this assumption can be checked by using various classical tests, like those in Johansen (1988), or model determination tests, like those in Phillips (1996). This assumption also implies that the variables Z_{*t} and Z_{2t} do not have any structural equilibrium relations in the long-run. Also, note that we allow Z_{*t} and Z_{2t} to be serially and contemporaneously correlated with the reduced form regression errors, unlike the standard SEM.

The identifying relations connecting the parameters of equations (2.1) and (2.2) are

$$\pi_1 - \Pi_1\beta = \gamma \tag{2.3}$$

$$\pi_2 - \Pi_2\beta = 0. \tag{2.4}$$

Because equation (2.1) signifies the true structural relation, these conditions are assumed to hold true. Tests for relation (2.4) have been called overidentification tests [see Anderson and Kunitomo (1992) for a unifying framework of many existing overidentification tests]. When Π_2 has full column rank [i.e., $r(\Pi_2) = n$], the coefficient vector β is fully identified because once the coefficient matrices π_2 and Π_2 (or their consistent estimates) are known the coefficient vector β is estimable as we can see from equation (2.4). When $\Pi_2 = 0$, equation (2.4) indicates that the coefficient vector β is totally unidentified. However, if $\Pi_1 = 0$ in addition (the case of $\{y_{2t}\}$ being stationary), $\gamma = \pi_1$ from equation (2.3) so that the coefficient vector γ is fully identified in this case.

In this paper, we are interested in the general case where Π_1 and Π_2 are of arbitrary ranks. Specifically, assume

Assumption 4 $r(\Pi_1) = k_{12} \leq k_1,\ r(\Pi_2) = n_1 \leq n.$

Because some coefficients are identified and some are not under this assumption, Assumption 4 is known as the partial identification condition [cf., Phillips (1989)].

Following Phillips (1989) and Choi and Phillips (1992), define the rotation matrix

$$S = [\overset{n_1}{S_1}, \ \overset{n_2}{S_2}] \in O(n),$$

where $O(n)$ denotes the orthogonal group of $n \times n$ matrices and S_2 spans the null space of Π_2 and $\Pi_{21} = \Pi_2 S_1$ has full column rank n_1. Note that the matrix S that satisfies these conditions is not unique. Furthermore, let

$$\beta_1 = S_1' \beta, \ \beta_2 = S_2' \beta$$

and

$$\Pi_{11} = \Pi_1 S_1, \ \Pi_{12} = \Pi_1 S_2, \ \Pi_{21} = \Pi_2 S_1, \ \Pi_{22} = \Pi_2 S_2 = 0.$$

The purpose of this coordinate rotation is to isolate the estimable part of β. Under this rotation, the identifying relations (2.3) and (2.4) become

$$\pi_1 - \Pi_{11}\beta_1 - \Pi_{12}\beta_2 = \gamma, \tag{2.5}$$

$$\pi_2 - \Pi_{21}\beta_1 = 0. \tag{2.6}$$

Equation (2.6) shows that the coefficient vector β_1 is identified because Π_{21} has full column rank. But, the coefficient vectors β_2 and γ are not identified, as is readily deduced from equation (2.5).

The estimable part of the coefficient vector γ can be isolated by rotating the coordinates in equation (2.5) in the same way as for equation (2.4). That is, define an orthogonal matrix

$$R = [\overset{k_{11}}{R_1}, \ \overset{k_{12}}{R_2}] \in O(k_1),$$

where $R_1'\Pi_1 = 0$, $R_2'\Pi_1$ has full column rank and $k_{11} + k_{12} = k_1$. Note that the matrix R is not unique. Multiplying R' to equation (2.5) gives

$$R'\pi_1 = \gamma_1 \tag{2.7}$$

$$R_2'\pi_1 - R_2'\Pi_{11}\beta_1 - R_2'\Pi_{12}\beta_2 = \gamma_2, \tag{2.8}$$

where $\gamma_1 = R_1'\gamma$ and $\gamma_2 = R_2'\gamma$. Here γ_1 is identified, but γ_2 is not. Equations (2.6), (2.7), and (2.8) constitute new identifying relations under the rotations undertaken so far.

These rotations provide a new structural equation

$$\begin{aligned} y_1 &= Y_2\beta + Z_1\gamma + u \\ &= Y_2 S S'\beta + Z_1 R R'\gamma + u \\ &= Y_{21}\beta_1 + Y_{22}\beta_2 + Z_{11}\gamma_1 + Z_{12}\gamma_2 + u, \end{aligned} \tag{2.9}$$

where $Y_{21} = Y_2 S_1$, $Y_{22} = Y_2 S_2$, $Z_{11} = Z_1 R_1$ and $Z_{12} = Z_1 R_2$. In equation (2.9), the coefficients (β_1, γ_1) are identified, but (β_2, γ_2) are not. Additionally, note that $Z_{11t}, Z_{12t} = I(1)$, due to Assumption 3, and, further, $Y_{21t}, Y_{22t} = I(1)$, as we see from equations (2.10) and (2.11) below. The original coefficients are recovered from the equations

$$\beta = S_1\beta_1 + S_2\beta_2, \ \gamma = R_1\gamma_1 + R_2\gamma_2.$$

We can study the effects of partial identification on the asymptotic properties of various estimators and tests by using these equations relating the different coordinate systems.

The reduced form model (2.2) can similarly be written in the new coordinate system as

$$y_1 = Z_1\pi_1 + Z_2\pi_2 + v_1,$$

and

$$\begin{aligned} Y_{21} &= Z_1\Pi_{11} + Z_2\Pi_{21} + V_{21}, \quad (2.10)\\ Y_{22} &= Z_1\Pi_{12} + V_{22}, \quad (2.11) \end{aligned}$$

where $V_{21} = V_2S_1$ and $V_{22} = V_2S_2$.

2.2 OLS Estimation

This subsection studies the asymptotic properties of OLS estimates for the structural regression equation (2.9). OLS estimates are known to be inconsistent in standard SEMs, but are consistent in regressions involving I(1) variables even when the regressors and error terms are contemporaneously and serially correlated. Here we study the effects of partial identification on these estimates.

The OLS estimates of β_1, β_2, γ_1 and γ_2 in equation (2.9) are written as

$$\begin{aligned} \hat\beta_1 &= (Y_{21}'EY_{21})^{-1}Y_{21}'Ey_1,\\ \hat\beta_2 &= (Y_{22}'FY_{22})^{-1}Y_{22}'Fy_1,\\ \hat\gamma_1 &= R_1'\hat\gamma,\ \hat\gamma_2 = R_2'\gamma,\\ \hat\gamma &= (Z_1'Z_1)^{-1}Z_1'y_1 - (Z_1'Z_1)^{-1}Z_1'[Y_{21}, Y_{22}]\begin{bmatrix}\hat\beta_1\\ \hat\beta_2\end{bmatrix}, \end{aligned}$$

where

$$E = Q_{z_1} - Q_{z_1}Y_{22}(Y_{22}'Q_{z_1}Y_{22})^{-1}Y_{22}'Q_{z_1}$$

and

$$F = Q_{z_1} - Q_{z_1}Y_{21}(Y_{21}'Q_{z_1}Y_{21})^{-1}Y_{21}'Q_{z_1}.$$

Using the above formulae and the weak convergence results in Phillips (1988), the following limit theory for the OLS estimates is obtained.

Theorem 1 *Define*

$$\begin{aligned} K[C, D, \lambda_1, \lambda_2, \lambda_3] &= \left[\textstyle\int_0^1 C(r)D(r)'dr + \lambda_1\right] - \left[\textstyle\int_0^1 C(r)R(r)'dr + \lambda_2\right]\\ &\quad\times \left[\textstyle\int_0^1 R(r)R(r)'dr\right]^{-1}\left[\textstyle\int_0^1 R(r)D(r)'dr + \lambda_3\right], \end{aligned}$$

$$\begin{aligned} K[C, dD, \lambda_1, \lambda_2, \lambda_3] &= \left[\textstyle\int_0^1 CdD + \lambda_1\right] - \left[\textstyle\int_0^1 C(r)R(r)'dr + \lambda_2\right]\\ &\quad\times \left[\textstyle\int_0^1 R(r)R(r)'dr\right]^{-1}\left[\textstyle\int_0^1 RdD + \lambda_3\right], \end{aligned}$$

$$
\begin{aligned}
K[dC, D, \lambda_1, \lambda_2, \lambda_3] \; = \; & \left[\int_0^1 dCD' + \lambda_1\right] - \left[\int_0^1 dCR' + \lambda_2\right] \\
& \times \; \left[\int_0^1 R(r)R(r)'dr\right]^{-1} \left[\int_0^1 R(r)D(r)'dr + \lambda_3\right],
\end{aligned}
$$

where $R(r) = [1, r, ..., r^p, B_1(r)']'$*. Then, under Assumptions 1, 2, and 3,*
(i)

$$
\begin{aligned}
T(\hat{\beta}_1 - \beta_1) \; \Rightarrow \; & \{\Pi_{21}' K[B_2, B_2, 0, 0, 0]\Pi_{21}\}^{-1} \Big[-\Pi_{21}' \Big\{K\Big[B_2, dB_4, \Gamma_{24}, 0, \tbinom{0}{\Gamma_{14}}\Big] \\
& - \; K\Big[B_2, dB_4, \Gamma_{24}, 0, \tbinom{0}{\Gamma_{14}}\Big] (S_2' \textstyle\sum_{44} S_2)^{-1} S_2' \textstyle\sum_{44}\Big\} S_1\beta_1 \\
& - \; S_1' \Big\{\textstyle\sum_{44} - \textstyle\sum_{44} S_2 (S_2' \textstyle\sum_{44} S_1)^{-1} S_2' \textstyle\sum_{44}\Big\} S_1\beta_1 \\
& + \; \Pi_{21}' \Big\{K\Big[B_2, dB_3, \Gamma_{23}, 0, \tbinom{0}{\Gamma_{13}}\Big] \\
& - \; K\Big[B_2, dB_4, \Gamma_{24}, 0, \tbinom{0}{\Gamma_{14}}\Big] (S_2' \textstyle\sum_{44} S_2)^{-1} S_2' \textstyle\sum_{43}\Big\} \\
& + \; S_1' \Big\{\textstyle\sum_{43} - \textstyle\sum_{44} S_2 (S_2' \textstyle\sum_{44} S_2)^{-1} S_2' \textstyle\sum_{43}\Big\}\Big] \\
= \; & \hat{b}_1, \; \textit{say.}
\end{aligned}
$$

(ii)

$$
\begin{aligned}
\hat{\beta}_2 \; \Rightarrow \; & (S_2' \textstyle\sum_{44} S_2)^{-1} [S_2' K[dB_4, B_2, \Gamma_{42}, [0, \Gamma_{41}], 0] \\
& - \{K[dB_4, B_2, \Gamma_{42}, [0, \Gamma_{41}], 0]\Pi_{21} + \textstyle\sum_{44} S_1\} \\
& \times \{\Pi_{21}' K[B_2, B_2, 0, 0, 0]\Pi_{21}\}^{-1} \Pi_{21}' K[B_2, B_2, 0, 0, 0]\}\Pi_{21}\beta_1 + S_2' \textstyle\sum_{43}] \\
= \; & \hat{b}_2, \; \textit{say.}
\end{aligned}
$$

(iii) *When time polynomials are present in* Z_{1t}*,*

$$
T^{1/2}(\hat{\gamma}_1 - \gamma_1) \Rightarrow -R_{11}'\delta,
$$

where R_{11}' *is the first column of* R_1' *and* δ *is the first element of the random vector*

$$
\begin{aligned}
& \left[\int_0^1 R(r)R(r)'dr\right]^{-1} \Big\{\Big[\int_0^1 R(r)dB_4(r)' + \tbinom{0}{\Gamma_{14}}\Big] S_1\beta_1 - \Big[\int_0^1 R(r)dB_3(r) + \tbinom{0}{\Gamma_{13}}\Big] \\
& \quad + \int_0^1 R(r)B_2(r)'dr\Pi_{21}\hat{b}_1 + \Big[\int_0^1 R(r)dB_4(r)' + \tbinom{0}{\Gamma_{14}}\Big] S_2\hat{b}_2\Big\}.
\end{aligned}
$$

(iv) *When time polynomials are not present in* Z_{1t}*,*

$$
\begin{aligned}
T(\hat{\gamma}_1 - \gamma_1) \; \Rightarrow \; & -R_1' \left[\int_0^1 B_1(r)B_1(r)'dr\right]^{-1} \Big\{\Big[\int_0^1 B_1(r)dB_4(r)' + \Gamma_{14}\Big] S_1\beta_1 \\
& - \; \Big[\int_0^1 B_1(r)dB_3(r) + \Gamma_{13}\Big] + \int_0^1 B_1(r)B_2(r)'dr\Pi_{21}\hat{b}_1 \\
& + \; \Big[\int_0^1 B_1(r)dB_4(r)' + \Gamma_{14}\Big] S_2\hat{b}_2\Big\}.
\end{aligned}
$$

(v) $\hat{\gamma}_2 \Rightarrow R_2'\pi_1 - R_2'\Pi_{11}\beta_1 - R_2'\Pi_{12}\hat{b}_2$.

Several remarks are in order for the results reported in Theorem 1. First, Theorem 1 shows that the OLS estimates for the identified coefficient vectors β_1 and γ_1 are consistent, but that those for the unidentified coefficient vectors β_2 and γ_2 are inconsistent and tend to random variables in the limit, unlike the stationary case where they converge to constants. Because $\hat{\beta} = S_1\hat{\beta}_1 + S_2\hat{\beta}_2$ and $\hat{\gamma} = R_1\hat{\gamma}_1 + R_2\hat{\gamma}_2$, these imply that the OLS estimates for the coefficient vectors β and γ are inconsistent when identification failure occurs. Thus, even in regressions involving I(1) variables, OLS estimates are inconsistent under identification failure.

Second, the OLS estimate for the identified coefficient vector β_1 is T-consistent, as is usual in cointegrating regressions. But when time polynomials like those specified in Assumption 1 are present in the regressor Z_{1t}, the OLS estimate $\hat{\gamma}_1$ is $\sqrt{T}$-consistent as show in part (iii) of Theorem 1. The reason for this reduction in the convergence rate is that the OLS estimate for the intercept term in the regression equation (2.1) converges most slowly (at the rate of $\sqrt{T}$). If the intercept term is omitted (that is, if $Z_{1t} = [t, ..., t^p, Z'_{*t}]'$) in the regression equation (2.1), though this is quite uncommon, the usual rate T-asymptotics will apply for the estimate $\hat{\gamma}_1$. When time polynomials are not present in the regressor Z_{1t}, part (iv) shows that the OLS estimate $\hat{\gamma}_1$ is T-consistent.

Third, the OLS estimates $\hat{\beta}_2$ and $\hat{\gamma}_2$ carry no information regarding the true coefficient vectors in the limit, which is of course a manifestation of identification failure.

Now we consider some special cases of Theorem 1. First, the totally unidentified case $\Pi_2 = 0$. In the following corollaries, $\hat{\beta}$ and $\hat{\gamma}$ denote the OLS estimates of the coefficient vectors β and γ, respectively.

Corollary 1 *Suppose that β is totally unidentified. Then, under Assumptions* 1, 2, *and* 3,
(i) $\hat{\beta} \xrightarrow{p} \sum_{44}^{-1} \sum_{43} = \tilde{b}$, *say.*
(ii) *When time polynomials are present in Z_{1t},*

$$T^{1/2}(\hat{\gamma}_1 - \gamma_1) \Rightarrow -R'_{11}\mu,$$

where μ is the first element of the random vector

$$\left[\int_0^1 R(r)R(r)'dr\right]^{-1} \left\{-\left[\int_0^1 R(r)dB_3(r) + \binom{0}{\Gamma_{13}}\right] + \left[\int_0^1 R(r)dB_4(r)' + \binom{0}{\Gamma_{14}}\right]\tilde{b}\right\}.$$

(iii) *When time polynomials are not present in Z_{1t},*

$$T(\hat{\gamma}_1 - \gamma_1) \Rightarrow -R'_1\left[\int_0^1 B_1(r)B_1(r)'dr\right]^{-1}\left\{-[\int_0^1 B_1(r)dB_3(r) + \Gamma_{13}] + \left[\int_0^1 B_1(r)dB_4(r)' + \Gamma_{14}\right]\tilde{b}\right\}.$$

(iv) $\hat{\gamma}_2 \xrightarrow{p} R'_2\pi_1 - R'_2\Pi_1\tilde{b}$.

This corollary shows that the OLS estimate $\hat{\beta}$ is inconsistent for β in the totally unidentified case. The OLS estimate $\hat{\gamma}_2$ is also inconsistent, due to the effect of $\hat{\beta}$. Both estimates converge to constants, as in the stationary case. Nonetheless, parts

(ii) and (iii) show that the estimate $\hat{\gamma}_1$ is consistent. Furthermore, as in Theorem 1, the rate of convergence of $\hat{\gamma}_1$ depends on the presence of time polynomials.

Next, consider the case where Π_2 has full rank, i.e., the case where the coefficient vector β is fully identified.

Corollary 2 *Suppose β is fully identified. Then, under Assumptions* 1, 2, *and* 3,
(i)

$$\begin{aligned} T(\hat{\beta}-\beta) \;&\Rightarrow\; \{\Pi_2' K[B_2, B_2, 0,0,0]\Pi_2\}^{-1}\left[-\Pi_2'\left\{K\left[B_2, dB_4, \Gamma_{24}, 0, \binom{0}{\Gamma_{14}}\right]\right\}\beta - \textstyle\sum_{44}\beta \right. \\ &\qquad \left. + \Pi_2'\left\{K\left[B_2, dB_3, \Gamma_{23}, 0, \binom{0}{\Gamma_{13}}\right]\right\} + \textstyle\sum_{43}\right] \\ &= \check{b},\ \textit{say}. \end{aligned}$$

(ii) *When time polynomials are present in Z_{1t},*

$$\begin{aligned} C_T(\hat{\gamma}-\gamma) \;&\Rightarrow\; -\mathrm{diag}[0,1,...,1]\Pi_1\check{b} + \mathrm{diag}[1,\overbrace{0,...0}^{p},\overbrace{1,...,1}^{k_1-p-1}]\left[\textstyle\int_0^1 R(r)R(r)'dr\right]^{-1} \\ &\quad \times\left\{-\textstyle\int_0^1 R_1(r)B_2(r)'dr\Pi_2\check{b} - \left[\int_0^1 R(r)dB_4(r)' + \binom{0}{\Gamma_{14}}\right]\beta \right. \\ &\quad \left. + \left[\textstyle\int_0^1 R(r)dB_3(r) + \binom{0}{\Gamma_{13}}\right]\right\} \end{aligned}$$

where $C_T = \mathrm{diag}(T^{1/2}, T, ..., T)$.
(iii) *When time polynomials are not present in Z_{1t},*

$$\begin{aligned} T(\hat{\gamma}-\gamma) \;&=\; -\Pi_1\check{b} + \left[\textstyle\int_0^1 B_1(r)B_1(r)'dr\right]^{-1}\left\{-\textstyle\int_0^1 B_1(r)B_2(r)'dr\Pi_2\check{b}\right. \\ &\quad \left. - \left[\textstyle\int_0^1 B_1(r)dB_4(r)' + \Gamma_{14}\right]\beta + \left[\textstyle\int_0^1 B_1(r)dB_3(r) + \Gamma_{13}\right]\right\}. \end{aligned}$$

This corollary shows that the OLS estimates $\hat{\beta}$ and $\hat{\gamma}$ are consistent when the coefficient vector β is fully identified. The rate of convergence for the estimate $\hat{\gamma}$ depends on whether or not Z_{1t} contains time polynomials, as parts (ii) and (iii) show. Furthermore, the asymptotic distributions derived in Corollary 2 are nonstandard and depend on many nuisance parameters, which makes it difficult to use these results for statistical inference.

However, when $\sum_{43} = 0$, $\sum_{44} = 0$, $\Gamma_{13} = 0$, $\Gamma_{14} = 0$, $\Gamma_{23} = 0$ and $\Gamma_{24} = 0$, i.e., when $V_2 = 0$ with probability 1 and Z_{*t} and Z_{2t} are totally exogenous, Corollary 2 implies that the limiting distribution of $\hat{\beta}$ is mixture normal. To show this, let $dB_3(r) - dB_4(r)'\beta = dB_5(r)$. Then,

$$T(\hat{\beta}-\beta) \Rightarrow \{\Pi_2' K[B_2, B_2, 0,0,0]\Pi_2\}^{-1}\Pi_2' K[B_2, dB_5, 0,0,0]. \tag{2.12}$$

But

$$\Pi_2' K[B_2, dB_5, 0,0,0]\big|_{\mathcal{F}} = \textstyle\int_0^1 \Gamma(r)dB_5(r)\Big|_{\mathcal{F}} \equiv N\left(0, \int_0^1 \Gamma(r)\Gamma(r)'dr\right), \tag{2.13}$$

where $\Gamma(r) = \Pi_2' \left\{ B_2(r) - \int_0^1 B_2(r)R(r)'dr \left[\int_0^1 R(r)R(r)'dr\right]^{-1} R(r) \right\}$ and the symbol "$\cdot|_{\mathcal{F}}$" signifies the conditional distribution given the σ-algebra $\mathcal{F} = \{\Gamma(s) : 0 \leq s \leq 1\}$. Furthermore,

$$\int_0^1 \Gamma(r)\Gamma(r)'dr = \Pi_2' K[B_2, B_2, 0, 0, 0]\Pi_2. \tag{2.14}$$

Therefore, it follows from (2.12), (2.13) and (2.14) that

$$T(\hat{\beta} - \beta) \Rightarrow \int_{\zeta \in R^n} N(0, \zeta) pdf(\zeta) d\zeta, \tag{2.15}$$

where $\zeta \equiv \left[\int_0^1 \Gamma(r)\Gamma(r)'dr\right]^{-1}$.

Under the same assumption as for (2.15), the OLS estimate $\hat{\gamma}$ has the mixture normal distribution regardless of the presence of time polynomials in Z_{1t}. We illustrate this result for the case where there are no time polynomials in Z_{1t}. Part (iii) of Corollary 2 gives

$$T(\hat{\gamma} - \gamma) = -(\Pi_1 + \Lambda)\check{b} + \left[\int_0^1 B_1(r)B_1(r)'dr\right]^{-1} \int_0^1 B_1(r)dB_5(r),$$

where $\Lambda = \left[\int_0^1 B_1(r)B_1(r)'dr\right]^{-1} \int_0^1 B_1(r)B_2(r)'dr\Pi_2$. Since both the first and second terms in the above equation are mixture normal variables, $\hat{\gamma}$ also has a mixture normal distribution.

Mixture normality results in nonstationary regressions are routinely used to develop a basis for statistical inference on cointegrating vectors [cf., Park and Phillips (1989), Phillips and Hansen (1990), inter alia]. But it is worth noting that the mixture normality results for OLS estimation hold under the very special assumptions that $\sum_{43} = 0$ and $\sum_{44} = 0$ in addition to the assumption that Z_{*t} and Z_{2t} are totally exogenous. These extra requirements limit the usefulness of OLS in cointegrated SEMs.

2.3 2SLS Estimation

For stationary SEMs, it is common to use estimators like 2SLS in place of OLS. This subsection therefore considers 2SLS estimation for the regression equation (2.9). If estimator consistency is the sole concern, Corollary 2 shows that using OLS will suffice, as long as the coefficients are identified. As will now be shown, 2SLS estimates have some desirable properties that are not shared by OLS at least in the fully identified case, but also some undesirable characteristics in partially identified cases.

The 2SLS estimates for the coefficient vectors in the structural equation (2.1) are defined as:

$$\begin{aligned} \bar{\beta}_1 &= (Y_{21}'GY_{21})^{-1}Y_{21}'Gy_1, \\ \bar{\beta}_2 &= (Y_{22}'HY_{22})^{-1}Y_{22}'Hy_1, \\ \bar{\gamma}_1 &= R_1'\bar{\gamma},\ \bar{\gamma}_2 = R_2'\bar{\gamma}, \\ \bar{\gamma} &= (Z_1'Z_1)^{-1}Z_1'y_1 - (Z_1'Z_1)^{-1}Z_1'[Y_{21}, Y_{22}] \begin{bmatrix} \bar{\beta}_1 \\ \bar{\beta}_2 \end{bmatrix}, \end{aligned}$$

where

$$
\begin{aligned}
G &= L - LY_{22}(Y_{22}'LY_{22})^{-1}Y_{22}'L, \\
H &= L - LY_{21}(Y_{21}'LY_{21})^{-1}Y_{21}'L, \\
L &= P_z - P_{z_1}.
\end{aligned}
$$

The following theorem reports the asymptotic distributions of these estimators.

Theorem 2 *Define*

$$
\begin{aligned}
J[A,B] &= A - BS_2\left(S_2'B'A^{-1}BS_2\right)^{-1}S_2'B' \\
M[A,B,C] &= A - AB^{-1}CS_2\left(S_2'C'B^{-1}CS_2\right)^{-1}S_2'C', \\
W[A,B,C] &= AB^{-1}C - A\Pi_{21}[\Pi_{21}'B\Pi_{21}]^{-1}\Pi_{21}'C,
\end{aligned}
$$

and

$$
N[A,B] = A - A\Pi_{21}(\Pi_{21}'B\Pi_{21})^{-1}\Pi_{21}'B.
$$

Then, under Assumptions 1, 2, *and* 3,
(i)

$$
\begin{aligned}
T(\bar{\beta}_1 - \beta_1) \Rightarrow\ & \left[\Pi_{21}'J\left\{K[B_2,B_2,0,0,0], K\left[B_2, dB_4, \Gamma_{24}, 0, \left(\begin{smallmatrix}0\\ \Gamma_{14}\end{smallmatrix}\right)\right]\right\}\Pi_{21}\right]^{-1} \\
& \times\left[-\Pi_{21}'M\left\{K[dB_4,B_2,\Gamma_{42},[0,\Gamma_{41}],0], K[B_2,B_2,0,0,0],\right.\right. \\
& \left. K\left[B_2, dB_4, \Gamma_{24}, 0, \left(\begin{smallmatrix}0\\ \Gamma_{14}\end{smallmatrix}\right)\right]\right\}' S_1\beta_1 \\
& +\Pi_{21}'M\left\{K[dB_3,B_2,\Gamma_{32},[0,\Gamma_{31}],0], K[B_2,B_2,0,0,0],\right. \\
& \left.\left. K\left[B_2, dB_4, \Gamma_{24}, 0, \left(\begin{smallmatrix}0\\ \Gamma_{14}\end{smallmatrix}\right)\right]'\right\}\right] \\
=\ & \bar{b}_1, \textit{ say.}
\end{aligned}
$$

(ii)

$$
\begin{aligned}
T^{-1}\bar{\beta}_2 \Rightarrow\ & [S_2W\{K[dB_4,B_2,\Gamma_{42},[0,\Gamma_{41}],0], K[B_2,B_2,0,0,0], \\
& K[dB_4,B_2,\Gamma_{42},[0,\Gamma_{41}],0]'\}S_2]^{-1} \\
& \times S_2'N\left\{K[dB_4,B_2,\Gamma_{42},[0,\Gamma_{41}],0], K[B_2,B_2,0,0,0]\right\}\Pi_{21}\beta_1 \\
=\ & \bar{b}_2, \textit{ say.}
\end{aligned}
$$

(iii) *When time polynomials are present in* Z_{1t},

$$
T^{-1/2}(\bar{\gamma}_1 - \gamma_1) \Rightarrow -R_{11}'\eta,
$$

where R_{11}' *is the first column of the matrix* R_1' *and* η *is the first element of the vector*

$$
\left[\int_0^1 R(r)R(r)'dr\right]^{-1}\left[\int_0^1 R(r)dB_4'(r) + \left(\begin{smallmatrix}0\\ \Gamma_{14}\end{smallmatrix}\right)\right]S_2\bar{b}_2.
$$

(iv) *When time polynomials are not present in* Z_{1t}*,*

$$\bar{\gamma}_1 \Rightarrow \gamma_1 - R_1' \left[\int_0^1 R(r)R(r)'dr\right]^{-1} \left\{\int_0^1 R(r)dB_4'dr + \binom{0}{\Gamma_{14}}\right\} \bar{b}_2.$$

(v) $T^{-1}\bar{\gamma}_2 \Rightarrow -R_2'\Pi_{12}\bar{b}_2$.

Several aspects of this theorem merit our attention. First, as in the case of the OLS estimation, the 2SLS estimate of β_1 is T-consistent. But unlike the corresponding OLS estimate, the 2SLS estimate of γ_1 is inconsistent. It even diverges in probability when the time polynomials specified in Assumption 1 are present.

Second, the 2SLS estimates for the unidentified coefficient vectors all diverge in probability at the rate of T, unlike the corresponding OLS estimates. An intuitive explanation for this behavior is that the 2SLS estimates of the unidentified structural coefficients of Y_{22} (viz. β_2) rely on instruments, Z_2, that are totally irrelevant (cf. equation (2.11)), due to the lack of identifiability of β_2. These instruments project the data Y_{22} into linear maps of the equation errors. In particular, $(P_z - P_{z_1})Y_{22} = (P_z - P_{z_1})V_{22}$ and then $Y_{22}'(P_z - P_{z_1})Y_{22} = V_{22}'(P_z - P_{z_1})V_{22} = O_p(1)$, due to the fact Z_2 is I(1). In consequence, there is less leverage in the instrumented regressors than there is in the original I(1) regressors Y_{22} themselves. It is this reduction in regressor excitation that leads to the divergence of the 2SLS estimates of the unidentified structural coefficients β_2. Naturally, the 2SLS estimates of the coefficients of the exogenous regressors are contaminated by these poor characteristics of the structural estimates. The effect is not only the divergence (at rate T) of the 2SLS estimates of the unidentified coefficients, γ_2, of the exogenous variables, but also the inconsistency of the 2SLS estimates of the identified coefficients, γ_1, of the exogenous variables, noted in the previous paragraph.

Third, the above features of the coefficient estimates in the transformed model imply that the 2SLS estimates of β and γ diverge in probability when identification failure occurs. Therefore, under identification failure, the 2SLS estimates may exhibit more erratic small-sample behavior than the corresponding OLS estimates because the OLS estimates under identification failure converge weakly to well-defined random variables as shown in Theorem 1.

The following corollary to Theorem 2 deals with the totally unidentified case. In Corollaries 3 and 4 that follow, the 2SLS estimates for β and γ are denoted as $\bar{\beta}$ and $\bar{\gamma}$, respectively.

Corollary 3 *Suppose that* β *is totally unidentified, and let* $\bar{W}[A, B, C] = AB^{-1}C$*. Then, under Assumptions 1, 2, and 3,*
(i)

$$\begin{aligned}\bar{\beta} \Rightarrow\ & [\bar{W}\{K[dB_4, B_2, \Gamma_{42}, [0, \Gamma_{41}], 0], K[B_2, B_2, 0, 0, 0], \\ & K[dB_4, B_2, \Gamma_{42}, [0, \Gamma_{41}], 0]'\}]^{-1} [\bar{W}\{K[dB_4, B_2, \Gamma_{42}, [0, \Gamma_{41}], 0], \\ & K[B_2, B_2, 0, 0, 0], K[B_2, dB_3, \Gamma_{23}, 0, \binom{0}{\Gamma_{13}}\}] \\ =\ & \bar{b},\ \textit{say}.\end{aligned}$$

(ii) *When time polynomials are present in* Z_{1t},

$$T^{1/2}(\bar{\gamma}_1 - \gamma_1) \Rightarrow R_{11}'\omega$$

where ω *is the first element of the random vector*

$$\left[\int_0^1 R(r)R(r)'dr\right]^{-1}\left\{\left[\int_0^1 R(r)dB_3'(r) + \binom{0}{\Gamma_{13}}\right] - \left[\int_0^1 R(r)dB_4'(r) + \binom{0}{\Gamma_{14}}\right]\bar{b}\right\}.$$

(iii) *When time polynomials are not present in* Z_{1t},

$$\begin{aligned} T(\bar{\gamma}_1 - \gamma_1) \;\Rightarrow\; & R_1'\left[\int_0^1 B_1(r)B_1(r)'dr\right]^{-1} \\ & \times\left\{\left[\int_0^1 B_1(r)dB_3'(r) + \Gamma_{13}\right] - \left[\int_0^1 B_1(r)dB_4'(r) + \Gamma_{14}\right]\bar{b}\right\}. \end{aligned}$$

(iv) $\bar{\gamma}_2 \Rightarrow R_2'\pi_1 - R_2'\Pi_1\bar{b}$.

This corollary shows that the 2SLS estimate $\bar{\beta}$ is inconsistent and contains no information regarding the coefficient vector β in the totally unidentified case. But, unlike the partially identified case studied in Theorem 2, the 2SLS estimate $\bar{\gamma}_1$ is consistent. Furthermore, the rate of consistency of $\bar{\gamma}_1$ depends on the presence of the time polynomials in Z_{1t}. Additionally, this corollary shows that $\bar{\gamma}_2$ is inconsistent but does not diverge, in contrast to Theorem 2. Note that in this fully unidentified case $\Pi_{21} = 0$, and hence $\bar{b}_2 = 0$ in part (ii) and $\eta = 0$ in part (iii) of Theorem 2, so that the limits given in these parts of the theorem are degenerate. In this way, the results of Theorem 2 calibrate with the specialized results of the corollary, which show the appropriate limits, after renormalization, for the case where β is totally unidentified

The next corollary considers the case of full identifiability. The results are important for empirical practice, because statistical inference on the coefficient vectors under full identifiability will naturally depend on these results.

Corollary 4 *Suppose that* β *is fully identified. Then, under Assumptions* 1, 2, *and* 3,
(i)

$$\begin{aligned} T(\bar{\beta} - \beta) \;\Rightarrow\; & \{\Pi_2'K[B_2, B_2, 0, 0, 0]\Pi_2\}^{-1} \\ & \times\left[-\Pi_2'\left\{K\left[B_2, dB_4, \Gamma_{24}, 0, \binom{0}{\Gamma_{14}}\right]\right\}\beta + \Pi_2'\left\{K\left[B_2, dB_3, \Gamma_{23}, 0, \binom{0}{\Gamma_{13}}\right]\right\}\right] \\ =\; & \ddot{b},\ \textit{say} \end{aligned}$$

(ii) *When time polynomials are present in* Z_{1t},

$$\begin{aligned} C_T(\bar{\gamma} - \gamma) \;\Rightarrow\; & -\mathrm{diag}[0, 1, ..., 1]\Pi_1\ddot{b} + \mathrm{diag}[1, \overbrace{0, ..., 0}^{p}, \overbrace{1, ..., 1}^{k_1 - p - 1}]\left[\int_0^1 R(r)R(r)'dr\right]^{-1} \\ & \times\left\{-\int_0^1 R_1(r)B_2(r)'dr\Pi_2\ddot{b} - \left[\int_0^1 R(r)dB_4(r)' + \binom{0}{\Gamma_{14}}\right]\beta \right. \\ & \left. + \left[\int_0^1 R(r)dB_3(r) + \binom{0}{\Gamma_{13}}\right]\right\} \end{aligned}$$

where $C_T = \mathrm{diag}(T^{1/2}, T, ..., T)$ *and* $\ddot{b}_i$ *denotes the i-th element of the random vector* $\ddot{b}$.

(iii) *When time polynomials are not present in* Z_{1t},

$$T(\bar{\gamma} - \gamma) \Rightarrow -\Pi_1 \ddot{b} + \left[\int_0^1 B_1(r) B_1(r)' dr\right]^{-1} \left\{ -\int_0^1 B_1(r) B_2(r)' dr \Pi_2 \ddot{b} - \left[\int_0^1 B_1(r) dB_4(r)' + \Gamma_{14}\right] \beta + \left[\int_0^1 B_1(r) dB_3(r) + \Gamma_{13}\right] \right\}.$$

This corollary shows that the 2SLS estimates $\bar{\beta}$ and $\bar{\gamma}$ are consistent when the coefficient vector β is fully identified. Comparing these results with those in Corollary 2, we find that the asymptotic distributions of the 2SLS estimates are nonstandard but depend on a lesser number of nuisance parameters than the OLS estimates (i.e., $\sum_{44}$ and $\sum_{43}$ do not enter the formulae for the asymptotic distributions of $\bar{\beta}$ and $\bar{\gamma}$).

When $\Gamma_{13} = 0$, $\Gamma_{14} = 0$, $\Gamma_{23} = 0$ and $\Gamma_{24} = 0$, i.e., when Z_{*t} and Z_{2t} are totally exogenous, Corollary 4 shows that the 2SLS estimate $\bar{\beta}$ has the same mixture normal distribution as the OLS estimate $\hat{\beta}$ which is given in (2.15). But note that the conditions $\sum_{43} = 0$ and $\sum_{44} = 0$ are not required for the 2SLS estimate $\bar{\beta}$ to have a mixture normal distribution. Furthermore, it is straightforward to show that $\bar{\gamma}$ has a mixture normal distribution in the limit irrespective of the presence of time polynomials in Z_{1t}. The mixture normality results obtained under the special assumptions imposed can profitably be used for statistical inference, as we will explore in Section 3.

3 Efficient Estimation for a Fully Identified, Cointegrated Structural Equation

Corollaries 2 and 4 in Section 2 showed that the OLS and 2SLS estimators have non-normal distributions in the limit and depend on nuisance parameters in a complicated manner even when the parameter vectors are fully identified. Under such circumstances, it is well known that Wald tests do not have chi-square distributions in the limit. This feature makes it quite difficult to perform tests on coefficient vectors by using OLS and 2SLS.

For the reduced form equations (2.2), various methods have been suggested in the literature for efficient estimation and standard chi-square asymptotics for Wald tests [e.g., Phillips and Hansen (1990), Phillips and Loretan (1991), Park (1992), Saikkonen (1991), Stock and Watson (1993)]. This section examines some of these methods when they are applied to a fully identified structural equation in the cointegrated SEM and adapts the methods so that efficient estimation of structural coefficient vectors and standard asymptotics for Wald tests become feasible. Among the various methods, we will examine Phillips and Hansen's fully modified (FM) regression and the leads-and-lags (LL) regression introduced in Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993). These methods are known to provide efficient estimation for reduced form cointegrating coefficients. In studying these methods, we will assume for simplicity that there are no time polynomials in the exogenous variable Z_{1t}.

Park (1990) and Saikkonen (1993) also considered estimating cointegrated SEMs. Park suggests transforming the variables in the model by running Park's (1992) canonical cointegrating regressions on the reduced form equations and then using these transformed variables to estimate structural form parameters. Saikkonen proposes indirect least squares methods which employ the efficient estimates of the reduced form parameters. These methods are similar in that estimating the reduced form equations are required, though their motivations are different. Unlike these methods, the methods we propose here do not require estimating reduced form equations. Estimation and hypothesis testing for the coefficient vectors will be based solely on structural form equations. This can result in improved finite sample precision in estimation, because the sampling error from the reduced form equation estimation is not imported into structural form estimation. However, full system estimation, which is covered in the aforementioned articles, will not be considered in this paper.

3.1 Phillips and Hansen's Fully Modified (FM) Regression

This subsection derives the asymptotic distribution of the Phillips–Hansen FM–OLS estimator applied to the structural equation (2.1) that is fully identified, and proposes the fully modified 2SLS (FM–2SLS) estimator. To set up the procedure, let $\alpha = [\beta', \gamma']'$ and $X = [Y_2, Z_1]$. Additionally, denote the long-run variance-covariance matrix of $p_t = (u_t', \Delta X_t')'$ by $\Psi = \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix}$. As in Section 2.1, we decompose Ψ as $\Psi = \Theta + \Upsilon + \Upsilon'$ where Θ is the probability limit of $T^{-1}\sum_{t=1}^{T} p_t p_t'$ and $\Upsilon = \sum_{k=1}^{\infty} E(p_0 p_{0+k}')$, and let $\Xi = \Theta + \Upsilon = \begin{bmatrix} \Xi_1 \\ \Xi_2 \end{bmatrix} \begin{matrix} 1 \\ n+k_1 \end{matrix}$. Furthermore, let $\hat{Q}$ be a consistent estimator of Q which uses the OLS residuals $\hat{u}_t = y_{1t} - \hat{\beta}' Y_{2t} - \hat{\gamma}' Z_{1t}$ and ΔX_t. Then, the FM–OLS estimator is defined as

$$\hat{\alpha}_{\text{FM}-\text{OLS}} = (X'X)^{-1}(X'y_1^+ - T\hat{\Xi}_2\hat{\kappa}) \tag{3.1}$$

where $y_1^+ = y_1 - \Delta X \hat{\Psi}_{22}^{-1}\hat{\Psi}_{12}'$, $\hat{\kappa} = [1, -\hat{\Psi}_{12}\hat{\Psi}_{22}^{-1}]'$ and $\Delta X = [X_1 - X_0, ..., X_T - X_{T-1}]'$.

The asymptotic distribution of $\hat{\alpha}_{FM-OLS}$ is reported in the following theorem.

Theorem 3 *Suppose that equation* (2.1) *is fully identified. Then, under Assumptions* 1, 2, *and* 3,

$$T(\hat{\alpha}_{FM-OLS} - \alpha) \Rightarrow \left[\int_0^1 D_2(r)D_2(r)'dr\right]^{-1} \left\{ \int_0^1 D_2(r)dD_{1.2}(r)dr + \begin{bmatrix} \sum_{43} - \sum_{44}\beta \\ 0 \end{bmatrix} \right.$$
$$\left. + [-J_2 \textstyle\sum_{-} K' + K \sum J_2' + K(\sum - \sum_{-})K']\Phi_{22}^{-1}\Phi_{12}', \right\}$$

where $D(r) = \begin{bmatrix} D_1(r) \\ D_2(r) \end{bmatrix} \begin{matrix} 1 \\ n+k_1 \end{matrix}$ *is a vector Brownian motion with the covariance matrix* $\Phi = J\Omega J'$*, the matrix* Φ *is partitioned conformal to the matrix* Ψ*,* $D_{1.2}(r) = D_1(r) - D_2(r)'\Phi_{22}^{-1}\Phi_{12}'$*,*

$$J = \begin{bmatrix} 0 & 0 & 1 & -\beta' \\ \Pi_1' & \Pi_2' & 0 & 0 \\ I & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} \begin{matrix} 1 \\ n+k_1 \end{matrix} \quad \textit{and } K = \begin{matrix} k_1 & k_2 & 1 & n & \\ \left[\begin{matrix} 0 \\ 0 \end{matrix}\right. & \begin{matrix} 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} & \left.\begin{matrix} I \\ 0 \end{matrix}\right] & \begin{matrix} n \\ k_1 \end{matrix} \end{matrix}.$$

This theorem shows that the weak limit of the FM–OLS estimator contains a second order bias which appears in the form of nuisance parameters. Therefore, inadvertently applying the FM–OLS to the structural form equation will cause non-standard asymptotics for Wald tests. Furthermore, this also brings an efficiency loss, as will be shown later. But once the nuisance parameters are eliminated, the weak limit is mixture normal because $D_{1.2}(r)$ is statistically independent of $D_2(r)$. Thus, we may further modify the FM–OLS estimator, so that the nuisance parameters are eliminated. But as will be shown below, it is simpler to use the 2SLS estimator to overcome this difficulty. Last, note that the covariance matrix Φ is nonsingular due to the assumption of full identifiability.

The fully modified 2SLS estimator for the structural form equation is defined by replacing X in (3.1) with $P_z X$. That is,

$$\hat{\alpha}_{\mathrm{FM-2SLS}} = (X'P_z X)^{-1}(X'P_z y_1^+ - T\hat{\Xi}_2\hat{\kappa}),$$

where y_1^+, $\hat{\Xi}_2$ and $\hat{\kappa}$ are defined in the same way as for (3.1).

The following theorem reports the asymptotic distribution of $\hat{\alpha}_{FM-2SLS}$.

Theorem 4 *Under the same assumptions as for Theorem* 3,

$$T(\hat{\alpha}_{\mathrm{FM-2SLS}} - \alpha) \Rightarrow \left[\int_0^1 D_2(r)D_2(r)'dr\right]^{-1}\left\{\int_0^1 D_2(r)dD_{1.2}(r)dr\right\},$$

where $D_2(r)$ and $D_{1.2}(r)$ are as defined in Theorem 3.

This theorem establishes that the weak limit of $\alpha_{\mathrm{FM-2SLS}}$ is mixture normal, because $D_2(r)$ and $D_{1.2}(r)$ are statistically independent. This also implies that the $\hat{\alpha}_{\mathrm{FM-2SLS}}$ is median unbiased in the limit.

Consider the null hypothesis

$$H_0 : D\alpha = d, \tag{3.2}$$

where D has full row rank. The Wald test for this hypothesis is defined as

$$W(\hat{\alpha}_{\mathrm{FM-2SLS}}) = (D\hat{\alpha}_{\mathrm{FM-2SLS}} - d)'\left[\hat{\Phi}_{1.2}(X'P_z X)^{-1}\right]^{-1}(D\hat{\alpha}_{\mathrm{FM-2SLS}} - d),$$

where $\hat{\Phi}_{1.2}$ is the consistent estimate of the long-run variance $\Phi_{1.2} = \Phi_{11} - \Phi_{12}\Phi_{22}^{-1}\Phi_{21}$ which uses the 2SLS regression residuals. Note that the long-run variance can be estimated by using conventional spectral density estimation methods [e.g., Andrews (1991)]. Using the mixture normality result of Theorem 4, we obtain in a straightforward way:

Corollary 5 *Under the same assumptions as for Theorem* 3,

$$W(\hat{\alpha}_{\mathrm{FM-2SLS}}) \Rightarrow \chi^2_{r(R)}.$$

3.2 Leads-and-Lags (LL) Regression

This subsection considers asymptotic properties of the LL regression for the structural equation (2.1). Letting $V_t = [V_{1t}', V_{2t}']'$ and $\Delta Z_t = [\Delta Z_{1t}', \Delta Z_{2t}']'$, we make the following assumption for the LL regression.

Assumption 5 *V_t can be represented as*

$$V_t = \sum_{j=-\infty}^{\infty} C_j \Delta Z_{t+j} + e_t,$$

where $\sum_{j=-\infty}^{\infty} ||C_j|| < \infty$ ($|| \cdot ||$ denotes the usual Euclidean norm), $E(\Delta Z_t e_{t+k}') = 0$ ($k = 0, \pm 1, \pm 2, ...$), e_t is a stationary process with the spectral density matrix $f_{ee}(\lambda) = f_{vv}(\lambda) - f_{v\Delta z}(\lambda) f_{\Delta z \Delta z}(\lambda)^{-1} f_{\Delta z v}(\lambda)$ and $E(e_t e_t') = \sum^e$, $E(e_{t-1} e_t') = \sum_-^e$.

More explicit conditions for these assumptions to hold can be found in Saikkonen (1991). Furthermore, we assume

Assumption 6 *$k^3/T \to 0$ and $\sqrt{T} \sum_{|j|>k} ||C_j|| \to 0$.*

Now, write

$$u_t = V_{1t} - \beta' V_{2t} = \sum_{j=-\infty}^{\infty} \rho_j' \Delta Z_{t+j} + e_{1t} - \beta' e_{2t}, \tag{3.3}$$

where e_{1t} and e_{2t} are conformal partitions of e_t. Inserting (3.3) into equation (2.1) yields the LL regression equation

$$y_{1t} = \alpha' X_t + \sum_{j=-k}^{k} \rho_j' \Delta Z_{t+j} + \dot{u}_t, \tag{3.4}$$

where $\dot{u}_t = e_{1t} - \beta' e_{2t} + \sum_{|j|>k} \rho_j' \Delta Z_{t+j}$. We denote the OLS estimates of α, $\rho_{-k}, ..., \rho_k$ from this regression equation as $\hat{\alpha}_{\text{LL}-\text{OLS}}, \hat{\rho}_{-k}, ..., \hat{\rho}_k$. Note that this regression equation is different from those in Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993) where the leads and lags are differences of regressors.

The asymptotic distribution of $\hat{\alpha}_{\text{LL}-\text{OLS}}$ is given in the following theorem.

Theorem 5 *Suppose that equation* (2.1) *is fully identified. Then, under Assumptions A1-A6*

$$T(\hat{\alpha}_{\text{LL}-\text{OLS}} - \alpha) \Rightarrow \left[\int_0^1 D_2(r) D_2(r)' dr\right]^{-1} \times \left\{\int_0^1 D_2(r) dD_c(r) dr + \begin{bmatrix} \sum_{21}^e - \sum_{-21}^e - (\sum_{22}^e + \sum_{-22}^e)\beta \\ 0 \end{bmatrix}\right\},$$

where $D_2(r)$ is as defined in Theorem 3, $D_c(r)$ is a vector Brownian motion with the covariance matrix $\Phi_{11} - \Phi_{12}\Phi_{22}^{-1}\Phi_{21}$, $\sum_{21}^e$ and $\sum_{22}^e$ ($\sum_{-21}^e$ and $\sum_{-22}^e$) are partitions of $\sum^e$ ($\sum_-^e$) conformal to e_{1t} and e_{2t}.

This theorem shows that the weak limit of $\hat{\alpha}_{\text{LL}-\text{OLS}}$ depends on nuisance parameters and that $\hat{\alpha}_{\text{LL}-\text{OLS}}$ is second-order biased. Conventional Wald tests on coefficient vectors therefore do not have chi-square distributions in the limit. But because $D_2(r)$ and $D_c(r)$ are statistically independent, the distribution is mixture normal once the nuisance parameters are eliminated. We may pursue further corrections of $\hat{\alpha}_{\text{LL}-\text{OLS}}$ such that its weak limit is free of nuisance parameters, but, again, using 2SLS estimation is more straightforward.

Let $M_t = [X_t', \Delta Z_{t-k}', ..., \Delta Z_{t+k}']'$ and $N_t = [Z_t', \Delta Z_{t-k}', ..., \Delta Z_{t+k}']'$. Then, the 2SLS estimate of the parameter vector $\phi = [\alpha', \rho_{-k}', ..., \rho_k']'$ from equation (3.4) is given by

$$\begin{aligned}\hat{\phi}_{\text{LL}-\text{2SLS}} &= \left[\left(\sum_{t=k+1}^{T-k} M_t N_t'\right)\left(\sum_{t=k+1}^{T-k} N_t N_t'\right)^{-1}\left(\sum_{t=k+1}^{T-k} N_t M_t\right)\right]^{-1} \\ &\quad \times \left(\sum_{t=k+1}^{T-k} M_t N_t'\right)\left(\sum_{t=k+1}^{T-k} N_t N_t'\right)^{-1}\left(\sum_{t=k+1}^{T-k} N_t y_{1t}\right).\end{aligned}$$

The asymptotic distribution of $\hat{\alpha}_{LL-2SLS}$ is:

Theorem 6 *Under the same assumptions as for Theorem* 5,

$$T(\hat{\alpha}_{\text{LL}-\text{2SLS}} - \alpha) \Rightarrow \left[\int_0^1 D_2(r) D_2(r)' dr\right]^{-1} \int_0^1 D_2(r) dD_c(r) dr.$$

where $D_2(r)$ and $D_c(r)$ are as defined in Theorem 5.

This theorem shows that the weak limit of $\hat{\alpha}_{\text{LL}-\text{2SLS}}$ is mixture normal. Note that $D_c(r) \equiv D_{1.2}(r)$ because these Gaussian processes have the same mean and covariance structure. Therefore, the weak limits of $\hat{\alpha}_{\text{FM}-\text{2SLS}}$ and $\hat{\alpha}_{\text{LL}-\text{2SLS}}$ are the same.

Furthermore, the Wald test for the null hypothesis (3.2) is defined by

$$\begin{aligned}W(\hat{\alpha}_{\text{LL}-\text{2SLS}}) &= (D\hat{\alpha}_{\text{LL}-\text{2SLS}} - d)' \\ &\quad \times \left\{\hat{\Phi}_{1.2}\left[\left(\sum_{t=k+1}^{T-k} M_t N_t'\right)\left(\sum_{t=k+1}^{T-k} N_t N_t'\right)^{-1}\left(\sum_{t=k+1}^{T-k} N_t M_t\right)\right]^{-1}_{n+k_1}\right\}^{-1} \\ &\quad \times (D\hat{\alpha}_{\text{LL}-\text{2SLS}} - d),\end{aligned}$$

where A_c denote a $c \times c$ block matrix in the north-west corner of matrix A and $\hat{\Phi}_{1.2}$ is the long-run variance estimate of $\Phi_{1.2}$ using the 2SLS regression residuals. Theorem 6 implies:

Corollary 6 *Under the same assumptions as for Theorem* 5,

$$W(\hat{\alpha}_{\text{LL}-\text{2SLS}}) \Rightarrow \chi^2_{r(R)}.$$

3.3 Comparison of FM-OLS and FM-2SLS

This subsection compares the efficiency of OLS and 2SLS for a fully identified structural equation. We will discuss only FM–OLS and FM–2SLS here, because it is easy to deduce that $\hat{\alpha}_{\mathrm{LL-2SLS}}$ is more efficient than $\hat{\alpha}_{\mathrm{LL-OLS}}$ once the efficiency comparison of FM–OLS and FM–2SLS is made.

Because the FM–OLS and FM–2SLS estimators have non-normal distributions in the limit, the efficiency comparison cannot be made in the conventional way of comparing their covariance matrices. However, we may evaluate the efficiency of estimators having non-normal distributions in the limit by the criterion of probability concentration used in Phillips (1991) and Saikkonen (1991). More specifically, when $\hat{a}$ and $\tilde{a}$ are T-consistent estimates of the $n \times 1$ parameter vector a, $\hat{a}$ is more efficient than $\tilde{a}$ if

$$\lim_T P\{T(\hat{a} - a) \in C\} \leq \lim_T P\{T(\tilde{a} - a) \in C\} \tag{3.5}$$

for any convex set $C \in R^n$, symmetric about the origin.

Using criterion (3.5), we can make the efficiency comparison of $\hat{\alpha}_{\mathrm{FM-OLS}}$ and $\hat{\alpha}_{\mathrm{FM-2SLS}}$ as follows. First of all, let $L = \left[\int_0^1 D_2(r)D_2(r)'dr\right]^{-1}\left\{\int_0^1 D_2(r)dD_{1.2}(r)dr\right\}$, which is the weak limit of $\hat{\alpha}_{\mathrm{LL-2SLS}}$. Then, denoting the weak limit of $\hat{\alpha}_{\mathrm{FM-OLS}}$ as N, we have $N = L + M$ where M is obvious from Theorem 3. Furthermore, define the σ-algebra $\mathcal{F} = \sigma\{D_2(s) : 0 \leq s \leq 1\}$. Then,

$$L|_{\mathcal{F}} \equiv N\left(0, \left[\int_0^1 D_2(r)D_2(r)'dr\right]^{-1}(\Phi_{11} - \Phi_{12}\Phi_{22}^{-1}\Phi_{21})\right)$$

and

$$E(LM'|\mathcal{F}) = 0 \text{ a.s.}$$

Because L and M are Gaussian processes, the latter equality implies that L and M are conditionally independent. Thus, using Lemma 1 in Basawa and Scott (1983) gives

$$P\{N \in C|\mathcal{F}\} \leq P\{L \in C|\mathcal{F}\}$$

for any convex set $C \in R^{n+k_1}$, symmetric about the origin. We deduce from this that

$$\begin{aligned}\lim_T P\{T(\hat{\alpha}_{FM-OLS} - \alpha) \in C\} &= P\{N \in C\} = E(P\{N \in C|\mathcal{F}\}) \\ &\leq E(P\{L \in C|\mathcal{F}\}) = P\{L \in C\} \\ &= \lim_T P\{T(\hat{\alpha}_{\mathrm{FM-2SLS}} - \alpha) \in C\},\end{aligned}$$

which implies that $\hat{\alpha}_{\mathrm{FM-2SLS}}$ is asymptotically more efficient than $\hat{\alpha}_{\mathrm{FM-OLS}}$ according to criterion (3.5).

3.4 Simulations

This subsection investigates the finite sample performance of the regression estimators we have so far studied by using simulation. Data were generated as:

$$y_{1t} = \beta Y_{2t} + \gamma Z_{1t} + u_t,\ \beta = 1,\ \gamma = 1,\ u_t = v_{1t} - V_{2t},$$

$$\begin{aligned} Y_{2t} &= \Pi_1 Z_{1t} + \Pi_2' Z_{2t} + V_{2t},\ \Pi_1 = 1,\ \Pi_2 = 1 \ or\ [1,1]', \\ w_t &= \delta A w_{t-1} + e_t, w_t = [\Delta Z_{1t}, \Delta Z_{2t}', v_{1t}, V_{2t}]',\ w_0 = 0, \end{aligned}$$

where A is a lower diagonal matrix whose non-zero elements are all 1's.The random vector e_t is iid normal, the diagonal elements of its covariance matrix are all 1's, and off-diagonal elements are all 0.3.

For simulation, we let $\delta = 0.1$, 0.3, and 0.5; and $T = 100$ and 200. Note that the structural coefficients β and γ are just-identified when $\Pi_2 = 1$, and over-identified when $\Pi_2 = [1,1]'$. Using these simulated data, squared biases and mean squared errors of OLS, 2SLS, FM–OLS, FM–2SLS, LL–OLS, LL–2SLS and Saikkonen's (1991) indirect least squares (OLS) estimators were calculated out of 5,000 iterations, the results of which are reported in Table 1. Part (1) of Table 1 contains the results for the just-identified case ($k_2 = 1$), and part (2) those for the over-identified case ($k_2 = 2$).

The long-run variances for the FM–OLS and FM–2SLS estimators were estimated by using Andrews' (1991) methods with a VAR(1) approximation for the prefilter. Furthermore, we used four leads and four lags for the leads-and-lags estimators. For the OLS estimator, we estimated the reduced form model by using multivariate FM–OLS and then calculated the OLS estimator as given in Section 3 of Saikkonen. The long-run variances for the multivariate FM–OLS were also estimated by using Andrews' (1991) methods with a VAR(1) approximation for the prefilter.

A few key findings from Table 1 can be summarized as follows.

(i) The LL–2SLS estimator has lower bias than the LL–OLS estimator in all cases, which is what we would predict from the asymptotic results in Section 3.2. But FM–2SLS is shown to be more biased in finite samples than FM–OLS when there is high serial dependence in the series $\{w_t\}$. Comparing OLS and 2SLS, we find that 2SLS is less biased than OLS as the asymptotic results in Section 2 presage. But, in most cases, 2SLS is more biased than at least one of FM–2SLS, LL–2SLS and OLS.

(ii) LL–2SLS has lower mean squared errors than LL–OLS in all cases, which confirms the efficiency comparison in Section 3.3. But FM–OLS has lower mean squared error than FM–2SLS when $\delta = 0.3$ and $\delta = 0.5$.

(iii) OLS shows higher mean squared errors than 2SLS except when $T = 100$ and $\delta = 0.5$.

(iv) Comparing the efficiency of LL–2SLS, FM–2SLS and ILS in the sense of mean squared errors, we find that LL–2SLS is most efficient except when $\delta = 0.1$ and the coefficients are over-identified. Furthermore, FM–2SLS is more efficient than ILS when $k_2 = 1$, $\delta = 0.1$ and $\delta = 0.5$. But, for the remaining cases, ILS is seen to be more efficient. But ILS has very high mean squared errors compared to LL–2SLS and FM–2SLS when the series $\{w_t\}$ is highly serially correlated. Surprisingly, 2SLS is more efficient and sometimes less biased than these modified estimators when there is low serial dependence in the data. But when the data are highly serially correlated, LL–2SLS outperforms 2SLS in terms of both biases and mean squared errors.

4 Summary and Further Remarks

We have derived the asymptotic distributions of the OLS and 2SLS estimators for a partially identified SEM involving I(1) variables. The two polar cases of total unidentifiability and full identification are also given. The results show that the rank condition for identification is as important in the nonstationary SEM as it is in the stationary SEM. But the OLS estimator is also consistent in the nonstationary SEM as long as the rank condition is satisfied.

We have also studied fully modified regression and the leads-and-lags regression applied to a fully identified single equation in the nonstationary SEM. The results show that these two estimators do not have mixture normal distributions. As alternatives, we proposed 2SLS versions of the two regression estimators, which have mixture normal distributions in the limit. Wald tests using the 2SLS versions of these estimators have chi-square distributions in the limit. Some simulations generally confirm the relevance of these asymptotic results in finite samples.

This paper assumes the existence of structural cointegrating relations in the SEM. In practice, these relations need to be tested. Devising cointegration tests for the SEM is, therefore, in order. Furthermore, testing the rank condition of identification and overidentifying restrictions in the cointegrated SEM has received little attention and a systematic study of these issues awaits further work.

Appendix: Proofs

Lemma A *Suppose Assumptions* 1, 2, *and* 3 *hold. Then, for* $i,j=1$ *or* 2,

$$\text{(i)}\quad T^{-2}Z_2'EZ_2 \Rightarrow K[B_2,B_2,0,0,0],$$

$$\text{(ii)}\quad T^{-1}V_i'EZ_2 \Rightarrow K[dB_{i+2},B_2,\Gamma_{(i+2)2},[0,\Gamma_{(i+2)1}],0] - \textstyle\sum_{(i+2)4} S_2(S_2'\sum_{44} S_2)^{-1}K[dB_4,B_2,\Gamma_{42},[0,\Gamma_{41}],0],$$

$$\text{(iii)}\quad T^{-1}V_i'EV_j \Rightarrow \textstyle\sum_{(i+2)(j+2)} - \sum_{(i+2)4} S_2(S_2'\sum_{44} S_2)^{-1}S_2'\sum_{4(j+2)},$$

$$\text{(iv)}\quad T^{-1}V_i'FV_j \xrightarrow{p} \textstyle\sum_{(i+2)(j+2)},$$

$$\text{(v)}\quad T^{-1}V_i'FZ_2 \Rightarrow K[dB_{i+2},B_2,\Gamma_{(i+2)2},[0,\Gamma_{(i+2)1}],0] - \left\{K[dB_{i+2},B_2,\Gamma_{(i+2)2},[0,\Gamma_{(i+2)1}],0]\Pi_{21} + \textstyle\sum_{(i+2)4} S_1\right\} \times \left\{\Pi_{21}'K[B_2,B_2,0,0,0]\Pi_{21}\right\}^{-1}\Pi_{21}'K[B_2,B_2,0,0,0],$$

where E *and* F *are defined in Section* 2.2*; and* $K[\cdot,\cdot,\cdot,\cdot,\cdot]$ *and* $R(r)$ *are defined in Theorem* 1.

Proof (i) Using equation (2.11), write $Z_2'EZ_2 = Z_2'Q_{z_1}Z_2 - Z_2'Q_{z_1}V_{22}(V_{22}'Q_{z_1}V_{22})^{-1}$ $\times V_{22}'Q_{z_1}Z_2$. Then, letting $D_T = \text{diag}[T^{1/2}, T^{1+1/2}, ..., T^{p+1/2}, T, ..., T]$, we obtain

$$\begin{aligned} T^{-2}Z_2'Q_{z_1}Z_2 &= T^{-2}Z_2'Z_2 - Z_2'Z_1T^{-1}D_T^{-1}(D_T^{-1}Z_1'Z_1D_T^{-1})^{-1}T^{-1}D_T^{-1}Z_1'Z_2 \\ &\Rightarrow K[B_2, B_2, 0, 0, 0], \end{aligned} \tag{A.1}$$

$$\begin{aligned} T^{-1}Z_2'Q_{z_1}V_{22} &= T^{-1}Z_2'V_{22} - T^{-1}Z_2'Z_1D_T^{-1}(D_T^{-1}Z_1'Z_1D_T^{-1})^{-1}D_T^{-1}Z_1'V_{22} \\ &\Rightarrow K\left[B_2, dB_4, \Gamma_{24}, 0, \binom{0}{\Gamma_{14}}\right] S_2, \end{aligned} \tag{A.2}$$

and

$$\begin{aligned} T^{-1}V_{22}'Q_{z_1}V_{22} &= T^{-1}V_{22}'V_{22} - T^{-1}V_{22}'Z_1D_T^{-1}(D_T^{-1}Z_1'Z_1D_T^{-1})^{-1}D_T^{-1}Z_1'V_{22} \\ &\Rightarrow S_2' \textstyle\sum_{44} S_2 \end{aligned} \tag{A.3}$$

by using the weak convergence results in Phillips (1988). The stated result follows from (A.1), (A.2) and (A.3).

(ii) Write $V_i'EZ_2 = V_i'Q_{z_1}Z_2 - V_i'Q_{z_1}V_{22}(V_{22}'Q_{z_1}V_{22})^{-1}V_{22}'Q_{z_1}Z_2$. Using the same methods as for (A.2) and (A.3), respectively, we obtain

$$T^{-1}V_i'Q_{z_1}Z_2 \Rightarrow K[dB_{i+2}, B_2, \Gamma_{(i+2)2}, [0, \Gamma_{(i+2)1}], 0], \tag{A.4}$$

$$T^{-1}V_i'Q_{z_1}V_{22} \Rightarrow \textstyle\sum_{(i+2)4} S_2. \tag{A.5}$$

The required result is obtained by using (A.2), (A.4) and (A.5).

(iii) Write $V_i'EV_j = V_i'Q_{z_1}V_j - V_i'Q_{z_1}V_{22}(V_{22}'Q_{z_1}V_{22})^{-1}V_{22}'Q_{z_1}V_j$. As for (A.3),

$$T^{-1}V_i'Q_{z_1}V_j \Rightarrow \textstyle\sum_{(i+2)(j+2)}. \tag{A.6}$$

Then (A.3), (A.5) and (A.6) yields the stated result.

(iv) We have $V_i'FV_j = V_i'Q_{z_1}V_j - V_i'Q_{z_1}Y_{21}(Y_{21}'Q_{z_1}Y_{21})^{-1}Y_{21}'Q_{z_1}V_j$. Furthermore, we deduce from (A.1), (A.4) and (A.5)

$$\begin{aligned} T^{-1}V_i'Q_{z_1}Y_{21} &= T^{-1}V_i'Q_{z_1}(Z_2\Pi_{21} + V_{21}) \\ &\Rightarrow K[dB_{i+2}, B_2, \Gamma_{(i+2)2}, [0, \Gamma_{(i+2)1}], 0]\Pi_{21} + \textstyle\sum_{(i+2)4} S_1 \end{aligned} \tag{A.7}$$

and

$$T^{-2}Y_{21}'Q_{z_1}Y_{21} = (Z_2\Pi_{21} + V_{21})'Q_{z_1}(Z_2\Pi_{21} + V_{21}) \Rightarrow \Pi_{21}'K[B_2, B_2, 0, 0, 0]\Pi_{21}. \tag{A.8}$$

Therefore, the stated result follows from (A.6), (A.7) and (A.8).

(v) Write $V_i'FZ_2 = V_i'Q_{z_1}Z_2 - V_i'Q_{z_1}Y_{21}(Y_{21}'Q_{z_1}Y_{21})^{-1}Y_{21}'Q_{z_1}Z_2$. But, due to (A.1) and (A.4),

$$T^{-2}Y_{21}'Q_{z_1}Z_2 = T^{-2}(Z_2\Pi_{21} + V_{21})'Q_{z_1}Z_2 \Rightarrow \Pi_{21}'K[B_2, B_2, 0, 0, 0]. \tag{A.9}$$

Using (A.4), (A.7), (A.8) and (A.9) provides the desired result.

Proof of Theorem 1 (i) Using equation (2.9); and the relations $u = V_1 - V_2\beta$ and $S_1S_1' + S_2S_2' = I$ gives

$$\begin{aligned}\hat{\beta}_1 &= (Y_{21}'EY_{21})^{-1}Y_{21}'Ey_1 \\ &= \beta_1 + (Y_{21}'EY_{21})^{-1}Y_{21}'EY_{22}\beta_2 + (Y_{21}'EY_{21})^{-1}Y_{21}'EZ_1\gamma \\ &\quad + (Y_{21}'EY_{21})^{-1}Y_{21}'E(V_1 - V_2\beta) \\ &= \beta_1 + (Y_{21}'EY_{21})^{-1}Y_{21}'EV_2S_2S_2'\beta + (Y_{21}'EY_{21})^{-1}Y_{21}'E(V_1 - V_2\beta) \\ &= \beta_1 + (Y_{21}'EY_{21})^{-1}Y_{21}'EV_2(S_2S_2' - I)\beta + (Y_{21}'EY_{21})^{-1}Y_{21}'EV_1 \\ &= \beta_1 - (Y_{21}'EY_{21})^{-1}Y_{21}'EV_{21}\beta_1 + (Y_{21}'EY_{21})^{-1}Y_{21}'EV_1. \qquad \text{(A.10)}\end{aligned}$$

Thus,

$$T(\hat{\beta}_1 - \beta_1) = -(T^{-2}Y_{21}'EY_{21})^{-1}T^{-1}Y_{21}'EV_{21}\beta_1 + (T^{-2}Y_{21}'EY_{21})^{-1}T^{-1}Y_{21}'EV_1.$$

But parts (i), (ii) and (iii) of Lemma A yields

$$T^{-1}Y_{21}'EY_{21} = T^{-2}(Z_2\Pi_{21}+V_{21})'E(Z_2\Pi_{21}+V_{21}) \Rightarrow \Pi_{21}'K[B_2, B_2, 0, 0, 0]\Pi_{21}, \quad \text{(A.11)}$$

$$\begin{aligned}T^{-1}Y_{21}'EV_{21} &= T^{-1}(Z_2\Pi_{21} + V_{21})'EV_2S_1 \\ &\Rightarrow \Pi_{21}'\Big\{K\Big[B_2, dB_4, \Gamma_{24}, 0, \tbinom{0}{\Gamma_{14}}\Big] \\ &\quad - K\Big[B_2, dB_4, \Gamma_{24}, 0, \tbinom{0}{\Gamma_{14}}\Big](S_2'\textstyle\sum_{44}S_2)^{-1}S_2'\sum_{44}\Big\}S_1 \\ &\quad + S_1'\Big\{\textstyle\sum_{44} - \sum_{44}S_2(S_2'\sum_{44}S_1)^{-1}S_2'\sum_{44}\Big\}S_1 \qquad \text{(A.12)}\end{aligned}$$

and

$$\begin{aligned}T^{-1}Y_{21}'EV_1 &= T^{-1}(Z_2\Pi_{21} + V_{21})'EV_1 \\ &\Rightarrow \Pi_{21}'\Big\{K\Big[B_2, dB_3, \Gamma_{23}, 0, \tbinom{0}{\Gamma_{13}}\Big] \\ &\quad - K\Big[B_2, dB_4, \Gamma_{24}, 0, \tbinom{0}{\Gamma_{14}}\Big](S_2'\textstyle\sum_{44}S_2)^{-1}S_2'\sum_{43}\Big\} \\ &\quad + S_1'\{\textstyle\sum_{43} - \sum_{44}S_2(S_2'\sum_{44}S_2)^{-1}S_2'\sum_{43}\}. \qquad \text{(A.13)}\end{aligned}$$

Therefore, the result follows from (A.11), (A.12) and (A.13).

(ii) We obtain $Y_{22}'FY_{22} = V_{22}'FV_{22}$ and $Y_{22}'Fy_1 = V_{22}'Fy_1$ by using equation (2.11). Thus,

$$\hat{\beta}_2 = (V_{22}'FV_{22})^{-1}V_{22}'FZ_2\Pi_{21}\beta_1 + (V_{22}'FV_{22})^{-1}V_{22}'FV_1. \qquad \text{(A.14)}$$

for which the relation $S_1S_1' + S_2S_2' = I$ is used as for (A.10). Now applying parts (iv) and (v) of Lemma A to (A.14) provides the desired result in a straightforward way.

(iii) Because $R_1'\pi_1 = \gamma_1$,

$$\begin{aligned}\hat{\gamma}_1 &= R_1'\hat{\gamma} \\ &= R_1'(Z_1'Z_1)^{-1}Z_1'y_1 - R_1'(Z_1'Z_1)^{-1}Z_1'[Y_{21}, Y_{22}]\begin{bmatrix}\hat{\beta}_1 \\ \hat{\beta}_2\end{bmatrix} \\ &= \gamma_1 + R_1'(Z_1'Z_1)^{-1}Z_1'Z_2\pi_2 + R_1'(Z_1'Z_1)^{-1}Z_1'V_1 - R_1'(Z_1'Z_1)^{-1}Z_1'Y_{21}\beta_1 \\ &\quad - R_1'(Z_1'Z_1)^{-1}Z_1'Y_{21}(\hat{\beta}_1 - \beta_1) - R_1'(Z_1'Z_1)^{-1}Z_1'Y_{22}\hat{\beta}_2 \\ &= \gamma_1 + a_1 + a_2 - a_3 - a_4 - a_5, \text{ say.} \qquad \text{(A.15)}\end{aligned}$$

But using $R_1'\Pi_1 = 0$ and $\pi_2 = \Pi_{21}\beta_1$ gives

$$a_1 - a_3 = -R_1'(Z_1'Z_1)^{-1}Z_1'V_{21}\beta_1, \tag{A.16}$$

$$a_4 = \{R_1'(Z_1'Z_1)^{-1}Z_1'Z_2\Pi_{21} + R_1'(Z_1'Z_1)^{-1}Z_1'V_{21}\}(\hat{\beta}_1 - \beta_1), \tag{A.17}$$

$$a_5 = R_1'(Z_1'Z_1)^{-1}Z_1'V_{22}\hat{\beta}_2. \tag{A.18}$$

Plugging (A.16), (A.17), (A.18) into (A.15) and standardizing by D_T for asymptotic analysis, we obtain

$$\begin{aligned} T^{1/2}(\hat{\gamma}_1 - \gamma_1) &= -T^{1/2}R_1'D_T^{-1}(D_T^{-1}Z_1'Z_1D_T^{-1})^{-1}\Big\{D_T^{-1}Z_1'V_{21}\beta_1 - D_T^{-1}Z_1'V_1 + \\ &\quad + D_T^{-1}T^{-1}(Z_1'Z_2\Pi_{21} + Z_1'V_{21})T(\hat{\beta}_1 - \beta_1) + D_T^{-1}Z_1'V_{22}\hat{\beta}_2\Big\} \\ &= -T^{1/2}R_1'D_T^{-1}\kappa_T = -R_{11}'\kappa_{1t} - R_{12}'\mathrm{diag}(T^{-1}, ..., T^{-p})\kappa_{2T} \\ &\quad - T^{-1/2}R_{13}'\kappa_{3T}, \end{aligned} \tag{A.19}$$

where $R_1' = [\overset{1}{R_{11}'}, \overset{p}{R_{12}'}, \overset{k_1-p-1}{R_{13}'}]\ k_{11}$ and $\kappa_T' = [\overset{1}{\kappa_{1T}'}, \overset{p}{\kappa_{2T}'}, \overset{k_1-p-1}{\kappa_{3T}'}]\ 1$. Now applying the weak convergence results in Phillips (1988) yields the stated result.

(iv) Replacing D_T by T and deleting $T^{1/2}$ in (A.19) and applying the weak convergence results in Phillips (1988) provides the desired result.

(v) Expressing $\hat{\gamma}_2$ by employing the same methods as for (A.15) and using the weak convergence results in Phillips (1988) gives

$$\begin{aligned} \hat{\gamma}_2 &= R_2'\pi_1 - R_2'\Pi_{11}\beta_1 - R_2'(Z_1'Z_1)^{-1}Z_1'V_{21}\beta_1 + R_2'(Z_1'Z_1)^{-1}Z_1'V_1 \\ &\quad - [R_2'\Pi_{11} + R_2'(Z_1'Z_1)^{-1}Z_1'Z_2\Pi_{21} + R_2'(Z_1'Z_1)^{-1}Z_1'V_{21}](\hat{\beta}_1 - \beta_1) \\ &\quad - [R_2'\Pi_{12} + R_2'(Z_1'Z_1)^{-1}Z_1'V_{22}]\hat{\beta}_2 \\ &\Rightarrow R_2'\pi_1 - R_2'\Pi_{11}\beta_1 - R_2'\Pi_{12}\hat{b}_2 \end{aligned} \tag{A.20}$$

as required.

Proof of Corollary 1 (i) Delete S_2 and put $\Pi_{21} = 0$ in part (ii) of Theorem 1. Then, the required result follows.

(ii) This follows from part (iii) of Theorem 1 once S_2 and the terms associated with β_1 are deleted.

(iii) This follows from part (iv) of Theorem 1 as in part (iii).

(iv) Replacing Π_{12} with Π_1 and deleting the term involving β_1 in (A.21) gives the wanted result.

Proof of Corollary 2 (i) Delete terms involving S_2, erase S_1 and replace Π_{21} and β_1 with Π_2 and β, respectively, in part (i) of Theorem 1. Then, the result follows.

(ii) Write

$$\begin{aligned} D_T(\hat{\gamma} - \gamma) &= -D_T\Pi_1T^{-1}T(\hat{\beta} - \beta) + \left(D_T^{-1}Z_1'Z_1D_T^{-1}\right)^{-1} \\ &\quad \times\{-D_T^{-1}Z_1'Z_2\Pi_2(\hat{\beta} - \beta) - D_T^{-1}Z_1'V_2\hat{\beta} + D_T^{-1}Z_1'V_1\}. \end{aligned} \tag{A.21}$$

Let $T(\hat{\beta}-\beta) = \check{b}_T$. Then $D_T \Pi_1 T^{-1} T(\hat{\beta}-\beta) = \text{diag}[T^{-1/2}, T^{1/2}, ..., T^{1/2+p-1}, 1, ..., 1]\Pi_1 \check{b}_T$. Therefore,

$$B_T D_T \Pi_1 T^{-1} T(\hat{\beta}-\beta) = \text{diag}[T^{-1/2}, 1, ..., 1]\Pi_1 \check{b}_T,$$

where $B_T = \text{diag}[1, T^{-1/2}, ..., T^{-(1/2+p-1)}, 1, ..., 1]$. But $B_T D_T = C_T$ and $\check{b}_T \Rightarrow \check{b}$, which implies

$$C_T \Pi_1 (\hat{\beta}-\beta) \Rightarrow \text{diag}[0, 1, ..., 1]\Pi_1 \check{b}. \quad \text{(A.22)}$$

Furthermore,

$$B_T \to \text{diag}[1, 0, ..., 0, 1, ..., 1]. \quad \text{(A.23)}$$

Now we obtain the required result from (A.21), (A.22), and (A.23) by applying the weak convergence results in Phillips (1988).

(iii) Replacing D_T with T in (A.21) and applying the weak convergence results in Phillips (1988) gives the required result.

Lemma B *Suppose that Assumptions* 1, 2, *and* 3 *hold. Then, for* $i, j = 1$ *or* 2,

(i) $T^{-2} Z_2' G Z_2 \Rightarrow J\left\{K[B_2, B_2, 0, 0, 0], K\left[B_2, dB_4, \Gamma_{24}, 0, \binom{0}{\Gamma_{14}}\right]\right\}$,

(ii) $T^{-1} V_i' G Z_2 \Rightarrow M\left\{K[dB_{i+2}, B_2, \Gamma_{(i+2)2}, [0, \Gamma_{(i+2)1}], 0], K[B_2, B_2, 0, 0, 0],\right.$
$\left. K\left[B_2, dB_4, \Gamma_{24}, 0, \binom{0}{\Gamma_{14}}\right]\right\}$,

(iii) $V_i' G V_j = O_p(1)$,

(iv) $V_i' H V_j \Rightarrow W\{K[dB_{i+2}, B_2, \Gamma_{(i+2)2}, [0, \Gamma_{(i+2)1}], 0], K[B_2, B_2, 0, 0, 0],$
$K[dB_{j+2}, B_2, \Gamma_{(j+2)2}, [0, \Gamma_{(j+2)1}], 0]'\}$,

(v) $T^{-1} V_i' H Z_2 \Rightarrow N\{K[dB_{i+2}, B_2, \Gamma_{(i+2)2}, [0, \Gamma_{(i+2)1}], 0], K[B_2, B_2, 0, 0, 0]\}$,

where G, H and L are defined in Section 2.3*;* $K[\cdot,\cdot,\cdot,\cdot,\cdot]$ *and* $R(r)$ *are defined in Theorem* 1*; and* $J[\cdot,\cdot]$, $M[\cdot,\cdot,\cdot]$, $W[\cdot,\cdot,\cdot]$ *and* $N[\cdot,\cdot]$ *are defined in Theorem* 2.

Proof (i) Since $L = Q_{z_1} Z_2 (Z_2' Q_{z_1} Z_2)^{-1} Z_2' Q_{z_1}$,

$$Z_2' L Z_2 = Z_2' Q_{z_1} Z_2, \quad \text{(A.24)}$$

$$Z_2' L V_i = Z_2' Q_{z_1} V_i \quad \text{(A.25)}$$

and

$$V_i' L V_j = V_i' Q_{z_1} Z_2 (Z_2' Q_{z_1} Z_2)^{-1} Z_2' Q_{z_1} V_j. \quad \text{(A.26)}$$

Now, writing $Z_2' G Z_2 = Z_2' L Z_2 - Z_2' L V_{22} (V_{22}' L V_{22})^{-1} V_{22}' L Z_2$ and using (A.1) and (A.2) gives the stated result.

(ii) We may write $V_i' G Z_2 = V_i' L Z_2 - V_i' L V_{22} (V_{22}' L V_{22})^{-1} V_{22}' L Z_2$. Rewriting this by using (A.25) and (A.26) and applying (A.1), (A.2), and (A.4) gives the wanted the result.

(iii) Using (A.1) and (A.4), we find $V_i' L V_j = O_p(1)$. But $V_i' G V_j = V_i' L V_j - V_i' L V_{22} (V_{22}' L V_{22})^{-1} V_{22}' L V_j$, from which the result follows.

(iv) Because $V_i' L V_j = O_p(1)$ and $V_i L Z_2 = O_p(T)$, we may write by using (A.24), (A.25), and (A.26)

$$\begin{aligned} V_i' H V_j &= V_i' Q_{z_1} Z_2 (Z_2' Q_{z_1} Z_2)^{-1} Z_2' Q_{z_1} V_j - [V_i' Q_{z_1} Z_2 \Pi_{21} + O_p(1)] \\ &\quad \times [\Pi_{21}' Z_2' Q_{z_1} Z_2 \Pi_{21} + O_p(T)]^{-1} [\Pi_{21}' Z_2' Q_{z_1} V_j + O_p(1)]. \end{aligned}$$

Now applying (A.1) and (A.4) gives the desired result.

(v) As in part (iv), write

$$\begin{aligned} V_i' H Z_2 &= V_i' Q_{z_1} Z_2 - [V_i' Q_{z_1} Z_2 \Pi_{21} + O_p(1)][\Pi_{21}' Z_2 Q_{z_1} Z_2 \Pi_{21} + O_p(T)]^{-1} \\ &\quad \times [\Pi_{21}' Z_2' Q_{z_1} Z_2 + O_p(T)] \end{aligned}$$

and use (A.1) and (A.4). Then, the desired result follows.

Proof of Theorem 2 (i) As for (A.10), we have

$$\bar\beta_1 = \beta_1 - (Y_{21}' G Y_{21})^{-1} Y_{21}' G V_{21} \beta_1 + (Y_{21}' G Y_{21})^{-1} Y_{21}' G V_1.$$

But due to parts (i), (ii) and (iii) of Lemma B

$$\begin{aligned} T^{-2} Y_{21}' G Y_{21} &= T^{-2} (Z_2 \Pi_{21} + V_{21})' G (Z_2 \Pi_{21} + V_{21}) = T^{-2} \Pi_{21}' Z_2' G Z_2 \Pi_{21} + o_p(1) \\ &\Rightarrow \Pi_{21}' J \left\{ K[B_2, B_2, 0, 0, 0], K\Big[B_2, dB_4, \Gamma_{24}, 0, \left(\begin{smallmatrix} 0 \\ \Gamma_{14} \end{smallmatrix}\right)\Big] \right\} \Pi_{21}, \end{aligned} \tag{A.27}$$

$$\begin{aligned} T^{-1} Y_{21}' G V_{21} &= T^{-1} (Z_2 \Pi_{21} + V_{21})' G V_2 S_1 = T^{-1} \Pi_{21}' Z_2' G V_2 S_1 + o_p(1) \\ &\Rightarrow \Pi_{21}' M \{ K[dB_4, B_2, \Gamma_{42}, [0, \Gamma_{41}], 0], K[B_2, B_2, 0, 0, 0], \\ &\quad K\Big[B_2, dB_4, \Gamma_{24}, 0, \left(\begin{smallmatrix} 0 \\ \Gamma_{14} \end{smallmatrix}\right)\Big] \Big\}' S_1 \end{aligned} \tag{A.28}$$

and

$$\begin{aligned} T^{-1} Y_{21}' G V_1 &= T^{-1} (Z_2 \Pi_{21} + V_{21})' G V_1 = T^{-1} \Pi_{21}' Z_2' G V_1 + o_p(1) \\ &\Rightarrow \Pi_{21}' M \{ K[dB_3, B_2, \Gamma_{32}, [0, \Gamma_{31}], 0], K[B_2, B_2, 0, 0, 0], \\ &\quad K\Big[B_2, dB_4, \Gamma_{24}, 0, \left(\begin{smallmatrix} 0 \\ \Gamma_{14} \end{smallmatrix}\right)\Big] \Big\}'. \end{aligned} \tag{A.29}$$

The result follows from (A.27), (A.28), and (A.29).

(ii) We have as for (A.14) $\bar\beta_2 = (V_{22}' H V_{22})^{-1} V_{22}' H Z_2 \Pi_{21} \beta_1 + (V_{22}' H V_{22})^{-1} V_{22}' H V_1$. Thus, the stated result follows from parts (iv) and (v) of Lemma B.

(iii) Using (A.19), we have

$$\begin{aligned} T^{-1/2}(\bar\gamma_1 - \gamma_1) &= -R_1' T^{1/2} D_T^{-1} (D_T^{-1} Z_1' Z_1 D_T^{-1})^{-1} \Big\{ T^{-1} D_T^{-1} Z_1' V_{21} \beta_1 - T^{-1} D_T^{-1} Z_1' V_1 \\ &\quad + D_T^{-1} T^{-2} (Z_1' Z_2 \Pi_{21} + Z_1' V_{21}) T (\bar\beta_1 - \beta_1) + D_T^{-1} Z_1' V_{22} T^{-1} \bar\beta_2 \Big\} \\ &= -R_1' T^{1/2} D_T^{-1} (D_T^{-1} Z_1' Z_1 D_T^{-1})^{-1} D_T^{-1} Z_1' V_{22} T^{-1} \bar\beta_2 + o_p(1) \\ &= -R_1' T^{1/2} D_T^{-1} \lambda_T + o_p(1), \text{ say.} \end{aligned} \tag{A.30}$$

Letting $R_1' = [\overset{1}{R_{11}'}, \overset{p}{R_{12}'}, \overset{k_1-p-1}{R_{13}'}]\ k_{11}$ and $\lambda_T' = [\overset{1}{\lambda_{1T}'}, \overset{p}{\lambda_{2T}'}, \overset{k_1-p-1}{\lambda_{3T}'}]\ 1$, we have

$$\begin{aligned} R_1' T^{1/2} D_T^{-1} \lambda_T &= R_{11}'\lambda_{1T} + R_{12}'\mathrm{diag}(T^{-1}, ..., T^{-p})\lambda_{2T} + R_{13}'\lambda_{3T}T^{-1/2} \\ &= R_{11}'\lambda_{1T} + o_p(1). \end{aligned} \tag{A.31}$$

The stated result is obtained from (A.30) and (A.31).

(iv) Because

$$\begin{aligned} \bar{\gamma}_1 - \gamma_1 &= -R_1'(Z_1'Z_1)^{-1}\{Z_1'V_{21}\beta_1 - Z_1'V_1 + (Z_1'Z_2\Pi_{21} + Z_1'V_{21})(\bar{\beta}_1 - \beta_1) + Z_1'V_{22}\bar{\beta}_2\} \\ &= -R_1'(T^{-2}Z_1'Z_1)^{-1}T^{-1}Z_1'V_{22}T^{-1}\bar{\beta}_2 + o_p(1), \end{aligned} \tag{A.32}$$

the required result follows in a straightforward way.

(v) We obtain by using (A.20

$$T^{-1}\bar{\gamma}_2 = -[R_2'\Pi_{12} + R_2'(Z_1'Z_1)^{-1}Z_1'V_{22}]T^{-1}\hat{\beta}_2 + o_p(1) \Rightarrow -R_2'\Pi_{12}\bar{b}_2,$$

as required.

Proof of Corollary 3 (i) We have $\bar{\beta} = (Y_2'LY_2)^{-1}Y_2'Ly_1$. But because $\Pi_2 = 0$ and $LZ_1 = 0$, $\bar{\beta} = (V_2'LV_2)^{-1}V_2'LV_1$. Now, using (A.26), (A.1) and (A.4) gives the required result.

(ii) As for (A.32), we may write

$$\bar{\gamma}_1 - \gamma_1 = R_1'(Z_1'Z_1)^{-1}Z_1'V_1 - R_1'(Z_1'Z_1)^{-1}Z_1'V_2\bar{\beta}. \tag{A.33}$$

But as in (A.19), $T^{1/2}(\bar{\gamma}_1 - \gamma_1) = R_{11}'\tau_{1T} + o_p(1)$ where τ_{1T} is the first element of the vector $(D_T^{-1}Z_1'Z_1D_T^{-1})^{-1}D_T^{-1}Z_1'V_1 - (D_T^{-1}Z_1'Z_1D_T^{-1})^{-1}D_T^{-1}Z_1'V_2\hat{\beta}$, which yields the stated result.

(iii) This follows straightforwardly from (A.33).

(iv) Because

$$\begin{aligned} \bar{\gamma}_2 &= R_2'\pi_1 + R_2'(Z_1'Z_1)^{-1}Z_1'V_1 - R_2'\Pi_1\bar{\beta} - R_2'(Z_1'Z_1)^{-1}Z_1'V_2\bar{\beta} \\ &= R_2'\pi_1 - R_2'\Pi_1\bar{\beta} + o_p(1), \end{aligned}$$

the required result follows.

Proof of Corollary 4 (i) Replace Y_{21}, V_{21}, G and β_1 with Y_2, V_2, L, and β, respectively, in the proof of part (i) of Theorem 2, and apply the weak convergence results in Phillips (1988) as in (A.27), (A.28), and (A.29). Then, the result follows.

Parts (ii), (iii) follow in exactly the same way as parts (ii) and (iii) of Corollary 2.

Proof of Theorem 3 Assume $Z_{10} = 0$, $Z_{20} = 0$ and $V_{20} = 0$ for simplicity. This brings no loss of generality for the asymptotic results we are to derive. Write

$$\hat{\alpha}_{FM-OLS} - \alpha = (X'X)^{-1}(X'u - X'\Delta X\hat{\Psi}_{22}^{-1}\hat{\Psi}_{12}' - T\hat{\Xi}_2\hat{\kappa}) \tag{A.34}$$

and let

$$p_t = \begin{bmatrix} u_t \\ \Delta X_t \end{bmatrix} = Jw_t + \begin{bmatrix} 0 \\ \Delta V_{2t} \\ 0 \end{bmatrix} = a_t + b_t, \; a_t = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \begin{matrix} 1 \\ n+k_1 \end{matrix}, \; b_t = \begin{bmatrix} b_{1t} \\ b_{2t} \end{bmatrix} \begin{matrix} 1 \\ n+k_1 \end{matrix},$$

where w_t and J are defined in Assumption 3 and Theorem 3, respectively. Assumption 3 gives

$$T^{-1/2} \sum_{t=1}^{[Tr]} p_t = T^{-1/2} \sum_{t=1}^{[Tr]} a_t + o_p(1) \Rightarrow JB(r) = D(r), \tag{A.35}$$

where $E[D(r)D(r)'] = J\Omega J' = \Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$. Using (A.35), we obtain

$$T^{-2} X'X = T^{-2} \sum_{t=1}^{T} S_t^{a_2} S_t^{a_2 \prime} + o_p(1) \Rightarrow \int_0^1 D_2(r) D_2(r)' dr, \tag{A.36}$$

$$\begin{aligned} T^{-1} X'u &= T^{-1} \sum_{t=1}^{T} S_t^{a_2} a_{1t} + T^{-1} \sum_{t=1}^{T} S_t^{b_2} (V_{1t} - \beta' V_{2t}) \\ &\Rightarrow \int_0^1 D_2(r) dD_1(r) + J_2 \Gamma J_1' + \begin{bmatrix} \sum_{43} - \sum_{44} \beta \\ 0 \end{bmatrix}, \end{aligned} \tag{A.37}$$

where $S_t^x = \sum_{i=1}^t x_i$. Furthermore, $\Psi = 2\pi f_{pp}(0) = 2\pi f_{aa}(0) = \Phi$, because $f_{bb}(0) = 0$ and

$$2\pi f_{ab}(0) = \sum_{t=-\infty}^{\infty} E(a_0 b_t') = [0, JE(w_0 v_{2\infty}') - JE(w_0 v_{2-\infty}'), 0] = 0.$$

Note that the latter equality holds due to the existence of the spectral density matrix $f_{ww}(0) = \Omega/(2\pi)$ assumed in Assumption 3. Therefore,

$$\hat{\Psi} \xrightarrow{p} \Phi. \tag{A.38}$$

Now, write

$$\begin{aligned} T^{-1} X' \Delta X \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' &= T^{-1} \sum_{t=1}^{T} S_t^{a_2} a_{2t}' \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' + T^{-1} \sum_{t=1}^{T} S_t^{a_2} b_{2t}' \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' \\ &\quad + T^{-1} \sum_{t=1}^{T} S_t^{b_2} a_{2t}' \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' + T^{-1} \sum_{t=1}^{T} S_t^{b_2} b_{2t}' \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' \\ &= d_1 + d_2 + d_3 + d_4, \text{ say.} \end{aligned}$$

Then, Assumption 3, (A.35) and (A.38) yield

$$d_1 \Rightarrow \int_0^1 D_2(r) dD_2(r)' \Phi_{22}^{-1} \Phi_{12}' + J_2 \Gamma J_2' \Phi_{22}^{-1} \Phi_{12}', \tag{A.39}$$

$$\begin{aligned} d_2 &= T^{-1} J_2 \sum_{t=1}^{T} S_t^w \Delta w_t' K' \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' = \\ &= J_2 T^{-1} (w_T S_T^{\prime w} - \sum_{t=1}^{T} w_t w_{t-1}') K' \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' \xrightarrow{p} -J_2 \textstyle\sum_{-} K' \Phi_{22}^{-1} \Phi_{12}', \end{aligned} \tag{A.40}$$

$$d_3 = T^{-1} K \sum_{t=1}^{T} w_t w_t' J_2' \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' \xrightarrow{p} K \textstyle\sum J_2' \Phi_{22}^{-1} \Phi_{12}' \tag{A.41}$$

and

$$d_4 = T^{-1} K \sum_{t=1}^{T} w_t \Delta w_t' K' \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' \xrightarrow{p} K(\textstyle\sum - \sum_{-}) K' \Phi_{22}^{-1} \Phi_{12}', \tag{A.42}$$

for which we used the relations $a_{2t} = J_2 \Delta w_t$ and $b_{2t} = K w_t$. Note that K is defined in Theorem 3. Moreover,

$$\hat{\Xi}_2 \hat{\kappa} \xrightarrow{p} \Xi_{21} - \Xi_{22} \Phi_{22}^{-1} \Phi_{12}' = J_2 \Gamma J_1' - J_2 \Gamma J_2' \Phi_{22}^{-1} \Phi_{12}', \tag{A.43}$$

where $\hat{\Xi}_2 = [\Xi_{21}, \Xi_{22}]$. Plugging (A.36)–(A.37) and (A.39)–(A.43) into (A.34) gives the required result.

Proof of Theorem 4 Assume $Z_{10} = 0$, $Z_{20} = 0$ and $V_{20} = 0$ for simplicity as in the proof of Theorem 3. Write

$$\hat{\alpha}_{FM-2SLS} - \alpha = (X' P_z X)^{-1} (X' P_z u - X' \Delta X \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' - T \hat{\Xi}_2 \hat{\kappa}). \tag{A.44}$$

Let $B_a(r) = [B_1(r)', B_2(r)']'$ and $Z_t = [Z_{1t}',, Z_{2t}']'$. Then, using $X_t = J_{21} Z_t + \begin{bmatrix} V_{2t} \\ 0 \end{bmatrix}$, where J_{21} is the first $(k_1 + k_2)$ columns of J_2 in Theorem 3, we obtain

$$\begin{aligned} T^{-2} X' P_z X \; &= \; \{J_{21} T^{-2} \textstyle\sum_{t=1}^T Z_t Z_t' + o_p(1)\} \left(T^{-2} \textstyle\sum_{t=1}^T Z_t Z_t'\right)^{-1} \\ &\quad \times \{T^{-2} \textstyle\sum_{t=1}^T Z_t Z_t' J_{21} + o_p(1)\} \\ &\Rightarrow J_{21} \textstyle\int_0^1 B_a(r) B_a(r)' dr J_{21}' = \int_0^1 D_2(r) D_2(r)' dr, \end{aligned} \tag{A.45}$$

$$\begin{aligned} T^{-1} X' P_z u \; &= \; \{J_{21} T^{-2} \textstyle\sum_{t=1}^T Z_t Z_t' + o_p(1)\} \left(T^{-2} \textstyle\sum_{t=1}^T Z_t Z_t'\right)^{-1} (T^{-1} \textstyle\sum_{t=1}^T Z_t w_t' J_1') \\ &\Rightarrow J_{21} \textstyle\int_0^1 B_a(r) dB(r)' J_1' + J_2 \Gamma J_1' = \int_0^1 D_2(r) dD_1(r) dr + J_2 \Gamma J_1', \end{aligned} \tag{A.46}$$

$$\begin{aligned} T^{-1} X' P_z \Delta X \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' \; &= \; \{J_{21} T^{-2} \textstyle\sum_{t=1}^T Z_t Z_t' + o_p(1)\} \left(T^{-2} \textstyle\sum_{t=1}^T Z_t Z_t'\right)^{-1} \\ &\quad (T^{-1} \textstyle\sum_{t=1}^T Z_t w_t' J_2') \hat{\Psi}_{22}^{-1} \hat{\Psi}_{12}' \\ &\Rightarrow J_{21} \textstyle\int_0^1 B_a(r) dB_a(r)' J_{21}' \Phi_{22}^{-1} \Phi_{12}' + J_2 \Gamma J_2' \Phi_{22}^{-1} \Phi_{12}' \\ &= \textstyle\int_0^1 D_2(r) dD_2(r)' \Phi_{22}^{-1} \Phi_{12}' + J_2 \Gamma J_2' \Phi_{22}^{-1} \Phi_{12}'. \end{aligned} \tag{A.47}$$

Plugging (A.43) and (A.45)–(A.47) into (A.44) gives the desired result.

Proof of Corollary 5 : This is an immediate consequence of Theorem 4.

Proof of Theorem 5 This theorem can be proved by using the same methods as in Saikkonen (1991) which uses the methods of Berk (1974), Lewis and Reinsel (1985) and Said and Dickey (1984). Therefore, only a brief outline will be given here. As in Saikkonen (1991, p.21), we have under the given assumptions

$$T(\hat{\alpha}_{LL-OLS} - \alpha) = (T^{-2} \textstyle\sum_{t=k+1}^{T-k} x_t x_t') [T^{-1} \textstyle\sum_{t=k+1}^{T-k} x_t (e_{1t} - \beta' e_{2t})] + o_p(1).$$

But

$$T^{-1} \textstyle\sum_{t=k+1}^{T-k} x_t (e_{1t} - \beta' e_{2t}) \Rightarrow \int_0^1 D_2(r) dD_c(r) + \begin{bmatrix} \sum_{21}^e - \sum_{-21}^e - (\sum_{22}^e + \sum_{-22}^e \beta \\ 0 \end{bmatrix}. \tag{A.48}$$

Note that the spectral density of $e_{1t} - \beta' e_{2t}$ at the zero frequency is 2π times $\Phi_{11} - \Phi_{12}\Phi_{22}^{-1}\Phi_{21}$ from Assumption 4. Now the required result follows from (A.36) and (A.48).

Proof of Theorem 6 As in the proof of Theorem 5, we have by using the same methods as in Saikkonen (1991)

$$\begin{aligned} T(\hat{\alpha}_{2SLS} - \alpha) = & \Big[\Big\{ T^{-2} \textstyle\sum_{t=k+1}^{T-k} X_t Z_t' + o_p(1) \Big\} \Big\{ T^{-2} \textstyle\sum_{t=k+1}^{T-k} Z_t Z_t' + o_p(1) \Big\}^{-1} \\ & \times \Big\{ T^{-2} \textstyle\sum_{t=k+1}^{T-k} Z_t X_t + o_p(1) \Big\} \Big]^{-1} \Big\{ T^{-2} \textstyle\sum_{t=k+1}^{T-k} X_t Z_t' + o_p(1) \Big\} \\ & \times \Big\{ T^{-2} \textstyle\sum_{t=k+1}^{T-k} Z_t Z_t' + o_p(1) \Big\}^{-1} T^{-1} \textstyle\sum_{t=k+1}^{T-k} Z_t (e_{1t} - \beta' e_{2t}) + o_p(1). \end{aligned}$$

Therefore, following the same line of argument as in (A.45), (A.46), and (A.48) yields the stated result.

Proof of Corollary 6 This follows from Theorem 6.

References

Adams, F. G. and L. R. Klein (1991): "Performance of Quarterly Econometric Models of the United States: A New Round of Model Comparisons," in L. R. Klein (ed.), *Comparative Performance of U.S. Econometric Models*. New York: Oxford University Press.

Anderson, T. W. and N. Kunitomo (1992): "Tests of Overidentification and Predeterminedness in Simultaneous Equation Models," *Journal of Econometrics*, 54, 49–78.

Andrews, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Basawa, I. V. and D. J. Scott (1983): *Asymptotic Optimal Inference for Non-ergodic Models*. New York: Springer–Verlag.

Berk, K. N. (1974): "Consistent Autoregressive Spectral Estimates," *Annals of Statistics*, 2, 489–502.

Choi, I. and P. C. B. Phillips (1992): "Asymptotic and Finite Sample Distribution Theory for IV Estimators and Tests in Partially Identified Structural Equations," *Journal of Econometrics*, 51, 113–150.

Engle, R. and C. W. J. Granger (1987): "Co-integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251–276.

Hausman, J. (1983): "Specification and Estimation of Simultaneous Equation Models," in Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, Vol. 1. Amsterdam: North–Holland.

Hsiao, C. (1983): "Identification," in Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, Vol. 1. Amsterdam: North–Holland.

Hsiao, C. (1995): "Cointegration and Dynamic Simultaneous Equations Model," mimeo, University of Southern California.

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

Kydland, F. E. and E. C. Prescott (1995): "The Econometrics of the General Equilibrium Approach to Business Cycles," in K. D. Hoover (ed.), *Macroeconometrics: Developments, Tensions and Prospects*. Massachusetts: Kluwer Academic Publishers.

Lewis, R. and G. C. Reinsel (1985): "Prediction of Multivariate Time Series by Autoregressive Model Fitting," *Journal of Multivariate Analysis*, 16, 393–411.

Lucas, G. (1976): "Econometric Policy Evaluation: A Critic," in K. Brunner and A. H. Meltzer (eds.), *Carnegie–Rochester Conference Series on Public Policy*, Vol. 1. Amsterdam: North–Holland.

Park, J. (1990): "Maximum Likelihood Estimation of simultaneous Cointegrated Models," mimeo, University of Toronto.

Park, J. (1992): "Canonical Cointegrating Regression," *Econometrica*, 60, 119–143.

Park, J. and P. C. B. Phillips (1989): "Statistical Inference in Regressions with Integrated Processes: Part 2," *Econometric Theory*, 5, 95–131.

Phillips, P. C. B. (1983): "Exact Small Sample Theory in the Simultaneous Equations Model," in Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, Vol. 1. Amsterdam: North–Holland.

Phillips, P. C. B. (1988): "Multiple Regression with Integrated Processes," in N. U. Prabhu (ed.), *Contemporary Mathematics*, Vol. 80. Providence: American Mathematical Society.

Phillips, P. C. B. (1989): "Partially Identified Econometric Models," *Econometric Theory*, 5, 181–240.

Phillips, P. C. B. (1991). "Optimal inference in cointegrated systems," *Econometrica* 59, 283–306

Phillips, P. C. B. (1996). "Econometric Model Determination," *Econometrica,* 64, 763-812.

Phillips, P. C. B. and B. Hansen (1990): "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies*, 57, 99–125.

Phillips, P. C. B. and M. Loretan (1991): "Estimating Long Run Economic Equilibria," *Review of Economic Studies*, 58, 407–436.

Phillips, P. C. B. and V. Solo (1992): "Asymptotics for linear processes," *Annals of Statistics*, 20, 971–1001.

Said, S. E. and D. A. Dickey (1984): "Testing for a Unit Root in Autoregressive-moving Average Models of Unknown Order," *Biometrika*, 71, 599–607.

Saikkonen, P. (1991): "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory*, 9, 1–21.

Saikkonen, P. (1993) "Estimation of Cointegration Vectors with Linear Restrictions," *Econometric Theory*, 9, 155–188.

Shoven, J. B. and J. Whalley (1984): "Applied General-Equilibrium Models of Taxation and International Trade: An Introduction and Survey," *Journal of Economic Literature*, 22, 1007–1051.

Sims, C. A. (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

Sims, C. A., J. H. Stock and M. W. Watson (1990): "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica*, 58, 113–144.

Stock, J. H. and M. W. Watson (1993): "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61, 783–820.