TECHNICAL APPENDIX

The specification and asymptotic distribution of each of the six estimators employed is presented in this technical appendix. Consider the p-variable system of I(1) variables $y_t$ partitioned into subvectors of dimension $p_1$ and $p_2$ with $p_1 + p_2 = p$.

\[
y_{1t} = \beta y_{2t} + u_{1t} \quad \text{(A1)}
\]

\[
\Delta y_{2t} = u_{2t} \quad \text{(A2)}
\]

where $\beta$ is a $p_1 \times p_2$ matrix of coefficients that represent the cointegrating relations and it is assumed that $u_t \equiv \text{iid}(0, \Sigma)$, where

\[
\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{21}' \\ \sigma_{21} & \Sigma_{22} \end{bmatrix} \quad \text{(A3)}
\]

with the partition of $\Sigma$ conformable with (A1) and (A2). The MLE of $\beta$ in (A1) is OLS on

\[
y_{1t} = \beta y_{2t} + \gamma \Delta y_{2t} + u_{1,2t} \quad \text{(A4)}
\]

where $u_{1,2t} = u_{1t} - \sigma_{12} \Sigma_{22}^{-1} u_{2t}$ and $\gamma = \Sigma_{22}^{-1} \sigma_{21}$, see Phillips and Loretan (1991).

The limiting distribution of this estimator is given by (see Phillips and Loretan, 1991, inter alia)

\[
T(\hat{\beta} - \beta) = \left( \int_0^1 S_2 \, dS_{1,2} \right)^{-1} \left( \int_0^1 S_2 \, dS_{1,2}' \right) \quad \text{(A5)}
\]

where $S_2$ and $S_{1,2}$ are independent Brownian motions. The distribution in (A5) is a Gaussian mixture of normals implying that standard Chi-Square inference may be used to conduct hypothesis tests. When the error term $u_t$ is weakly dependent the variance $\Sigma$ is now the long-run variance, estimates of which can be obtained using spectral methods or a kernel estimator.
The distribution of OLS on (A1) is given by,

$$T(\hat{\beta} - \beta) = \left( \int_{0}^{1} S_2 S_2' \right)^{-1} \int_{0}^{1} S_2 dS_{12} + \left( \int_{0}^{1} S_2 S_2' \right)^{-1} \int_{0}^{1} S_2 dS_2' \Sigma_{22}^{-1} \sigma_{21} + \left( \int_{0}^{1} S_2 S_2' \right)^{-1} \sigma_{21} \tag{A6}$$

where the first term on the right hand side (RHS) is the mixture of normals. The second term on the RHS is a matrix unit root distribution. The third term on the RHS is a bias term arising from the contemporaneous correlation between $u_{1t}$ and $u_{2t}$ and thus the regressor $y_{2t}$. Note that when $\sigma_{21} = 0$ the OLS and MLE estimators are equivalent.

One way, then, to eliminate the asymptotic bias in the OLS estimator is to transform the model in such a way as to eliminate the endogeneity of the regressor. Phillips and Hansen (1990) fully modified OLS (FM-OLS) does just that. Let $\hat{\sigma}_{21}$ be a consistent estimate of $\sigma_{21}$, then a modified OLS estimator of $\beta$ is given by

$$\hat{\beta}^* = (y_{2t}'y_{2t})^{-1} (y_{2t}'y_{1t} - T \hat{\sigma}_{21}) \tag{A7}$$

which eliminates the third RHS term of (A6). To remove the second RHS term of (A6) a further modification is achieved by constructing

$$y_{1t}^* = y_{1t} - \hat{\sigma}_{22} \Delta y_{2t} \tag{A8}$$

where $\hat{\sigma}_{22}$ is a consistent estimate of $\Sigma_{22}$. A FM-OLS estimator is then easily constructed and standard inference may proceed.

If it is assumed that the system given by (A1) and (A2) have a finite order vector autoregressive representation (VAR) as in (A9),

$$Y_t = \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + ... + \Pi_k Y_{t-k} + \epsilon_t \tag{A9}$$

then the system may be rewritten as n error correction model (ECM) as in (A10).
where $\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \Gamma_2 \Delta Y_{t-2} + \ldots + \Gamma_{k-1} \Delta Y_{t-k+1} + \Pi Y_{t-k} + \varepsilon_t$ \hspace{1cm} (A10)

If there is cointegration among the variables in $Y$, then $\Pi$, the coefficient matrix on the lagged levels term in (A10), will have reduced rank and can be decomposed into two $p \times r$ full rank matrices $\alpha$ and $\beta$ such that $\Pi = \alpha \beta'$. The columns spanned by $\beta$ represent the cointegrating vectors. Maximum likelihood estimation of (A10) can be achieved by reduced rank regression, which involves the following steps: (1) regress $\Delta Y_t$ on $\Delta Y_{t-1}, \ldots, \Delta Y_{t-k+1}$ and save the residuals $R_{0t}$, (2) regress $Y_{t-k}$ on $Y_{t-1}, \ldots, \Delta Y_{t-k+1}$ and save the residuals $R_{kt}$, (3) the estimate of the cointegrating vector(s) is given by the first $r$ canonical variate of $R_{kt}$ with respect to $R_{0t}$. Johansen (1991) demonstrates that this estimator has the same asymptotic distribution as that given in (A5).

The dynamic OLS (DOLS) estimator of Stock and Watson (1993) uses leads and lags of the RHS variables to eliminate the correlation of the RHS variables with the error. It is estimated using OLS on (A11),

$$y_{1t} = \gamma_0 + \beta y_{2t} + \delta(L) \Delta y_{2t} + \varepsilon_t$$

where $\delta(L)$ is a two-sided polynomial in the lag operator. In general, the residuals from such DOLS regression will be serially correlated. This has no effect on the asymptotic distribution of the DOLS estimator, which is identical to that given in (A5), but a consistent estimate of the long-run covariance must be obtained to generate appropriate small sample test statistics. A Parzen kernel estimator is employed to generate the results presented in the paper.

The autoregressive distributed lag (ARDL) estimator of Pesaran and Shin (1995) is based on OLS estimation of equation (A12),
where the estimate of the cointegration vector(s) is given by $\hat{\beta} = \psi/\phi(1)$, where $\phi(1) = 1 - \phi_1 - \phi_2 - ... - \phi_p$. The asymptotic distribution of this estimator is also given by (A5).

The Engle and Yoo (1991) (EY) three-step estimator is based upon the ECM in (A10) (the second step) with $\beta$ set equal to the value generated by estimating (A1) by OLS (the first step). The last step in the procedure is to estimate an auxiliary regression of the weighted residuals, given by $\hat{\epsilon}_it / \hat{\sigma}_i^2$, on the weighted regressors given by $\hat{\alpha}_i y_{it} / \hat{\sigma}_i^2$, where $\hat{\alpha}_i$ is the estimate of the error correction coefficient in equation $i$ and $\hat{\sigma}_i^2$ is the variance of the residuals from equation $i$ in the ECM. The parameter estimate from this auxiliary regression is then used to update the OLS estimate of $\beta$ and the standard error from the auxiliary regression is asymptotically valid such that the large sample distribution of the EY estimator is the same as that given in (A5).

REFERENCES