FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

TWEEKERKENSTRAAT 2
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

Estimating Long-Run Relationships between Observed Integrated Variables by Unobserved Component Methods

Gerdie Everaert*

February 2007

2007/452

* SHERPPA, Ghent University, Gerdie.Everaert@UGent.be, http://www.sherppa.be

D/2007/7012/23

Estimating Long-Run Relationships between Observed Integrated Variables by Unobserved Component Methods *

Gerdie Everaert†

February 26, 2007

Abstract

A regression including integrated variables yields spurious results if the residuals contain a unit root. Although the obtained estimates are unreliable, this does not automatically imply that there is no long-run relation between the included variables as the unit root in the residuals may be induced by omitted or unobserved integrated variables. This paper uses an unobserved component model to estimate the partial long-run relation between observed integrated variables. This provides an alternative to standard cointegration analysis. The proposed methodology is described using a Monte Carlo simulation and applied to investigate purchasing-power parity.

JEL Classification: C15, C32
Keywords: Spurious Regression, Cointegration, Unobserved Component Model, PPP.

1 Introduction

Since the seminal articles of Engle and Granger (1987) and Johansen (1988), cointegration analysis has become a standard econometric tool for estimating relationships involving integrated variables. An important drawback of cointegration analysis is that it requires all variables constituting an equilibrium relation to be included in the analysis. The implication of omitting relevant integrated variables is double. First, any economic theory indicating a long-run relation between a vector of integrated variables will fail to yield a cointegrated regression in the presence of omitted integrated variables. Examples of this problem are legion. Using post-war aggregate U.S. data Rudd and Whelan (2006), for instance, fail to reject the null hypothesis of no cointegration between consumption, labour income and financial wealth. They argue that a non-stationary component in the relation between these variables might be induced by non-stationarity of the expected return on wealth. Second, Engel (2000) shows that standard cointegration tests are seriously biased towards finding cointegration in the presence of omitted integrated variables. Simulating samples of 100 observations, cointegration tests with a nominal size of 5% are found to have true sizes that range from 90% to 99% even though there is a non-stationary component that accounts for 42% of the 100-year forecast variance. Combining these two arguments suggests that cointegration analysis is

*I thank Lorenzo Pozzi for useful comments and suggestions. I acknowledge financial support from the Interuniversity Attraction Poles Program - Belgian Science Policy, contract no. P5/21.

†SHERPPA, Ghent University, Gerdie.Everaert@UGent.be, http://www.sherppa.be

not always conclusive in terms of detecting a long-run relation between integrated variables, both in cases where cointegration is found and in cases where it is rejected.

The interest of this paper is to estimate the long-run relation between integrated variables when possibly not all of the integrated variables constituting an equilibrium relation are included in the analysis. As an alternative to standard cointegration analysis, we propose to use the unobserved component (UC) approach as outlined in, for instance, Harvey (1989) and Durbin and Koopman (2001). In the UC framework, omitted variables can be treated as unobserved components which can be inferred from the observed data using the Kalman filter. This allows for estimation of the long-run relationship between the observed integrated variables using maximum likelihood and for inference on the long-run parameters using a Wald or a likelihood ratio test even if the variables are not cointegrated.

Although UC models have recently become very popular to decompose time series into a number of unobserved components (like trend, cycle, seasonal, ...), they are only rarely used to filter omitted variables from a long-run relation between integrated variables. Examples are Harvey et al. (1986), who add an unobserved component to the employment-output relation to account for the underlying productivity trend, and Sarantis and Stewart (2001) who add an unobserved component to the consumption-income relationship to account for omitted variables such as wealth. A major obstruction to its use is that a general UC model is not necessarily identified. Nelson and Plosser (1982), for instance, show that a difference-stationary process can be decomposed into a permanent and transitory component in an infinite number of ways depending on the assumed correlation between innovations to these two unobserved components. Typically identification is achieved by assuming innovations to be mutually independent. In the context of this paper, this restriction would imply innovations to the observed and to the unobserved/omitted variables to be uncorrelated. Obviously, this is a strong restriction. In a number of recent papers (see e.g. Morley et al., 2003; Morley, 2007) it is shown that an UC model with correlated innovations is identified provided that it has sufficiently rich dynamics. Therefore, this paper allows for correlated innovations and checks under which conditions the model is identified.

The performance of the UC framework relative to standard cointegration analysis, both in terms of estimation and inference, is studied using a Monte Carlo experiment. We consider a simple dynamic bivariate triangular process similar to the one in Inder (1993). Concerning the relation between the 2 observed integrated variables, the experiment nests three important cases: (i) independent random walks, (ii) cointegrated variables and (iii) a long-run relation with an integrated missing component. In all of these cases, the UC framework provides a consistent and asymptotically normally distributed estimate for the long-run relation between the 2 observed variables. In the first and third case, this entails an important improvement over standard estimators, for instance a static ordinary least squares (OLS) estimator, as these yield spurious results in both of these cases. In the second case, the performance of the UC model is similar to that of the static OLS estimator, which is superconsistent in this case. The nice properties of the UC model in the third case are of important practical relevance especially as under the various parameter settings considered in the experiment, standard cointegration tests with a nominal size of 5% have actual sizes up to 89.1%. This indicates that spurious regressions will wrongly be considered to be cointegrating regressions in far too many cases. Combining the results for the three cases suggests that the UC model is a valuable alternative to standard cointegration analysis.

The paper is structured as follows. Section 2 introduces the correlated UC framework as an alternative to standard single equation cointegration analysis in a simple dynamic bivariate process. Section 3 presents a Monte Carlo comparison. In section 4, the proposed UC methodology is applied to testing purchasing-power parity (PPP).

2 Cointegration versus UC analysis

2.1 A simple bivariate process

Consider a dynamic bivariate triangular process

$$\alpha(L)y_{1t} = \mu_1 + \beta(L)y_{2t} + \mu_t, \tag{1}$$

$$\phi(L)\Delta y_{2t} = \mu_2 + \varepsilon_{2t}, \tag{2}$$

$$\mu_t = \nu_t + \varepsilon_{1t}, \tag{3}$$

$$\nu_t = \nu_{t-1} + \eta_{1t}, \tag{4}$$

where y_{1t} and y_{2t} are scalar variables, μ_1 and μ_2 are constants and the error terms ε_{1t}, ε_{2t} and η_{1t} are zero mean Gaussian white noise processes with covariance matrix Ω

$$\Omega = \begin{bmatrix} \sigma^2_{\varepsilon_1} & \sigma_{\varepsilon_1\varepsilon_2} & \sigma_{\varepsilon_1\eta_1} \\ \sigma_{\varepsilon_1\varepsilon_2} & \sigma^2_{\varepsilon_2} & \sigma_{\varepsilon_2\eta_1} \\ \sigma_{\varepsilon_1\eta_1} & \sigma_{\varepsilon_2\eta_1} & \sigma^2_{\eta_1} \end{bmatrix}. \tag{5}$$

The lag polynomials are defined as $\alpha(L) = 1 - \alpha_1 L - \ldots - \alpha_p L^p$, $\beta(L) = \beta_0 + \beta_1 L + \ldots + \beta_q L^q$ and $\phi(L) = 1 - \phi_1 L + \ldots + \phi_r L^r$. They all have roots outside the unit circle. The choice for such a simple model is purely for expositional purposes. A highly similar model is used in, among others, Kremers et al. (1992) and Inder (1993). The main difference is that the error structure of y_{1t} includes an I(1) component when $\sigma^2_{\eta_1} \neq 0$. The non-zero off-diagonal elements of Ω allow for endogeneity of y_{2t} and collinarity between innovations to ν_t and innovations to y_{1t} and y_{2t}.

The long-run relation between y_{2t} and y_{1t} can be obtained from rewriting equation (1) as

$$y_{1t} = \lambda_1 + \lambda_2 y_{2t} + \gamma_1(L)\Delta y_{1t} + \gamma_2(L)\Delta y_{2t} + \omega_t, \tag{6}$$

where $\lambda_2 = \beta(1)/\alpha(1)$ measures the long-run impact of y_{2t} on y_{1t}, with $\alpha(1) = 1 - \alpha_1 - \ldots - \alpha_p$

and $\beta(1) = \beta_0 + \beta_1 + \ldots + \beta_q$, and where

$$\lambda_1 = \frac{\mu}{\alpha(1)}, \quad \gamma_1(L) = \frac{\alpha(1) - \alpha(L)}{\alpha(1)(1-L)}, \quad \gamma_2(L) = \frac{\beta(L) - \beta(1)}{\alpha(1)(1-L)}, \quad \omega_t = \frac{\mu_t}{\alpha(1)}.$$

As equation (2) implies that y_{2t} is I(1), y_{1t} and y_{2t} are said to be cointegrated if $\lambda_2 \neq 0$ and $\sigma^2_{\eta_1} = 0$ such that y_{1t} and y_{2t} are I(1) while μ_t in equation (1) and ω_t in equation (6) are I(0). If $\lambda_2 = 0$ and/or $\sigma^2_{\eta_1} \neq 0$, y_{1t} and y_{2t} are not cointegrated.

2.2 Static cointegration analysis

If λ_2 is the primary parameter of interest, Engle and Granger (1987) suggest to use OLS to estimate a static version of the model in (1)

$$y_{1t} = \lambda_1 + \lambda_2 y_{2t} + \upsilon_t, \tag{7}$$

where from using (6) $\upsilon_t = \gamma_1(L)\Delta y_{1t} + \gamma_2(L)\Delta y_{2t} + \omega_t$. In estimating (7) the dynamic terms in Δy_{1t} and Δy_{2t} and possible endogeneity of y_{2t} can be ignored due to the superconsistency property of the OLS estimator when y_{1t} and y_{2t} are cointegrated (Stock, 1987). When y_{1t} and y_{2t} are not cointegrated, estimating (7) yields spurious results. Testing the null hypothesis that y_{1t} and y_{2t} are not cointegrated amounts to testing whether $\upsilon_t \sim I(1)$ against the alternative that $\upsilon_t \sim I(0)$. Typically this is done using the cointegrating regression Augmented Dickey-Fuller (CRADF) test. As the CRADF test has low power against stationary alternatives with a root close to unity, Shin (1994) and Harris and Inder (1994) suggest to test for the null of cointegration by testing for stationarity of υ_t building on the Kwiatkowski et al. (1992) unit root test. Leybourne and McCabe (1994) test the null of cointegration by directly testing whether the variance of the shocks to the random walk component ν_t is zero, i.e. $H_0 : \sigma^2_{\eta_1} = 0$.

2.3 Dynamic cointegration analysis

Although the OLS estimator for λ_2 in a static equation like (7) is superconsistent, its asymptotic distribution depends on nuisance parameters arising from serial correlation in ν_t and endogeneity of y_{2t} (Phillips and Durlauf, 1986; Phillips and Hansen, 1990). Moreover, omitting dynamic terms and ignoring endogeneity leads to substantial small sample biases (Banerjee et al., 1986) and results in low power of the residual based cointegration tests mentioned above (Kremers et al., 1992; Zivot, 1994; Banerjee et al., 1996).

Dynamic OLS/GLS estimator

The dynamic OLS (DOLS) estimator suggested by Saikkonen (1991) eliminates nuisance terms that stem from endogeneity and serial correlation by augmenting (7) with leads and lags of Δy_{2t}

$$y_{1t} = \lambda_1 + \lambda_2 y_{2t} + \sum_{j=-k_1}^{k_2} b_j \Delta y_{2t-j} + \epsilon_t, \tag{8}$$

where the error term ϵ_t is uncorrelated with the regressors at all leads and lags but is in general serially correlated. Therefore, Stock and Watson (1993) suggest to estimate (8) using feasible generalised least squares (GLS), referred to as dynamic GLS (DGLS). Monte Carlo evidence in Stock and Watson (1993) shows that the DOLS and DGLS estimators yield an important improvement over the static OLS estimator both in terms of estimation and inference.

ECM estimator

Banerjee et al. (1986) suggest to estimate λ_2 by estimating the dynamic equation (1), which is conveniently written in error correction model (ECM) form as

$$\delta_1(L)\Delta y_{1t} = \delta_2'(L)\Delta y_{2t} - \alpha(1)\left(y_{1t-1} - \lambda_1 - \lambda_2' y_{2t-1}\right) + \mu_t, \tag{9}$$

where

$$\delta_1(L) = \frac{\alpha(L) - \alpha(1)L}{(1-L)}, \quad \delta_2(L) = \frac{\beta(L) - \beta(1)L}{(1-L)}.$$

Using Monte Carlo simulations Inder (1993) demonstrates that in a dynamic setting the ECM estimator provides precise estimates and valid inference even in the presence of endogenous variables. To test for the null of no cointegration Kremers et al. (1992), Zivot (1994) and Banerjee et al. (1996) suggest to use a t-test for $\alpha(1) = 0$. Using Monte Carlo simulations, this ECM test for cointegration is found to be more powerful than residual based tests.

2.4 An unobserved component framework

The results of the above mentioned estimation and test procedures are not always conclusive in terms of detecting the long-run relation between y_{1t} and y_{2t}. Upon finding no cointegration one cannot automatically conclude that there is no long-run relation between y_{1t} and y_{2t}, as the unit root in μ_t may be induced by omitted or unobserved I(1) variables. In this case ν_t represents (a linear combination of) omitted or unobserved I(1) variables which ought to be included in (1) for μ_t to be I(0). Although the presence of a unit root in μ_t implies that the results obtained from estimating equation (7), (8) or (9) using OLS/GLS are spurious, the model in equations (1)-(4) can be cast into a linear Gaussian state-space (SS) representation and estimated using maximum likelihood (ML). As such, it is possible to obtain a non-spurious estimate for the long-run relation between y_{1t} and y_{2t} even if they are not cointegrated.

State-space representations

The general SS form[1] is given by

$$y_t = Ax_t + Z\alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0,H), \qquad t = 1,\ldots,T, \tag{10}$$

$$\alpha_{t+1} = S\alpha_t + R\eta_t, \qquad \eta_t \sim N(0,Q), \qquad E\left[\varepsilon_t'\eta_t\right] = G, \tag{11}$$

[1] See e.g. Harvey (1989) or Durbin and Koopman (2001) for an extensive overview of SS models.

where y_t is a $p \times 1$ vector of p observed endogenous variables modelled in the observation equation (10), x_t is a $k \times 1$ vector of k observed exogenous or predetermined variables and α_t is a $m \times 1$ vector of m unobserved states modelled in the state equation (11). The matrices A, Z, S, R, H, Q and G are time-invariant but generally depend on an unknown parameter vector ψ.

If the explanatory variable y_{2t} is weakly exogenous in equation (1), the SS representation of the UC model in equations (1)-(5) is given by the observation equation

$$y_{1t} = \begin{bmatrix} \frac{(1-\alpha(L))}{L} & \beta(L) \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t} \end{bmatrix} + \nu_t + \varepsilon_{1t}, \tag{12}$$

with the state equation being equation (4) and where $H = \sigma^2_{\varepsilon_1}, Q = \sigma^2_{\eta_1}$ and $G = \sigma_{\varepsilon_1 \eta_1}$. Without loss of generality, the constant μ_1 is included in the state variable ν_t. The condition that y_{2t} is weakly exogenous means that no relevant information to the estimation of the unknown parameters in equation (1) is lost by conditioning on y_{2t} (Engle et al., 1983), i.e. it is not necessary to construct a model for y_{2t}. This is the case if $\sigma_{\varepsilon_1 \varepsilon_2} = 0$ and $\sigma_{\varepsilon_2 \eta_1} = 0$.

If the explanatory variable y_{2t} is not weakly exogenous, the reduced form of the UC model in equations (1)-(2) is given by

$$y_{1t} = \mu_1 + \beta_0 \mu_2 + \alpha'(L) y_{1t-1} + \beta'(L) y_{2t-1} + \nu_t + \varepsilon_{1t} + \beta_0 \varepsilon_{2t}, \tag{13}$$

$$y_{2t} = \mu_2 + \phi'(L) y_{2t-1} + \varepsilon_{2t}, \tag{14}$$

where

$$\alpha'(L) = \frac{1-\alpha(L)}{L}; \quad \beta'(L) = \frac{\beta(L) - \beta_0(1-\phi'(L)L)}{L}; \quad \phi'(L) = \frac{1+(L-1)\phi(L)}{L}.$$

This reduced form UC model can be cast in SS form with observation equation

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha'(L) & \beta'(L) \\ 0 & \phi'(L) \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \beta_0 & 1 & \beta_0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{2t} \\ \nu_t \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ 0 \end{bmatrix}, \tag{15}$$

and state equation

$$\begin{bmatrix} \varepsilon_{2t} \\ \nu_t \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{2t-1} \\ \nu_{t-1} \\ \mu_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{2t} \\ \eta_{1t} \end{bmatrix}, \tag{16}$$

where $H = \sigma^2_{\varepsilon_1}, Q = \begin{bmatrix} \sigma^2_{\varepsilon_2} & \sigma_{\varepsilon_2 \eta_1} \\ \sigma_{\varepsilon_2 \eta_1} & \sigma^2_{\eta_1} \end{bmatrix}$ and $G = \begin{bmatrix} \sigma_{\varepsilon_1 \varepsilon_2} & \sigma_{\varepsilon_1 \eta_1} \end{bmatrix}'$. Without loss of generality, the constant μ_1 is again included in the state variable ν_t.

Identification

It is not immediately obvious that the SS models presented above are identified. First consider the single equation model in equations (12) and (4). In order to check identification, first derive the stationary UC autoregressive moving-average (ARMA) representation of the model which is, under the assumption that y_{2t} is weakly exogenous, given by

$$\alpha(L)\Delta y_{1t} - \beta(L)\Delta y_{2t} = \eta_{1t} + \Delta\varepsilon_{1t}. \tag{17}$$

The right-hand side of equation (17) will have nonzero autocovariances through lag 1. Then, by Granger's lemma (Granger and Newbold, 1986) the reduced-form ARMA representation of the model is given by

$$\alpha(L)\Delta y_{1t} - \beta(L)\Delta y_{2t} = \xi(L)e_t \equiv \varsigma_t, \tag{18}$$

where the MA lag polynomial $\xi(L)$ is of order 1. After normalising $\xi_0 = 1$, ξ_1 and σ_e^2, are obtained such that the autocovariances of the right-hand side of (18) match with those of (17). The SS model in (17) is identified if its parameters can be inferred from the reduced form in (18). As the left-hand side of both equations is equal, the $(p+q+1)$ parameters in the AR lag polynomials $\alpha(L)$ and $\beta(L)$ are always identified. The remaining 3 structural coefficients $(\sigma^2_{\varepsilon_1}, \sigma^2_{\eta_1}, \sigma_{\varepsilon_1\eta_1})$ cannot be inferred from the 2 coefficients ξ_1 and σ_e^2 on the right-hand side of (18), though. The relation between the reduced form and the UC parameters is given by

$$\begin{aligned} \gamma_0 &= 2(\sigma^2_{\varepsilon_1} + \sigma_{\varepsilon_1\eta_1}) + \sigma^2_{\eta_1}, \\ \gamma_1 &= -(\sigma^2_{\varepsilon_1} + \sigma_{\varepsilon_1\eta_1}), \\ \gamma_\tau &= 0 \quad \text{for} \quad \tau \geqslant 2, \end{aligned}$$

where the values for the autocovariances $\gamma_\tau = \operatorname{cov}(\varsigma_t, \varsigma_{t-\tau})$ for $\tau = 0, 1$ can be estimated from the data through the reduced-form ARMA representation of the model. This shows that $\sigma^2_{\varepsilon_1}$ and $\sigma_{\varepsilon_1\eta_1}$ are not individually identified but only their sum is. Morley et al. (2003) point out that a separate identification of $\sigma^2_{\varepsilon_1}$ and $\sigma_{\varepsilon_1\eta_1}$ requires ε_{1t} to exhibit rich enough dynamics. In stead of generalising equation (3) by allowing ε_{1t} to be generated from a more general ARMA process we impose $\sigma_{\varepsilon_1\eta_1} = 0$. This identifying restriction affects the estimate for $\sigma^2_{\varepsilon_1}$ but leaves the estimates for the other structural coefficients unaffected. As we are only interested in estimating the long-run relation between y_{1t} and y_{2t}, this restriction is without loss of generality.

Next, consider the stationary UC vector autoregressive moving-average (VARMA) representation of the multiple equation framework in(15)-(16)

$$\begin{aligned} \phi(L)\alpha(L)\Delta y_{1t} - \phi(L)(\beta(L) - \beta_0)\Delta y_{2t} &= \beta_0\mu_2 + \phi(L)\Delta\varepsilon_{1t} + \beta_0\varepsilon_{2t} + \phi(L)\eta_{1t}, \quad (19) \\ \phi(L)\Delta y_{2t} &= \mu_2 + \varepsilon_{2t}. \quad (20) \end{aligned}$$

The right-hand side of (19) will have nonzero autocovariances through lag (r+1). Then, by Granger's lemma (Granger and Newbold, 1986) the reduced-form VARMA representation of the model is given by

$$\phi(L)\alpha(L)\Delta y_{1t} - \phi(L)(\beta(L) - \beta_0)\Delta y_{2t} = \omega_1 + \xi_{11}(L)e_{1t} + \xi_{12}(L)e_{2t} \equiv \varsigma_{1t}, \tag{21}$$

$$\phi(L)\Delta y_{2t} = \omega_2 + e_{2t} \equiv \varsigma_{2t}, \tag{22}$$

where the MA lag polynomials $\xi_{11}(L)$ and $\xi_{12}(L)$ are both of order $(r+1)$ with normalisations $\xi_{11,0} = 1$ and $\xi_{12,0} = 0$. The remaining coefficients in $\xi_{11}(L)$ and $\xi_{12}(L)$ and $\sigma^2_{e_1}, \sigma^2_{e_2}, \sigma_{e_1 e_2}$ are obtained such that the autocovariances of the right-hand sides of (21)-(22) match with those of (19)-(20). The UC model is identified if its parameters can be inferred from the reduced form. First note that the left-hand sides of (19)-(20) and (21)-(22) are equal. This shows that the $(p+q+r)$ parameters in the AR lag polynomials $\alpha(L)$, $(\beta(L)-\beta_0)$ and $\phi(L)$ are always identified. Second, the coefficients μ_2 and β_0 on the right-hand side of (19)-(20) and the 6 variance-covariances in Ω need to be inferred from the $2(r+1)$ MA coefficients, the two intercept terms ω_1 and ω_2 on the right-hand side of (21)-(22) and the 3 variance-covariance parameters $\sigma^2_{e_1}, \sigma^2_{e_2}, \sigma_{e_1 e_2}$. A necessary condition for identification is that $r \geqslant 1$, i.e. the reduced form contains as least as much parameters as the structural UC model. In order to check whether this necessary condition is also sufficient for identification, a one-to-one mapping between the reduced-form and the structural parameters should be established, though. Consider for instance a model with $\mu_2 \neq 0$ and $r = 1$. The parameters in the AR part of this model can be estimated directly from the data while the observable moments of the MA part of the model are the intercept terms ω_1 and ω_2 and the autocovariances

$$\begin{aligned}
\gamma_{0,11} &= 2(1+\phi_1+\phi_1^2)(\sigma^2_{\varepsilon_1} + \sigma_{\varepsilon_1\eta_1}) + \beta_0^2\sigma^2_{\varepsilon_2} + (1+\phi_1^2)\sigma^2_{\eta_1} + 2\beta_0\sigma_{\varepsilon_1\varepsilon_2} + 2\beta_0\sigma_{\varepsilon_2\eta_1}, \\
\gamma_{0,12} &= \gamma_{0,21} = \beta_0\sigma^2_{\varepsilon_2} + \sigma_{\varepsilon_1\varepsilon_2} + \sigma_{\varepsilon_2\eta_1}, \\
\gamma_{0,22} &= \sigma^2_{\varepsilon_2}, \\
\gamma_{1,11} &= -(1+\phi_1)^2(\sigma^2_{\varepsilon_1} + \sigma_{\varepsilon_1\eta_1}) - \phi_1\sigma^2_{\eta_1} - \beta_0(1+\phi_1)\sigma_{\varepsilon_1\varepsilon_2} - \beta_0\phi_1\sigma_{\varepsilon_2\eta_1}, \\
\gamma_{1,12} &= \beta_1\sigma^2_{\varepsilon_2} - (1+\phi_1)\sigma_{\varepsilon_1\varepsilon_2} - \phi_1\sigma_{\varepsilon_2\eta_1}, \\
\gamma_{2,11} &= \phi_1(\sigma^2_{\varepsilon_1} + \sigma_{\varepsilon_1\eta_1}) + \beta_0\phi_1\sigma_{\varepsilon_1\varepsilon_2}, \\
\gamma_{2,12} &= \phi_1\sigma_{\varepsilon_1\varepsilon_2}, \\
\gamma_{\tau,2j} &= 0 \quad \text{for} \quad \tau \geqslant 1 \quad \text{and} \quad \gamma_{\tau,1j} = 0 \quad \text{for} \quad \tau \geqslant 2,
\end{aligned} \tag{23}$$

where $\gamma_{\tau,ij} = \operatorname{cov}(\varsigma_{it}, \varsigma_{jt-\tau})$ with $i,j = 1,2$. The 7 non-zero autocovariances together with the intercepts terms ω_1 and ω_2 provide 9 pieces of information from which in principle the remaining 8 UC coefficients $(\mu_2, \beta_0, \sigma^2_{\varepsilon_1}, \sigma^2_{\varepsilon_2}, \sigma^2_{\eta_1}, \sigma_{\varepsilon_1\varepsilon_2}, \sigma_{\varepsilon_1\eta_1}$ and $\sigma_{\varepsilon_2\eta_1})$ can be inferred. Inspection of the autocovariances reveals that $\sigma^2_{\varepsilon_1}$ and $\sigma_{\varepsilon_1\eta_1}$ are again not individually identified, though. Without loss of generality the model is again identified by setting $\sigma_{\varepsilon_1\eta_1} = 0$. Imposing this identifying restriction implies that the remaining structural coefficients are identified either if $\mu_2 \neq 0$ or if

$r \geqslant 1$. In the former, the non-zero drift term μ_2 implies that both μ_2 and β_0 can be identified from the intercept terms ω_1 and ω_2 while, even if $r = 0$, the remaining 5 parameters in Ω can be identified from the 5 non-zero autocovariances in (23). In the latter, the zero drift term μ_2 implies that β_0 should also be identified from the autocovariances in (23). This requires at least 6 non-zero autocovariances, i.e. $r \geqslant 1$.

Intuitively, identification of the long-run impact of y_{2t} on y_{1t} requires (i) y_{2t} to be weakly exogenous such that it is orthogonal to the unobserved component ν_t or (ii) the dynamic behaviour of y_{2t} to differ from that of the unobserved component ν_t once innovations to these components are allowed to be cross-correlated.

ML estimation and inference

Provided that the UC model is identified, the likelihood for the linear Gaussian SS model in (10)-(11) can be calculated by a routine application of the Kalman filter and maximised with respect to the unknown parameter vector ψ using an iterative numerical procedure. Pagan (1980) shows that the resulting ML estimator $\widehat{\psi}$ is consistent and asymptotically normally distributed provided that (i) ψ is an interior point of the parameter space and (ii) the transition matrix S does not contain unknown polynomials with roots inside the unit circle. As λ_2 is an interior point and S does not contain unknown parameters, inference on the long-run relation between y_{1t} and y_{2t} is possible using a standard Wald or likelihood ratio (LR) test. Note that testing whether y_{1t} and y_{2t} are cointegrated implies testing whether $\sigma^2_{\eta_1} = 0$, which is on the boundary of the parameter space. In this case, the distribution of a Wald or LR test is non-standard but can in principle be obtained using Monte Carlo simulation. As the interest of this paper is to estimate the long-run relationship between y_{1t} and y_{2t} irrespectively of whether they are cointegrated or not, we don't want to test this hypothesis, though.

3 Monte Carlo experiment

In this section, the performance of 5 alternative estimators for the long-run relationship between integrated variables is compared using a Monte Carlo simulation.

3.1 Design

As before, we consider the model in (1)-(5), setting $p = q = r = 1$. Data are generated under a variety of settings for the parameter vector $\psi = \left(\alpha_1, \beta_0, \lambda_2, \phi_1, \mu_2, \sigma^2_{\varepsilon_1}, \sigma^2_{\varepsilon_2}, \sigma^2_{\eta_1}, \sigma_{\varepsilon_1\varepsilon_2}, \sigma_{\varepsilon_1\eta_1}, \sigma_{\varepsilon_2\eta_1}\right)$. For each parameter setting we generate 2500 series of length T. First, we draw a sample of T observations for ε_{1t}, ε_{2t} and η_{1t} out of a multivariate normal distribution with variance-covariance matrix Ω. Next, we generate y_{2t} and ν_t by first drawing initial values for Δy_{2t} and $\Delta\nu_t$ from their stationary distribution and then calculating initial values for y_{2t} and ν_t from a diffuse initialisation on the level

$$
\begin{aligned}
y_{2,1} &= y_{2,0} + \Delta y_{2,1}, \qquad \text{where} \quad \Delta y_{2,1} = (1-\phi_1^2)^{-1/2}\varepsilon_{2,1} \quad \text{and} \quad y_{2,0} \sim N(0,\infty), && (24)\\
\nu_1 &= \nu_0 + \Delta\nu_1, \qquad \text{where} \quad \Delta\nu_1 = \eta_1 \quad \text{and} \quad \nu_0 \sim N(0,\infty). && (25)
\end{aligned}
$$

Data for $t = 2,\ldots,T$ are then generated using (2) and (4) respectively. In order to generate data for y_{1t}, first define the latent variable ζ_t to be the deviation of y_{1t} from its equilibrium level implied by the levels of μ_1, y_{2t} and ν_t

$$
\zeta_t = y_{1t} - \frac{1}{(1-\alpha_1)}\left(\mu_1 + (\beta_0+\beta_1)y_{2t} + \nu_t\right). \tag{26}
$$

Using (1), ζ_t can be written as

$$
\begin{aligned}
\zeta_t &= \alpha_1\zeta_{t-1} + \xi_t, && (27)\\
\text{where} \quad \xi_t &= \varepsilon_{1t} - \frac{\alpha_1\beta_0+\beta_1}{(1-\alpha_1)}\Delta y_{2t} - \frac{\alpha_1}{(1-\alpha_1)}\eta_{1t}.
\end{aligned}
$$

Setting $\Delta y_{2t} = (1-\phi_1^2)^{-1/2}\varepsilon_{2,1}$ from (24), ζ_t can be initialised from (27) as

$$
\zeta_1 = \xi_1(1-\alpha_1^2)^{-1/2}. \tag{28}
$$

After generating ζ_t for $t = 2,\ldots,T$ using (27) and given data on y_{2t} and ν_t, y_{1t} can then be calculated from (26).

We consider five alternative estimators: (i) the static OLS estimator in equation (6), (ii) the DGLS estimator in equation (8), (iii) the ECM estimator in equation (9), (iv) the univariate UC model (UC_U) in equations (12) and (4) and (v) the multivariate UC model (UC_M) in equations (15)-(16). The DGLS estimator is implemented by setting as a rule of thumb $k_1 = k_2 = \text{int}\left[T^{1/3}\right]$ (see Saikkonen, 1991). Selecting an alternative number of leads and lags did not have a noteworthy impact on the main results. The ECM, UC_U and UC_M estimators are implemented by setting, if relevant, $p = q = r = 1$.

The design of the Monte Carlo experiment nests three special cases. First (Case I), when $\lambda_2 = 0$ and $\sigma^2_{\eta_1} \neq 0$, there is no long-run relationship between the random walk processes y_{1t} and y_{2t}. Second (Case II), when $\lambda_2 \neq 0$ and $\sigma^2_{\eta_1} = 0$, y_{1t} and y_{2t} are cointegrated. Third (case III), when $\lambda_2 \neq 0$ and $\sigma^2_{\eta_1} \neq 0$, there is a long-run relationship between y_{1t} and y_{2t} but they are not cointegrated. For each of these cases we analyse the performance of the 5 alternative estimators in terms of estimation and inference. Next to the median bias and the root mean squared error (Rmse), we report the percentiles of the empirical distribution of λ_2.[2] We also check the performance of standard cointegration tests, i.e. a CRADF test on the residuals from the OLS

[2] We report the median bias instead of the mean bias and percentiles along with the Rmse as the dynamic estimators ECM, UC_U and UC_M for $\lambda_2 = (\beta_0+\beta_1)/(1-\alpha_1)$ do not have finite moments, especially when α_1 is close to 1. It should be noted that the percentiles are not necessarily finite either but they should be less vulnerable to large outliers in the distribution of λ_2.

and DGLS regressions and a t-test on the error-correction term in the ECM regression. Critical values for these cointegration test statistics are obtained by simulating their distributions under the null hypothesis of no cointegration, i.e by first drawing y_{1t} and y_{2t} from a standardised multivariate random walk as $\Delta y_t \sim IN(0, I_k)$ and next calculating the considered cointegration test statistics. The obtained critical values for the OLS and the ECM-based cointegration tests coincide with those reported in Mackinnon (1996) and Ericsson and MacKinnon (2002) respectively. The reason why critical values are simulated is that no such values are available for a cointegration test based on DGLS residuals.

3.2 Results

Results for Case I are reported in Table 1. It is a well-known fact that the OLS estimator yields spurious results in this case. This can be seen from the fact that, although the median bias is negligible, the null hypothesis that $\widehat{\lambda}_2 = 0$ is rejected (with a nominal size of 5%) in 74.9% of the cases for $T = 50$. This problem even aggravates as T grows large. The reason for this huge size bias is that the OLS estimator does not converge in probability and the t- statistic does not have a well-defined asymptotic distribution. Note that the DGLS and the ECM estimators face the same problem, although somewhat less pronounced. The UC_U and UC_M estimators are both unbiased and more or less correctly sized. This shows that these estimators are not vulnerable to the spurious regression problem. In terms of estimation, the UC_U estimator is much more precise than the UC_M estimator, which is overparameterised in this case.

Results for the case where y_{1t} and y_{2t} are cointegrated (Case II) are reported in Table 2. The performance of the estimators is explored both in a static setting (see Panels (a) and (b) where $\alpha_1 = 0$ and $\beta_0 = \lambda_2 = 1$) and in a dynamic setting (see Panels (c) and (d) where $\alpha_1 = 0.5$, $\beta_0 = 0.2$). In both settings, we discriminate between a case where y_{2t} is exogenous ($\sigma_{\varepsilon_1\varepsilon_2} = 0$) and a case where y_{2t} is endogenous ($\sigma_{\varepsilon_1\varepsilon_2} = 0.5$). In terms of estimation, the OLS estimator is superconsistent, i.e. it converges to the true population value at a rate faster than in normal asymptotics even if dynamic terms and endogeneity issues are ignored. The results in Panels (b), (c) and (d) show that even in a small sample ($T = 50$) the bias induced by endogeneity and ignoring dynamic terms is small and, as implied by the superconsistency property, disappears quickly as T increases. In terms of inference, the OLS estimator has a correct size in the static setting without endogeneity. The effects of endogeneity on the distribution of the OLS estimator are minimal, while in the dynamic setting the size is unacceptably high. Turning to the DGLS and ECM estimators, only the latter yields an improvement on the OLS estimator in terms of estimation. Despite its overparameterisation in the static setting, the ECM estimator has a bias and Rmse which are highly similar to those of the OLS estimator. In the dynamic setting, the ECM estimator has a notable lower bias and Rmse. In terms of inference, both the DGSL and

Table 1: Monte Carlo results case I: $\lambda_2 = 0$, $\sigma^2_{\eta_1} = 1$
($\alpha_1 = 0$, $\beta_0 = 0$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\varepsilon_1} = 0$, $\sigma^2_{\varepsilon_2} = 1$, $\sigma_{\varepsilon_1\varepsilon_2} = 0$, $\sigma_{\varepsilon_1\eta_1} = 0$, $\sigma_{\varepsilon_2\eta_1} = 0$)

T	Estimator	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$	$P(CI)$
50	OLS	-0.001	0.278	-0.581	0.568	0.749	0.040
	DGLS	0.005	0.362	-0.750	0.769	0.381	0.076
	ECM	0.001	> 10	-1.307	1.518	0.291	0.062
	UC_U	-0.003	0.138	-0.281	0.271	0.062	-
	UC_M	-0.003	0.346	-0.615	0.548	0.045	-
100	OLS	-0.006	0.216	-0.436	0.389	0.858	0.048
	DGLS	-0.001	0.219	-0.443	0.418	0.395	0.084
	ECM	-0.004	4.531	-0.815	0.741	0.359	0.071
	UC_U	0.002	0.095	-0.187	0.186	0.062	-
	UC_M	-0.004	0.263	-0.404	0.385	0.045	-
250	OLS	-0.002	0.143	-0.273	0.290	0.911	0.050
	DGLS	0.004	0.129	-0.246	0.266	0.454	0.098
	ECM	0.001	1.657	-0.399	0.437	0.418	0.068
	UC_U	0.002	0.058	-0.113	0.115	0.055	-
	UC_M	0.002	0.125	-0.246	0.244	0.054	-
500	OLS	-0.002	0.101	-0.199	0.199	0.941	0.062
	DGLS	-0.001	0.095	-0.195	0.190	0.504	0.120
	ECM	-0.003	2.755	-0.252	0.274	0.434	0.078
	UC_U	0.000	0.041	-0.078	0.082	0.052	-
	UC_M	-0.001	0.090	-0.181	0.176	0.076	-

Notes: bias($\widehat{\lambda}_2$) and Rmse($\widehat{\lambda}_2$) are the Monte Carlo median bias and root mean squared error of $\widehat{\lambda}_2$, respectively. $p_{2,5\%}$ and $p_{97,5\%}$ are the 2,5% and 97,5% percentiles, respectively, of the Monte Carlo distribution of $\widehat{\lambda}_2$. $P(\widehat{\lambda}_2)$ is the rejection rate at the 5% level of the test statistic testing $\widehat{\lambda}_2 = \lambda_2$. For the OLS, DGLS and ECM estimators this is a standard t-test. For the UC_U and UC_M this is a Wald test. $P(CI)$ is the rejection rate at the 5% level for a test for the null hypothesis of no cointegration. For the OLS and DGLS estimators, this is a DF test on the estimated residuals. For the ECM estimator this is a t-test on the error-correction term. The simulated 5% critical values for the OLS, DGLS and ECM estimators respectively are (i) -3.47, -4.34 and -3.30 for $T = 50$, (ii) -3.40, -3.64 and -3.25 for $T = 100$, (iii) -3.36, -3.32 and -3.24 for $T = 250$ and (iv) -3.34, -3.24 and -3.23 for $T = 500$.

ECM estimator have a more or less correct size in all cases. Only in the case of endogeneity, the size of the ECM estimator is somewhat different from 5%. Consistent with the finding of Inder (1993), the relative performance of the ECM estimator suggests that it is better to overspecify the dynamics of the model than to underspecify. Interestingly, the performance of the UC_U and the UC_M estimators are almost identical to that of the ECM estimator. This shows that adding an unobserved I(1) component to the error structure of the ECM does not impair performance in the special case of cointegration between y_{1y} and y_{2t}.

Results for the case where y_{1t} and y_{2t} have a long-run relationship but are not cointegrated (Case III) are reported in Table 3. The OLS, DGLS and ECM estimators all yield spurious results. As in Case I, the median bias is negligible but the size of testing the hypothesis that $\widehat{\lambda}_2 = \lambda_2 = 1$ is unacceptably high. Again, the problem aggravates as T grows large. Important to note is that cointegration tests also have a large size bias, especially in the static case and for the ECM estimator in the dynamic case. This is consistent with the conclusion of Engel (2000) that standard cointegration tests are strongly biased towards rejecting the null hypothesis of no cointegration in the presence of an omitted permanent component. The UC_U estimator yields an improvement both in terms of estimation precision and inference. However, its size only converges slowly to 5% as T grows large, especially for $\sigma_{\varepsilon_1\varepsilon_2} \neq 0$. The UC_M estimator is less precise compared to the UC_U estimator, i.e. its Rmse is about the same as for the OLS, DGLS and ECM estimators, but has better size properties compared to the UC_U estimator. Note that in a small sample the size is also too large, though.

Tables 4-5 further explore the performance of the UC_U and UC_M estimators in Case III for a number of alternative values for the parameters α_1, β_0, ϕ_1, μ_2 and $\sigma^2_{\eta_1}$. Table 4 reports results for a setting where innovations to the dependent variable y_{1t} are correlated with innovations to the unobserved component ν_t (i.e. $\sigma_{\varepsilon_1\eta_1} = 0.5$). Table 5 reports results for a setting where innovations to the explanatory variable y_{2t} are correlated with innovations to the unobserved component ν_t (i.e. $\sigma_{\varepsilon_2\eta_1} = 0.5$). As the OLS, DGLS and ECM estimators all yield spurious results for $\sigma^2_{\eta_1} \neq 0$, comparable to the results in Table 3, they were excluded from the tables to economise on space (results available on request). Comparing Panel (b) from Table 4 with Panel (c) from Table 3 shows that, although $\sigma_{\varepsilon_1\eta_1}$ is set to zero as an identifying restriction for both the UC_U and UC_M estimator, the performance of both estimators is not affected, as indicated in section 2.4, by simulating data with $\sigma_{\varepsilon_1\eta_1} = 0.5$. The results in Table 4 further show that the estimation precision deteriorates as (i) α_1 increases (compare Panels (a), (b) and (c)), (ii) the dynamics in the explanatory variable y_{2t} are less rich (see Panel (e) for less inertia in y_{2t}, i.e. lower value of ϕ_1, and Panel (f) for zero drift, i.e. $\mu_2 = 0$) and (iii) as the variance of the unobserved component ν_t increases (compare Panels (b), (g) and (h)). Given its overparameterisation, UC_M is again less precise than UC_U. Inference is similar over all parameter settings included in Table 4. Turning to Table 5, it is immediately clear that the UC_U estimator performs poorly, both in terms of estimation and inference when $\sigma_{\varepsilon_2\eta_1} \neq 0$. The performance of the UC_M estimator deteriorates a little bit, especially for small T, but is still satisfactory for larger T.

Summarising, the Monte Carlo simulation shows that standard estimators and cointegration tests can only deal with cases where y_{1t} and y_{2t} are either independent random walks or cointegrated. In the case of a long-run relation with an integrated missing component, these estimators yield spurious results with standard cointegration tests indicating these results to be a cointegration regression in far too many cases. The UC approach yields results similar to those of standard estimators in the case of cointegration and does not yield spurious results in the cases where y_{1t} and y_{2t} are independent random walks or related in the long-run with an unobserved component. This suggests that it is better to overspecify than to underspecify the error structure of the model.

Table 2: Monte Carlo results for case II: $\lambda_2 = 1$, $\sigma^2_{\eta_1} = 0$
($\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\varepsilon_1} = 1$, $\sigma^2_{\varepsilon_2} = 1$, $\sigma_{\varepsilon_1\eta_1} = 0$, $\sigma_{\varepsilon_2\eta_1} = 0$)

T	Estimator	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$	$P(CI)$	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$	$P(CI)$
		(a) $\alpha_1 = 0$, $\beta_0 = 1$, $\sigma_{\varepsilon_1\varepsilon_2} = 0$						(b) $\alpha_1 = 0$, $\beta_0 = 1$, $\sigma_{\varepsilon_1\varepsilon_2} = 0.5$					
50	OLS	−0.001	0.025	0.946	1.049	0.049	1.000	−0.001	0.023	0.948	1.046	0.039	1.000
	DGLS	−0.001	0.053	0.894	1.097	0.067	0.996	−0.001	0.043	0.908	1.081	0.074	0.994
	ECM	−0.001	0.026	0.944	1.053	0.064	1.000	−0.005	0.025	0.934	1.033	0.090	1.000
	UC_U	−0.001	0.027	0.941	1.053	0.094	-	−0.005	0.029	0.921	1.033	0.112	-
	UC_M	−0.001	0.029	0.937	1.055	0.060	-	−0.001	0.025	0.948	1.049	0.064	-
100	OLS	0.000	0.009	0.983	1.017	0.042	1.000	0.000	0.008	0.984	1.017	0.039	1.000
	DGLS	0.000	0.014	0.975	1.029	0.057	1.000	0.000	0.012	0.980	1.023	0.054	1.000
	ECM	0.000	0.009	0.982	1.018	0.052	1.000	−0.001	0.008	0.980	1.012	0.086	1.000
	UC_U	0.000	0.010	0.982	1.018	0.066	-	−0.001	0.011	0.972	1.012	0.098	-
	UC_M	0.000	0.010	0.981	1.019	0.042	-	0.000	0.008	0.984	1.016	0.048	-
250	OLS	0.000	0.002	0.996	1.004	0.042	1.000	0.000	0.002	0.996	1.004	0.043	1.000
	DGLS	0.000	0.002	0.996	1.004	0.043	1.000	0.000	0.002	0.996	1.004	0.044	1.000
	ECM	0.000	0.002	0.996	1.004	0.045	1.000	0.000	0.002	0.996	1.003	0.061	1.000
	UC_U	0.000	0.002	0.996	1.004	0.052	-	0.000	0.003	0.994	1.003	0.072	-
	UC_M	0.000	0.002	0.996	1.004	0.031	-	0.000	0.002	0.996	1.004	0.028	-
		(c) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\sigma_{\varepsilon_1\varepsilon_2} = 0$						(d) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\sigma_{\varepsilon_1\varepsilon_2} = 0.5$					
50	OLS	−0.036	0.093	0.761	1.080	0.412	0.463	−0.028	0.075	0.797	1.056	0.410	0.516
	DGLS	−0.004	0.097	0.803	1.177	0.110	0.490	−0.003	0.079	0.823	1.146	0.122	0.496
	ECM	−0.004	0.054	0.873	1.099	0.075	0.998	−0.008	0.052	0.862	1.070	0.060	0.942
	UC_U	−0.004	0.055	0.870	1.098	0.100	-	−0.009	0.057	0.845	1.070	0.085	-
	UC_M	−0.002	0.059	0.865	1.110	0.075	-	−0.003	0.049	0.888	1.096	0.076	-
100	OLS	−0.010	0.033	0.910	1.035	0.364	0.988	−0.008	0.026	0.930	1.025	0.370	0.992
	DGLS	0.000	0.026	0.951	1.053	0.076	0.999	−0.002	0.023	0.962	1.044	0.078	0.998
	ECM	0.000	0.018	0.963	1.034	0.056	1.000	−0.002	0.018	0.962	1.026	0.037	1.000
	UC_U	0.000	0.020	0.962	1.037	0.079	-	−0.002	0.018	0.961	1.026	0.061	-
	UC_M	0.000	0.023	0.961	1.037	0.050	-	0.000	0.017	0.967	1.032	0.048	-
250	OLS	−0.002	0.007	0.983	1.011	0.342	1.000	−0.002	0.006	0.986	1.008	0.368	1.000
	DGLS	0.000	0.004	0.992	1.008	0.048	1.000	0.000	0.004	0.993	1.008	0.058	1.000
	ECM	0.000	0.004	0.993	1.007	0.049	1.000	0.000	0.004	0.993	1.006	0.019	1.000
	UC_U	0.000	0.004	0.993	1.007	0.064	-	0.000	0.004	0.992	1.006	0.029	-
	UC_M	0.000	0.004	0.992	1.008	0.037	-	0.000	0.004	0.993	1.007	0.027	-

Notes: See Table 1.

Table 3: Monte Carlo results for case III: $\lambda_2 = 1$, $\sigma^2_{\eta_1} = 0.5$
($\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\varepsilon_1} = 1$, $\sigma^2_{\varepsilon_2} = 1$, $\sigma_{\varepsilon_1\eta_1} = 0$, $\sigma_{\varepsilon_2\eta_1} = 0$)

T	Estimator	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$	$P(CI)$	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$	$P(CI)$
		(a) $\alpha_1 = 0$, $\beta_0 = 1$, $\sigma_{\varepsilon_1\varepsilon_2} = 0$						(b) $\alpha_1 = 0$, $\beta_0 = 1$, $\sigma_{\varepsilon_1\varepsilon_2} = 0.5$					
50	OLS	−0.002	0.198	0.591	1.405	0.692	0.676	−0.002	0.200	0.587	1.400	0.700	0.670
	DGLS	0.003	0.267	0.468	1.563	0.452	0.498	0.008	0.268	0.449	1.538	0.447	0.423
	ECM	−0.001	0.225	0.526	1.416	0.453	0.702	−0.015	0.230	0.491	1.370	0.486	0.756
	UC_U	−0.005	0.129	0.746	1.262	0.171	-	−0.043	0.126	0.705	1.173	0.137	-
	UC_M	−0.004	0.242	0.560	1.431	0.182	-	0.001	0.228	0.580	1.418	0.165	-
100	OLS	−0.002	0.152	0.689	1.274	0.838	0.730	−0.002	0.154	0.678	1.291	0.842	0.734
	DGLS	−0.003	0.166	0.656	1.309	0.576	0.696	0.000	0.164	0.655	1.316	0.559	0.593
	ECM	−0.004	0.158	0.665	1.292	0.586	0.775	−0.011	0.163	0.643	1.288	0.613	0.816
	UC_U	−0.001	0.082	0.831	1.168	0.107	-	−0.044	0.089	0.792	1.099	0.104	-
	UC_M	−0.003	0.163	0.707	1.281	0.092	-	−0.003	0.160	0.699	1.283	0.084	-
250	OLS	−0.001	0.101	0.807	1.206	0.906	0.845	−0.001	0.102	0.793	1.215	0.909	0.840
	DGLS	−0.001	0.101	0.803	1.205	0.666	0.840	−0.001	0.103	0.798	1.221	0.639	0.739
	ECM	−0.002	0.103	0.807	1.212	0.671	0.861	−0.004	0.104	0.792	1.205	0.683	0.891
	UC_U	0.001	0.049	0.904	1.094	0.062	-	−0.043	0.064	0.857	1.047	0.137	-
	UC_M	0.001	0.089	0.826	1.176	0.048	-	0.001	0.090	0.815	1.180	0.042	-
		(c) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\sigma_{\varepsilon_1\varepsilon_2} = 0$						(d) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\sigma_{\varepsilon_1\varepsilon_2} = 0.5$					
50	OLS	−0.038	0.399	0.089	1.744	0.739	0.043	−0.038	0.401	0.107	1.756	0.749	0.028
	DGLS	−0.005	0.507	−0.091	2.023	0.372	0.074	−0.003	0.490	−0.053	1.962	0.369	0.080
	ECM	0.076	2.249	−0.021	3.128	0.389	0.330	0.056	7.607	−0.876	3.334	0.297	0.084
	UC_U	0.007	0.353	0.493	1.739	0.192	-	−0.041	> 10	0.349	1.964	0.279	-
	UC_M	0.012	0.499	0.062	1.989	0.182	-	0.015	0.918	0.135	2.006	0.167	-
100	OLS	−0.015	0.307	0.347	1.533	0.857	0.068	−0.013	0.308	0.344	1.558	0.855	0.028
	DGLS	−0.007	0.305	0.342	1.582	0.394	0.086	−0.007	0.300	0.359	1.578	0.388	0.087
	ECM	0.042	1.518	0.289	2.033	0.484	0.432	0.032	6.681	−0.171	2.315	0.368	0.130
	UC_U	0.006	0.196	0.636	1.418	0.124	-	−0.003	> 10	0.530	1.535	0.234	-
	UC_M	0.001	> 10	0.409	1.599	0.122	-	0.002	0.320	0.372	1.621	0.098	-
250	OLS	−0.004	0.203	0.604	1.402	0.913	0.111	−0.004	0.204	0.587	1.425	0.915	0.034
	DGLS	0.003	0.182	0.655	1.369	0.449	0.098	0.004	0.182	0.656	1.368	0.432	0.084
	ECM	0.018	0.351	0.542	1.550	0.571	0.516	0.016	2.898	0.423	1.677	0.437	0.162
	UC_U	0.003	0.111	0.778	1.224	0.075	-	0.022	0.188	0.740	1.306	0.125	-
	UC_M	0.005	0.181	0.654	1.361	0.054	-	0.004	0.182	0.630	1.366	0.050	-

Notes: See Table 1.

Table 4: Monte Carlo results for case III: $\lambda_2 = 1$, $\sigma^2_{\varepsilon_1} = 1$, $\sigma^2_{\varepsilon_2} = 1$, $\sigma_{\varepsilon_1\varepsilon_2} = 0$, $\sigma_{\varepsilon_1\eta_1} = 0.5$, $\sigma_{\varepsilon_2\eta_1} = 0$

T	Estimator	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$
		(a) $\alpha_1 = 0.25$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$					(b) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$				
50	UC_U	−0.019	0.198	0.618	1.413	0.196	−0.010	0.412	0.449	1.730	0.219
	UC_M	−0.001	0.323	0.384	1.612	0.212	0.011	0.495	0.021	1.975	0.203
100	UC_U	0.001	0.120	0.765	1.242	0.117	0.002	0.203	0.629	1.429	0.137
	UC_M	−0.002	0.196	0.601	1.375	0.129	0.007	0.322	0.372	1.614	0.139
250	UC_U	0.000	0.072	0.854	1.144	0.073	0.000	0.116	0.768	1.239	0.079
	UC_M	0.004	0.123	0.747	1.229	0.065	0.010	0.187	0.621	1.346	0.071
		(c) $\alpha_1 = 0.75$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$					(d) $\alpha_1 = 0.5$, $\beta_0 = 0$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$				
50	UC_U	−0.049	> 10	−0.111	3.132	0.213	−0.008	0.325	0.436	1.705	0.218
	UC_M	0.027	1.583	−1.133	3.497	0.171	0.013	0.497	0.061	1.994	0.205
100	UC_U	−0.018	1.527	0.282	1.937	0.137	0.003	0.198	0.633	1.403	0.124
	UC_M	0.013	0.879	−0.231	2.311	0.081	0.007	0.320	0.391	1.617	0.128
250	UC_U	0.000	0.232	0.560	1.477	0.080	0.000	0.115	0.768	1.239	0.073
	UC_M	0.018	0.370	0.234	1.695	0.054	0.004	0.123	0.746	1.346	0.063
		(e) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\phi_1 = 0.25$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$					(f) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0$, $\sigma^2_{\eta_1} = 0.5$				
50	UC_U	−0.014	> 10	0.276	2.042	0.218	−0.006	0.495	0.362	1.923	0.188
	UC_M	0.021	0.766	−0.391	2.463	0.214	0.062	0.876	−0.416	2.503	0.162
100	UC_U	0.000	0.290	0.514	1.627	0.146	−0.003	0.258	0.578	1.582	0.124
	UC_M	0.009	0.502	0.070	1.915	0.139	0.032	1.224	−0.252	2.272	0.099
250	UC_U	0.000	0.166	0.684	1.347	0.089	0.002	0.147	0.721	1.307	0.084
	UC_M	0.012	0.281	0.428	1.527	0.072	0.044	0.575	−0.080	2.168	0.060
		(g) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.25$					(h) $\alpha_1 = 0.5$, $\beta_0 = 0$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.75$				
50	UC_U	−0.009	0.240	0.561	1.513	0.233	−0.002	1.006	0.379	1.894	0.196
	UC_M	0.003	0.348	0.350	1.689	0.218	0.015	1.935	−0.176	2.226	0.187
100	UC_U	0.000	0.147	0.721	1.306	0.142	0.006	0.243	0.571	1.506	0.130
	UC_M	0.003	0.224	0.568	1.435	0.164	0.011	0.428	0.257	1.759	0.117
250	UC_U	0.000	0.084	0.830	1.169	0.078	0.002	0.139	0.724	1.288	0.082
	UC_M	0.005	0.132	0.733	1.244	0.071	0.011	0.229	0.529	1.427	0.064

Notes: See Table 1.

Table 5: Monte Carlo results for case III: $\lambda_2 = 1$, $\sigma^2_{\varepsilon_1} = 1$, $\sigma^2_{\varepsilon_2} = 1$ $\sigma_{\varepsilon_1\varepsilon_2} = 0$, $\sigma_{\varepsilon_1\eta_1} = 0$, $\sigma_{\varepsilon_2\eta_1} = 0.5$

T	Estimator	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$	bias($\widehat{\lambda}_2$)	Rmse ($\widehat{\lambda}_2$)	$p_{2,5\%}$	$p_{97,5\%}$	$P(\widehat{\lambda}_2)$
		(a) $\alpha_1 = 0.25$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$					(b) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$				
50	UC_U	0.160	0.232	0.849	1.496	0.321	0.230	0.385	0.770	1.826	0.296
	UC_M	0.062	0.294	0.372	1.498	0.220	0.104	3.193	0.041	1.901	0.205
100	UC_U	0.161	0.191	0.959	1.357	0.427	0.231	0.281	0.914	1.548	0.386
	UC_M	0.031	0.213	0.539	1.315	0.136	0.053	0.343	0.315	1.495	0.133
250	UC_U	0.163	0.173	1.044	1.282	0.779	0.232	0.250	1.045	1.417	0.708
	UC_M	0.015	0.120	0.740	1.205	0.072	0.026	0.181	0.614	1.312	0.070
		(c) $\alpha_1 = 0.75$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$					(d) $\alpha_1 = 0.5$, $\beta_0 = 0$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$				
50	UC_U	0.431	> 10	0.573	3.206	0.247	0.224	0.372	0.762	1.807	0.296
	UC_M	0.213	2.966	−0.764	3.356	0.173	0.106	0.494	0.067	1.883	0.206
100	UC_U	0.440	> 10	0.819	2.214	0.309	0.230	0.281	0.907	1.550	0.370
	UC_M	0.109	0.737	−0.323	2.048	0.099	0.054	0.346	0.339	1.498	0.130
250	UC_U	0.447	0.496	1.082	1.851	0.671	0.231	0.249	1.041	1.417	0.696
	UC_M	0.051	0.363	0.230	1.619	0.065	0.025	0.180	0.619	1.311	0.071
		(e) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\phi_1 = 0.25$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.5$					(f) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0$, $\sigma^2_{\eta_1} = 0.5$				
50	UC_U	0.359	0.584	0.762	2.193	0.315	0.356	3.153	0.796	2.212	0.369
	UC_M	0.166	0.698	−0.278	2.371	0.215	0.319	1.475	−0.052	2.425	0.260
100	UC_U	0.366	0.205	0.931	1.816	0.455	0.360	0.435	0.973	1.851	0.518
	UC_M	0.087	0.407	0.040	1.756	0.139	0.301	0.664	−0.140	2.158	0.190
250	UC_U	0.366	0.390	1.109	1.632	0.805	0.355	0.381	1.122	1.609	0.866
	UC_M	0.040	0.268	0.421	1.465	0.068	0.286	0.608	−0.064	1.964	0.122
		(g) $\alpha_1 = 0.5$, $\beta_0 = 0.2$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.25$					(h) $\alpha_1 = 0.5$, $\beta_0 = 0$, $\phi_1 = 0.5$, $\mu_2 = 0.25$, $\sigma^2_{\eta_1} = 0.75$				
50	UC_U	0.147	0.255	0.788	1.566	0.285	0.296	1.157	0.759	2.032	0.311
	UC_M	0.069	0.562	0.326	1.610	0.230	0.131	0.652	−0.143	2.140	0.199
100	UC_U	0.152	0.194	0.910	1.389	0.325	0.297	0.354	0.923	1.673	0.429
	UC_M	0.036	0.226	0.521	1.351	0.158	0.066	0.412	0.179	1.618	0.117
250	UC_U	0.156	0.170	1.019	1.292	0.611	0.291	0.314	1.074	1.519	0.769
	UC_M	0.016	0.128	0.724	1.219	0.081	0.034	0.221	0.526	1.382	0.070

Notes: See Table 1.

4 An empirical application: testing PPP

The PPP proposition states that once converted to a common currency, national price levels should be equal

$$p_t - s_t = p_t^*, \tag{29}$$

where p_t is the log of the price level in the home country, s_t is the log of the nominal exchange rate, measured as the home currency price of one unit of foreign exchange, and p_t^* is the log of the price level in the foreign country. Defining the real exchange rate q_t as the relative price of foreign goods

$$q_t = s_t - p_t + p_t^*, \tag{30}$$

PPP holds if $q_t = 0$. While numerous empirical studies have shown that PPP does not hold in the short run (see e.g. Frenkel, 1981), the question whether PPP serves as an anchor for the long-run real exchange rate has been an area of animated research. Empirical studies in the 1980s typically concluded that PPP failed to hold even in the long run as (i) q_t was found to follow a random walk (see e.g. Adler and Lehman, 1983) and (ii) even after relaxing the assumption of long-run homogeneity s_t, p_t and p_t^* were not found to be cointegrated (see e.g. Corbae and Ouliaris, 1988). This failure to find evidence in favour of PPP is often attributed to low power of standard (ADF type) unit root and cointegration tests to stationary alternatives with a root close to unity. Using longer data series and higher-powered techniques, more recent tests do find evidence in favour of long-run PPP (see e.g. Kim, 1990; Ardeni and Lubian, 1991; Glen, 1992; Taylor, 2002). Using data on the real exchange rate, relative to the U.S. Dollar, for 19 countries over a period of more than 100 years Taylor (2002), for instance, shows that the unit root hypothesis can be rejected for most countries using a generalised-least-squares (GLS) version of the DF unit root test.

Engel (2000) argues that these unit root and cointegration tests in favour of PPP may have reached the wrong conclusion, though. To see why, consider the price indices p_t and p_t^* to be weighted averages of traded and non-traded goods prices

$$\begin{aligned} p_t &= (1-\kappa)p_t^T + \kappa p_t^N, \\ p_t^* &= (1-\kappa^*)p_t^{T*} + \kappa^* p_t^{N*}, \end{aligned}$$

where the superscripts T and N indicate traded and non-traded goods respectively and κ and κ^* are the shares that non-traded goods take in the overall price index in the home and the foreign country respectively. The real exchange rate can now be decomposed into the relative price of traded goods q_t^T and a component q_t^N which is a weighted difference of the relative price of non-traded to traded goods prices

$$q_t = q_t^T + q_t^N, \tag{31}$$

where

$$q_t^T = s_t + p_t^{T*} - p_t^T, \quad (32)$$

$$q_t^N = \kappa^*(p_t^{N*} - p_t^{T*}) - \kappa(p_t^N - p_t^T). \quad (33)$$

As almost any theory of international price determination implies that deviations from the law of one price for traded goods are stationary, Engel (2000) argues that if q_t is found to be non-stationary this should be because the relative price of non-traded goods q_t^N is non-stationary, which can be due to permanent shocks to productivity for instance. As there are no reliable long time series available for q_t^N, non-stationarity of q_t^N is tested using unit root tests on q_t. This hypothesis has been rejected in the recent literature. Using a simulation experiment, Engel (2000) shows that both unit root and ECM cointegration tests on q_t have a considerable size bias in the presence of a (large) permanent component in q_t, i.e. for a nominal size of 5% the true size ranges from 90% to 99% in 100-year long data series when there is a non-stationary component that accounts for 42% of the 100-year forecast variance of q_t. This size bias is confirmed by the simulation results in the previous section.

Combining equations (30) and (31)

$$p_t - s_t = p_t^* - q_t^T - q_t^N, \quad (34)$$

shows that there is a one-to-one relation between $p_t - s_t$ and p_t^* but that this is not a cointegrating relation if q_t^N is non-stationary. As standard estimators yield spurious results in this case, we will estimate the long-run relation between $p_t - s_t$ and p_t^* using the UC framework outlined in section 2.4. The model in equation (32)-(34) can be written is the format of the model in equations (1)-(4) by setting $y_{1t} = p_t - s_t$, $y_{2t} = p_t^*$ and assuming (see equation (6))

$$q_t^T = -\left(\gamma_1(L)\Delta y_{1t} + \gamma_2(L)\Delta y_{2t} + \varepsilon_{1t}/\alpha(1)\right), \quad (35)$$

$$\Delta q_t^N = -\eta_{1t}/\alpha(1). \quad (36)$$

Consistent with the assumptions in Engel (2000), equation (35) allows q_t^T to be a stationary process with rich dynamics while equation (36) restricts q_t^N to be a simple random walk. The covariance matrix Ω in equation (5) allows shocks to the observed variables $p_t - s_t$ and p_t^* and to the unobserved component q_t^N to be contemporaneously correlated.

Data are taken from Taylor (2002). They consist of the log of annual nominal exchange rates s_{it} measured as domestic currency units per U.S. Dollar and the log of consumer price deflators p_{it}, where the index $i = 1, \ldots, 20$ covers a set of 20 countries and the index $t = 1892, \ldots, 1996$ covers a set of 105 years. The variables y_{1t} and y_{2t} in equations (1)-(2) are taken to be the U.S. Dollar-denominated price levels in the domestic country, $p_{it} - s_{it}$, and abroad, p_{it}^*, respectively. In order to avoid the choice of a base country, p_{it}^* is constructed by aggregating for each country i the price levels of the other 19 countries in the dataset into a 'world' basket as $p_{it}^* = (N-1)^{-1} \sum_{j \neq i} (p_{jt} - s_{jt})$.

Given that p_{it}^* is a 'world' basket, the model's assumption that y_{2t} is weakly exogenous for the long-run parameters seems reasonable.

A standard ADF unit root test shows that the unit root hypothesis cannot be rejected for $p_{it} - s_{it}$ and p_{it}^* in any of the included countries (results available on request). As a first step, we therefore estimate the long-run relation between $p_{it} - s_{it}$ and p_{it}^* using the OLS, DGLS and ECM estimators and test for cointegration using an ADF test on the residuals of the OLS and DGLS regression and a t-test on the error-correction term in the ECM regression. Panel (a) in Table 6 reports the results for the ECM estimator (the results for the OLS and DGLS estimators are qualitatively similar and therefore not reported). At the 5% level of significance, the null hypothesis of no cointegration is rejected for only 10 out of the 20 countries. Moreover, using a standard t-test λ_2 is found to be significantly different from 1 in 4 out of the 10 countries where cointegration is found. So, the evidence in favour of PPP is at most modest.

Next, we estimate the long-run relation between $p_{it} - s_{it}$ and p_{it}^* using the UC model in equations (15)-(16). As the full model is not necessarily identified, we first estimate a univariate AR model to check whether p_t^* has dynamics that are sufficiently rich for identification.[3] As using an ADF unit root test p_t^* is found to be non-stationary, we impose a unit root by estimating the model in first differences. The results (standard errors in brackets)

$$\Delta p_t^* = \underset{(0.009)}{0.026} + \underset{(0.097)}{0.463} \Delta p_{t-1}^* - \underset{(0.097)}{0.258} \Delta p_{t-2}^*. \tag{37}$$

indicate that p_t^* is a non-stationary AR(3) model with drift, which is sufficiently rich for identification. The results of applying the UC_M estimator are reported in Panel (b) of Table 6. The most important result is that the PPP hypothesis is rejected in only 3 out of 20 countries, i.e. Canada, Denmark and Japan. While $\widehat{\lambda}_2$ is not too far from 1 for Canada and Denmark, it equals 1.47 for Japan. The strong rejection of the PPP hypothesis in Japan may be due to the fact that the secular increase in the relative price of non-traded goods induces a drift in the unobserved component, which is not allowed for in the model. Note that out of the 10 countries for which we could not reject the null hypothesis of no cointegration based on the ECM, 7 countries (Brazil, Germany, the Netherlands, Portugal, Spain, Sweden and Switzerland) have $\widehat{\sigma}_{\eta_1} \neq 0$. Further note that in 2 countries (Finland and Mexico) for which we found cointegration based on the ECM, we also have $\widehat{\sigma}_{\eta_1} \neq 0$. This indicates that in these countries the conclusion that $p_{it} - s_{it}$ and p_{it}^* are cointegrated is wrong. Consistent with the line of argumentation in Engel (2000), the UC framework shows that U.S. Dollar-denominated price levels converge to their PPP level in most countries once a permanent component in the real exchange rate, which may be induced by non-stationarity of the relative price of non-traded goods, is modelled explicitly.

[3] Note that the 'world' basket p_{it}^* is slightly different for each country i. Therefore, in this explorative estimation p_t^* is calculated as the average over the 20 countries in the sample, i.e. $p_t^* = (N)^{-1} \sum_j (p_{jt} - s_{jt})$.

Table 6: Testing the PPP hypothesis using the ECM and UC_M estimators

	(a) ECM					(b) UC_M												
	p	q	$\widehat{\lambda}_2$		t_{CI}	p	q	$\widehat{\alpha}(1)$		$\widehat{\lambda}_2$		$\widehat{\sigma}^2_{\varepsilon_1}$	$\widehat{\sigma}^2_{\varepsilon_2}$	$\widehat{\sigma}^2_{\eta_1}$	$\widehat{\sigma}_{\varepsilon_1\varepsilon_2}$	$\widehat{\sigma}_{\varepsilon_2\eta_1}$	$Q(1)$	$Q(4)$
Argentina	1	1	0.92	(0.08)	−5.18***	1	1	−0.45	(0.08)	0.91	(0.09)	0.096	0.005	0.000	−0.003	0.000	0.84	0.32
Australia	1	1	0.92	(0.05)	−3.17*	1	1	−0.17	(0.05)	0.92	(0.05)	0.005	0.006	0.000	−0.001	0.000	0.29	0.67
Belgium	1	2	1.10	(0.06)	−3.93**	1	1	−0.34	(0.07)	1.08	(0.08)	0.030	0.006	0.000	0.005	0.000	0.07	0.29
Brazil	1	1	1.08	(0.27)	−2.18	1	1	−0.89	(0.26)	0.96	(0.80)	0.010	0.006	0.006	−0.005	−0.003	0.90	0.02
Canada	2	1	0.89	(0.04)	−3.54**	2	1	−0.15	(0.04)	0.88	(0.05)	0.002	0.007	0.000	0.000	0.000	0.35	0.91
Denmark	2	2	1.13	(0.04)	−4.63***	2	2	−0.28	(0.05)	1.13	(0.04)	0.008	0.006	0.000	−0.002	0.000	0.35	0.48
Finland	2	1	0.95	(0.04)	−4.81***	2	1	−0.70	(0.21)	0.99	(0.17)	0.018	0.006	0.002	0.006	0.002	0.64	0.92
France	1	1	0.94	(0.04)	−3.41**	1	1	−0.20	(0.05)	0.94	(0.05)	0.005	0.006	0.000	0.001	0.000	0.14	0.22
Germany	2	1	1.03	(0.10)	−2.04	1	1	−0.74	(0.11)	0.92	(0.21)	0.000	0.006	0.004	0.001	−0.001	0.62	0.29
Italy	1	1	0.96	(0.09)	−2.90	1	1	−0.19	(0.06)	0.95	(0.11)	0.023	0.007	0.000	−0.006	0.000	0.23	0.42
Japan	2	1	1.47	(0.05)	−3.62**	2	1	−0.21	(0.05)	1.47	(0.05)	0.006	0.006	0.000	0.000	0.000	0.67	0.86
Mexico	1	1	0.76	(0.06)	−4.46***	2	1	−0.78	(0.16)	0.76	(0.24)	0.019	0.006	0.004	0.002	0.002	0.94	1.00
Netherlands	1	2	1.13	(0.09)	−1.77	1	1	−0.81	(0.10)	1.02	(0.15)	0.000	0.006	0.004	0.001	−0.002	0.78	0.96
Norway	2	2	1.04	(0.04)	−3.42**	2	1	−0.20	(0.06)	1.02	(0.04)	0.007	0.006	0.000	0.004	0.000	0.34	0.88
Portugal	1	2	0.86	(0.06)	−3.39*	1	2	−0.51	(0.11)	0.88	(0.14)	0.016	0.006	0.001	0.007	0.000	0.43	0.71
Spain	1	1	0.91	(0.11)	−1.91	1	1	−1.08	(0.15)	1.11	(0.20)	0.000	0.006	0.006	0.000	0.002	0.20	0.11
Sweden	1	1	0.98	(0.05)	−2.49	1	1	−1.03	(0.12)	1.00	(0.14)	0.000	0.006	0.003	−0.001	0.000	0.75	0.99
Switzerland	1	1	1.27	(0.06)	−2.41	1	1	−0.83	(0.10)	1.21	(0.18)	0.000	0.006	0.004	0.001	−0.002	0.25	0.75
UK	1	1	0.93	(0.04)	−3.00	1	1	−0.18	(0.05)	0.93	(0.05)	0.003	0.006	0.000	0.001	0.000	0.40	0.21
US	2	1	0.98	(0.03)	−4.45***	2	1	−0.15	(0.03)	0.96	(0.05)	0.001	0.007	0.000	0.000	0.000	0.87	0.49

Note: Standard errors in brackets. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level. The lag orders p and q in the ECM and UC_M are chosen by setting a maximum lag order of 2 and deleting insignificant lags combined with a standard test for serial correlation in the residuals based on the Box-Ljung statistic. The lag order r in the UC_M is set to 2 (see results equation (37)). For the ECM, t_{CI} is a test for the null of no cointegration, i.e. a standard t-statistic on the error-correction coefficient $\widehat{\alpha}(1)$ in the ECM. The simulated critical values for this test are -4.03, -3.41 and -3.09 at the 1%, 5% and 10% level respectively. For the UC_M, $Q(1)$ and $Q(4)$ are p-values of Box-Ljung test statistics for serial correlation in the one-step forecast errors of y_{2t} for a lag order of 1 and 4 respectively.

5 Concluding comments

The concept of cointegration requires all variables constituting a long-run equilibrium relation to be included in the analysis. Integrated variables that are among a set of variables constituting an equilibrium relation but that are omitted from the analysis imply that (i) the remaining variables are not cointegrated and (ii) standard cointegration tests have a serious size bias. As such, standard estimators for a long-run relation between integrated variables may yield spurious results with standard cointegration tests indicating these results to be a cointegration regression. This paper uses an UC framework to estimate the long-run relationship between integrated variables when possibly not all relevant variables are included in the analysis. A Monte Carlo experiment shows that for a simple dynamic bivariate triangular process, this approach yields results similar to those of standard estimators in the case of cointegration and does not yield spurious results in the cases where the included variables are independent random walks or related in the long-run with an unobserved component. Based on a Monte Carlo simulation, Inder (1993) concludes that 'estimates which include dynamics are much more reliable, even if the dynamic structure is overspecified'. The Monte Carlo simulation in this paper suggests that estimates that include unobserved integrated components may be even more reliable, even if both the dynamic structure and the error structure are overspecified. As the general UC model allows shocks to the observed and to the unobserved variables to be contemporaneously correlated, one important issue is that of identification. In this paper, identification stems from the assumption that the unobserved component is a simple random walk, with the observed explanatory variable having richer dynamics. In practise, this assumption is not necessarily fulfilled. Allowing for richer dynamics in the unobserved component is possible although for the model to be identified they should differ from the type of dynamics in the observed components.

The proposed methodology is applied to testing PPP. Using data on consumer prices and nominal exchange rates for a set of 20 countries over 105 years taken from Taylor (2002), standard cointegration analysis provides only weak evidence in favour of PPP. Consistent with the line of argumentation in Engel (2000), the UC framework shows that U.S. Dollar-denominated price levels converge to their PPP level in most countries once a permanent component in the real exchange rate, which may be induced by non-stationarity of the relative price of non-traded goods, is modelled explicitly. This suggests that PPP holds for traded goods. The strong rejection of the PPP hypothesis in Japan suggests that the assumption that the unobserved component is a random walk may be too simplistic in some countries. A more detailed analysis with richer dynamics in the unobserved component is left for future research.

References

Adler, M., Lehman, B., 1983. Deviations from purchasing power parity in the long run. Journal of Finance 38 (5), 1471–1487.

Ardeni, P. G., Lubian, D., July 1991. Is there trend reversion in purchasing power parity? European Economic Review 35 (5), 1035–1055.

Banerjee, A., Dolado, J. J., Hendry, D. F., Smith, G. W., Aug. 1986. Exploring equilibrium relationships in econometrics through static models: Some monte-carlo evidence. Oxford Bulletin of Economics and Statistics 48 (3), 253–277.

Banerjee, A., Donald, J., Mestre, R., 1996. Ecm tests for cointegration in a single equation framework. Tech. rep., Oxford University.

Corbae, D., Ouliaris, S., August 1988. Cointegration and tests of purchasing power parity. The Review of Economics and Statistics 70 (3), 508–11.

Durbin, J., Koopman, S., 2001. Time Series Analysis by State Space Methods. Oxford University Press, Oxford.

Engel, C., August 2000. Long-run ppp may not hold after all. Journal of International Economics 51 (2), 243–273.

Engle, R., Hendry, D., Richard, J., 1983. Exogeneity. Econometrica 51, 277–304.

Engle, R. F., Granger, C. W. J., March 1987. Co-integration and error correction: Representation, estimation, and testing. Econometrica 55 (2), 251–76.

Ericsson, N. R., MacKinnon, J. G., 2002. Distributions of error correction tests for cointegration. Econometrics Journal 5 (2), 285–318.

Frenkel, J. A., 1981. The collapse of purchasing power parities during the 1970s. European Economic Review 16 (1), 145–165.

Glen, J. D., Aug. 1992. Real exchange rates in the short, medium and long run. Journal of International Economics 33 (1-2), 147–166.

Granger, C., Newbold, P., 1986. Forecasting Economic Time Series, 2nd Edition. Academic Press, Orlando.

Harris, D., Inder, B., 1994. A test of the null of cointegration. In: Hargreaves, C. (Ed.), Non-Stationary Time Series Analysis and Cointegration. Oxford University Press, Oxford, pp. 133–152.

Harvey, A., 1989. Forecasting structural time series models and the Kalman filter. Cambridge University Press.

Harvey, A. C., Henry, S. G. B., Peters, S., Wren-Lewis, S., Dec. 1986. Stochastic trends in dynamic regression models: An application to the employment-output equation. Economic Journal 96 (384), 975–985.

Inder, B., 1993. Estimating long-run relationships in economics: A comparison of different approaches. Journal of Econometrics 57 (1-3), 53–68.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics & Control 12 (2-3), 231–254.

Kim, Y. B., Nov. 1990. Purchasing power parity in the long run: A cointegration approach. Journal of Money, Credit and Banking 22 (4), 491–503.

Kremers, J. J. M., Ericsson, N. R., Dolado, J. J., August 1992. The power of cointegration tests. Oxford Bulletin of Economics and Statistics 54 (3), 325–48.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root? Journal of Econometrics 54 (1-3), 159–178.

Leybourne, S. J., McCabe, B. P. M., February 1994. A simple test for cointegration. Oxford Bulletin of Economics and Statistics 56 (1), 97–103.

Mackinnon, J. G., 1996. Numerical distribution functions for unit root and cointegration tests. Journal of Applied Econometrics 11 (6), 601–618.

Morley, J. C., 2007. The slow adjustment of aggregate consumption to permanent income. Journal of Money, Credit, and Banking (in press).

Morley, J. C., Nelson, C. R., Zivot, E., 2003. Why are the beveridge-nelson and unobserved-components decompositions of gdp so different? Review of Economics and Statistics 85 (2), 235–243.

Nelson, C. R., Plosser, C. I., 1982. Trends and random-walks in macroeconomic time-series: Some evidence and implications. Journal of Monetary Economics 10 (2), 139–162.

Pagan, A., 1980. Some identification and estimation rresults for regression models with stochastically varying coefficients. Journal of Econometrics 13 (3), 341–363.

Phillips, P. C. B., Durlauf, S. N., Aug. 1986. Multiple time-series regression with integrated porcesses. Review of Economic Studies 53 (4), 473–495.

Phillips, P. C. B., Hansen, B. E., January 1990. Statistical inference in instrumental variables regression with i(1) processes. Review of Economic Studies 57 (1), 99–125.

Rudd, J., Whelan, K., 2006. Empirical proxies for the consumption-wealth ratio. Review of Economic Dynamics 9, 34–51.

Saikkonen, P., Mar. 1991. Asymptotically efficient estimation of cointegration regressions. Econometric Theory 7 (1), 1–21.

Sarantis, N., Stewart, C., 2001. Unobserved components in an error-correction model of consumption for southern european countries. Empirical Economics 26, 391–405.

Shin, Y. C., Mar. 1994. A residual-based test of the null of cointegration against the alternative of no cointegration. Econometric Theory 10 (1), 91–115.

Stock, J., September 1987. Asymptotic properties of least squares estimators of cointegrating vectors. Econometrica 55 (5), 1035–56.

Stock, J. H., Watson, M. W., July 1993. A simple estimator of cointegrating vectors in higher order integrated systems. Econometrica 61 (4), 783–820.

Taylor, A. M., 2002. A century of purchasing-power parity. The Review of Economics and Statistics 84 (1), 139–150.

Zivot, E., 1994. Single equation conditional error correction model based tests for cointegration. Discussion Papers in Economics at the University of Washington 94-12, Department of Economics at the University of Washington.