Efficient Estimation of the SUR Cointegration Regression Model and Testing for Purchasing Power Parity*

Hyungsik Roger Moon†
Department of Economics
University of Southern California
moonr@usc.edu

Benoit Perron‡
Département de sciences économiques, CIREQ, and CIRANO
Université de Montréal
benoit.perron@umontreal.ca

April 2004

Abstract

This paper studies the efficient estimation of seemingly unrelated linear models with integrated regressors and stationary errors. We consider two cases. The first one has no common regressor among the equations. In this case, we show that by adding leads and lags of the first differences of the regressors and estimating this augmented dynamic regression model by generalized least squares using the long-run covariance matrix, we obtain an efficient estimator of the cointegrating vector that has a limiting mixed normal distribution. In the second case we consider, there is a common regressor to all equations, and we discuss efficient minimum distance estimation in this context. Simulation results suggests that our new estimator compares favorably with others already proposed in the literature. We apply these new estimators to the testing of the proportionality and symmetry conditions implied by purchasing power parity (PPP) among the G-7 countries. The tests based on the efficient estimates easily reject the joint hypotheses of proportionality and symmetry for all countries with either the United States or Germany as numeraire. Based on individual tests, our results suggest that Canada and Germany are the most likely countries for which the proportionality condition holds, and that Italy and Japan for the symmetry condition relative to the United States.

Keywords: Seemingly Unrelated Regressions, Efficient Estimation, Purchasing Power Parity, Minimum distance

JEL classification: C32, F31

*We thank Menzie Chinn, Elena Pesavento, Lynda Khalaf, and two anonymous referees for their valuable comments on an earlier version of the paper which was circulated under the title, The Seemingly Unrelated Dynamic Cointegration Regression Model and Testing for Purchasing Power Parity. We also thank participants at the 2001 Canadian Economic Association meeting in Montreal and the 2000 Meeting of the Société candienne de science économique (SCSE) for their comments.

†Financial assistance from USC Faculty Development Fund is gratefully acknowledged.

‡Financial assistance from the Fonds FCAR and the Mathematics of Information Technology and Complex Systems (MITACS) network is gratefully acknowledged.

1 Introduction

The purchasing power parity (PPP) doctrine plays a central role in most open-economy macroeconomic models. It states that nominal exchange rates should reflect relative price behavior. There is by now an enormous literature on testing PPP. A survey of earlier tests can be found in Froot and Rogoff (1995) , and a recent comprehensive survey of recent research can be found in Sarno and Taylor (2002) . Most of these tests are univariate and use data for a single country at a time in testing for PPP. Denote by s_{it} the nominal exchange rate of country i (defined as the quantity of domestic currency of country i per unit of foreign currency), p_t^* is the foreign price level , and p_{it} is the domestic price level of country i. Purchasing power parity implies that if we run a regression for country i of the type :

$$\ln s_{it} = \beta_{0i} + \beta_{1i} \ln (p_{it}) + \beta_{2i} \ln (p_t^*) + \varepsilon_{it}, \tag{1}$$

we should have $\beta_{1i} = -\beta_{2i}$ (the so-called symmetry condition) and $\beta_{1i} = 1$ (also called the proportionality condition). The bulk of the tests impose these two conditions and test whether the the real exchange rate, defined as:

$$q_{it} = \frac{s_{it} p_t^*}{p_{it}}$$

is stationary around a constant. The large majority of these tests support a unit root in real exchange rate which is evidence against PPP.

In this paper, we use (1) directly to test the symmetry and proportionality conditions to investigate *sources of rejection of PPP.* To improve *efficiency* of the test, we use cross-country data to take advantage of correlations across countries.

In order to do this, we study efficient estimation of a system of seemingly unrelated regression (SUR) equations with integrated regressors. We distinguish two cases: the case where there is a common regressor in each equation (as there is in equation (1) since each equation involves the price level of the numeraire country) and the case where there is no such common regressor. The case with a common regressor allows us to test directly the symmetry condition, *i.e.* $\beta_{1i} = -\beta_{2i}$, while the case without a common regressor will be used to test proportionality (the coefficient on relative price levels should be 1).

Park and Ogaki (1991) demonstrated that the conventional feasible GLS estimation of a system of regression equations with integrated regressors has a nonstandard limit distribution that is skewed and shifted away from the true parameter. This renders inference in these systems difficult. To solve this problem, Park and Ogaki (1991) suggested using the *seemingly unrelated canonical cointegrating regression* (SUCCR) estimator as an extension of the original CCR estimator first presented in Park (1992). On the other hand, Moon (1999) suggested using *fully-modified* (FM) estimators (*e.g.*, see Phillips and Hansen, 1990, Phillips, 1991, and Phillips, 1995) on these systems. These two sets of estimators are efficient relative to single-equation estimators, and have a limiting mixed normal distributions, rendering inference straightforward.

This paper suggests an alternative approach for obtaining efficient estimators with a mixed normal limiting distribution based on results of Saikkonen (1991) and Stock and Watson (1993). The approach (which we call dynamic FGLS) consists of adding leads and lags of the first differences of the regressors and using feasible generalized least squares on this augmented dynamic regression model using the long-run covariance matrix. As the other two estimators mentioned above, our estimator is more efficient than equation-by-equation or system-wide ordinary least squares. [1]

When there is a common regressor, GLS is not feasible since the long-run covariance matrix is singular and cannot be inverted. For this case, we propose to use a minimum distance estimator suggested by Moon and Schorfheide (2002) and Elliott (2000). This estimator is known to be efficient and is asymptotically equivalent to our DGLS without common regressors.

We apply our new estimators to testing for purchasing power parity (PPP) among the G-7 countries over the recent float. Our new methodology allows us to test separately the symmetry condition (the coefficients on the domestic and price levels are of the same magnitude but of different sign) and proportionality condition (this common coefficient is 1). Contrary to most tests used in the literature, this test treats PPP as the null hypothesis and is asymptotically invariant to the choice of numeraire *if PPP holds*. Moreover, it easily handles multiple countries simultaneously and thus can take advantage of correlations among countries to construct more *efficient* tests, while allowing inference on individual countries. Using our efficient estimators, we can reject the presence of PPP for several countries in our sample when using both the United States and Germany as the numeraire country. The rejections come from the failure of both the symmetry and proportionality conditions.

The most closely paper is that by Li (1999) . Using a Bayesian approach, she tests the symmetry and proportionality conditions within a SUR framework. The parameters β_{1i} and β_{2i} are modelled as independent draws from normal distributions in each period. The symmetry condition in this case is that the magnitude of the mean of these two distributions is the same, while the proportionality condition is that the magnitude of both means is 1. Using data for a larger sample of OECD data, she finds little support for these two conditions, although the evidence against them diminishes at the sampling frequency diminishes.

The outline for the rest of the paper is as follows. Section 2 introduces our estimators of the SUR model with integrated regressors and derives their limiting asymptotic distribution. Section 3 presents results from a simulation experiment comparing estimators of the integrated SUR model. Section 4 presents our empirical methodology and results for testing PPP among industrialized countries over the recent float, while section 5 concludes.

2 Efficient Estimation of the SUR Cointegration Model

In this section we study a SUR model with integrated regressors. Suppose that there are N individual linear cointegration regression equations,

$$\begin{aligned} y_{it} &= \beta_{0i} + \beta_{1i}' x_{it} + u_{it} \qquad (2) \\ x_{it} &= x_{it-1} + v_{it}, \end{aligned}$$

[1] A similar analysis has been recently proposed independently by Mark, Ogaki, and Sul (2000) who apply it to tests of forward exchange rate unbiasedness and international correlation between saving and investment.

where u_{it} and v_{it} are scalar and L-vector valued stationary processes, respectively, for $i = 1, \cdots, N$, and $t = 1, ..., T$. Assume that $x_{i0} = O_p(1)$ as $T \to \infty$ for all $i = 1, ..., N$. Let $y_t = (y_{1t}, \ldots, y_{Nt})'$, $\tilde{x}_t = diag(\tilde{x}_{1t}, \ldots, \tilde{x}_{Nt})$, $\tilde{x}_{it} = (1, x_{it}')'$, $x_t = (x_{1t}', ..., x_{Nt}')'$, $u_t = (u_{1t}, \ldots, u_{Nt})'$, and $v_t = (v_{1t}', \ldots, v_{Nt}')'$. Then, using vector notation, we rewrite (2) as

$$\begin{aligned} y_t &= \tilde{x}_t' \beta + u_t, \\ x_t &= x_{t-1} + v_t, \end{aligned} \tag{3}$$

where $\beta = \left(\beta_1', ..., \beta_N'\right)'$ and $\beta_i = \left(\beta_{0i}, \beta_{1i}'\right)'$.

Define $w_t = \begin{pmatrix} u_t' & v_t' \end{pmatrix}'$. Our analysis will be for $T \to \infty$ for N fixed. We assume that the partial sum process of w_t converges in distribution to a Brownian Motion as $T \to \infty$,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} w_t \Rightarrow B(r) \equiv BM(\Omega), \tag{4}$$

where $\Omega = \sum_{h=-\infty}^{\infty} E(w_0 w_h') = \begin{pmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{pmatrix}$, $B(r) = \left(B_u(r)', B_v(r)'\right)'$, $B_u(r) = (B_{u,1}(r), ..., B_{u,N}(r))'$, $B_v(r) = \left(B_{v,1}(r)', ..., B_{v,N}(r)'\right)'$, and the partitions of Ω and $B(r)$ are conformable with u_t and v_t. The functional central limit theorem assumed in (4) is satisfied under mild regularity conditions on w_t (*e.g.*, see Phillips and Solo, 1992). In the next subsection, we assume that the long-run covariance matrix Ω of w_t is positive definite, which excludes the possibility that there exists a cointegrating relation among the elements of x_t. The case where Ω is singular because there is a common regressor will be discussed in the following subsection.

Our setup is a SUR version of a triangular system as developed by Phillips (1991). In complementary work, Groen and Kleibergen (2003) look at maximum likelihood estimation and testing in a SUR extension of a system of error correction models with correlation across observational units, but they do not allow for common regressors. Their analysis is for large T and fixed N like ours. Larsson, Lyhagen, and Lothgren (2001) use the same framework as Groen and Kleibergen to develop tests of cointegrating rank. However, they assume independence across observational units and use large N, large T approximations.

2.1 Dynamic GLS Estimation of the Integrated Seemingly Unrelated Regression Model

The main purpose of this section is to develop an efficient estimation method for the cointegration parameter β in the SUR cointegration model (2) when the long-run covariance matrix Ω is of full rank and establish its asymptotic properties. When there is (long-run) correlation between u_t and v_t (as in a simultaneous equation model) and/or serial correlation, it is well known that the limiting distribution of the OLS estimator of model (2) is miscentered and skewed, and this causes difficulties in statistical inference. To overcome the problem, we modify the regression model and make the transformed error asymptotically uncorrelated with the regressors and uncorrelated through time.

To do so, we decompose the stacked vector u_t into two components, the projection onto the sigma field generated by $\{v_t\}_{t=-\infty}^{\infty}$, $\sum_{j=-\infty}^{\infty} \pi_j v_{t-j}$, and a residual, namely

$$u_t = \sum_{j=-\infty}^{\infty} \pi_j v_{t-j} + \xi_t. \tag{5}$$

and in this case, ξ_t is uncorrelated with $\{v_t\}_{t=-\infty}^{\infty}$. Denote the long-run covariance of ξ_t by $\Omega_{uu.v}$. Under mild regularity conditions [2], we know that

$$\Omega_{uu.v} = \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}. \tag{6}$$

Now, in view of the decomposition (5) , we can write model (3) as

$$y_t = \tilde{x}_t'\beta + \sum_{j=-\infty}^{\infty} \pi_j v_{t-j} + \xi_t. \tag{7}$$

In equation (7) we have added an infinite number of regressors $\{v_t\}_{t=-\infty}^{\infty}$, which makes estimating it impossible with a finite number of observations. We propose a feasible version of (7) obtained by truncating the infinite sum and replacing it by $\sum_{j=-K}^{K} \pi_j \Delta x_{t-j}$, where K is assumed to tend to infinity as $T \to \infty$ at an appropriate rate to be specified below. Hence the estimable version of model (7) is

$$\begin{aligned} y_t &= \tilde{x}_t'\beta + \sum_{j=-K}^{K} \pi_j \Delta x_{t-j} + \xi_t^* = \tilde{x}_t'\beta + \sum_{j=-K}^{K} \left(\Delta x_{t-j}' \otimes I_M\right) vec(\pi_j) + \xi_t^* \\ &= z_t' b + \xi_t^*, \end{aligned} \tag{8}$$

where we have defined $\xi_t^* = \xi_t + e_t$, $e_t = \sum_{|j|>K} \pi_j v_{t-j}$, $z_t = \left(\tilde{x}_t',\ \Delta x_{t-K}' \otimes I_N, ..., \Delta x_{t+K}' \otimes I_N\right)'$, $b = \left(\beta', \Pi_K'\right)'$, and $\Pi_K = (vec(\pi_{-K})', ..., vec(\pi_K)')$. Note that the regression model (8) is simply an augmented version of the original SUR model (2) obtained by adding leads and lags of Δx_t.

We may consider estimating the regression model (8) in two ways, either by OLS or by GLS using the long-run correlation in the error ξ_t as is appropriate when the error term has serial correlation. Let $Y_t = (y_{1,t}, ..., y_{N,t})'$, $Y = \left(Y_{K+1}', ..., Y_{T-K}'\right)'$, $\tilde{X} = (\tilde{x}_{K+1}, ..., \tilde{x}_{T-K})'$, $\Delta_{K,t} = \left(\Delta x_{t-K}', ..., \Delta x_{t+K}'\right)' \otimes I_N$, $\Delta_K = (\Delta_{K,K+1}, ..., \Delta_{K,T-K})'$, and $Z = \left(\tilde{X}, \Delta_K\right)$. The system dynamic OLS (hereafter SDOLS) and the dynamic (feasible) GLS (hereafter DGLS) estimators are defined, respectively, as:

$$\begin{aligned} \hat{b}_{SDOLS} &= \left(Z'Z\right)^{-1} Z'Y, \\ \hat{b}_{DGLS} &= \left(Z'\left(I_{T-2K} \otimes \hat{\Omega}_{uu.v}\right)^{-1} Z\right)^{-1} \left(Z'\left(I_{T-2K} \otimes \hat{\Omega}_{uu.v}\right)^{-1} Y\right) \end{aligned}$$

or in summation notation,

$$\hat{b}_{SDOLS} = \left(\sum_{t=K+1}^{T-K} z_t z_t'\right)^{-1} \left(\sum_{t=K+1}^{T-K} z_t y_t\right) = b + \left(\sum_{t=K+1}^{T-K} z_t z_t'\right)^{-1} \left(\sum_{t=K+1}^{T-K} z_t \xi_t^*\right),$$

[2] See, for example, Brillinger (1975) and Saikkonen (1991).

and

$$\hat{b}_{DGLS} = \left(\sum_{t=K+1}^{T-K} z_t \hat{\Omega}_{uu.v}^{-1} z_t' \right)^{-1} \left(\sum_{t=K+1}^{T-K} z_t \hat{\Omega}_{uu.v}^{-1} y_t \right) = b + \left(\sum_{t=K+1}^{T-K} z_t \hat{\Omega}_{uu.v}^{-1} z_t' \right)^{-1} \left(\sum_{t=K+1}^{T-K} z_t \hat{\Omega}_{uu.v}^{-1} \xi_t^* \right),$$

where $\hat{\Omega}_{uu.v}$ is a consistent estimate of $\Omega_{uu.v}$. Note that the $DGLS$ estimator $\hat{\beta}_{DGLS}$ is a GLS estimator using the long-run correlation information in the system (8).

An alternative estimator is the dynamic OLS estimator for the cointegrating vector β_i (hereafter IDOLS) in the i^{th} individual regression model of (2). Now define $w_{it} = (u_{it}, v_{it}')'$ and $\Omega^{i,j} = \begin{pmatrix} \Omega_{uu}^{i,j} & \Omega_{uv}^{i,j} \\ \Omega_{vu}^{i,j} & \Omega_{vv}^{i,j} \end{pmatrix} = \sum_{h=-\infty}^{\infty} E\left(w_{i,0} w_{j,h}'\right)$. Let

$$\theta_{it} = u_{it} - \sum_{s=-\infty}^{\infty} \pi_{ij} v_{it-s}, \tag{9}$$

be the linear projection residual of the i^{th} individual equation regression error u_{it} on the closed linear space of $\{v_{it}\}_{t=-\infty}^{\infty}$ for $i = 1, ..., N$. Denote the long-run variance of θ_{it} by $\Omega_{uu.v}^{i}$. Under mild regularity conditions,

$$\Omega_{uu.v}^{i} = \Omega_{uu}^{i,i} - \Omega_{uv}^{i,i} \left(\Omega_{vv}^{i,i}\right)^{-1} \Omega_{vu}^{i,i}. \tag{10}$$

Now we define the following i^{th} individual dynamic regression

$$y_{it} = \tilde{x}_{it}' \beta_i + \sum_{s=-K}^{K} \pi_{is} \Delta x_{it-s} + \theta_{it}^*, \tag{11}$$

where $\theta_{it}^* = \theta_{it} + \epsilon_{it}$ and $\epsilon_{it} = \sum_{|s|>K} \pi_{is} v_{it-s}$. We denote $\hat{\beta}_i$ the OLS regression estimator for β_i of (11) and $\hat{\beta}_{IDOLS} = \left(\hat{\beta}_1', ..., \hat{\beta}_N'\right)'$.

Let $\hat{\beta}_{DGLS}$ and $\widehat{\Pi}_{DGLS}$ be sub-vectors of $\hat{b}_{DGLS}$ whose sizes are conformable to those of β and Π. In a similar fashion, define $\hat{\beta}_{SDOLS}$, $\widehat{\Pi}_{SDOLS}$, $\hat{\beta}_{IDOLS}$, and $\widehat{\Pi}_{IDOLS}$. Define $F_{T^*} = I_N \otimes diag\left(\sqrt{T^*}, T^* I_L\right)$, where $T^* = T - 2K$. We collect the asymptotic distributions of the above estimators in the following proposition.

Proposition Suppose that the functional limit theorem (4) holds and ξ_t in (5) and θ_{it} in (9) have long-run covariance matrices $\Omega_{uu.v}$ in (6) and $\Omega_{uu.v}^{i}$ in (10) respectively. Further suppose that $\frac{K^3}{T} \to 0$ and $\sqrt{T} \sum_{s>|K|} \|\pi_s\| = o(1)$. Then, as $T \to \infty$,

(a) $F_{T^*}\left(\hat{\beta}_{DGLS} - \beta\right) \Rightarrow MN\left(0, \left(\int_0^1 \tilde{B}_v(r) \Omega_{uu.v}^{-1} \tilde{B}_v'(r) dr\right)^{-1}\right)$,

(b) $F_{T^*}\left(\hat{\beta}_{SDOLS} - \beta\right) \Rightarrow MN\left(0, \left(\int_0^1 \tilde{B}_v(r) \tilde{B}_v'(r) dr\right)^{-1} \left(\int_0^1 \tilde{B}_v(r) \Omega_{uu.v} \tilde{B}_v'(r)\right) \left(\int_0^1 \tilde{B}_v(r) \tilde{B}_v'(r) dr\right)^{-1}\right)$,

(c) $F_{T^*}\left(\hat{\beta}_{IDOLS} - \beta\right) \Rightarrow MN\left(0, \left(\int_0^1 \tilde{B}_v(r) \tilde{B}_v'(r) dr\right)^{-1} \left(\int_0^1 \tilde{B}_v(r) \tilde{\Omega}_{uu.v} \tilde{B}_v'(r)\right) \left(\int_0^1 \tilde{B}_v(r) \tilde{B}_v'(r) dr\right)^{-1}\right)$

where $MN(\cdot,\cdot)$ denote a mixed normal distribution, $\tilde{B}_v(r) = diag\left(\tilde{B}_{v,1}(r), ..., \tilde{B}_{v,N}(r)\right)$, $\tilde{B}_{v,i}(r) = \left(1, B_{v,i}(r)'\right)'$, $\tilde{\Omega}_{uu.v} = diag\left(\Omega_{uu.v}^{1}, \ldots, \Omega_{uu.v}^{N}\right)$.

Moreover, it is easy to show that the limiting variance of $\hat{\beta}_{DGLS}$ is smaller than that of $\hat{\beta}_{SDOLS}$ with probability one because

$$\begin{aligned} & \left(\int_0^1 \tilde{B}_v(r)\,\Omega_{uu.v}^{-1} \tilde{B}_v'(r)\,dr \right)^{-1} \\ < \; & \left(\int_0^1 \tilde{B}_v(r)\,\tilde{B}_v'(r)\,dr \right)^{-1} \left(\int_0^1 \tilde{B}_v(r)\,\Omega_{uu.v} \tilde{B}_v'(r) \right) \left(\int_0^1 \tilde{B}_v(r)\,\tilde{B}_v'(r)\,dr \right)^{-1} \end{aligned} \tag{12}$$

with probability one. Therefore, $\hat{\beta}_{DGLS}$ is asymptotically more efficient than $\hat{\beta}_{SDOLS}$.

Moreover, since $\Omega_{uu.v} \leq \tilde{\Omega}_{uu.v}$, we can conclude that $F_{T^*}\left(\hat{\beta}_{SDOLS} - \beta\right)$ is asymptotically more efficient than $F_{T^*}\left(\hat{\beta}_{IDOLS} - \beta\right)$. Therefore, among the three estimators, we can conclude that the DGLS estimator $\hat{\beta}_{DGLS}$ is the most efficient and the individual dynamic OLS estimator $\hat{\beta}_{IDOLS}$ is the least efficient. This result is also obtained by Park and Ogaki (1991) and Moon (1999) with the CCR method and the FM method respectively.

In a linear cointegration regression model with no restriction on the cointegrating vectors, it is well known that there is no efficiency gain in GLS estimation over OLS estimation (*e.g.*, Phillips and Park, 1988 and Stock and Watson, 1993). This asymptotic equivalence result is similar in spirit to the classical Grenander and Rosenblatt theorem (1957) which states that polynomial time trends with stationary error can be estimated efficiently by OLS.

However, as we verify through (12), in the dynamic augmented model (8) the GLS estimator is asymptotically more efficient than the OLS estimator. As discussed in Park and Ogaki (1991), the asymptotic efficiency gain of GLS in the SUR model comes from overidentification parameter restrictions.

To see this, write the SUR model (3) in a multivariate regression form,

$$y_t = \tilde{b}\tilde{X}_t + u_t, \tag{13}$$

where

$$\tilde{b} = diag\left(\beta_1, ..., \beta_N\right) \tag{14}$$

and $\tilde{X}_t = \left(\tilde{x}_{1t}', ..., \tilde{x}_{Nt}'\right)'$. In model (13), overidentifying parameter restrictions are imposed by restricting the off-block diagonal elements of $\tilde{b}$ to be zero, while in the models studied in Phillips and Park (1988) and Stock and Watson (1993), the regression coefficients are exactly identified.[3]

The mixed normality property of the limiting distributions of $\hat{\beta}_{SDOLS}$, $\hat{\beta}_{DGLS}$, and $\hat{\beta}_{IDOLS}$ enables us to use conventional chi-square tests for the null hypothesis of parameter restrictions. Suppose that we are interested in testing a null hypothesis on the parameters,

$$H_0 : \varphi(\beta) = r, \tag{15}$$

[3] Once the SUR model (1) is expressed in the form of a multivariate cointegration model with restrictions on the coefficients as in (13) and (14), we could think of employing other estimation procedures to efficiently estimate the restricted parameters $(\beta_1, ..., \beta_N)'$. For example, we could apply the efficient minimum distance (MD) estimation method proposed by Elliott (2000) and Moon and Schorfheide (2002). Moon and Schorfheide (2002) show that the asymptotic distributions of the efficient MD estimator and the DGLS estimators are identical. Details of the MD estimation procedures will be discussed in the following section.

where φ is a $(q \times 1)$ vector-valued, continuously differentiable function and the first derivative of φ at the true parameter β, $\Phi(\beta)$, has full rank q. The Wald test statistic W_T using $\hat{\beta}_{DGLS}$ is then defined as

$$W_T \equiv (\varphi(\hat{\beta}_{DGLS}) - r)' \left(\Phi(\hat{\beta}_{DGLS}) \left(\sum_{t=1}^{T} \tilde{x}_t \hat{\Omega}_{uu.v}^{-1} \tilde{x}_t' \right)^{-1} \Phi(\hat{\beta}_{DGLS})' \right)^{-1} (\varphi(\hat{\beta}_{DGLS}) - r), \tag{16}$$

and it is easy to verify that under the assumptions stated above, as $T \to \infty$, $W_T \Rightarrow \chi_q^2$, a chi-square distribution with q degrees of freedom.

2.2 Efficient Estimation with Common Regressors

The asymptotic results in the previous section rest on the assumption that the elements of x_t have a positive definite long-run covariance matrix which excludes the possibility that there exists a cointegrating relation among the elements of $x_t = (x_{1t}', ..., x_{Nt}')'$. This restriction is violated if some variables in x_{it} are common for all $i = 1, ..., N$. Existence of common regressors in a system of equation like (2) is not rare in empirical applications, as discussed in our PPP application below. In this section we discuss how to estimate the coefficients efficiently in the presence of common regressors. The procedure summarized here is the minimum distance estimation method proposed by Moon and Schorfheide (2002) and by Elliott (2000).

Suppose that the integrated regressors in model (2) consist of two components,

$$x_{it} = (x_{it1}', x_{t2}')', \tag{17}$$

where $x_{it1,}$ is an L_1- vector regressor varying across individuals and x_{t2} is an L_2- vector common regressor across individuals ($L = L_1 + L_2$). Let β_{1i1} and β_{1i2} denote the coefficients of x_{it1} and x_{t2}, respectively. Now defining $X_t = (x_{1t1}', ..., x_{Nt1}', x_{t2}')'$ and $\tilde{X}_t = (1, X_t')'$, we express model (2) in the following multivariate regression form,

$$y_t = \Gamma \tilde{X}_t + u_t, \tag{18}$$

where the coefficient matrix Γ is an $(N \times (NL_1 + L_2 + 1))$ matrix. Then, the restricted coefficient matrix Γ in (18) is

$$\begin{aligned} \Gamma &= \left(\Gamma_1 \vdots \Gamma_2' \vdots \Gamma_3' \right), \\ \underset{(N\times 1)}{\Gamma_1} &= (\beta_{01}, ..., \beta_{0i}, ..., \beta_{0N})', \\ \underset{(NL_1\times N)}{\Gamma_2} &= diag\,(\beta_{111}, ..., \beta_{1i1}, ..., \beta_{1N1}), \\ \underset{(L_2\times N)}{\Gamma_3} &= (\beta_{112}, ..., \beta_{1i2}, ..., \beta_{1N2}). \end{aligned}$$

Denote

$$\underset{N(1+L_1+L_2)}{\delta} = \left(\beta_0', \beta_{11}', \beta_{12}'\right)',$$

where

$$\beta_0 = (\beta_{01}, ..., \beta_{0i}, ..., \beta_{0N})', \ \beta_{11} = \left(\beta_{111}', ..., \beta_{1i1}', ..., \beta_{1N1}'\right)', \ \beta_{12} = \left(\beta_{112}', ..., \beta_{1i2}', ..., \beta_{1N2}'\right)'.$$

Denote $\gamma = vec\,(\Gamma)$ [4]. Then, we can find a linear relationship between the unrestricted parameters γ and the restricted parameter δ, which is of interest:

$$\underset{N(NL_1+L_2+1)\times 1}{\gamma} = G\delta, \tag{19}$$

where

$$\underset{(N(NL_1+L_2+1)\times N(1+L_1+L_2))}{G} = \begin{pmatrix} I_N & 0 & 0 \\ 0 & K_{NL_1,N}\left(diag(i_1,\ldots,i_N)\otimes I_{L_1}\right) & 0 \\ 0 & 0 & K_{L_2,N} \end{pmatrix},$$

$K_{NL_1,N}$ and $K_{L_2,N}$ are the $N^2L_1\times N^2L_1$ and $NL_2\times NL_2$ commutation matrices, respectively, (*i.e.*, when A is an $(m\times n)$ matrix, $vec(A') = K_{m,n}vec(A)$), and i_k denotes the k^{th} column of the $N\times N$ identity matrix I_N.

To estimate the restricted parameter of interest δ efficiently, in this paper we suggest to follow the minimum distance (MD) method proposed by Moon and Schorfheide (2002) and Elliott (2000). For this, we assume the following. Denote $V_t = \Delta X_t$. Assume that

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T} V_t \Rightarrow B_V(r) \equiv BM(\Omega_V),$$

where $\Omega_V = \sum_{h=-\infty}^{\infty} E(V_0V_h')$, a positive definite matrix. Let $\Omega_{uu.V} = \Omega_{uu} - \Omega_{Vu}'\Omega_{VV}^{-1}\Omega_{Vu}$, where $\Omega_{Vu} = \sum_{h=-\infty}^{\infty} E(V_0u_h')$. Let $\tilde{B}_V(r) = \left(1, B_V(r)'\right)'$. Define $D_{T^*} = diag\left(\sqrt{T^*}, T^*I_{NL_1+L_2}\right)$. $F_{T^*} = diag\left(\sqrt{T^*}I_N, T^*I_{NL}\right)$.

Step 1 We first estimate Γ efficiently by running some efficient system cointegrating regression of y_t on $\tilde{X}_t$. For this, one can use the FM method (*e.g.*, Phillips and Hansen, 1989), the CCR method (*e.g.*, Park, 1992), or the DOLS method (*e.g.*, Saikonnen, 1991, Stock and Watson, 1993). Denote this estimator $\hat{\Gamma}$ and let $\hat{\gamma} = vec\left(\hat{\Gamma}\right)$. Under regularity conditions, $\hat{\Gamma}$ is the efficient estimator of the unrestricted coefficients Γ (see Phillips, 1991), and its limit is

$$\left(D_{T^*}\otimes I_N\right)(\hat{\gamma}-\gamma) \Rightarrow MN\left(0, \left(\int_0^1 \tilde{B}_V(r)\,\tilde{B}_V(r)'\,dr\right)^{-1}\otimes \Omega_{uu.V}\right).$$

Step 2 Estimate δ using the minimum distance (MD) method. For this, assume that $\hat{\Omega}_{uu.V}$ is a consistent estimator of $\Omega_{uu.V}$. Let $\hat{W} = \left(\sum_{t=1}^{T^*} D_{T^*}^{-1}\tilde{X}_t\tilde{X}_t'D_{T^*}^{-1}\right)\otimes\hat{\Omega}_{uu.V}^{-1}$. The efficient MD estimator $\hat{\delta}$ is found by minimizing

$$\min_{\delta}\,(\hat{\gamma}-G\delta)'\,\hat{W}\,(\hat{\gamma}-G\delta),$$

which yields the GLS solution

$$\hat{\delta} = \left(G'\hat{W}G\right)^{-1}G'\hat{W}\hat{\gamma}$$

Notice that $\hat{W} \Rightarrow W = \int_0^1 \tilde{B}_V(r)\,\tilde{B}_V(r)'\,dr\otimes\Omega_{uu.V}^{-1}$. According Moon and Schorfheide (2002), we have

$$F_{T^*}\left(\hat{\delta}-\delta\right) \Rightarrow MN\left(0, (G'WG)^{-1}\right). \tag{20}$$

Having the limit of $\hat{\delta}$ in (20), it is straightforward to construct a Wald test for restrictions on parameter δ, and we omit it.

[4] $vec\,(A)$ denotes a vector that stacks the columns of the matrix A.

3 Simulation comparison

In this section, we want to compare the relative merit of the various estimators of the integrated SUR model with no common regressor. The data generating process that will be used for this purpose contains N equations with L integrated regressors in each equation (for a total of NL integrated regressors):

$$\begin{aligned} y_{1t} &= \beta_{01} + \beta_{11,1}x_{1t,1} + \ldots + \beta_{11,L}x_{1t,L} + u_{1t} \\ &\vdots \\ y_{Nt} &= \beta_{0N} + \beta_{1N,1}x_{Nt,1} + \ldots + \beta_{1N,L}x_{Nt,L} + u_{Nt} \end{aligned}$$

where x_{it} are correlated random walks:

$$\Delta x_{it} = v_{it}, \quad i = 1, \ldots, N.$$

We assume that $u_t = (u_{1t}, ..., u_{Nt})'$ is correlated with $v_t = (v'_{1t}, ..., v'_{Nt})'$:

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim i.i.d.N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \varrho & \Sigma \\ \Sigma & \Phi \end{pmatrix} \right).$$

where each submatrix has a particular form. The covariance matrix is of dimension $(N + NL) \times (N + NL)$. The first submatrix, ϱ is the covariance matrix of u_t and has the form:

$$\underset{N \times N}{\varrho} = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

that is the parameter ρ controls the degree of correlation between the equations and will therefore affect the efficiency gains of GLS relative to OLS estimators.

On the other hand, the submatrix Σ is the covariance matrix between the regressor innovations and regression disturbances. This will therefore affect the behavior of the static OLS and GLS estimators of (3). We suppose that there is a constant correlation between each integrated regressor and the disturbance of the equation in which this regressor enters. The submatrix Σ of size $N \times NL$ therefore takes the form:

$$\Sigma = \begin{pmatrix} \sigma & \cdots & \sigma & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \sigma & \cdots & \sigma & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \sigma & \cdots & \sigma \\ & \underbrace{\qquad}_{L} & & & \underbrace{\qquad}_{L} & & & \underbrace{\qquad}_{L} & \end{pmatrix}$$

Finally, the covariance matrix of v_t, denoted by Φ, assumes that the regressors are uncorrelated within an equation, but that regressors across equations have a constant correlation ϕ.This $NL \times NL$ matrix therefore takes the form:

$$\Phi = \begin{pmatrix} & & & \cdots & \phi & \cdots & \phi \\ & I_L & & \cdots & \vdots & \ddots & \vdots \\ & & & \cdots & \phi & \cdots & \phi \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \phi & \cdots & \phi & \cdots & & & \\ \vdots & \ddots & \vdots & \cdots & & I_L & \\ \phi & \cdots & \phi & \cdots & & & \end{pmatrix}$$

The parameter ϕ will control the relative efficiency of system versus individual equation OLS.

We consider three combinations of N and L. We look at the case of 2 equations with a single regressor in each equation ($N = 2,\ L = 1$), 2 equations with 2 regressors ($N = 2,\ L = 2$), and 4 equations with a single regressor. ($N = 4,\ L = 1$) This will allow us to look at the impact of increasing the number of equations in the system and the number of variables per equation. We set the three covariances (ρ, σ, ϕ) to extreme values. We choose the largest values that will keep the full covariance matrix Ω positive definite for all three combinations of N and L we consider. Therefore, we look at the cases $(0, 0, 0)$, $(0.3, 0.3, 0.3)$, $(0.9, 0, 0)$, $(0, 0.7, 0)$, and $(0, 0, 0.4)$.

Without loss of generality, we set all coefficients to 1. We want to compare the constant and slope coefficient estimates $IDOLS$, $SDOLS$, $DGLS$, $FM-OLS$ [5], $FM-GLS$, and $SUCCR$.[6] The sample sizes chosen are $T = 100$ and $T = 300$, and the number of replications is 10,000. The number of leads and lags in the dynamic estimators (K) is selected by BIC,and the maximum number of lags is set at 4 for $T = 100$ and 7 for $T = 300$.[7] We also report the rejection frequency of the hypotheses on the first constant $\beta_{01} = 1$ and slope $\beta_{11,1} = 1$ for each estimator based on individual t-tests using the asymptotic critical values at level $\alpha = 5\%$. All long-run covariance matrices here and in the rest of the paper are estimated using the Andrews (1991) procedure with data-based bandwidth and quadratic spectral kernel. Because it made very little difference, we did not recompute the long-run covariance matrices after estimating each regression. In other words, the long-run covariance matrices are computed using the residuals from the static OLS regression. There is only one case where recomputing the covariance matrix made a difference is with 2 endogenous integrated regressors ($N = 2$, $L = 2$, $\sigma = 0.7$). In this case, the static OLS estimator is very biased and this leads to test statistics that are too small and gross underrejection. We recomputed the covariance matrix in this case.

[5] There is an error in $\hat{\beta}_{FM-SOLS}$ defined in Moon (1999). The component $\hat{\pi}_m$ should be

$$\hat{\pi}_m = \begin{pmatrix} 0 \\ \hat{\Delta}_{vu}^{m,m} - \left(\hat{\Delta}_{vv}^{m,\cdot}\right)\left(\left(\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\right)_{m,\cdot}\right)' \end{pmatrix}.$$

[6] We thank Masao Ogaki for providing us with his code for SUCCR.

[7] The number of leads and lags, K, is chosen to be close to $T^{1/3}$, considering the restriction in the proposition.

Since the efficiency gains due to *GLS*-type estimators come from a non-diagonal long-run covariance matrix, it is important to see how each parameter affects it. For the special case with 2 equations and a single integrated regressor the long-run covariance matrix $\Omega_{uu.v}$ is

$$\Omega_{uu.v} = \begin{pmatrix} 1 - \frac{\sigma^2}{1-\phi^2} & \rho + \frac{\phi\sigma^2}{1-\phi^2} \\ \rho + \frac{\phi\sigma^2}{1-\phi^2} & 1 - \frac{\sigma^2}{1-\phi^2} \end{pmatrix}$$

This expression highlights the role played by the three parameters of the DGP as described above. For the last four parameter designs, this matrix is therefore:

$$\begin{pmatrix} 0.67 & 0.33 \\ 0.33 & 0.67 \end{pmatrix}, \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}, \begin{pmatrix} 0.51 & 0 \\ 0 & 0.51 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so that we can expect *GLS* to be more efficient than *OLS* for the second and particularly the third design since the off-diagonal elements are much larger in this case. The variance in the fourth scenario should be cut by about half for all estimators with no change to the relative merit of each relative to the base case. Finally, the last design should give identical results to the baseline case.

The results of the experiments are presented in table 1 for $T = 100$ and table 2 for $T = 300$. The first line for each set of parameters gives the results for the constant the second line gives the results for the first slope parameter.

The main features of the results are:

- All estimators are essentially unbiased even with the smaller sample size. The only exception is the FM estimators in the endogenous regressor case ($\sigma = 0.7$) that have a bias of about 10% upward for $T = 100$. This bias disappears with $T = 300$;

- The convergence of the estimates of the slope coefficient $\left(\hat{\beta}_{11}\right)$ is much faster than that of the constant, as expected given the different rates of convergence;

- There are large efficiency gains in using GLS in cases where ρ is high. These are cases where the cross-equation correlation is high;

- All estimators suffer from some mild size distortions with $T = 100$. The size of the tests is much improved with $T = 300$. FM-OLS and FM-GLS have more severe distortions, and our DGLS and SDOLS estimators do very well in controlling the size of the tests;

- There are notable efficiency gains of using system methods relative to individual equation methods when the regressors are endogenous ($\sigma \neq 0$) and correlated ($\phi \neq 0$). The only place where this is relevant is for our case 2 with 2 integrated regressors. The magnitude of either σ or ϕ individually is not important as the expression for the long-run covariance matrix above highlights. However, a high value of σ on its own would bias static estimators;

- The two fully-modified estimators do not handle well the case of endogenous regressors ($\sigma = 0.7$). These two estimators are biased, more variable, and have large size distortions;
- *SUCCR* seems to provide the most efficient estimator among those studied here. It does suffer from more severe size distortions than our estimators. Its performance in the endogenous regressors case is better than the fully-modified case, but not as good as that of our DOLS or DGLS estimators;
- Increasing the number of regressors and equations leads to size distortions. The effect of doubling the number of regressors or doubling the number of equations is about the same.

Overall, our DGLS estimator performs well in all cases relative to its competitors and has substantial advantages in some cases. In particular, it seems to control size well and to handle endogenous regressors admirably.

4 Empirical Application: Testing for PPP

As already discussed, there is a large literature on testing PPP. The tests carried out in this paper are based on the system of regressions:

$$\ln s_{it} = \beta_{0i} + \beta_{1i} \ln (p_{it}) + \beta_{2i} \ln (p_t^*) + \varepsilon_{it} \tag{21}$$

We test PPP in two steps, first looking at the proportionality condition while assuming that symmetry holds, and then testing the symmetry condition directly. We see four main advantages to this approach:

1. Separate tests of symmetry and proportionality: looking at these two conditions separately will provide us with evidence on the sources of rejections of PPP in most other studies. Our new tools provide efficient tests of the symmetry condition in (21) which consists of testing $\beta_{1i} = -\beta_{2i}$ and of the proportionality condition which consists of testing that the coefficient on relative prices is 1. In tests of symmetry, there is a regressor that is common to all equations (the price level of the numeraire country) which makes it necessary to use the minimum distance estimator of section 2.2. There is no common regressor in our tests of proportionality and so the efficient *DGLS* estimator of section 2.1 can be applied;
2. PPP is the null hypothesis: in classical testing methodology, there is a strong bias towards the null hypothesis and we do not reject it unless there is strong evidence against it. Hence, it seems natural to specify the null hypothesis as the theory to be tested, in this case PPP. Most studies reverse the role of the null and alternative hypotheses and in our opinion this alters the interpretation of test results in undesirable ways. If we cannot reject the null hypothesis of real exchange rate nonstationarity for example, it is unclear whether that is because PPP does not hold or because the selected tests have low power;

3. The class of tests is invariant to the choice of numeraire country: under the null hypothesis of PPP, a change in the numeraire country still leads to a regression of the form (21) with $\beta_{1i} = 1$ and $\beta_{2i} = -1$. Of course, the properties of the error term ε_t depend on the choice of numeraire and therefore for a given sample, our evidence will differ on the presence or absence of PPP. However, asymptotically, the inference obtained will remain valid regardless of the choice of numeraire if the null hypothesis is true. Under the alternative hypothesis, this is not the case, and the power of our test will depend on which country is used as numeraire;

4. Individual tests: our methodology has the distinct advantage of using multivariate information to improve efficiency while preserving the ability of carrying out tests of PPP on individual currencies. Panel unit root tests do not allow for this possibility as the usual null hypothesis is that all real exchange rates in the panel have a unit root. When this null hypothesis is rejected, it is not clear how to interpret the results and in particular how to decide which countries have stationary real exchange rates and which ones do not.

5. No need for cointegration tests: a further advantage is that it bypasses the difficulties associated with unit root and cointegration tests since if (21) is not a cointegration regression, that is if ε_{it} is non-stationary for any value of β_{0i}, β_{1i} and β_{2i}, our test statistics will diverge to infinity (equation (21) becomes a spurious regression, see theorem 1(d) in Phillips (1986)). In this case, our test will correctly conclude that PPP is not supported by the data;

6. Efficient estimators: finally, our tests are based on efficient estimators. Other multivariate approaches, for example Flôres et al. (1999), cannot claim efficiency since the estimator of an AR(1) model does not have a mixed normal distribution. Only with estimators that have a limiting mixed normal distribution can meaningful efficiency statements be made.

Other work using this framework to test PPP include Edison, Gagnon, and Melick (1995) who find that tests of PPP based on the coefficients of the cointegrating vector tend to be more powerful than simple cointegration tests. This is intuitive since we are testing against a more specific alternative than a general cointegration test. Cheung and Lai (1993) also look at tests on the coefficients of the cointegrating vector and find that these are more powerful than cointegration tests. Finally, Li (1999) used a Bayesian approach to conclude that the data does not support PPP at the monthly frequency but is more supportive at the quarterly and annual frequencies.

The data we employ covers the entire recent float for the G-7 countries (1974-1998, 100 quarterly observations per country). We end the sample in 1998 to avoid potential impacts of the introduction of the Euro. The data was obtained from IFS and consists of monthly averages of the bilateral exchange rates relative to the US dollar and national price indices. We consider three separate prices indices: consumer prices, producer prices, and GDP deflators. [8]

Table 3 presents the short-run correlation matrix of changes in the 6 log of real exchange rates relative to the US dollar using consumer prices and is similar to table 1 in Flores et al. (1999). Three explanations can account for the

[8] Producer prices are not available for France and starts in 1981 for Italy in IFS. We have deleted these two countries from the analysis when using producer prices.

large correlation, in particular among European countries. The first one is the presence of various nominal exchange rate co-ordination mechanisms among European countries over the period. The second source of correlation, as pointed out by O'Connell (1998), is the use of a numeraire country (in this case the United States). Finally, the international transmission of shocks among these countries leads to correlation among their real exchange rates.

Table 3. Short-run correlation matrix of (log) real exchange rate changes

Canada	1.000					
France	0.031	1.000				
Germany	0.097	0.909	1.000			
Italy	0.121	0.835	0.818	1.000		
Japan	0.035	0.466	0.494	0.330	1.000	
United Kingdom	0.187	0.655	0.644	0.701	0.390	1.000

Table 4 is identical to the previous table, but it compares the *long-run* correlation of changes in the log real exchange rates among the countries involved. There are large off-diagonal entries, especially among European countries, but also between Japan and Europe. Canada is the only country with small (and sometimes negative) correlation with the other countries in the sample. These long-run correlations were neglected by previous authors and will be used to gain efficiency in testing for PPP.

Table 4. Long-run correlation matrix of log real exchange rate changes

Canada	1.000					
France	−0.156	1.000				
Germany	−0.110	0.929	1.000			
Italy	0.001	0.851	0.841	1.000		
Japan	−0.155	0.557	0.615	0.380	1.000	
United Kingdom	0.007	0.713	0.563	0.690	0.268	1.000

4.1 Tests of proportionality

To begin with, we impose symmetry ($\beta_{1i} = -\beta_{2i}$) to directly test the proportionality condition. We thus look at regressions:

$$\ln s_{it} = \beta_{0i} + \beta_{1i} \ln (p_{it}/p_t^*) + \varepsilon_{it} \quad i = 1, ..., N \tag{22}$$

and test the null hypothesis that $\beta_{1i} = 1$ for each country i. Since there is no common regressor in this case, we can apply our efficient estimator of section 2.1 to this set of regressions. The condition that no cointegration exists among regressors reduces to the condition that relative prices are not cointegrated. We therefore perform residual-based tests for cointegration on our data, and the results are presented in table 5 below along with the appropriate MacKinnon (1996) critical values for all possible normalizations (we have included a linear trend as some relative prices exhibit trending behavior). It is not necessary to repeat most tests with Germany as numeraire since the test statistics are identical. To see this, suppose that we test for cointegration with normalization on country 1's relative price relative to the United States. The estimated regression is:

$$\ln\left(\frac{p_t^{CAN}}{p_t^{US}}\right) = \gamma_0 + \gamma_1 t + \gamma_2 \ln\left(\frac{p_t^{FR}}{p_t^{US}}\right) + \gamma_3 \ln\left(\frac{p_t^{GE}}{p_t^{US}}\right) + \gamma_4 \ln\left(\frac{p_t^{IT}}{p_t^{US}}\right) + \gamma_5 \ln\left(\frac{p_t^{JA}}{p_t^{US}}\right) + \gamma_6 \ln\left(\frac{p_t^{UK}}{p_t^{US}}\right) + \varepsilon_{it}$$

If we subtract $\ln\left(p_t^{GE}\right)$ and add $\ln\left(p_t^{US}\right)$ to each side of this equation and collect terms, we obtain:

$$\ln\left(\frac{p_t^{CAN}}{p_t^{GE}}\right) = \gamma_0 + \gamma_1 t + \gamma_2 \ln\left(\frac{p_t^{FR}}{p_t^{GE}}\right) + (1 - \gamma_2 - \gamma_3 - \gamma_4 - \gamma_5) \ln\left(\frac{p_t^{US}}{p_t^{GE}}\right) + \gamma_4 \ln\left(\frac{p_t^{IT}}{p_t^{GE}}\right) + \gamma_5 \ln\left(\frac{p_t^{JA}}{p_t^{GE}}\right) + \gamma_6 \ln\left(\frac{p_t^{UK}}{p_t^{GE}}\right) + \varepsilon_{it}$$

which is the same regression and gives numerically identical test statistics. Therefore, the only test that we need to carry out with Germany as numeraire is the one that normalizes on the US relative price level.

Table 5. Cointegration tests of (log) relative price levels
(* indicates significant at the 5% level)

		Consumer prices			Producer prices			GDP deflators		
Numeraire	Normalization	Z_t	Z_α	ADF	Z_t	Z_α	ADF	Z_t	Z_α	ADF
	Canada	-2.471	-13.085	-2.393	-2.015	-9.157	-1.933	-2.226	-11.625	-2.297
	France	-3.766	-26.028	-3.351	-	-	-	-3.521	-21.351	-3.889
	Germany	-4.012	-27.584	-3.117	-2.580	-14.962	-2.724	-3.418	-22.302	-3.365
U. S.	Italy	-4.030	-28.696	-3.797	-	-	-	-3.561	-22.638	-4.157
	Japan	-4.618	-33.202	-3.081	-3.203	-23.060	-3.475	-10.472*	-84.004*	-2.966
	U.K.	-4.529	-31.598	-5.650*	-2.644	-11.598	-3.247	-4.726	-23.065	-5.842*
Germany	U.S.	-2.956	-16.634	-3.588	-3.039	-19.810	-3.078	-2.629	-14.209	-2.970
5% critical value		-5.188	-42.141	-5.188	-4.576	-34.147	-4.576	-5.188	-42.141	-5.188

There are a few rejections of the null hypothesis of no cointegration, but these represent the exception rather than the rule. The only rejections occur with the U.K. CPI and GDP deflator using the ADF test and the Japanese GDP deflator using the Z_t and Z_α tests. It thus appears reasonable to assume that relative prices are not cointegrated in the following.

Table 6 presents the results of estimating equation (22) using all 6 estimators in the Monte Carlo experiment above and using the three different sets of price indices with the United States as the numeraire. The number of leads and lags was selected using BIC. It turns out that the criterion often prefers a small number of leads and lags, especially for system methods.

Looking at first at the results for consumer prices, we can easily reject the joint null hypothesis of proportionality at any reasonable significance level with all three system methods ($SDOLS$, $DGLS$, and $SUCCR$). The Wald statistics are reported in the last column of the table and each has a $\chi^2(6)$ distribution under the null hypothesis. Turning to individual tests, we cannot reject proportionality for Canada, Italy, and the UK when using our efficient GLS estimator. The non-rejection for these three countries is robust across methods as well. We can reject the null hypothesis for Italy and UK only with SUCCR and for Canada with FM-OLS. On the other hand, Germany is only rejected with our DGLS estimator. On the basis of the reported standard errors, dynamic GLS generally provides more precise estimates than any either individual OLS or DOLS, suggestive of efficiency gains by accounting for cross-country correlations. The use of system OLS is not particularly effective in providing more precise estimates because the much higher number of additional parameters that must be estimated dominate the efficiency gains from the cross-country correlations. The reported standard errors for SUCCR are much smaller than for those for any other method, and this leads to a very large Wald statistic for the joint test.

As emphasized by other authors, it is important to look at system methods in testing for PPP. However, it is not sufficient to reject proportionality for Italy and Germany. It is important to also look at cross-country correlations through the use of GLS. This provides efficiency gains and increased power sufficient to reject the hypothesis that the proportionality condition implied by purchasing power parity is a reasonable description of long-run exchange rate behavior.

Table 6. Tests of proportionality condition - Estimates of β_{1i} with United States as numeraire (standard errors in parentheses, * indicates significantly different from 1 at 5% level)

		Canada	France	Germany	Italy	Japan	U.K.	Joint test (p-value)
Consumer prices	$IDOLS$	1.083 (1.044)	2.390* (0.417)	1.059 (0.301)	1.132 (0.161)	1.891* (0.245)	0.628 (0.325)	
	$SDOLS$	−0.345 (1.516)	2.484* (0.489)	0.296 (0.589)	1.205 (0.278)	1.085 (0.539)	0.777 (0.560)	18.43 (0.005)
	$DGLS$	1.279 (0.414)	1.629* (0.078)	0.677* (0.120)	0.898 (0.082)	1.580* (0.247)	0.638 (0.273)	143.42 (0.000)
	$FM-OLS$	−1.112* (1.021)	1.645 (0.329)	0.925 (0.265)	0.846 (0.116)	2.099* (0.239)	0.869 (0.234)	
	$FM-GLS$	1.106 (0.333)	1.567* (0.067)	0.998 (0.069)	0.963 (0.043)	2.153* (0.125)	0.923 (0.129)	
	$SUCCR$	1.098 (0.182)	1.608* (0.051)	0.921 (0.044)	0.912* (0.026)	1.934* (0.104)	0.694* (0.063)	591.45 (0.000)
Producer prices	$IDOLS$	1.100 (0.114)	-	1.814* (0.406)	-	1.901* (0.177)	0.600 (0.206)	
	$SDOLS$	0.933 (0.163)	-	2.369 (0.827)	-	2.004* (0.269)	0.586* (0.288)	25.07 (0.000)
	$DGLS$	1.012 (0.114)	-	1.541 (0.328)	-	1.679* (0.144)	0.833* (0.201)	54.60 (0.000)
	$FM-OLS$	1.036 (0.107)	-	1.681 (0.378)	-	1.833* (0.157)	0.542* (0.171)	
	$FM-GLS$	1.066 (0.082)	-	1.249 (0.206)	-	1.616* (0.102)	0.659* (0.131)	
	$SUCCR$	1.076* (0.070)	-	1.175* (0.155)	-	1.578* (0.088)	0.699* (0.124)	155.66 (0.000)
GDP deflators	$IDOLS$	1.548 (1.356)	1.942* (0.449)	0.741 (0.573)	0.833 (0.112)	2.286* (0.268)	0.113* (0.278)	
	$SDOLS$	0.054 (1.823)	1.926 (0.631)	−0.269 (0.832)	0.814 (.205)	1.248 (0.637)	−0.022* (0.430)	16.75 (0.010)
	$DGLS$	1.381 (0.647)	1.027 (0.135)	0.700 (0.196)	0.694* (0.071)	1.175 (0.217)	0.206* (0.191)	69.85 (0.000)
	$FM-OLS$	1.142 (1.125)	0.929 (0.285)	0.733 (0.481)	0.685* (0.078)	1.548 (0.285)	0.240* (0.181)	
	$FM-GLS$	0.686 (0.469)	1.085 (0.092)	0.753 (0.150)	0.779* (0.037)	2.685* (0.119)	0.863 (0.110)	
	$SUCCR$	1.083 (0.261)	0.915 (0.071)	0.827 (0.126)	0.696* (0.026)	2.440* (0.100)	0.513* (0.058)	839.71 (0.000)

The results with producer prices are slightly different from those with consumer prices for the four countries for which we have data. We reject proportionality for Japan and the UK and fail to reject for Canada and Germany. As emphasized by other researchers, the use of this price index is more favorable to PPP. The statistics for the joint null hypothesis tend to be smaller. Note that the appropriate distribution for these statistics is $\chi^2(4)$ since we have data for only 4 countries. Finally, the results with GDP deflators are similar. As a rule, we reject the proportionality null hypothesis for

Italy, Japan, and the United Kingdom, irrespective of the method used, while we fail to reject for Canada, France, and Germany.

Table 7 reports the same information as table 6 but using Germany as the numeraire country. The results are more sympathetic to proportionality. There are fewer rejections based on individual tests, and the joint null of proportionality cannot be rejected at the 5% level using the DOLS estimator for the consumer prices and GDP deflators. Another difference is that the use of producer prices is less supportive of proportionality than consumer prices or GDP deflators as evidenced by the higher value of the Wald statistics (despite the lower degrees of freedom).

Table 7. Tests of proportionality condition - Estimates of β_{1i} with Germany as numeraire
(standard errors in parentheses, * indicates significantly different from 1 at 5% level)

		Canada	France	U.S.	Italy	Japan	U.K.	Joint test (p-value)
Consumer prices	$IDOLS$	1.051 (0.327)	0.924 (0.050)	1.059 (0.301)	0.850 (0.121)	2.374* (0.650)	0.765 (0.166)	
	$SDOLS$	0.319 (0.600)	0.879 (0.071)	0.296 (0.589)	0.756 (0.168)	1.004 (0.921)	0.442* (0.251)	5.28 (0.509)
	$DGLS$	1.360 (0.304)	0.890* (0.049)	1.129 (0.285)	0.842 (0.096)	1.328 (0.363)	0.547* (0.161)	33.03 (0.000)
	$FM-OLS$	1.385 (0.265)	0.944 (0.035)	1.072 (0.265)	0.934 (0.071)	2.875* (0.614)	0.841 (0.106)	
	$FM-GLS$	1.195 (0.178)	0.956 (0.028)	0.774 (0.150)	0.920 (0.055)	2.354* (0.234)	0.831* (0.085)	
	$SUCCR$	1.067 (0.099)	0.942* (0.018)	0.803* (0.090)	0.857* (0.025)	1.651* (0.084)	0.756* (0.033)	206.90 (0.000)
Producer prices	$IDOLS$	1.515* (0.204)	-	1.814* (0.408)	-	1.645* (0.153)	1.401 (0.258)	
	$SDOLS$	1.607 (0.365)	-	2.369 (0.831)	-	1.486* (0.200)	1.077 (0.180)	23.03 (0.000)
	$DGLS$	1.108 (0.129)	-	0.883 (0.293)	-	1.538* (0.147)	0.778 (0.133)	53.23 (0.000)
	$FM-OLS$	1.353 (0.183)	-	1.616 (0.380)	-	1.590* (0.149)	0.868 (0.094)	
	$FM-GLS$	1.192 (0.100)	-	1.150 (0.210)	-	1.551* (0.114)	0.683* (0.074)	
	$SUCCR$	1.274* (0.072)	-	1.367* (0.126)	-	1.564* (0.109)	0.741* (0.066)	168.26 (0.000)
GDP deflators	$IDOLS$	1.074 (0.606)	0.867 (0.096)	0.741 (0.632)	0.828 (0.113)	1.542 (0.377)	0.713 (0.271)	
	$SDOLS$	0.074 (0.894)	0.784 (0.126)	−0.269 (0.916)	0.725 (0.172)	0.667 (0.516)	0.332 (0.347)	4.29 (0.638)
	$DGLS$	1.379 (0.425)	0.911 (0.061)	0.908 (0.490)	0.779* (0.077)	1.168 (0.335)	0.453* (0.185)	20.94 (0.002)
	$FM-OLS$	1.703 (0.520)	0.921 (0.063)	1.172 (0.530)	0.859* (0.070)	1.590 (0.313)	0.700 (0.155)	
	$FM-GLS$	1.384 (0.284)	0.929 (0.041)	0.985 (0.316)	0.819* (0.041)	2.161* (0.209)	0.763* (0.110)	
	$SUCCR$	1.301 (0.222)	0.925* (0.031)	0.837 (0.274)	0.818* (0.035)	2.368* (0.202)	0.778* (0.075)	138.84 (0.000)

The results we obtain are therefore sensitive to the choice of numeraire. This is to be expected given that the null hypothesis of PPP is rejected by the data and numeraire invariance holds only under the null hypothesis. The various estimators seem to give similar results. However, our system dynamic OLS seems to be heavily penalized in terms of precision by the large number of parameters it has to estimate. The efficiency gain coming from the addition of leads and lags of first differences of relative prices from other countries is outweighed by the burden of having to estimate the additional parameters. Based on the more efficient system estimators (DGLS and SUCCR), the joint null proportionality is easily rejected regardless of the choice of numeraire or price index. Individual tests suggest that Japan is the country where we reject proportionality the most emphatically, while Canada seems to be the country for which proportionality is most difficult to reject.

4.2 Tests of symmetry

The second part in our investigation of sources of rejection of PPP is to test the symmetry condition. The general formulation of our system of equation is:

$$\ln s_{it} = \beta_{0i} + \beta_{1i} \ln (p_{it}) + \beta_{2i} \ln (p_t^*) + \varepsilon_{it}. \tag{23}$$

Symmetry implies that $\beta_{1i} = -\beta_{2i}$. Since we do not impose this condition, there is a common regressor in each equation in the form of the price level for the numeraire country, p_t^*. We therefore use the minimum distance estimator of section 2.2 to estimate the parameters of this system. The method first estimates the unrestricted form of (23) that is log nominal exchange rates are regressed on *all* log price levels using the Stock and Watson (1993) leads and lags estimator. The number of leads and lags to add is estimated using BIC again. In a second step, the restrictions that only the domestic and numeraire price levels appear in each equation are imposed using minimum distance. This is the efficient estimator in this context. We then test the null hypothesis of symmetry for each country individually and jointly. The results are presented in table 8 with the United States as numeraire and table 9 with Germany as numeraire.

Table 8. Tests of symmetry condition - Estimates of β_{1i} and β_{2i} with United States as numeraire (standard errors in parentheses, * indicates symmetry rejected at 5% level)

		Canada	France	Germany	Italy	Japan	U.K.	Joint test (p-value)
Consumer prices	β_{1i}	−2.520* (0.325)	1.704* (0.103)	−2.018* (0.200)	2.205 (0.165)	6.518 (0.963)	0.540* (1.121)	665.26 (0.000)
	β_{2i}	3.749* (0.359)	−0.405* (0.151)	2.307* (0.147)	−2.417 (0.310)	−6.346 (0.495)	2.464* (1.498)	
Producer prices	β_{1i}	0.516 (0.163)	–	1.042* (0.408)	–	2.895* (0.284)	1.127 (0.149)	23.16 (0.000)
	β_{2i}	−0.522 (0.215)	–	−0.540* (0.289)	–	−1.792* (0.132)	−1.376 (0.277)	
GDP deflators	β_{1i}	−1.347* (0.279)	2.249* (0.236)	−0.990* (0.187)	0.371* (0.291)	4.767* (0.518)	−2.929* (1.028)	1672.0 (0.000)
	β_{2i}	0.965* (0.307)	−5.937* (0.340)	−2.763* (0.162)	−4.608* (0.677)	−1.830* (0.284)	−2.875* (1.649)	

In almost all cases, symmetry is easily rejected. There are six non-rejections in the tables: with the United States as numeraire, we cannot reject symmetry for Italian and Japanese CPI, and Canadian and UK PPI. With Germany as numeraire, the only two non-rejections are for Canadian and Italian consumer prices.

Table 9. Tests of symmetry condition - Estimates of β_{1i} and β_{2i} with Germany as numeraire (standard errors in parentheses, * indicates symmetry rejected at 5% level)

		Canada	France	U.S.	Italy	Japan	U.K.	Joint test (p-value)
Consumer prices	β_{1i}	−0.299 (0.265)	1.180* (0.079)	2.648* (0.256)	0.510 (0.108)	9.786* (0.857)	−3.165* (0.490)	693.77 (0.000)
	β_{2i}	0.290 (0.515)	−1.642* (0.172)	−6.670* (0.482)	−0.734 (0.361)	−15.930* (0.861)	10.938* (1.157)	
Producer prices	β_{1i}	1.144* (0.160)	-	1.539* (0.230)	-	0.640* (0.294)	−0.040* (128)	41.24 (0.000)
	β_{2i}	−2.608* (0.517)	-	−2.897* (0.507)	-	−2.545* (0.487)	0.786* (0.371)	
GDP deflators	β_{1i}	−1.192* (0.090)	0.537* (0.032)	−1.684* (0.147)	−0.119* (0.047)	0.060* (0.120)	−0.516* (0.063)	2179.3 (0.000)
	β_{2i}	3.574* (0.136)	0.019* (0.061)	3.637* (0.200)	2.429* (0.146)	−0.622* (0.097)	2.234* (0.139)	

It is important to note that our rejection of PPP for a subset of countries based on failures of both the symmetry and proportionality conditions is perfectly consistent with recent findings of evidence for PPP based on panel unit root tests. These tests impose the symmetry and proportionality conditions, and a rejection of the null hypothesis that all real exchange rates have a unit root simply suggests that for *some* countries, PPP is a reasonable approximation to the behavior of their exchange rates. However, that methodology does not inform us on which country are likely to be supportive of PPP and which ones are not and what is the source of rejection of PPP. Our methodology suggests that on balance, Canada is the most likely country for which PPP holds relative to the United States. We also find that proportionality has, on the whole, more support in the data than symmetry, as evidenced by the higher Wald statistics for the joint hypothesis of symmetry.

4.3 A model with a proxy for nontraded goods

Since PPP does not seem to be a reasonable approximation for most countries, we investigated an alternative model of exchange rate determination. One common explanation for the failure of PPP is the presence of nontraded goods for which the law of one price does not need to hold. This is also why some authors prefer to use producer prices rather than consumer prices when testing for PPP since the former is thought to contain a smaller proportion of goods that are not traded internationally. We try to control for the presence of nontraded goods by considering the model proposed by Kakkar and Ogaki (1999) . In essence, their model consists of using the ratio of producer to consumer prices as a proxy for the relative price of tradable versus non-tradable goods. The model to be estimated is therefore:

$$\ln q_{it} = \gamma_{0i} + \gamma_{1i,1} \ln w_{it} + \gamma_{1i,2} \ln w_t^* + u_{it} \tag{24}$$

where q_{it} is the real exchange rate for country i, w_{it} is the proxy for the relative price of tradable goods for the home country i and w_t^* is the same variable for the foreign country (the numeraire). As in Kakkar and Ogaki, we will use the ratio of producer to consumer prices as our definition of w_{it}. The real exchange rate is computed using our most comprehensive price index, the GDP deflator. The coefficients $\gamma_{1i,k}$ have the interpretation of the share of tradable goods in the overall price index for either the country of interest or the numeraire. We therefore expect these coefficients to lie between 0 and 1 in magnitude (the coefficient for the numeraire country should have a negative sign).

In the same way as before, the use of a numeraire country implies that there is a common regressor to all equations in the system. Thus, Kakkar and Ogaki could only estimate their model for one country at a time, Instead, we use the minimum distance estimator with the Stock and Watson estimator in the first stage and number of leads and lags selected by BIC. Table 8 reports results of the estimation of this system of equations.

Table 9. Estimates of $\gamma_{1i,1}$ and $\gamma_{1i,2}$ in Kakkar-Ogaki model with United States as numeraire (standard errors in parentheses)

	Canada	Germany	Japan	U.K.
$\gamma_{1i,1}$	1.664 (0.164)	3.690 (0.650)	2.108 (0.709)	2.985 (0.767)
$\gamma_{1i,2}$	-1.339 (0.137)	-1.038 (0.591)	0.682 (0.916)	-2.035 (0.526)

The coefficients that we find are close to those in Kakkar and Ogaki with the exception of the foreign price of tradables for Japan (Kakkar and Ogaki found -4.09 using post Bretton-Woods data with real exchange rates defined using consumption deflators). All coefficients have the expected sign, again with the exception of the coefficient on the foreign relative price of tradables for Japan. However, most lie outside of the unit interval although it would not be possible to reject that they lie within that interval in some cases. Note that for the cases that are comparable to Kakkar and Ogaki (they have post Bretton Woods data for Canada, UK and Japan), our estimates have smaller standard errors, reflecting the efficiency gain in estimating the model as a system. We must however agree with Kakkar and Ogaki and conclude that there is little support for their model.

5 Conclusion

This paper has proposed new estimators of the SUR model with integrated regressors based on an augmented regression model. We have derived the asymptotic distributions of our estimators and indicated that a feasible generalized least squares estimator of the augmented model using the long-run covariance matrix is the most efficient among them. We have also suggested the use of a minimum distance estimator in the case where there exists a common regressor in each equation in the system.

Monte Carlo results suggest that our dynamic GLS estimator compares favorably with other estimators and improves noticeably upon them in some situations. Moreover, inference with this estimator has size close to its nominal level. Fully-modified estimators suffer from more noticeable size distortions.

An application of the methods to testing of purchasing power parity among G-7 countries demonstrates the importance of analyzing this issue in a system framework and taking into account cross-country correlations. Our approach allows for direct tests of the symmetry and proportionality conditions implied by PPP. With our dynamic GLS system method, we are able to reject proportionality for PPP for 3 of the 6 countries in our analysis when using the United States as numeraire and consumer prices. In this same case, we reject symmetry for 4 out of 6 countries. This casts doubts on the validity of PPP as a reasonable description of long-run exchange rate behavior and suggests that rejections of PPP are due to failure of both the symmetry and proportionality conditions.

6 Appendix

Proof of Proposition

The proof uses standard arguments and will be mostly omitted. It suffices to notice that we can write

$$G_{T^*}\left(\hat{b}_{DGLS}-b\right)=\left(\begin{array}{c} F_{T^*}\left(\hat{\beta}_{DGLS}-\beta\right) \\ \sqrt{T^*}\left(\widehat{\Pi}_{DGLS}-\Pi\right) \end{array}\right)=\left(\begin{array}{cc} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{array}\right)^{-1}\left(\begin{array}{c} B_1 \\ B_2 \end{array}\right),$$

where

$$\begin{aligned}
A_{11} &= \sum_{t=K+1}^{T-K} F_{T^*}^{-1} \tilde{x}_t \hat{\Omega}_{uu.v}^{-1} \tilde{x}_t' F_{T^*}^{-1} \\
A_{12} &= \begin{pmatrix} \frac{1}{\sqrt{T^*}} \sum_{t=K+1}^{T-K} F_{T^*}^{-1} \tilde{x}_t \hat{\Omega}_{uu.v}^{-1} \left(\Delta x_{t-K}' \otimes I_N \right) & \cdots & \frac{1}{\sqrt{T^*}} \sum_{t=K+1}^{T-K} F_{T^*}^{-1} \tilde{x}_t \hat{\Omega}_{uu.v}^{-1} \left(\Delta x_{t+K}' \otimes I_N \right) \end{pmatrix} \\
A_{22} &= \begin{pmatrix} \frac{1}{T^*} \sum_{t=K+1}^{T-K} \left(\Delta x_{t-K} \Delta x_{t-K}' \otimes \hat{\Omega}_{uu.v}^{-1} \right) & \cdots & \frac{1}{T^*} \sum_{t=K+1}^{T-K} \left(\Delta x_{t-K} \Delta x_{t+K}' \otimes \hat{\Omega}_{uu.v}^{-1} \right) \\ \vdots & \ddots & \vdots \\ \frac{1}{T^*} \sum_{t=K+1}^{T-K} \left(\Delta x_{t+K} \Delta x_{t-K}' \otimes \hat{\Omega}_{uu.v}^{-1} \right) & \cdots & \frac{1}{T^*} \sum_{t=K+1}^{T-K} \left(\Delta x_{t+K} \Delta x_{t+K}' \otimes \hat{\Omega}_{uu.v}^{-1} \right) \end{pmatrix},
\end{aligned}$$

and

$$\begin{aligned}
B_1 &= F_{T^*}^{-1} \sum_{t=K+1}^{T-K} \tilde{x}_t \hat{\Omega}_{uu.v}^{-1} \xi_t^* \\
B_1 &= \begin{pmatrix} \frac{1}{\sqrt{T^*}} \sum_{t=K+1}^{T-K} \left(\Delta x_{t-K} \otimes \hat{\Omega}_{uu.v}^{-1} \xi_t^* \right) \\ \vdots \\ \frac{1}{\sqrt{T^*}} \sum_{t=K+1}^{T-K} \left(\Delta x_{t+K} \otimes \hat{\Omega}_{uu.v}^{-1} \xi_t^* \right) \end{pmatrix}.
\end{aligned}$$

Then, with $\frac{K^3}{T} \to 0$ and $\sqrt{T} \sum_{s>|K|} \|\pi_s\| = o(1)$, we can use similar arguments to those in Saikkonen(1991) to show that as $T \to \infty$

$$\begin{aligned}
& F_{T^*} \left(\hat{\beta}_{DGLS} - \beta \right) \\
\Rightarrow & \left(\int_0^1 \tilde{B}_v(r) \Omega_{uu.v}^{-1} \tilde{B}_v'(r) \, dr \right)^{-1} \int_0^1 \tilde{B}_v(r) \Omega_{uu.v}^{-1} dB_{uu.v}(r) \\
\equiv & MN \left(0, \left(\int_0^1 \tilde{B}_v(r) \Omega_{uu.v}^{-1} \tilde{B}_v'(r) \, dr \right)^{-1} \right),
\end{aligned}$$

where $B_{uu.v}(r) = B_u(r) - \Omega_{uv} \Omega_{vv}^{-1} B_v(r)$. The derivation of the other two results is similar. ■

References

[1] Andrews, D.W.K. (1991): Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica*, 59, 817-858.

[2] Brillinger, D. (1975): *Time Series: Data Analysis and Theory*, Holt, Rinehart, and Winston.

[3] Cheung, Y.-W. and K. S. Lai (1993): Long-run Purchasing Power Parity During the Recent Float, *Journal of International Economics*, 34, 181-192.

[4] Edison, H. J., J. E. Gagnon, and W. R. Melick (1997): Understanding the Empirical Literature on Purchasing Power Parity: the post_Bretton Woods Era, *Journal of International Money and Finance*, 16, 1-17.

[5] Elliott, G. (2000): Estimating Restricted Cointegrating Vectors, *Journal of Business & Economic Statistics*, 18, 91–99.

[6] Flôres, R., P. Jorion, P. Preumont, and A. Szafarz (1999): Multivariate Unit Root Tests of the PPP Hypothesis, *Journal of Empirical Finance*, 6, 335-353.

[7] Froot, K. A. and K. Rogoff (1995): Perspectives on PPP and Long-Run Real Exchange Rates, in Grossman, G. M. and K. Rogoff, eds., *Handbook of International Economics, vol. III* , Elsevier Science.

[8] Grenander, U. and M. Rosenblatt (1957): *Statistical Analysis of Stationary Time Series*, New York, John Wiley and Sons.

[9] Kim, Y. (1990): Purchasing Power parity in the Long Run: A Cointegration Approach, *Journal of Money, Credit, and Banking*, 22, 491-503.

[10] Li, K. (1999): Testing Symmetry and Proportionality in PPP: A Panel-data Approach, *Journal of Business and Economics Statistics*, 17, 409-418.

[11] MacKinnon, J. G. (1996): Numerical Distribution Functions for Unit Root and Cointegration Tests, *Journal of Applied Econometrics*, 11, 601-618.

[12] Mark, N., M. Ogaki, and D. Sul (2000): Dynamic Seemingly Unrelated Cointegrating Regression, Mimeo.

[13] Moon, H.R. (1999): A Note on Fully-Modified Estimation of Seemingly Unrelated Regressions Models with Integrated Regressors, *Economics Letters*, 65, 25–31.

[14] Moon, H. R. and F. Schorfheide (2002): Minimum Distance Estimation of Time Series Models, *Econometric Theory*, 18, 1385-1407.

[15] O'Connell, P. G. J. (1998): The Overvaluation of Purchasing Power Parity, *Journal of International Economics*, 44, 1-20.

[16] Park, J. (1992): Canonical Cointegrating Regressions, *Econometrica*, 60, 119-143.

[17] Park, J. and M. Ogaki (1991): Seemingly Unrelated Canonical Cointegrating Regressions, University of Rochester working paper 280.

[18] Park, J. and P. C.B. Phillips (1988): Statistical Inference in Regressions with Integrated Processes: Part I, *Econometric Theory*, 4, 468-497.

[19] Pedroni, P. (2002): Purchasing Power parity: Tests in Cointegrated Panels, *Review of Economics and Statistics*, 83, 727-731.

[20] Pedroni, P. (2004): Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the Purchasing Power Parity Hypothesis, *Econometric Theory*, forthcoming.

[21] Phillips, P. C.B. (1986): Understanding Spurious Regressions in Econometrics, *Journal of Econometrics*, 33, 311-340.

[22] Phillips, P. C.B. (1991): Optimal Inference in Cointegrated Systems, *Econometrica*, 59, 283-306.

[23] Phillips, P. C.B. (1995): Fully Modified Least Squares and Vector Autoregressions, *Econometrica*, 63, 1023-1078.

[24] Phillips, P. C.B. and B. Hansen (1990): Statistical Inference in Instrumental Regressions with I(1) Processes, *Review of Economic Studies*, 57, 99-125.

[25] Phillips, P. C.B. and J. Park (1988): Asymptotic Equivalence of OLS and GLS in Regression with Integrated Regressors, *Journal of the American Statistical Association*, 83, 111–115.

[26] Phillips, P. C.B. and V. Solo (1992): Asymptotics for Linear Processes, *Annals of Statistics*, 20, 971-1001.

[27] Saikkonen, P. (1991): Asymptotically Efficient Estimation of Cointegration Regressions, *Econometric Theory*, 7, 1–21.

[28] Saikkonen, P. (1993): Estimation of Cointegration Vectors with Linear Restrictions, *Econometric Theory*, 9, 19–35.

[29] Saikkonen, P. (1995): Problems with the Asymptotic Theory of Maximum Likelihood Estimation in Integrated and Cointegrated systems, *Econometric Theory*, 11, 888–911.

[30] Sarno, L. and M. P. Taylor (2002): Purchasing Power Parity and the Real Exchange Rate, *IMF Staff Papers*, 49, 65-105.

[31] Stock, J. and M. Watson (1993): A Simple Estimator of Cointegrating Vectors in Higher Order Integrated System, *Econometrica*, 61, 783–820.

[32] Zellner A. (1962): An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests of Aggregation Bias, *Journal of the American Statistical Association*, 57, 500-509.

Table 1. Simulation results, $T = 100$

2 equations, 1 integrated regressor per equation

	Mean bias						Std. error						Size of 5% test					
$(\rho, \sigma, \phi) =$	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR
(0,0,0)	.001	.002	.000	.000	.001	-.002	.192	.193	.194	.193	.194	.185	7.3	7.2	7.4	8,1	8.7	8.6
	-.001	.000	.001	.001	.000	.000	.037	.037	.037	.036	.036	.036	7.5	7.3	7.9	8.6	9.4	9.4
(0.3,0.3,0.3)	.000	-.001	-.001	-.002	-.002	-.001	.184	.184	.179	.187	.182	.167	7.6	7.4	7.4	8.3	9.1	7.8
	.000	.000	.000	.002	.003	.003	.036	.035	.034	.035	.032	.031	7.8	7.3	7.7	8.7	9.5	8.9
(0.9,0,0)	.001	.001	.000	.000	.000	.000	.186	.188	.133	.189	.134	.132	6.7	6.4	6.4	7.4	7.9	8.1
	-.001	-.001	.000	-.001	.000	.000	.036	.036	.019	.035	.019	.018	7.1	7.0	6.4	8.3	9.1	9.1
(0,0.7,0)	.003	.003	.003	.002	-.002	.002	.137	.138	.139	.148	.150	.135	6.1	6.0	6.4	8.0	8.9	7.7
	.000	.000	.001	.004	.005	.006	.026	.026	.026	.028	.028	.027	5.9	5.9	6.6	9.0	10.1	9.0
(0,0,0.4)	.000	.000	-.001	-.001	-.002	.000	.190	.191	.192	.194	.196	.186	7.3	7.4	7.7	8.4	9.1	8.1
	.000	.000	.000	.000	.000	.001	.036	.036	.037	.036	.036	.035	7.0	7.0	7.7	8.5	9.7	8.7

2 equations, 2 integrated regressors per equation

	Mean bias						Std. error						Size of 5% test					
$(\rho, \sigma, \phi) =$	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR
(0,0,0)	-.003	-.003	-.003	-.002	-.002	.002	.263	.264	.266	.270	.270	.248	8.5	8.7	9.1	10.9	12.2	10.7
	.000	.000	.000	.000	.000	.000	.045	.045	.045	.045	.045	.043	8.8	8.9	9.3	12.3	13.3	12.6
(0.3,0.3,0.3)	-.001	-.001	.000	.000	.000	.002	.235	.219	.195	.233	.220	.191	9.8	7.7	7.3	10.5	12.8	9.8
	-.001	.000	.000	.004	.005	.004	.040	.037	.031	.039	.036	.031	9.8	7.6	7.1	12.0	15.4	11.7
(0.9,0,0)	-.002	-.002	-.001	-.003	-.001	.000	.265	.264	.175	.269	.184	.173	8.8	8.8	7.6	10.3	11.4	10.1
	.000	.000	.000	.000	.000	.000	.045	.044	.023	.043	.024	.024	8.5	8.3	6.5	11.1	12.6	12.4
(0,0.7,0)	.000	.000	.000	.003	.004	.000	.037	.037	.041	.179	.188	.106	2.3	8.0	10.6	11.4	17.4	3.5
	.000	.000	.000	.010	.012	.016	.006	.006	.007	.032	.035	.025	2.1	7.8	11.8	15.8	25.4	13.5
(0,0,0.4)	.001	.001	.001	.003	.003	-.001	.263	.264	.265	.270	.271	.249	8.3	8.3	8.6	10.6	11.8	10.4
	.000	.000	.000	-.001	-.001	.000	.045	.045	.045	.045	.045	.042	8.6	8.8	9.1	12.4	13.0	11.1

4 equations, 1 integrated regressor per equation

	Mean bias						Std. error						Size of 5% test					
$(\rho, \sigma, \phi) =$	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR
(0,0,0)	.000	.000	.000	.000	.000	-.001	.191	.200	.203	.197	.199	.188	8.1	8.2	9.1	9.7	11.2	10.5
	-.001	-.001	-.001	-.001	-.001	.001	.037	.039	.039	.037	.038	.036	8.9	8.9	10.4	10.5	13.0	12.2
(0.3,0.3,0.3)	.000	-.001	-.001	-.001	-.001	.004	.186	.193	.183	.192	.183	.170	8.3	8.3	9.2	9.8	11.5	10.6
	.000	.000	.002	.003	.004	.004	.035	.036	.033	.036	.033	.032	8.8	8.2	10.0	10.7	13.6	12.8
(0.9,0,0)	.000	-.001	.000	-.001	-.001	.000	.193	.201	.128	.200	.130	.127	8.2	8.2	8.0	10.0	10.8	10.7
	.000	.000	.000	.000	.000	.000	.037	.038	.016	.037	.017	.016	8.7	8.4	9.0	10.7	14.8	15.0
(0,0.7,0)	.001	.001	.001	.002	.002	.001	.138	.144	.146	.153	.157	.140	7.0	6.5	7.5	9.7	11.5	10.5
	.000	.000	.002	.006	.009	.007	.026	.027	.028	.030	.031	.028	7.2	6.6	8.5	11.1	14.6	12.7
(0,0,0.4)	-.001	-.002	-.002	-.001	-.001	-.001	.194	.201	.204	.199	.202	.191	8.4	8.0	9.1	9.5	11.1	10.9
	.000	.000	.000	.000	.000	.000	.037	.038	.039	.036	.037	.037	8.7	8.5	10.3	10.4	12.7	12.4

Results are computed over 10,000 replications. The first row for each parameter combination is for the constants and the second row provides results for the slope parameters. Mean estimate is the mean of the coefficient estimates obtained by each method, std. error is the standard deviation of those estimates over the replications, and the size is the frequency of rejection of the (correct) null hypothesis $H_0 : \beta_{1,1} = 1$ using a two-sided test. The parameter ρ is the correlation between disturbances across equations, ω is the correlation between each regressor and the disturbance in its own equation, and ϕ is the correlation between regressors in different equations.

Table 2. Simulation results, $T = 300$

2 equations, 1 integrated regressor per equation

	Mean bias						Std. error ($\times 10^{-1}$)						Size of 5% test					
$(\rho,\sigma,\phi)=$	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR
(0,0,0)	.000	.000	.000	.000	.000	-.001	1.030	1.034	1.035	1.039	1.040	1.026	5.6	5.7	5.9	5.9	6.3	6.3
	.000	.000	.000	.000	.000	.000	.113	.113	.113	.113	.113	.111	5.8	5.9	6.1	6.2	6.4	6.2
(0.3,0.3,0.3)	.000	.000	.000	.000	.000	.001	1.005	1.003	.971	1.012	.981	.941	5.9	5.6	5.8	5.9	6.0	5.9
	.000	.000	.000	.000	.000	.001	.110	.109	.104	.109	.104	.102	5.9	5.8	6.2	6.5	7.1	6.4
(0.9,0,0)	-.002	-.002	-.001	-.002	-.001	.000	1.038	1.039	.729	1.043	.733	.726	5.7	5.7	5.3	6.1	6.0	5.9
	.000	.000	.000	.000	.000	.000	.114	.114	.056	.113	.056	.056	5.6	5.6	5.2	6.1	6.1	6.3
(0,0.7,0)	-.001	-.001	-.001	-.001	-.001	.001	.738	.741	.744	.755	.757	.734	5.1	5.1	5.2	5.9	6.2	5.9
	.000	.000	.000	.001	.001	.001	.081	.081	.081	.083	.083	.081	5.1	5.0	5.1	6.0	6.4	5.8
(0,0,0.4)	-.001	-.001	-.001	-.001	-.001	-.001	1.053	1.056	1.060	1.059	1.063	1.007	6.2	6.2	6.2	6.4	6.6	5.8
	.000	.000	.000	.000	.000	.000	.113	.113	.114	.112	.112	.113	5.5	5.6	5.6	6.1	6.1	6.6

2 equations, 2 integrated regressors per equation

	Mean bias						Std. error ($\times 10^{-1}$)						Size of 5% test					
$(\rho,\sigma,\phi)=$	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR
(0,0,0)	.001	.001	.001	.001	.001	-.002	1.344	1.344	1.348	1.361	1.365	1.340	5.6	5.6	5.8	6.3	6.7	6.6
	.000	.000	.000	.000	.000	.000	.134	.134	.134	.133	.134	.134	6.0	5.9	6.0	6.6	6.8	7.4
(0.3,0.3,0.3)	.000	.000	.000	.001	.000	.001	1.210	1.135	.989	1.160	1.043	1.002	7.9	5.7	5.2	6.7	6.8	6.4
	.000	.000	.000	.000	.001	.001	.121	.113	.092	.115	.099	.092	8.1	6.1	5.7	7.4	8.7	7.2
(0.9,0,0)	.002	.001	.001	.001	.001	-.002	1.362	1.364	.907	1.385	.924	.906	5.8	5.8	5.8	6.7	6.9	6.5
	.000	.000	.000	.000	.000	.000	.134	.134	.068	.134	.070	.070	5.8	5.8	5.5	6.7	7.2	7.4
(0,0.7,0)	.000	.000	.000	.000	-.001	.001	.191	.192	.202	.462	.492	.266	2.2	6.0	7.0	6.4	9.0	0.5
	.000	.000	.000	.001	.002	.002	.019	.019	.020	.051	.056	.033	2.1	5.8	7.5	9.2	13.5	1.8
(0,0,0.4)	-.001	-.001	-.001	-.001	-.001	.001	1.358	1.361	1.364	1.377	1.380	1.325	6.1	6.2	6.2	6.6	6.9	6.9
	.000	.000	.000	.000	.000	.000	.132	.132	.133	.132	.133	.134	6.2	6.1	6.2	7.2	7.4	6.9

4 equations, 1 integrated regressor per equation

	Mean bias						Std. error ($\times 10^{-1}$)						Size of 5% test					
$(\rho, \sigma, \phi) =$	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR	IDOLS	SDOLS	DGLS	FM-OLS	FM-GLS	SUCCR
(0,0,0)	-.001	.000	.000	-.001	-.001	-.001	1.040	1.053	1.059	1.060	1.065	1.020	5.9	5.8	6.0	6.4	6.8	6.5
	.000	.000	.000	.000	.000	.000	.112	.113	.114	.114	.114	.114	6.2	6.0	6.4	7.0	7.6	7.2
(0.3,0.3,0.3)	.000	.000	.000	.000	.000	.001	.977	.984	.927	.989	.931	.907	6.2	6.3	6.5	6.7	7.3	6.8
	.000	.000	.000	.000	.001	.000	.108	.108	.098	.107	.098	.096	6.2	5.9	6.2	6.8	7.0	6.9
(0.9,0,0)	-.001	-.001	-.001	.000	-.001	.000	1.027	1.041	.679	1.043	.682	.683	5.9	5.7	5.7	6.1	6.4	6.3
	.000	.000	.000	.000	.000	.000	.112	.113	.044	.112	.045	.045	6.0	6.0	6.2	6.8	7.6	7.3
(0,0.7,0)	-.001	-.001	-.001	-.001	-.001	.001	.741	.752	.757	.773	.779	.735	5.3	5.4	5.8	6.4	7.0	6.3
	.000	.000	.000	.001	.001	.001	.081	.082	.083	.085	.087	.082	5.6	5.6	6.3	7.2	7.9	6.7
(0,0,0.4)	.000	.000	.000	.000	.000	.000	1.023	1.036	1.044	1.039	1.046	1.023	5.6	5.7	5.8	6.4	6.7	6.6
	.000	.000	.000	.000	.000	.000	.112	.113	.114	.114	.115	.113	5.7	5.7	6.3	6.4	7.2	7.1