A Spurious Regression Approach to Estimating Structural Parameters (Incomplete and Preliminary)

Chi-Young Choi
Department of Economics
University of New Hampshire

Ling Hu
Department of Economics
Ohio State University

Masao Ogaki
Department of Economics
Ohio State University

March 19, 2004

Abstract

Economic models often imply that certain variables are cointegrated. However, tests often fail to reject the null hypothesis of no cointegration for these variables. One possible explanation of these test results is that the error is unit root nonstationary due to a nonstationary measurement error in one variable. For example, currency held by the domestic economic agents for legitimate transactions is very hard to measure due to currency held by foreign residents and black market transactions. Therefore, money may be measured with a nonstationary error. If the money demand function is stable in the long-run, we have a cointegrating regression when money is measured with a stationary measurement error, but have a spurious regression when money is measured with a nonstationary measurement error. We can still recover structural parameters under certain conditions for the nonstationary measurement error. This paper proposes econometric methods based on asymptotic theory to estimate structural parameters with spurious regressions involving unit root nonstaionary variables.

Keywords: Spurious regression, GLS correction method.

JEL Classification Numbers: C10, C15

1 Introduction

Economic models often imply that certain variables are cointegrated. However, tests often fail to reject the null hypothesis of no cointegration for these variables. One possible explanation of these test results is that the error is unit root nonstationary due to a nonstationary measurement error in one variable. A nonstationary error in one variable leads to a spurious regression when the true value of the variable and the other variables are cointegrated. In the unit root literature, when the stochastic error of a regression is unit root nonstationary, the regression is called a spurious regression. This is because the standard t test tends to be spuriously significant even when the regressor is statistically independent of the regressand in Ordinary Least Squares. Monte Carlo simulations have often been used to show that the spurious regression phenomenon occurs with regressions involving unit root nonstationary variables (see, e.g., Granger and Newbold (1974), Nelson and Kang (1981, 1983)). Asymptotic properties of estimators and test statistics for regression coefficients of these spurious regressions have been studied by Phillips (1986, 1998) and Durlauf and Phillips (1988) among others. For example, currency held by the domestic economic agents for legitimate transactions is very hard to measure due to currency held by foreign residents and black market transactions. Therefore, money may be measured with a nonstationary error. As shown by Stock and Watson (1993) among others, if the money demand function is stable in the long-run, we have a cointegrating regression when all variables are measured without error. If the variables are measured with stationary measurement errors, we still have a cointegrating regression. However, if money is measured with a nonstationary measurement error, we have a spurious regression. We can still recover structural parameters under certain conditions for the nonstationary measurement error.

This paper proposes a new approach to estimating structural parameters with spurious regressions. Our approach is based on the Generalized Least Squares solution of the spurious regression problem analyzed by Ogaki and Choi (2002), who use an exact small sample analysis based on the conditional probability version of the Gauss-Markov Theorem. We have developed asymptotic theory for some estimators motivated by the GLS correction. The GLS estimation is shown to be consistent in spurious regressions and in drawing inferences based on this GLS estimator. We have developed a Hausman type specification test that is a consistent test for cointegration against the alternative hypothesis of no cointegration (or a spurious regression). We construct this test as we note that both the dynamic OLS and GLS corrected dynamic regression estimators are consistent in cointegration estimation while the dynamic OLS estimator is more efficient. On the other hand, when the regression is spurious, only the GLS corrected dynamic regression estimator is consistent. Hence we could do a cointegration test based on the specifications on the error. We show that under the null hypothesis of cointegration, the test statistics has a usual χ^2 limit distribution; while under the alternative hypothesis of a spurious regression, the test statistic diverges.

In the unit root literature, asymptotic theory and Monte Carlo simulations have been the main tools to analyze econometric methods. The exact small sample analysis based on the Gauss-Markov Theorem has not been used in general. There seem at least two reasons for this. First, in the unit root literature, most applications involve stochastic regressors, and the conditional expectation version of the Gauss-Markov Theorem is necessary. The standard measure theory definition of the conditional expectation assumes that the random vector's unconditional expectation exists and is finite. As the textbook of Judge et al. (1985) explains, this severely limits the usefulness of the conditional expectation version of the Gauss-Markov Theorem, because it is not possible to prove the existence of the unconditional expectation of the OLS estimator in most applications and simulations (due to the fact that the inverse of $X'X$ is involved in the OLS estimator where X is the design matrix). Second, the strict exogeneity assumption is usually violated in time series applications.

Ogaki and Choi (2002) propose to overcome the first difficulty by considering a definition of the conditional expectation based on the conditional probability measure. The conditional expectation based on the conditional probability measure can be defined even when the unconditional expectation does not exist as in Billingsley (1986). The Law of Iterated Expectations may not be satisfied when this definition is employed, but this does not cause problems for our particular application for spurious regressions. The second difficulty can be dealt with by adding leads and lags of the first difference of the stochastic regressors, leading to dynamic regressions proposed by Phillips and Loretan (1991) and Stock and Watson (1993) among others. The idea of the dynamic regressions can also be used for spurious regressions.

Using the conditional probability version of the Gauss-Markov Theorem, Ogaki and Choi (2002) study the exact small sample properties of spurious regressions. For the case of a classic spurious regression of a random walk onto a random walk that is independent of the regressand, they find that only the spherical variance assumption is violated. Therefore, they propose a GLS correction for the spurious regression. This solution is essentially the same as the well known solution of taking the first difference of all variables in this case, but the solution can also be used for the case with endogeneity, as long as the dynamic regression technique solves the endogeneity problem.

However, the stringent assumptions such as known covariance matrices employed by Ogaki and Choi (2002) for the exact small sample analysis are not sastisfied in applications. For this reason, in order to apply the GLS correction, it is necessary to relax some of their assumptions. Because the exact small sample properties cannot be analyzed when these assumptions are relaxed, we use asymptotic theory to ananlyze large sample properties of esimtators and test statistics based on the GLS correction. These methods are applied to estimate the long-run intertemporal elasticity of substitution and parameters of the money demand function.

2 The econometric model

We consider the following data generating process for observations $\{x_t, y_t\}$,

$$y_t = \beta_0 x_t + g_t + e_t \tag{1}$$

$$\Delta x_t = v_t \tag{2}$$

where g_t is generated by finite number of leads and lags of Δx_t,

$$g_t = C(L^{-1})\Delta x_t + D(L)\Delta x_t.$$

In our following analysis, to simplify notations, we usually take $g_t = \alpha_0 \Delta x_t = \alpha_0 v_t$. Without loss of generality, we can set $x_0 = 0$.

Now suppose that β_0 in (1) is the parameter of interest. The inference procedure about β_0 differs according to different assumptions on the error term e_t in (1). When e_t is stationary, the regression (1) is a cointegration regression; when e_t is a unit root nonstationary process, the regression is spurious. The latter case is motivated by our empirical studies in Macroeconomic modeling and it is the main interest in the project.

Assumption 1 *Let both v_t and u_t be zero mean stationary processes with $E|v_t|^\gamma < \infty$, $E|u_t|^\gamma < \infty$ for some $\gamma > 2$. Also assume that v_t and u_t are statistically independent, and they are both strong mixing with size $-\gamma/(\gamma - 2)$. Consider two situations:*

1. $\Delta e_t = u_t$.

2. $e_t = u_t$.

The conditions on v_t and u_t ensure the invariance principles: for $r \in [0,1]$, $n^{-1/2}\sum_{t=1}^{[nr]} v_t \to_d \sigma_1 V(r)$, $n^{-1/2}\sum_{t=1}^{[nr]} u_t \to_d \sigma_2 U(r)$ where $V(r)$ and $U(r)$ are independent standard Brownian motions and σ_1^2 and σ_2^2 are long run variances of the sequences $\{v_t\}$ and $\{u_t\}$ respectively. In fact de Jong and Davidson (2000) provides more general conditions for FCLT, but the conditions listed above are general enough to include many stationary Gaussian or non-Gaussian ARMA processes, which are commonly assumed in empirical modeling.

In the next two sections, we will summarize asymptotic properties of different estimation procedures under these two assumptions. Under assumption 1, the regression will be spurious and in this situation, OLS is not consistent and either GLS or Feasible GLS estimation will give consistent estimates. Under assumption 2, GLS estimator is not efficient as it is $\sqrt{n}$ convergent, but the feasible GLS estimator is n convergent and asymptotically equivalent to the OLS estimator.

2.1 Regressions with I(1) error

2.1.1 The dynamic OLS spurious estimation

Consider the OLS estimation of the regression

$$y_t = \beta x_t + \alpha v_t + error. \tag{3}$$

This is a spurious regression since for any value of β, the error term is always I(1). The OLS estimator $\hat{\beta}_n$ has the following limit distribution

$$\hat{\beta}_n - \beta_0 \to \frac{\sigma_2 \int_0^1 V(r)U(r)dr}{\sigma_1 \int_0^1 V(r)^2 dr} \equiv \xi \tag{4}$$

which can be written as a mixture of normal distributions centered at zero (Phillips, 1989). As discussed in Phillips (1986), in spurious regressions, the noise is as strong as the signal, hence uncertainty about β persists in the limiting distributions.

2.1.2 GLS corrected estimation

When we think of the problem with spurious regressions, it is the persistence in the error. In this problem, the error

$$e_t = \rho e_{t-1} + u_t.$$

where $\rho = 1$. Then we can filter all the variables by taking full difference, and use OLS to estimate

$$\Delta y_t = \beta \Delta x_t + \alpha \Delta v_t + u_t. \tag{5}$$

This procedure can be viewed as a GLS corrected estimation. We will also use GLS to refer to this procedure. Note that in regression (3), the estimator of β and α are asymptotically independent, hence we don't have to include v_t in the regression if we are only interested in β. But here we need to include the differences of the leads and lags (who are I(-1)) that are correlated with v_t to produce consistent estimate for β.

Define $\theta = (\beta, \alpha)'$, we can show that

$$\sqrt{n}(\tilde{\theta}_n - \theta_0) \to_d (\sigma_v^2 Q)^{-1} N(0, \sigma_v^2 \sigma_u^2 Q) = N\left(0, \frac{\sigma_u^2}{\sigma_v^2} Q^{-1}\right), \tag{6}$$

where

$$Q = \begin{bmatrix} 1 & 1-\psi_v \\ 1-\psi_v & 2(1-\psi_v) \end{bmatrix}, \tag{7}$$

where $\sigma_v^2 = E(v_t^2)$, $\sigma_u^2 = E(u_t^2)$, and ψ_v is the first order correlation coefficient of sequence $\{v_t\}$. We can see that now β can be consistently estimated (jointly with α).

2.1.3 The Feasible Cochrane-Orcutt GLS estimation

In a spurious regression, even if we don't know that the autocorrelation coefficient P of the error is unity, a feasible GLS procedure based on the Cochrane-Orcutt transformation will give asymptotically equivalent results as in (6). This has been shown by Blough (1992) and Phillips and Hodgson (1994). In this section, we will show that for the structural spurious regression problem we consider, Cochrane-Orcutt GLS could also provide consistent estimates.

Let the residual from OLS regression (3) denoted by $\hat{R}_t$, i.e.

$$\hat{R}_t = y_t - \hat{\beta}_n x_t - \hat{\alpha}_n v_t.$$

To conduct the Cochrane-Orcutt GLS estimation, first we run an AR(1) regression of $\hat{R}_t$,

$$\hat{R}_t = \rho \hat{R}_{t-1} + error. \tag{8}$$

In the appendix, we show that

$$\hat{\rho}_n - 1 = o_p(1) \quad \text{and} \quad n(\hat{\rho}_n - 1) = O_p(1).$$

Next, consider the following Cochrane-Orcutt transformation of the data:

$$\tilde{y}_t = y_t - \hat{\rho}_n y_{t-1}, \quad \tilde{x}_t = x_t - \hat{\rho}_n x_{t-1}, \quad \tilde{v}_t = v_t - \hat{\rho}_n v_{t-1}. \tag{9}$$

Now consider OLS estimation in the regression

$$\tilde{y}_t = \tilde{\beta}\tilde{x}_t + \tilde{\alpha}\tilde{v}_t + error \tag{10}$$

Define $z_t = (\tilde{x}_t, \tilde{v}_t)'$ and $\theta = (\beta, \alpha)'$, then the OLS estimator of θ is

$$\tilde{\theta}_n = \left[\sum_{t=1}^{n} z_t z_t'\right]^{-1} \left[\sum_{t=1}^{n} z_t \tilde{y}_t\right]. \tag{11}$$

The limit distribution for $\tilde{\theta}$ can be shown to be the same as in (6).

$$\sqrt{n}(\tilde{\theta}_n - \theta_0) \to_d (\sigma_v^2 Q)^{-1} N(0, \sigma_v^2 \sigma_u^2 Q) = N\left(0, \frac{\sigma_u^2}{\sigma_v^2} Q^{-1}\right), \tag{12}$$

Note that if in regression (3), v_t is not included, we still have the same limit distribution based on the residual from the OLS estimation.

In Appendix B, we described some extensions of the model. We show that if a constant is included in the DGP, the GLS or FGLS corrected estimation give results that are asymptotically equivalent as given in (6).

2.2 Regressions with I(0) error

In this section, we will consider the following problem. When the error term in (1) is I(0) instead of I(1), so that the regression is a cointegration rather than a spurious regression, while we apply the same procedure as under the spurious assumption, then how does the estimator behave asymptotically.

2.2.1 The dynamic OLS estimation when the error is I(0)

Suppose that the error term in the data generating process of y_t is actually I(0), $e_t = u_t$ while we keep all other assumptions. The DGP of y_t can be written as

$$y_t = \beta_0 x_t + \alpha_0 v_t + u_t.$$

Use OLS to estimate the regression

$$y_t = \beta x_t + \alpha v_t + error. \tag{13}$$

Clearly, this is a cointegration regression, and the limit distribution of the OLS estimator of β can be written as

$$n(\hat{\beta}_n - \beta_0) \to_d \frac{\sigma_2 \int_0^1 V(r)dU(r)}{\sigma_1 \int_0^1 V(r)^2 dr}. \tag{14}$$

If we don't include v_t in regression (13), there will be a bias in the OLS estimator. Denote this estimator by $\hat{\beta}_n^+$,

$$n(\hat{\beta}_n^+ - \beta_0) \to_d \frac{\sigma_2 \int_0^1 V(r)dU(r)}{\sigma_1 \int_0^1 V(r)^2 dr} + \frac{\frac{1}{2}(V(1)^2 + \sigma_v^2/\sigma_1^2)}{\int_0^1 V(r)^2 dr}.$$

However, the bias is $O_p(n^{-1})$ hence the OLS estimator $\hat{\beta}_n^+$ is still consistent.

2.2.2 Taking full difference when the error is I(0)

Now, if we take a full difference as we did in the I(1) case, and estimate the regression

$$\Delta y_t = \beta \Delta x_t + \alpha \Delta v_t + u_t - u_{t-1} = z_t'\theta + u_t - u_{t-1}.$$

Note that we lose efficiency in this transformation as the estimator $\tilde{\beta}$ is now $\sqrt{n}$ convergent rather than n convergent in the cointegration. With some minor revision of equation (6), we got results for the limit distribution of estimator in this case:

$$\sqrt{n}(\tilde{\theta}_n - \theta_0) \to_d N\left(\mathbf{0}, \frac{2\sigma_u^2(1-\psi_u)}{\sigma_v^2} Q^{-1}\right).$$

where ψ_u is the first order correlation coefficient of sequence $\{u_t\}$.

Therefore, the GLS correction or differencing is not efficient if the variables are actually cointegrated.

2.2.3 The Cochrane-Orcutt feasible GLS estimator when the error is I(0)

Instead of taking full difference, if we estimate the autoregression coefficient in the error and use this estimator to filter all sequences, then we will obtain an estimator that is asymptotically equivalent to the OLS estimator. Intuitively, in the case that the error is *i.i.d.* ($e_t = u_t$ is i.i.d.), then the AR(1) coefficient $\hat{P}_n$ will converge to zero, hence the transformed regression will be asymptotically equivalent to the original regression. Or, if the error is stationary and serially correlated, then the AR(1) coefficient

will be less than unit, and as has been shown in Park and Phillips (1988), the GLS estimator and OLS estimator in a cointegration are asymptotically equivalent.

First run OLS estimation of

$$y_t = \hat{\beta}_n x_t + \hat{\alpha}_n v_t + \hat{u}_t = z_t'\hat{\theta} + \hat{u}_t.$$

Then run an AR(1) regression of $\hat{u}_t$,

$$\hat{u}_t = \hat{\rho}_n \hat{u}_{t-1} + error.$$

Write

$$\hat{u}_t = y_t - z_t'\theta_0 + z_t'(\theta_0 - \hat{\theta}) = u_t + z_t'(\theta_0 - \hat{\theta}).$$

Now, consider the Cochrane-Orcutt transformation (9) and estimate

$$\tilde{y}_t = \tilde{\beta}_n \tilde{x}_t + \tilde{\alpha}\tilde{v}_t + error. \tag{15}$$

The limit distribution of $\tilde{\beta}_n$ can be shown to be

$$n(\tilde{\beta}_n - \beta_0) \to_d \frac{(1-\psi_u)^2\sigma_1\sigma_2 \int_0^1 V(r)dU(r)}{(1-\psi_u)^2\sigma_2^2 \int_0^1 V(r)^2 dr} = \frac{\sigma_2 \int_0^1 V(r)dU(r)}{\sigma_1 \int_0^1 V(r)^2 dr}. \tag{16}$$

which is exactly the same as the limit of the OLS estimator given in (14). In summary, the feasible GLS is not only valid in spurious regression, but also harmless to the estimator in the limit when the regression is actually a cointegration. In particular, FGLS can be regarded as a robust procedure with respect to error specifications. It is asymptotically equivalent to GLS estimator in a spurious regressions and it is asymptotically equivalent to OLS estimator in a cointegration regressions.

2.3 Hausman specification test

Now if we compare two estimators: an OLS estimator ($\hat{\beta}_n$) and an GLS estimator ($\tilde{\beta}_n$) corresponding to $\rho = 1$. This is equivalent to a question like: take difference or not. We let the error be I(0) under the null and the error be I(1) under the alternative. Our above discussions show that under the null, both OLS and GLS are consistent but OLS estimator is more efficient; while under the alternative, which corresponds to spurious regression, only GLS estimator is consistent. Since including the leads and lags doesn't have critical effect, we assume that there are no leads and lags in the DGP for simplicity. Hence the DGP under the null is

$$y_t = \beta_0 x_t + u_t, \tag{17}$$

and under the alternative is

$$y_t = \beta_0 x_t + e_t, \quad \Delta e_t = u_t. \tag{18}$$

Now under the null hypothesis, the OLS and GLS estimator are

$$\hat{\beta}_n = \frac{\sum_{t=1}^n x_t y_t}{\sum_{t=1}^n x_t^2} = \beta_0 + \frac{\sum_{t=1}^n x_t u_t}{\sum_{t=1}^n x_t^2} \tag{19}$$

and

$$\tilde{\beta}_n = \frac{\sum_{t=1}^n \Delta x_t \Delta y_t}{\sum_{t=1}^n \Delta x_t^2} = \beta_0 + \frac{\sum_{t=1}^n v_t(u_t - u_{t-1})}{\sum_{t=1}^n v_t^2} \tag{20}$$

Let q_n denote the difference between these two estimators

$$\begin{aligned}\sqrt{n}q_n = \sqrt{n}(\tilde{\beta}_n - \hat{\beta}_n) &= \frac{n^{-1/2}\sum_{t=1}^n v_t(u_t - u_{t-1})}{n^{-1}\sum_{t=1}^n v_t^2} - \frac{n^{-3/2}\sum_{t=1}^n x_t u_t}{n^{-2}\sum_{t=1}^n x_t^2} \\ &= \frac{n^{-1/2}\sum_{t=1}^n v_t(u_t - u_{t-1})}{n^{-1}\sum_{t=1}^n v_t^2} - o_p(1) \\ &\to_p N(0, \tau^2).\end{aligned}$$

where $\tau^2 = 2\sigma_u^2(1-\psi_u)/\sigma_v^2$. A consistent estimator for τ^2 can be computed based on the sample moments:

$$\hat{\tau}_n^2 = \frac{2s_n^2(1-\hat{\psi}_u)}{\hat{\sigma}_n^2}, \quad \text{where} \quad \hat{\sigma}_n^2 = \frac{1}{n}\sum_{t=1}^n \Delta x^2, \quad s_n^2 = \frac{1}{n}\sum_{t=1}^n (\Delta y_t - \tilde{\beta}_n \Delta x_t)^2$$

$$\hat{\psi}_n = \frac{1}{s_n^2 n}\sum_{t=2}^n (\Delta y_t - \tilde{\beta}_n \Delta x_t)(\Delta y_{t-1} - \tilde{\beta}_n \Delta x_{t-1})$$

Define the test statistics

$$h_n = \frac{nq_n^2}{\hat{\tau}_n^2} \to \frac{[N(0,\tau^2)]^2}{\tau^2} = \chi^2(1). \tag{21}$$

Hence h_n has a limiting $\chi^2(1)$ distribution under the null hypothesis. Now, under the alternative of I(1) errors, the inconsistent OLS estimator dominates.

$$q_n = \tilde{\beta}_n - \hat{\beta}_n \to \xi \tag{22}$$

where ξ is bounded in probability.

So under the alternative, for the statistics defined in (21), $q_n = O_p(1)$, $\hat{\tau}_n^2$ still converges to τ, hence h_n diverges. In summary, the Hausman test statistics has a limiting $\chi^2(1)$ distribution under the null and diverges under the alternative.

Note that in this test the null hypothesis is cointegration, while usually the null of a cointegration test is that no cointegrating relationship presents. Both type of tests will be useful to empirical researchers. For instance, if the results reject the null of no cointegration and also reject the null of cointegration, we may need to seek alternative model specifications.

2.4 Simulations

We know that in spurious regressions, FGLS estimator is asymptotic equivalent to GLS estimator; and in cointegration regressions, FGLS estimator is asymptotically equivalent to OLS estimator. In the simulation, we show the finite sample performance of FGLS in these two situations. In the simulation, we generate v_t and ϵ_t from independent standard normal distribution and let $u_t = \epsilon_t + 0.5\epsilon_t$. The parameters are set to be $\beta = 2, \alpha = 0.5$. Figure 1 shows the finite sample distribution when the error term is unit root nonstationary. The left figure plots the distribution of the GLS estimator and the right figure plots that of the feasible GLS estimator. We can see that although FGLS is consistent, it has larger variance even when the sample is relatively large. Figure 2 shows the finite sample distribution when the error term is I(0). The left figure plots the distribution of OLS estimator and the right figure plots that of FGLS estimator. These two estimators both converge very fast and the difference between them is almost invisible.

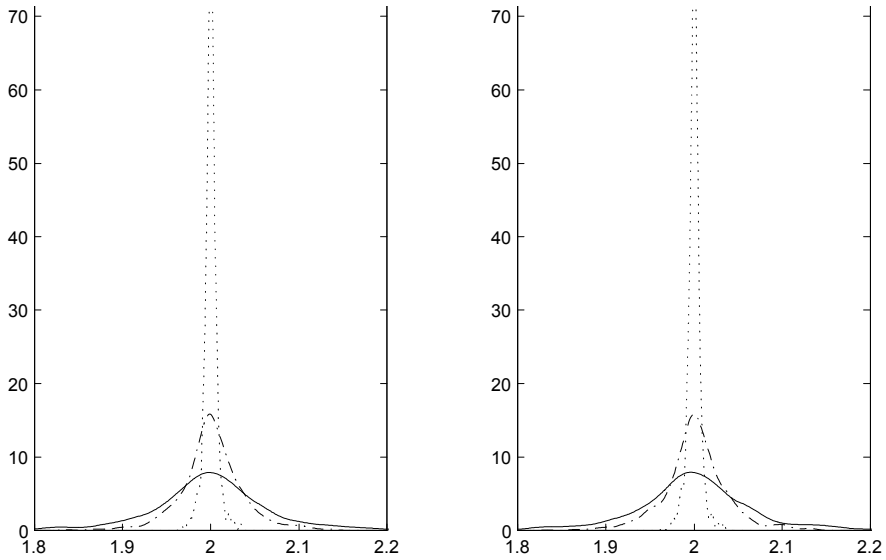

Figure 1: Distributions of OLS and FGLS estimators when the error is I(0)

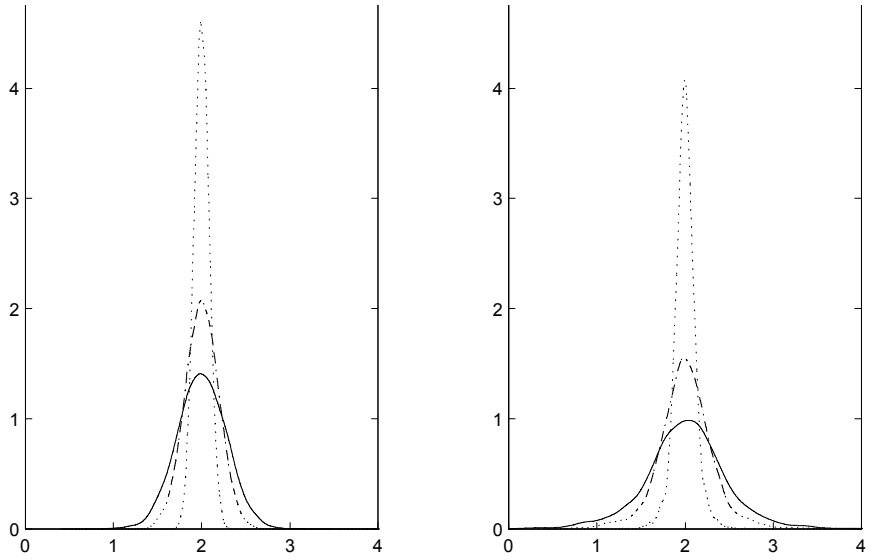

Figure 2: Distributions of GLS and FGLS estimators when the error is I(1)

3 Empirical Applications

In this section we apply the GLS correction methods to analyze three macroeconomic issues: the long-run intertemporal elasticity of substitution (IES), long-run money demand in the U.S., and the Purchasing Power Parity.

3.1 The long-run IES for consumption

We consider a simplified version of Cooley and Ogaki's (1996) model of consumption and leisure for the purpose of illustrating the project. In the model, the representative household maximizes

$$U = E_0[\sum_{t=0}^{\infty} \delta^t u(t)]$$

where E_t denotes the expectation conditioned on the information available at t. Consider a simple intraperiod utility function that is assumed to be time- and state-separable and separable in nondurable consumption, durable consumption, and leisure

$$u(t) = \frac{C(t)^{1-\beta} - 1}{1-\beta} + v(l(t))$$

where $v(\cdot)$ represents a continuously differentiable concave function, $C(t)$ is nondurable consumption, and $l(t)$ is leisure.

The usual first order condition for a household that equates the real wage rate with the marginal rate of substitution between leisure and consumption is:

$$W(t) = \frac{v'(l(t))}{C(t)^{-\beta}}$$

where $W(t)$ is the real wage rate. We assume that the stochastic process of leisure is (strictly) stationary in the equilibrium as in Eichenbaum, Hansen, and Singleton (1988). Then an implication of the first order condition is that $\ln(W(t)) - \beta \ln(C(t)) = \ln(v'(l(t)))$ is stationary. When we assume that the log of consumption is difference stationary, this implies that the log of the real wage rate and the log of consumption are cointegrated with a cointegrating vector $(1, -\beta)'$. Now assume that $\ln(W(t))$ and $\ln(C(t))$ are measured with errors. Imagine that the $\ln(C(t))$ is measured with a stationary measurement error, $\xi(t)$, and that $\ln(W(t))$ is measured with a difference stationary measurement error, $\epsilon(t)$ (perhaps because of the difficulty in measuring fringe benefits). Assume that $\epsilon(t)$ is independent of $\ln(C(t))$ and $\xi(t)$ at all leads and lags. Consider a regression

$$\ln(W^m(t)) = a + \beta \ln(C^m(t)) + u(t), \tag{23}$$

where $W^m(t)$ is the measured real wage rate, $C^m(t)$ is the measured consumption, and $u(t) = -\epsilon(t) + \beta\xi(t) + \ln(v'(l(t))) - a$. If $\epsilon(t)$ is stationary, then $u(t)$ is stationary, and Regression (23) is a cointegrating regression as in Cooley and Ogaki.

If $\epsilon(t)$ is unit root nonstationary, then Regression (23) is a spurious regression because $u(t)$ is nonstationary in this case. Hence the standard methods for cointegrating regressions cannot be used. However, the preference parameter β can still be estimated by the spurious regression method.

Table 1 presents the estimation results for preference parameter (β) based on various estimators: dynamic OLS, GLS, and dynamic feasible GLS.[1] We follow the empirical analysis in Cooley and Ogaki

[1] In dynamic GLS, the serial correlation coefficient in error term is estimated before being applied to the Cochrane-Orcutt transformation while it is assumed to be unity in GLS estimation which is analogous to regressing the first difference of variables without constant term.

except that GDP deflators are used in lieu of the implicit deflator based on consumption measures to deflate nominal wages.[2] The results in Table 1 illustrate several points. First, using non-durable consumption for $C(t)$ all point estimates for β have theoretically correct positive sign and they seem to be robust to the choice of lead and lag terms. This is readily confirmed by our results using BIC rule for lead and lag selection as the point estimates of the structural parameter is stable over various choice of leads and lags. Among the three estimators under study, DOLS estimates are the largest while DGLS estimates are the smallest. Interestingly, GLS estimates fall between the two. Our estimates are in general smaller than those found by Cooley and Ogaki based on the Canonical Cointegration Regression (CCR). The picture changes slightly when we add service consumption to non-durable consumption as consumption measure. Point estimates are consistently smaller by 0.14~0.23 than when only non-durable consumption is used, which is in sharp contrast with the results of Cooley and Ogaki who found favorable evidence for cointegration for ND rather than NDS. This time FGLS estimates appear to be smaller than the other two though the difference among them is not as large as before. This is reflected in our Hausman specification test results reported in Table 1. Overall our estiamtes throw additional light on the long-run relation between real wage and consumption in the absence of cointegrating restriction.

3.2 Elasticities of Money Demand in the U.S.

The long-run income and interest elasticities of money demand has often been estimated under the cointegrating restriction among real balances, real income, and interest rate. The restriction is legitimate if the money demand function is stable in the long-run and if all variables are measured without error. Indeed Stock and Watson (1993) found a supportive evidence of stable long-run M1 demand by estimating cointegrating vectors. However, if money is measured with a nonstationary measurement error, we have a spurious regression and the estimation results based on cointegration regression become questionable. We apply our GLS correction methods to estimate long-run income and interest elasticities of M1 demand. To this end, regression equations are set up with real money balance ($\frac{M}{P}$) as regressand and income (y) and interest (i) as regressors. Following Stock and Watson (1993), the annual time series for M1 deflated by the net national product price deflator is used for $\frac{M}{P}$, real net national product for y and the six month commercial paper rate in percentage for i. $\frac{M}{P}$ and y are in logarithms while three different regression equations are considered depending on the measures of interest. We have tried the following three functional forms. Equation 1 has been studied by Stock and Watson (1993).

$$\begin{aligned}
\ln\left(\frac{M}{P}\right)_t &= \alpha + \beta \ln(y_t) + \gamma i_t + u_t, \quad \text{(equation 1)} \\
\ln\left(\frac{M}{P}\right)_t &= \alpha + \beta \ln(y_t) + \gamma \ln(i_t) + u_t, \quad \text{(equation 2)} \\
\ln\left(\frac{M}{P}\right)_t &= \alpha + \beta \ln(y_t) + \gamma \ln\left[\frac{1+i_t}{i_t}\right] + u_t. \quad \text{(equation 3)}
\end{aligned}$$

It is worth noting that the liquidity trap is possible for the latter two functional forms. When the data contain the periods with very low nominal interest rates, the latter two functional forms may be more appropriate.

Table 2 presents the results. In all cases considered, point estimates for income elasticity of money demand (β) and inerest semi-elasticity (γ) have correct signs.[3] The point estimates of β are quite similar

[2] See Cooley and Ogaki (1996, page 127) for the detailed description of data.

[3] The point estimates for γ have positive sign in equation 3 because the regressor is approximately equivalent to the inverse of the counterparts in the other two equations.

across estimators but not necessarily for point estimates of γ in which GLS estimates are consistently lower than DOLS and FGLS estimates. Point estimate of β is largest in equation 1 and smallest in equation 3, but it appears to be stable in most cases. Estimated income elasticities range from 0.8 (in equation 3) through close to one in equation 1. As a consequence, long run U.S. money demand appears to be stable during 1900-1989, consistent with the finding by Stock and Watson.

3.3 Purchasing Power Parity

Let p_t and p_t^* denote the logarithms of the consumer price indices in the base country and foreign country respectively, and s_t be the logarithm of the price of foreign country's currency in terms of the base country's currency. Long-run PPP requires that a linear combination of these three variables be stationary. To be more specific, long-run PPP is said to hold if $f_t = s_t + p_t^*$ is cointegrated with p_t such that $\epsilon_t \sim I(0)$ in

$$f_t = \alpha + \beta p_t + \epsilon_t,$$

where $\beta = 1$. It is widely agreed that the empirical evidence on long-run PPP remains inconclusive despite extensive research. We apply our techniques to the long-standing issue in international economics among 21 OECD countries during the post Bretton Woods period.

Appendix

Appendix A: Proof of results in section 2.1

To show the distribution of the OLS estimator in regression (3), define

$$H_n = \begin{bmatrix} n & 0 \\ 0 & n^{1/2} \end{bmatrix}. \tag{24}$$

The OLS estimator for β and α can be written as

$$\begin{bmatrix} \hat{\beta}_n - \beta_0 \\ \hat{\alpha} - \alpha_0 \end{bmatrix} = \left[H_n^{-1} \begin{bmatrix} \sum_{t=1}^n x_t^2 & \sum_{t=1}^n x_t v_t \\ \sum_{t=1}^n x_t v_t & \sum_{t=1}^n v_t^2 \end{bmatrix} H_n^{-1} \right]^{-1} \begin{bmatrix} n^{-2}\sum_{t=1}^n x_t e_t \\ n^{-1}\sum_{t=1}^n v_t e_t \end{bmatrix}$$

For the first term,

$$\begin{bmatrix} n^{-2}\sum_{t=1}^n x_t^2 & n^{-3/2}\sum_{t=1}^n x_t v_t \\ n^{-3/2}\sum_{t=1}^n x_t v_t & n^{-1}\sum_{t=1}^n v_t^2 \end{bmatrix} \to \begin{bmatrix} \sigma_1^2 \int_0^1 V(r)^2 dr & 0 \\ 0 & \sigma_v^2 \end{bmatrix}. \tag{25}$$

where $\sigma_v^2 = E(v_t^2)$. For the second term,

$$\begin{bmatrix} n^{-2}\sum_{t=1}^n x_t e_t \\ n^{-1}\sum_{t=1}^n v_t e_t \end{bmatrix} \to_d \begin{bmatrix} \sigma_1\sigma_2 \int_0^1 V(r)U(r)dr \\ \sigma_1\sigma_2 \int_0^1 U(r)dV(r) \end{bmatrix}.$$

Equation (4) then follows.

To show the limit distribution of the GLS corrected estimator in regression (5), let $z_t = (\Delta x_t, \Delta v_t)$[4], then

$$\sqrt{n}(\tilde{\theta}_n - \theta_0) = \left[n^{-1} \sum_{t=1}^n z_t z_t' \right]^{-1} \left[n^{-1/2} \sum_{t=1}^n z_t u_t \right]. \tag{26}$$

For the first term,

$$n^{-1} \sum_{t=1}^n z_t z_t' = \begin{bmatrix} n^{-1}\sum v_t^2 & n^{-1}\sum v_t \Delta v_t \\ n^{-1}\sum v_t \Delta v_t & n^{-1}\sum \Delta v_t^2 \end{bmatrix} \to \sigma_v^2 \begin{bmatrix} 1 & 1-\psi_v \\ 1-\psi_v & 2(1-\psi_v) \end{bmatrix} = \sigma_v^2 Q, \quad say,$$

where ψ_v is the first order correlation coefficient of $\{v_t\}$. If v_t is i.i.d, $\psi_v = 0$ and $\sigma_v^2 = \sigma_1^2$.

For the second term, we want to show that

$$n^{-1/2} \sum_{t=1}^n z_t u_t = \begin{bmatrix} n^{-1/2}\sum v_t u_t \\ n^{-1/2}\sum \Delta v_t u_t \end{bmatrix} \to N(0, \sigma_v^2 \sigma_u^2 Q)$$

To show this, let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)'$ be an arbitrary vector of real numbers.

$$\begin{aligned} n^{-1/2} \sum_{t=1}^n \boldsymbol{\lambda}' z_t u_t &= n^{-1/2} \sum_{t=1}^n (\lambda_1 v_t + \lambda_2 \Delta v_t) u_t \\ &= \lambda_1 n^{-1/2} \sum_{t=1}^n v_t u_t + \lambda_2 n^{-1/2} \sum_{t=1}^n \Delta v_t u_t \\ &\to N(0, \sigma_v^2 \sigma_u^2 (\lambda_1^2 + 2\lambda_1\lambda_2(1-\psi_v) + 2\lambda_2^2(1-\psi_v)) \\ &= N(0, \sigma_v^2 \sigma_u^2 \boldsymbol{\lambda}' Q \boldsymbol{\lambda}) \end{aligned}$$

[4] Through this paper, we always use z_t to denote the vector of independent variables and let θ to denote the vector of parameters (but keep in mind that we are mostly interested in β). Note that in different regressions, those symbols denote different variables.

Hence, for the quantity defined in (26), we have the limit distribution given in (6).

To derive the limit distribution for the FGLS estimator, we first derive the limit distribution for $\hat{\rho}_n$ in regression (8). Write the process of $\hat{R}_t$ as

$$\begin{aligned}\hat{R}_t &= y_t - \hat{\beta}x_t - \hat{\alpha}v_t \\ &= y_{t-1} - \hat{\beta}x_{t-1} - \hat{\alpha}v_{t-1} + [(y_t - y_{t-1}) - \hat{\beta}_n(x_t - x_{t-1}) - \hat{\alpha}(v_t - v_{t-1})] \\ &= \hat{R}_{t-1} + [(\beta_0 - \hat{\beta}_n)v_t + (\alpha_0 - \hat{\alpha})(v_t - v_{t-1}) + u_t] \\ &= \hat{R}_{t-1} + h_t, \quad say.\end{aligned}$$

From this expression, we can see that $\hat{R}_t$ is a unit root process with serially correlation error h_t. Then the OLS estimator $\hat{\rho}_n$ can be written as

$$\hat{\rho}_n = \frac{\sum_{t=1}^n \hat{R}_t\hat{R}_{t-1}}{\sum_{t=1}^n \hat{R}_{t-1}^2} = 1 + \frac{\sum_{t=1}^n \hat{R}_{t-1}h_t}{\sum_{t=1}^n \hat{R}_{t-1}^2}$$

To derive the limit of $\hat{\rho}_n$, write

$$\hat{R}_t = y_t - \hat{\beta}_n x_t - \hat{\alpha}_n v_t = (\beta_0 - \hat{\beta}_n)x_t + (\alpha_0 - \hat{\alpha}_n)v_t + e_t. \tag{27}$$

Hence the denominator

$$\begin{aligned}\hat{R}_t^2 &= (\beta_0 - \hat{\beta}_n)x_t^2 + e_t^2 + 2(\beta_0 - \hat{\beta}_n)x_t e_t \\ &\quad +(\alpha_0 - \hat{\alpha}_n)^2 v_t^2 + 2(\alpha_0 - \hat{\alpha}_n)(\beta_0 - \hat{\beta}_n)x_t v_t + 2(\alpha_0 - \hat{\alpha}_n)e_t v_t,\end{aligned}$$

where we can see that the sum of the first line diverges faster as they are products of I(1) variables. In particular,

$$\begin{aligned}n^{-2}\sum_{t=1}^n \hat{R}_t^2 &= (\beta_0 - \hat{\beta}_n)^2 n^{-2}\sum_{t=1}^n x_t^2 + n^{-2}\sum_{t=1}^n e_t^2 + 2(\beta_0 - \hat{\beta}_n)n^{-2}\sum_{t=1}^n x_t e_t + o_p(1) \\ &\to_d \xi^2\sigma_1^2\int_0^1 V(r)^2 dr + \sigma_2^2\int_0^1 U(r)^2 dr + 2\xi\sigma_1\sigma_2\int_0^1 V(r)U(r)dr \equiv \zeta.\end{aligned}$$

For the numerator,

$$\hat{R}_{t-1}h_t = [(\beta_0 - \hat{\beta}_n)x_{t-1} + (\alpha_0 - \hat{\alpha})v_{t-1} + e_{t-1}][(\beta_0 - \hat{\beta}_n + \alpha_0 - \hat{\alpha})v_t - (\alpha_0 - \hat{\alpha})v_{t-1} + u_t]$$

The sum of all the terms of products in this expression converges when normed with n^{-1}. We omit the details here as we will not make use of the exact distribution of $\hat{\rho}_n$. Plug in the limits of all the terms, we can write

$$n^{-1}\sum_{t=1}^n \hat{R}_{t-1}h_t \to \eta, \quad say.$$

Hence,

$$n(\hat{\rho}_n - 1) = \frac{n^{-1}\sum_{t=1}^n \hat{R}_{t-1}h_t}{n^{-2}\sum_{t=1}^n \hat{R}_{t-1}^2} \to_d \frac{\eta}{\zeta}. \tag{28}$$

Actually, in our following computations, all we need to know is that

$$\hat{\rho}_n - 1 = o_p(1) \quad \text{and} \quad n(\hat{\rho}_n - 1) = O_p(1).$$

Below, we show how to derive the limit distribution for $\tilde{\theta}$. For the sequence of $\tilde{y}_t$, we can write it as

$$\begin{aligned}
\tilde{y}_t &= y_t - \hat{\rho}_n y_{t-1} \\
&= \beta_0 x_t + \alpha_0 v_t + e_t - \hat{\rho}_n(\beta_0 x_{t-1} + \alpha_0 v_{t-1} + e_{t-1}) \\
&= \beta_0(x_t - \hat{\rho}_n x_{t-1}) + \alpha_0(v_t - \hat{\rho}_n v_{t-1}) + (e_t - e_{t-1}) + (1-\hat{\rho}_n)e_{t-1} \\
&= \beta_0 \tilde{x}_t + \alpha_0 \tilde{v}_t + u_t + (1-\hat{\rho}_n)e_{t-1} \\
&= z_t'\theta_0 + u_t + (1-\hat{\rho}_n)e_{t-1}
\end{aligned}$$

Now, we can write

$$\hat{\theta} - \theta_0 = \left[\sum_{t=1}^n z_t z_t'\right]^{-1} \left[\sum_{t=1}^n z_t[u_t + (1-\hat{\rho}_n)e_{t-1}]\right]. \tag{29}$$

For the first term:

$$\sum_{t=1}^n z_t z_t' = \begin{bmatrix} \sum \tilde{x}_t^2 & \sum \tilde{x}_t\tilde{v}_t \\ \sum \tilde{x}_t\tilde{v}_t & \sum \tilde{v}_t^2 \end{bmatrix}$$

The asymptotics of each term follows. First,

$$\begin{aligned}
\sum_{t=1}^n \tilde{x}_t^2 &= \sum_{t=1}^n (x_t - \hat{\rho}_n x_{t-1})^2 \\
&= \sum_{t=1}^n [(1-\hat{\rho}_n)x_{t-1} + v_t]^2 \\
&= (1-\hat{\rho}_n)^2 \sum_{t=1}^n x_{t-1}^2 + 2(1-\hat{\rho}_n)\sum_{t=1}^n x_{t-1}v_t + \sum_{t=1}^n v_t^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
n^{-1}\sum_{t=1}^n \tilde{x}_t^2 &= n(1-\hat{\rho}_n)^2\left(n^{-2}\sum_{t=1}^n x_{t-1}^2\right) + 2(1-\hat{\rho}_n)\left(n^{-1}\sum_{t=1}^n x_{t-1}v_t\right) + n^{-1}\sum_{t=1}^n v_t^2 \\
&= n^{-1}\sum_{t=1}^n v_t^2 + o_p(1) \\
&\to \sigma_v^2
\end{aligned}$$

Similarly,

$$\begin{aligned}
n^{-1}\sum \tilde{x}_t\tilde{v}_t &= n^{-1}\sum_{t=1}^n v_t^2 - \hat{\rho}_n\left(n^{-1}\sum_{t=1}^n v_t v_{t-1}\right) + o_p(1) \\
&\to_p \sigma_v^2(1-\psi_v).
\end{aligned}$$

Finally, $n^{-1}\sum_{t=1}^n \tilde{v}_t^2 \to 2\sigma_v^2(1-\psi_v)$. Hence,

$$n^{-1}\sum_{t=1}^n z_t z_t' \to_p \sigma_v^2 \begin{bmatrix} 1 & 1-\psi_v \\ 1-\psi_v & 2(1-\psi_v) \end{bmatrix} = \sigma^2 Q. \tag{30}$$

Now, consider the second term in (29)

$$\sum_{t=1}^n z_t[u_t + (1-\hat{\rho}_n)e_{t-1}] = \begin{bmatrix} \sum_{t=1}^n \tilde{x}_t[u_t + (1-\hat{\rho}_n)e_{t-1}] \\ \sum_{t=1}^n \tilde{v}_t[u_t + (1-\hat{\rho}_n)e_{t-1}] \end{bmatrix}.$$

It is not hard to see that $n^{-1}\sum_{t=1}^{n} z_t[u_t + (1-\hat{\rho}_n)e_{t-1}] \to_p 0$. Intuitively, $\tilde{x}_t$ is asymptotically like v_t, while u and v are independent by assumption. Again, our remaining task is to show that

$$n^{-1/2}\sum_{t=1}^{n} z_t[u_t + (1-\hat{\rho}_n)e_{t-1}] \to N(\mathbf{0}, \sigma_v^2\sigma_u^2 Q). \tag{31}$$

This can be shown in the same way as in the proof for (6). Combine (31 with (30), we obtain the limit distribution for $\tilde{\theta}$, as given in (11).

Appendix B: Some extensions

So far we have assumed that there is no constant term or deterministic time trends in the DGP of y_t. If there is a constant term, e.g.

$$y_t = \delta_0 + \beta_0 x_t + \alpha_0 v_t + u_t.$$

Correspondingly, in the OLS estimation, we also include a constant,

$$y_t = \delta + \beta x_t + \alpha v_t + error. \tag{32}$$

The limit of $\hat{\beta}_n$ and $\hat{\alpha}_n$ are similar as in the case without constant, except that we have demeaned Brownian motions instead of standard Brownian motions in the limit. Since this is still a spurious regression, the estimator of the constant term diverges as was shown in Phillips (1986). Here

$$\begin{aligned}\hat{\delta}_n &= \bar{y} - \hat{\beta}_n\bar{x} - \hat{\alpha}_n\bar{v}_t \\ &= \delta_0 + (\hat{\beta}_n - \beta_0)\bar{x} + (\hat{\alpha}_n - \alpha_0)\bar{v} + \bar{e}\end{aligned}$$

Hence

$$\begin{aligned}n^{-1/2}\hat{\delta}_n &= n^{-1/2}\delta_0 + (\hat{\beta}_n - \beta_0)n^{-3/2}\sum_{t=1}^{n} x_t + (\hat{\alpha}_n - \alpha_0 n^{-3/2}\sum_{t=1}^{n} v_t + n^{-3/2}\sum_{t=1}^{n} e_t \\ &\to_d \bar{\xi}\sigma_1\int_0^1 V(r)dr + \sigma_2\int_0^1 U(r)dr.\end{aligned}$$

where we let $\bar{\xi}$ to denote the limit of $\hat{\beta}_n - \beta_0$. Next, if we do GLS or the differenced regression, the constant is canceled so we could have the same limit result as is given by (6). Finally, consider the Cochrane-Orcutt feasible GLS estimation. Still let $\hat{R}_t$ denote the OLS residual

$$\hat{R}_t = y_t - \hat{\delta}_n - \hat{\beta}_n x_t - \hat{\alpha}_n v_t.$$

Then do another OLS estimation in

$$\hat{R}_t = \hat{\rho}_n\hat{R}_{t-1} + error.$$

Write

$$\begin{aligned}\hat{R}_t &= y_t - \hat{\delta}_n - \hat{\beta}_n x_t - \hat{\alpha}_n v_t \\ &= \hat{R}_{t-1} + [(\beta_0 - \hat{\beta}_n)v_t + (\alpha_0 - \hat{\alpha}_n)(v_t - v_{t-1}) + u_t] \\ &= \hat{R}_{t-1} + h_t\end{aligned}$$

which takes the same form as in the previous section where no constant is included. Hence we still have

$$\hat{\rho}_n - 1 = \frac{\sum_{t=1}^n \hat{R}_{t-1} h_t}{\sum_{t=1}^n \hat{R}_{t-1}^2}.$$

Write the process of $\hat{R}_t$ as:

$$\hat{R}_t = y_t - \hat{\delta}_n - \hat{\beta}_n x_t - \hat{\alpha}_n v_t = (\beta_0 - \hat{\beta}_n)(x_t - \bar{x}) + (\alpha_0 - \hat{\alpha}_n)(v_t - \bar{v}) + (e_t - \bar{e}). \tag{33}$$

Comparing equation (33) with (27), the only difference in (33) is that all terms are subtracted by their sample means. This will correspond to demeaned Brownian motions instead of standard Brownian motions in the limit of the distribution of $\hat{\rho}_n$. Using similar methods as in the previous section, we can show that

$$\hat{\rho}_n - 1 = o_p(1) \quad \text{and} \quad n(\hat{\rho}_n - 1) = O_p(1).$$

Next, conduct the Cochrane-Orcutt transformation as in (9), and consider the OLS estimator in the regression

$$\tilde{y}_t = \tilde{\beta}_n \tilde{x}_t + \tilde{\alpha}_n \tilde{v}_t + error.$$

Define $z_t = (\tilde{x}_t, \tilde{v}_t)'$ and $\theta = (\beta, \alpha)'$, then

$$\tilde{\theta}_n = \left[(\sum_{t=1}^n z_t z_t')^{-1} \right] \left[\sum_{t=1}^n z_t \tilde{y}_t \right].$$

For $\tilde{y}_t$, write

$$\begin{aligned} \tilde{y}_t &= y_t - \hat{\rho}_n y_{t-1} \\ &= (1 - \hat{\rho}_n)\delta_0 + \beta_0 \tilde{x}_t + \alpha_0 \tilde{v}_t + u_t + (1 - \hat{\rho}_n) e_{t-1} \\ &= z_t' \theta_0 + (1 - \hat{\rho}_n)\delta_0 + u_t + (1 - \hat{\rho}_n) e_{t-1} \end{aligned}$$

Hence we can write

$$\hat{\theta}_n - \theta_0 = \left[\sum_{t=1}^n z_t z_t' \right]^{-1} \left[\sum_{t=1}^n z_t [(1 - \hat{\rho}_n)\delta_0 + u_t + (1 - \hat{\rho}_n) e_{t-1}] \right]. \tag{34}$$

The only difference of (34) with (29) is that we have a term $(1 - \hat{\rho}_n)\delta_0$ here. However, since δ_0 is just a finite constant and $n(1 - \hat{\rho}_n) = O_p(1)$, this term does not affect the asymptotics of $\tilde{\theta}_n$. Therefore, using the Cochrane-Orcutt transformation, the limit distribution of the estimators are the same no matter whether we have or have not a constant in the data generating process of the data. So we have the same result as given by (6).

Appendix C: Proof of results in section 2.2

To show the limit distribution of the dynamic OLS estimator in the cointegration, using the matrix H_n defined in (24), we can write

$$\begin{bmatrix} n(\hat{\beta}_n - \beta_0) \\ n^{1/2}(\hat{\alpha}_n - \alpha_0) \end{bmatrix} = \left[H_n^{-1} \begin{bmatrix} \sum_{t=1}^n x_t^2 & \sum_{t=1}^n x_t v_t \\ \sum_{t=1}^n x_t v_t & \sum_{t=1}^n v_t^2 \end{bmatrix} H_n^{-1} \right]^{-1} \begin{bmatrix} n^{-1} \sum_{t=1}^n x_t u_t \\ n^{-1/2} \sum_{t=1}^n v_t u_t \end{bmatrix}$$

For the first term on the right hand side,

$$\left[\begin{array}{cc} n^{-2}\sum_{t=1}^{n} x_t^2 & n^{-3/2}\sum_{t=1}^{n} x_t v_t \\ n^{-3/2}\sum_{t=1}^{n} x_t v_t & n^{-1}\sum_{t=1}^{n} v_t^2 \end{array} \right] \rightarrow \left[\begin{array}{cc} \sigma_1^2 \int_0^1 V(r)^2 dr & 0 \\ 0 & \sigma_v^2 \end{array} \right]. \tag{35}$$

so the estimator of the I(1) and I(0) components are asymptotically independent. For the second term on the right hand side,

$$\left[\begin{array}{c} n^{-1}\sum_{t=1}^{n} x_t u_t \\ n^{-1/2}\sum_{t=1}^{n} v_t u_t \end{array} \right] \rightarrow_d \left[\begin{array}{c} \sigma_1\sigma_2 \int_0^1 V(r) dU(r) \\ N(0, \sigma_v^2 \sigma_u^2) \end{array} \right]. \tag{36}$$

Equation (14) then follows.

To show the limit distribution for FGLS estimator in regression (15), write

$$\begin{aligned} & n^{-1}\sum_{t=1}^{n} \hat{u}_t^2 \\ = \; & n^{-1}\sum_{t=1}^{n} u_t^2 + 2\left(n^{-1}H_n^{-1}\sum_{t=1}^{n} u_t z_t'\right) H_n(\theta - \hat{\theta}) + H_n(\theta - \hat{\theta})' \left(n^{-1}H_n^{-1}\sum_{t=1}^{n} z_t z_t'\right) H_n(\theta - \hat{\theta}) \\ = \; & n^{-1}\sum_{t=1}^{n} u_t^2 + o_p(1) \rightarrow \sigma_u^2 \end{aligned}$$

Similarly, we can show that

$$n^{-1}\sum_{t=1}^{n} \hat{u}_t \hat{u}_{t-1} = n^{-1}\sum_{t=1}^{n} u_t u_{t-1} + o_p(1) \rightarrow \psi_u \sigma_u^2.$$

$$\hat{\rho}_n = \frac{n^{-1}\sum_{t=1}^{n} \hat{u}_t u_{t-1}}{n^{-1}\sum_{t=1}^{n} \hat{u}_t^2} \rightarrow_p \psi_u$$

Conduct the Cochrane-Orcutt transformation (9) and estimate

$$\tilde{y}_t = \tilde{\beta}\tilde{x}_t + \tilde{\alpha}\tilde{v}_t + error.$$

For the sequence of $\tilde{y}_t$, we can write it as

$$\tilde{y}_t = \beta_0 \tilde{x}_t + \alpha_0 \tilde{v}_t + u_t - \hat{\rho}_n u_{t-1}.$$

Using the same weight matrix H_n, write

$$\left[\begin{array}{c} n(\tilde{\beta} - \beta_0) \\ n^{1/2}(\tilde{\alpha} - \alpha_0) \end{array} \right] = \left[H_n^{-1} \left[\begin{array}{cc} \sum_{t=1}^{n} \tilde{x}_t^2 & \sum_{t=1}^{n} \tilde{x}_t \tilde{v}_t \\ \sum_{t=1}^{n} \tilde{x}_t \tilde{v}_t & \sum_{t=1}^{n} \tilde{v}_t^2 \end{array} \right] H_n^{-1} \right]^{-1} \left[\begin{array}{c} n^{-1}\sum_{t=1}^{n} \tilde{x}_t (u_t - \hat{\rho}_n u_{t-1}) \\ n^{-1/2}\sum_{t=1}^{n} \tilde{v}_t (u_t - \hat{\rho}_n u_{t-1}) \end{array} \right] \tag{37}$$

There are three different elements in the first term. Using similar methods as in the earlier proofs, we can show that

$$\begin{aligned} n^{-2}\sum_{t=1}^{n} \tilde{x}_t^2 & \rightarrow_d (1 - \psi_u)^2 \sigma_1^2 \int_0^1 V(r)^2 dr \\ n^{-2/3}\sum_{t=1}^{n} \tilde{x}_t \tilde{v}_t & \rightarrow_p 0 \\ n^{-1}\sum_{t=1}^{n} \tilde{v}^2 & \rightarrow (1 - 2\psi_u\psi_v + \psi_u^2)\sigma_v^2 \end{aligned}$$

Hence, the limit of the first item in (37) is

$$H_n^{-1}\left[\sum_{t=1}^{n} z_t z_t'\right] H_n^{-1} = \begin{bmatrix} n^{-2}\sum_{t=1}^{n}\tilde{x}^2 & n^{-3/2}\sum_{t=1}^{n}\tilde{x}\tilde{v} \\ n^{-3/2}\sum_{t=1}^{n}\tilde{x}\tilde{v} & n^{-1}\sum_{t=1}^{n}\tilde{v}^2 \end{bmatrix}$$
$$\rightarrow_d \begin{bmatrix} (1-\psi_u)^2\sigma_1^2\int_0^1 V(r)^2 dr & 0 \\ 0 & (1-2\psi_u\psi_v+\psi_u^2)\sigma_v^2 \end{bmatrix}.$$

Next, consider the second term in (37). Actually, we are only interested in the first element,

$$\begin{aligned} & n^{-1}\sum_{t=1}^{n}\tilde{x}(u_t-\hat{\rho}_n u_{t-1}) \\ = \;& n^{-1}\sum_{t=1}^{n}(v_t+(1-\hat{\rho}_n)x_{t-1})(u_t-\hat{\rho}_n u_{t-1}) \\ = \;& n^{-1}\sum_{t=1}^{n} v_t u_t - \hat{\rho}_n\sum_{t=1}^{n} v_t u_{t-1} + (1-\hat{\rho}_n)n^{-1}\sum_{t=1}^{n} x_{t-1}u_t - \hat{\rho}_n(1-\hat{\rho}_n)n^{-1}\sum_{t=1}^{n} x_{t-1}u_{t-1} \\ \rightarrow \;& (1-\psi_u)^2\sigma_1\sigma_2\int_0^1 V(r)dU(r). \end{aligned}$$

Therefore, we obtain the limit distribution for $\tilde{\beta}_n$ as given in (16).

Appendix D: Data Descriptions

For the purpose of comparisons, we use the same data set as in Cooley and Ogaki (1996) for the long-run intertemporal elasticity of substitution, and the data set of Stock and Watson (1993) for the U.S. money demand. Readers are referred to the original work for further details on data.

In the PPP application, two data sets have been employed. The first data set is comprised of quarterly (??) nominal exchange rates and consumer price indices during the post Bretton Woods period over 1973:1 - 1998:4 for 21 industrial countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States. They were retrieved from the International Monetary Fund's *International Financial Statistics (IFS)* and nominal exchange rates are end-of-quarter observations (IFS line code AE) while CPIs are quarterly averages (IFS line code 64). The second data set consists of monthly CPIs for durable goods and nominal exchange rates for Canada, Japan, and the U.S. They were obtained from the Organisation for Economic Cooperation and Development's (OECD) *Main Economic Indicators* CD-ROM for the period of 1965:1 - 2001:6.

Table 1: Application to Preference Parameter (β) Estimation

Estimator	k	AR(1) error		AR(2) error	
		ND	NDS	ND	NDS
DOLS	0	0.897 (0.041)	0.660 (0.039)	0.897 (0.043)	0.660 (0.049)
	1	0.897 (0.043)	0.661 (0.047)	0.897 (0.036)	0.661 (0.035)
	2	0.896 (0.040)	0.661 (0.041)	0.896 (0.041)	0.661 (0.042)
	3	0.895 (0.040)	0.662 (0.040)	0.895 (0.039)	0.662 (0.040)
	4	0.894 (0.043)	0.661 (0.045)	0.894 (0.040)	0.661 (0.046)
	5	0.892 (0.038)	0.660 (0.045)	0.892 (0.038)	0.660 (0.046)
	BIC	0.890 (0.034)	0.658 (0.044)	0.886 (0.042)	0.658 (0.053)
	[lag]	[6]	[10]	[9]	[10]
GLS	0	0.372 (0.050)	0.529 (0.049)	0.372 (0.050)	0.529 (0.049)
	1	0.637 (0.067)	0.620 (0.057)	0.637 (0.067)	0.620 (0.057)
	2	0.694 (0.072)	0.642 (0.059)	0.694 (0.072)	0.642 (0.059)
	3	0.765 (0.073)	0.671 (0.058)	0.765 (0.073)	0.671 (0.058)
	4	0.814 (0.075)	0.681 (0.059)	0.814 (0.075)	0.681 (0.059)
	5	0.827 (0.078)	0.683 (0.059)	0.827 (0.078)	0.683 (0.059)
	BIC	0.814 (0.075)	0.642 (0.059)	0.814 (0.075)	0.642 (0.059)
	[lag]	[4]	[2]	[4]	[2]
FGLS	0	0.901 (0.042)	0.662 (0.028)	0.901 (0.042)	0.662 (0.028)
	1	0.731 (0.030)	0.533 (0.025)	0.771 (0.026)	0.576 (0.021)
	2	0.749 (0.027)	0.550 (0.021)	0.736 (0.028)	0.536 (0.022)
	3	0.753 (0.030)	0.549 (0.023)	0.752 (0.030)	0.558 (0.023)
	4	0.744 (0.032)	0.531 (0.026)	0.752 (0.031)	0.518 (0.027)
	5	0.757 (0.031)	0.519 (0.030)	0.752 (0.031)	0.528 (0.031)
	BIC	0.744 (0.032)	0.531 (0.026)	0.752 (0.031)	0.558 (0.023)
	[lag]	[4]	[4]	[4]	[3]
Hausman-Test	0	7.003**	110.972**	7.003**	110.972**
	1	0.520	15.249**	0.520	15.249**
	2	0.111	7.950**	0.111	7.950**
	3	0.029	3.195*	0.029	3.195*
	4	0.119	1.134	0.119	1.134
	5	0.139	0.716	0.139	0.716

Note: Figures in the parenthesis represent standard errors. 'k' denotes the maximum length of leads and lags. GLS refers to the GLS corrected estimator which is obtained through regressing the first difference of variables without constant term. FGLS represents the FGLS estimator based on iterative Cochrane-Orcutt method. AR(1) error term is structured as $u_t = \rho u_{t-1} + \epsilon_t$ while AR(2) error term is $u_t = \delta_1 u_{t-1} + \delta_2 u_{t-2} + \epsilon_t$. The Hausman test statistic is $\frac{(\hat{\beta}_{OLS} - \tilde{\beta}_{GLS})^2)}{Var(\tilde{\beta}_{GLS})} \rightarrow \chi^2(1)$. The critical values of $\chi^2(1)$ are 2.71, 3.84 and 6.63 for ten, five, and one percent significance level. Single (double) asterisk represent that the null hypothesis of $\hat{\beta}_{OLS} = \tilde{\beta}_{GLS}$ can be rejected at 5% (1%) significance level.

Table 2: Application to Long Run U.S. Money Demand

Estimator	k	Equation 1		Equation 2		Equation 3	
		$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$
				AR(1) Error Term			
	0	0.944 (0.054)	-0.090 (0.015)	0.889 (0.057)	-0.308 (0.058)	0.850 (0.085)	0.906 (0.280)
	1	0.958 (0.048)	-0.096 (0.014)	0.884 (0.046)	-0.313 (0.045)	0.843 (0.066)	0.915 (0.216)
	2	0.970 (0.051)	-0.101 (0.014)	0.879 (0.044)	-0.320 (0.043)	0.837 (0.072)	0.941 (0.233)
DOLS	3	0.975 (0.055)	-0.104 (0.015)	0.871 (0.036)	-0.328 (0.035)	0.832 (0.062)	0.975 (0.205)
	4	0.967 (0.054)	-0.108 (0.015)	0.855 (0.029)	-0.334 (0.028)	0.824 (0.065)	0.995 (0.215)
	BIC	0.967 (0.054)	-0.108 (0.015)	0.851 (0.026)	-0.353 (0.026)	0.836 (0.054)	1.114 (0.193)
	[lag]	[4]		[6]		[6]	
	0	0.407 (0.081)	-0.014 (0.004)	0.419 (0.079)	-0.086 (0.022)	0.388 (0.078)	0.300 (0.082)
	1	0.654 (0.119)	-0.025 (0.010)	0.685 (0.115)	-0.177 (0.046)	0.643 (0.115)	0.506 (0.148)
	2	0.837 (0.134)	-0.050 (0.013)	0.848 (0.130)	-0.248 (0.053)	0.787 (0.133)	0.620 (0.161)
GLS	3	0.856 (0.145)	-0.067 (0.017)	0.884 (0.140)	-0.289 (0.061)	0.816 (0.146)	0.725 (0.185)
	4	0.962 (0.161)	-0.086 (0.022)	0.898 (0.151)	-0.283 (0.067)	0.811 (0.153)	0.654 (0.195)
	BIC	0.856 (0.145)	-0.067 (0.017)	0.884 (0.140)	-0.289 (0.061)	0.718 (0.155)	0.699 (0.229)
	[lag]	[3]		[3]		[7]	
	0	0.942 (0.052)	-0.083 (0.023)	0.893 (0.049)	-0.290 (0.079)	0.858 (0.071)	0.850 (0.435)
	1	0.888 (0.040)	-0.065 (0.009)	0.872 (0.035)	-0.278 (0.030)	0.815 (0.045)	0.744 (0.115)
	2	0.940 (0.045)	-0.081 (0.010)	0.901 (0.036)	-0.309 (0.031)	0.840 (0.054)	0.797 (0.128)
FGLS	3	0.980 (0.050)	-0.096 (0.011)	0.905 (0.029)	-0.330 (0.026)	0.851 (0.046)	0.912 (0.124)
	4	1.010 (0.045)	-0.108 (0.011)	0.886 (0.025)	-0.333 (0.023)	0.833 (0.051)	0.895 (0.133)
	BIC	1.006 (0.037)	-0.112 (0.009)	0.867 (0.024)	-0.347 (0.022)	0.824 (0.041)	0.986 (0.120)
	[lag]	[5]		[6]		[6]	
	0	289.892**		113.485**		79.188**	
	1	54.427**		10.203		9.946**	
Hausman-	2	15.059**		1.867		3.978	
Test	3	4.690		0.460		1.852	
	4	1.112		0.820		3.102	
				AR(2) Error Term			
	0	0.944 (0.034)	-0.090 (0.009)	0.889 (0.050)	-0.308 (0.050)	0.850 (0.076)	0.906 (0.251)
	1	0.958 (0.046)	-0.096 (0.013)	0.884 (0.048)	-0.313 (0.047)	0.843 (0.071)	0.915 (0.233)
	2	0.970 (0.046)	-0.101 (0.013)	0.879 (0.038)	-0.320 (0.037)	0.837 (0.062)	0.941 (0.201)
DOLS	3	0.975 (0.048)	-0.104 (0.013)	0.871 (0.035)	-0.328 (0.034)	0.832 (0.066)	0.975 (0.218)
	4	0.967 (0.050)	-0.108 (0.014)	0.855 (0.030)	-0.334 (0.028)	0.824 (0.077)	0.995 (0.255)
	BIC	0.967 (0.050)	-0.108 (0.014)	0.851 (0.026)	-0.353 (0.026)	0.832 (0.066)	0.975 (0.218)
	[lag]	[4]		[6]		[3]	
	0	0.407 (0.081)	-0.014 (0.004)	0.419 (0.079)	-0.086 (0.022)	0.388 (0.078)	0.300 (0.082)
	1	0.654 (0.119)	-0.025 (0.010)	0.685 (0.115)	-0.177 (0.046)	0.643 (0.115)	0.506 (0.148)
	2	0.837 (0.134)	-0.050 (0.013)	0.848 (0.130)	-0.248 (0.053)	0.787 (0.133)	0.620 (0.161)
GLS	3	0.856 (0.145)	-0.067 (0.017)	0.884 (0.140)	-0.289 (0.061)	0.816 (0.146)	0.725 (0.185)
	4	0.962 (0.161)	-0.086 (0.022)	0.898 (0.151)	-0.283 (0.067)	0.811 (0.153)	0.654 (0.195)
	BIC	0.856 (0.145)	-0.067 (0.017)	0.884 (0.140)	-0.289 (0.061)	0.718 (0.155)	0.699 (0.229)
	[lag]	[3]		[3]		[7]	
	0	0.942 (0.052)	-0.083 (0.023)	0.893 (0.049)	-0.290 (0.079)	0.858 (0.071)	0.850 (0.435)
	1	0.900 (0.039)	-0.069 (0.009)	0.872 (0.038)	-0.276 (0.031)	0.809 (0.049)	0.722 (0.118)
	2	0.948 (0.042)	-0.086 (0.010)	0.894 (0.033)	-0.312 (0.029)	0.839 (0.049)	0.830 (0.131)
FGLS	3	0.991 (0.044)	-0.100 (0.011)	0.907 (0.029)	-0.332 (0.026)	0.853 (0.050)	0.903 (0.128)
	4	1.012 (0.042)	-0.109 (0.010)	0.889 (0.026)	-0.335 (0.023)	0.827 (0.061)	0.856 (0.142)
	BIC	1.012 (0.042)	-0.109 (0.010)	0.857 (0.024)	-0.346 (0.021)	0.806 (0.043)	0.970 (0.121)
	[lag]	[4]		[6]		[6]	
	0	289.892**		113.485**		79.188**	
	1	54.427**		10.203		9.946**	
Hausman-	2	15.059**		1.867		3.978	
Test	3	4.690		0.460		1.852	
	4	1.112		0.820		3.102	

Note: Figures in the parenthesis represent standard errors. 'k' denotes the maximum length of leads and lags. In dynamic GLS, the serial correlation coefficient in error term is estimated before being applied to the Cochrane-Orcutt transformation whereas it is assumed to be unity in GLS estimation which is analogous to regressing the first difference of variables without constant term. AR(1) error term is structured as $u_t = \rho u_{t-1} + \epsilon_t$ while AR(2) error term is $u_t = \delta_1 u_{t-1} + \delta_2 u_{t-2} + \epsilon_t$. The Hausman test statistic is $(\hat{\Gamma}_{OLS} - \tilde{\Gamma}_{GLS})\Sigma(\hat{\Gamma}_{OLS} - \tilde{\Gamma}_{GLS})' \rightarrow \chi^2(2)$ where $\Gamma = [\beta, \gamma]$ and $\Sigma = \begin{bmatrix} var(\tilde{\beta}_{GLS}) & cov(\tilde{\beta}_{GLS}, \tilde{\gamma}_{GLS}) \\ cov(\tilde{\beta}_{GLS}, \tilde{\gamma}_{GLS}) & var(\tilde{\gamma}_{GLS}) \end{bmatrix}$. The critical values of $\chi^2(2)$ are 4.61, 5.99 and 9.21 for 10%, 5%, and 1%

significance levels. Single (double) asterisk represent that the null hypothesis of $\hat{\beta}_{OLS} = \tilde{\beta}_{GLS}$ can be rejected at 5% (1%).

References

[1] Billingsley, P. (1986), *Probability and Measure*, *Second Edition*, New York: Wiley.

[2] Blough, S.R. (1992), "Spurious Regressions, with AR(1) Correction and Unit Root Pretest," *Mimeo*, Johns Hopkins University.

[3] Cooley, T.F. and M. Ogaki (1996): "A Time Series Analysis of Real Wages, Consumption, and Asset Returns," *Journal of Applied Econometrics*, $11, 119 - 134$.

[4] de Jong, R. and J. Davidson (2000) The functional central limit theorem and weak convergence to stochastic integrals I: weakly dependent processes, *Econometric Theory*, Vol. 16, 621-642.

[5] Durlauf, S.N. and P.C.B. Phillips (1988): "Trends versus Random Walks in Time Series Analysis," *Econometrica*, $56, 1333 - 1354$.

[6] Granger, C.W.J. and P. Newbold (1974): "Spurious Regressions in Econometrics," *Journal of Econometrics*, $74, 111 - 120$.

[7] Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl, and T.-C. Lee. 1985. *The Theory And Practice of Econometrics*, Wiley.

[8] Nelson, C.R. and H. Kang (1981): "Spurious Periodicity in Inappropriately Detrended Time Series," *Econometrica* $49, 741 - 751$.

[9] _____ (1983) "Pitfalls in the Use of Time as an Explanatory Variable in Regression," *Journal of Business and Economic Statistics* $2, 73 - 82$.

[10] Ogaki, M. and C.Y. Choi (2001), "The Gauss-Marcov Theorem and Spurious Regressions," WP 01-13, Department of Economics, The Ohio State University.

[11] Park, J.Y. and P.C.B. Phillips (1988) "Asymptotic Equivalence of Ordinary Least Squares and Generalized least Squares in Regressions with Integrated Regressors," *Journal of the American Statistical Association*,$83 : 111 - 115$.

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

[13] _____ (1989) "Partially Identified Econometric Models," *Econometric Theory*, $5 : 95 - 132$.

[14] _____ (1998) "New Tools for Understanding Spurious Regression," *Econometrica* $66, 1299 - 1325$.

[15] Phillips, P.C.B. and D.J. Hodgson (1994) "Spurious Regression and Generalized Least Squares," *Econometric Theory*, $10 : 957 - 958$.

[16] Phillips, P.C.B. and M. Loretan (1991) "Estimating Long-run Economic Equilibria," *Review of Economic Studies*, $58 : 407 - 436$.

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