

Robust estimation for structural spurious regressions and a Hausman-type cointegration test

Chi-Young Choi^a, Ling Hu^b, Masao Ogaki^{b,*}

^a*Department of Economics, University of Texas at Arlington, USA*

^b*Department of Economics, Ohio State University, Columbus, OH 43210-1172, USA*

Received 23 February 2006; received in revised form 12 June 2007; accepted 20 June 2007

Available online 6 July 2007

Abstract

This paper analyzes an approach to correcting spurious regressions involving unit-root nonstationary variables by generalized least squares (GLS) using asymptotic theory. This analysis leads to a new robust estimator and a new test for dynamic regressions. The robust estimator is consistent for structural parameters not just when the regression error is stationary but also when it is unit-root nonstationary under certain conditions. We also develop a Hausman-type test for the null hypothesis of cointegration for dynamic ordinary least squares (OLS) estimation. We demonstrate our estimation and testing methods in three applications: (i) long-run money demand in the U.S., (ii) output convergence among industrial and developing countries, and (iii) purchasing power parity (PPP) for traded and non-traded goods.

© 2007 Elsevier B.V. All rights reserved.

JEL classification: C10; C15

Keywords: Spurious regression; GLS correction method; Dynamic regression; Test for cointegration

1. Introduction

In the unit-root literature, a regression is technically called a spurious regression when its stochastic error is unit-root nonstationary. 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 (OLS). 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). Phillips (1986, 1998) and Durlauf and Phillips (1988) among others have studied the asymptotic properties of estimators and test statistics for regression coefficients of these spurious regressions. This paper analyzes an approach to correct spurious regressions involving unit-root nonstationary variables by generalized least squares (GLS) using asymptotic theory. This analysis leads to a new robust estimator and a new test for dynamic regressions. The robust estimator is consistent for structural parameters not just when

*Corresponding author. Tel.: +1 614 292 5842; fax: +1 614 292 3906.

E-mail address: ogaki.1@osu.edu (M. Ogaki).

the regression error is stationary but also when it is unit-root nonstationary under certain conditions. We also develop a Hausman-type test for the null hypothesis of cointegration for dynamic OLS estimation.

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 because of a nonstationary measurement error in one variable or nonstationary omitted variables. In such cases, it is still possible to consistently estimate structural variables under certain conditions. When the error is unit-root nonstationary but structural parameters can be recovered, the regression is called a structural spurious regression.

As an example of a structural spurious regression, consider a regression to estimate the money demand function when money is measured with a nonstationary error. Currency held by domestic economic agents for legitimate transactions is very hard to measure, since currency is held by foreign residents and is also used for black market transactions. Therefore money may be measured with a nonstationary error. As shown by [Stock and Watson \(1993\)](#) among others, when the money demand function is stable in the long run, we have a cointegrating regression if all variables are measured without error. If the variables are measured with stationary measurement errors, we still have a cointegrating regression. If money is measured with a nonstationary measurement error, however, we have a spurious regression. We can still recover structural parameters under certain conditions. The crucial assumption is that the nonstationary measurement error is not cointegrated with the regressors.

Another example of a structural spurious regression is a regression of money demand with nonstationary omitted variables. Consider the case in which money demand is stable in the long run and a measure of shoe leather costs of holding money is included as an argument. If an econometrician omits the measure of the shoe leather costs from the money demand regression and the measure is nonstationary, then the regression error is nonstationary. The shoe leather costs of holding money are related to the value of time and, therefore, to the real wage rate. Because the real wage rate is nonstationary in standard dynamic stochastic general equilibrium models with a nonstationary technological shock, the omitted measure of the shoe leather costs is likely to be nonstationary. In this case, the money demand regression that omits the measure is spurious, but we can still recover structural parameters under certain conditions. The crucial assumption is that the omitted variable is not cointegrated with the regressors.

Our structural spurious regression approach is based on the GLS solution of the spurious regression problem analyzed by [Ogaki and Choi \(2001\)](#),¹ who use an exact small sample analysis based on the conditional probability version of the Gauss–Markov Theorem. We develop asymptotic theory for two estimators motivated by the GLS correction: the GLS corrected dynamic regression estimator and the feasible GLS (FGLS) corrected dynamic regression estimator. Because Ogaki and Choi only used an exact small sample analysis, they did not consider the FGLS corrected estimator. We will show these estimators to be consistent and asymptotically normally distributed in spurious regressions. When the error term is in fact stationary and hence the variables are cointegrated, the GLS corrected estimator is not efficient, but the FGLS corrected estimator, like the OLS estimator, is superconsistent. Hence, FGLS estimation is a robust procedure with respect to the error specification. The FGLS corrected estimator is asymptotically equivalent to the GLS corrected estimator in spurious regressions, and it is asymptotically equivalent to the OLS estimator in cointegrating regressions.

In some applications, it is hard to determine whether or not the error in the regression is stationary or unit-root nonstationary because test results are inconclusive. In such applications, the FGLS corrected estimator is attractive because it is consistent in both situations as long as the method of the dynamic regression removes the endogeneity problem.

This approach naturally motivates a Hausman-type test² for the null hypothesis of cointegration against the alternative hypothesis of no cointegration (or a spurious regression) in the dynamic OLS framework.

¹Another approach would be to take the first difference to induce stationarity and then use instrumental variables. This is the approach proposed by [Lewbel and Ng \(2005\)](#) for their nonstationary translog demand system. Our approach exploits the particular form of endogeneity assumed by many authors in the cointegration literature and avoids the use of instrumental variables. Our approach yields more efficient estimators as long as the particular form of endogeneity is correctly specified. This is especially important when weak instruments cause problems.

²This test can also be called a Durbin–Wu–Hausman type test as it is closely related to ideas and tests in [Durbin \(1954\)](#) and [Wu \(1973\)](#) as well as a family of tests proposed by [Hausman \(1978\)](#).

We construct this test by noting that while both the dynamic OLS and GLS corrected dynamic regression estimators are consistent in cointegration estimation, the dynamic OLS estimator is more efficient.³ 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 specification on the error. We show that under the null hypothesis of cointegration the test statistics have a χ^2 limit distribution, while under the alternative hypothesis of a spurious regression the test statistics diverge.

In some applications the assumption that the spurious regression is structural under the alternative hypothesis is not very attractive. If the violation of cointegration arises for reasons other than nonstationary measurement error or omitted variables, it is hard to believe that the resulting spurious regression is structural. For this reason we relax the assumption that the spurious regression is structural and show that the Hausman-type cointegration test statistic still diverges under the alternative hypothesis.

Dynamic OLS is used in many applications of cointegration. However, few tests for cointegration have been developed for dynamic OLS, with the exception of Shin's (1994) test. As in Phillips and Ouliaris (1990), the popular augmented Dickey–Fuller (ADF) test for the null hypothesis of no cointegration was originally designed to be applied to the residual from static OLS rather than the residual from dynamic OLS. Because the static OLS and dynamic OLS estimates are often substantially different, it is desirable to have a test for cointegration applicable to dynamic OLS. Another aspect of our Hausman-type test is that it is for the null hypothesis of cointegration. Ogaki and Park (1998) argue that it is desirable to test the null hypothesis of cointegration rather than that of no cointegration in many applications where economic models imply cointegration.

Using Monte Carlo experiments, we compare the finite sample performance of the Hausman-type test with the test proposed by Shin (1994), which is a locally best invariant test for the null of zero variance of a random walk component in the disturbances. According to the experiment results, the Hausman-type test is dominant in both size and power up to the sample size of 300. Shin's test becomes more powerful when the sample size increases, but only at the cost of higher size distortion.

In some applications, it is appropriate to consider the possibility that measurement error is $I(1)$ and is not cointegrated with the regressors. For these applications, the ADF test is applicable under the null hypothesis of a structural spurious regression, as shown by Hu (2006). For such applications, we recommend that both the ADF test and the Hausman-type test be applied because it is not clear which null hypothesis is more appropriate.

We demonstrate our estimation and testing methods in three applications: (i) long-run money demand in the U.S., (ii) output convergence among industrial and developing countries, and (iii) purchasing power parity (PPP) for traded and non-traded goods. In the first application, we focus on estimating unknown structural parameters, while in the last two applications we purport to testing for cointegration with the Hausman-type cointegration test where we relax the assumption that the spurious regression under the alternative hypothesis is structural.

The rest of the paper is organized as follows. Section 2 gives econometric analysis of the model, including asymptotic theories and finite sample simulation studies. Section 3 presents models of nonstationary measurement error and nonstationary omitted variables as well as the empirical results of three applications. Section 4 contains concluding remarks.

2. The econometric model

Consider the regression model

$$y_t = \beta' x_t + \eta_t, \quad (1)$$

where $\{x_t\}$ is an m -vector integrated process generated by

$$\Delta x_{it} = v_{it}.$$

³After completing the first draft, it has come to our attention that the Hausman-type test was originally proposed by Fernández-Macho and Mariel (1994) for the static OLS cointegrating regression with strict exogeneity and without any serial correlation. The test has not been popular probably because these assumptions are hard to justify in applications, and because the test was not developed for dynamic regressions.

The error term in (1) is assumed to be

$$\eta_t = \sum_{i=1}^m \sum_{j=-k}^k \gamma_{ij} v_{i,t-j} + e_t, \quad (2)$$

$$e_t = \rho e_{t-1} + u_t. \quad (3)$$

Assumption 1. Assume that $v_t = (v_{1t}, \dots, v_{mt})'$ and u_t are zero mean stationary processes with $E|v_{it}|^\alpha < \infty$, $E|u_t|^\alpha < \infty$ for some $\alpha > 2$, and strong mixing with size $-\alpha/(\alpha - 2)$. We also assume that the dynamic regression method removes the endogeneity problem. That is, $E(u_t v_s) = \mathbf{0}$ for all t, s . We call this the strict exogeneity assumption for the dynamic regression.

The conditions on v_t and u_t ensure the invariance principles: for $r \in [0, 1]$, $n^{-1/2} \sum_{t=1}^{[nr]} v_t \rightarrow_d V(r)$ and $n^{-1/2} \sum_{t=1}^{[nr]} u_t \rightarrow_d U(r)$, where $V(r)$ is an m -vector Brownian motion with covariance $\sum_{j=-\infty}^{\infty} E(v_t v_{t-j}')$ and $U(r)$ is a Brownian motion with variance $\sum_{j=-\infty}^{\infty} E(u_t u_{t-j})$. The functional central limit theorem holds for weaker assumptions than assumed here (de Jong and Davidson, 2000), but the conditions assumed above are general enough to include many stationary Gaussian or non-Gaussian ARMA processes that are commonly assumed in empirical modeling.

Let $\mathbf{v}_t = (\Delta x_{1,t-k}, \dots, \Delta x_{1,t}, \dots, \Delta x_{1,t+k}, \dots, \Delta x_{m,t-k}, \dots, \Delta x_{m,t}, \dots, \Delta x_{m,t+k})'$, and $\boldsymbol{\gamma} = (\gamma_{1,-k}, \dots, \gamma_{1,0}, \dots, \gamma_{1,k}, \dots, \gamma_{m,-k}, \dots, \gamma_{m,0}, \dots, \gamma_{m,k})'$. We estimate the structural parameter β in the regression

$$y_t = \beta' x_t + \boldsymbol{\gamma}' \mathbf{v}_t + e_t. \quad (4)$$

The inference procedure about β differs according to the different assumptions on the error term e_t in (3). When $|\rho| < 1$, e_t is stationary, and hence regression (4) is a cointegration regression with serially correlated error. When $\rho = 1$, e_t is a unit-root nonstationary process and the OLS regression is spurious. Both models are important in empirical studies in macroeconomics and finance.

In the next two sections, we will study the asymptotic properties of different estimation procedures under these two assumptions. Under the assumption that $\rho = 1$, OLS is not consistent while both the GLS correction and FGLS correction will give consistent and asymptotically equivalent estimators. Under the assumption that $|\rho| < 1$, the GLS corrected estimator is not efficient as it is \sqrt{n} convergent, but the FGLS estimator is n convergent and asymptotically equivalent to the OLS estimator. Therefore, FGLS is robust with respect to the error specifications ($\rho = 1$ or $|\rho| < 1$).

2.1. Regressions with $I(1)$ error

In this section we consider the situation when the error term is $I(1)$, i.e., $\rho = 1$ in (3). The estimation methods we study are dynamic OLS, the GLS correction, and the FGLS correction.

2.1.1. The dynamic OLS spurious estimation

We start with the dynamic OLS estimation of regression (4). Under the assumption of $\rho = 1$, this regression is spurious since for any value of β the error term is always $I(1)$. In Appendix A, we show that the DOLS estimator $\hat{\beta}_{\text{dols}}$ has the following limit distribution:

$$(\hat{\beta}_{\text{dols}} - \beta_0) \rightarrow_d \left[\int_0^1 V(r) V(r)' dr \right]^{-1} \left[\int_0^1 V(r) U(r) dr \right]. \quad (5)$$

$\hat{\boldsymbol{\gamma}}$ in the estimation is also inconsistent with $\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0 = O_p(1)$. As remarked in Phillips (1986, 1989), 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 $\rho = 1$, we can filter all variables in regression (4) by taking the full first difference and use OLS to estimate

$$\Delta y_t = \beta' \Delta x_t + \gamma' \Delta v_t + u_t = \theta' \Delta z_t + u_t, \tag{6}$$

where $\theta = (\beta', \gamma)'$ and $z_t = (x_t', v_t)'$. This procedure can be viewed as GLS corrected estimation.⁴

If we let $\tilde{\theta}_{\text{dglS}}$ denote the GLS corrected estimator, then we can show that

$$\sqrt{n}(\tilde{\theta}_{\text{dglS}} - \theta_0) \rightarrow_d N(\mathbf{0}, \Omega), \tag{7}$$

where $\Omega = Q^{-1} A Q^{-1}$ with $Q = E(\Delta z_t \Delta z_t')$ and A being the long-run variance matrix of $\Delta z_t u_t$. Thus β in a structural spurious regression can be consistently estimated (jointly with γ), and the estimators are asymptotically normal. In the special case when $m = 1$, $\{v_{1t}\}$ and $\{u_t\}$ are *i.i.d.* sequences and $\eta_t = e_t$, (7) gives that

$$\sqrt{n}(\tilde{\theta}_{\text{dglS}} - \theta_0) \rightarrow_d N(0, \sigma_u^2 / \sigma_{1v}^2),$$

where σ_u^2 and σ_{1v}^2 are the variances of u_t and v_{1t} , respectively.

2.1.3. The FGLS estimation

To use GLS to estimate a regression with serial correlation in empirical work, a Cochrane–Orcutt FGLS procedure is usually adopted. This procedure also works for spurious regressions as shown by Phillips and Hodgson (1994). They show that the FGLS estimator is asymptotically equivalent to that in the differenced regression when the error is unit-root nonstationary. In the present paper, we will show that the FGLS correction to the dynamic regression provides a consistent and robust estimator for structural spurious regressions.

Let the residual from the OLS regression (4) be denoted by \hat{e}_t ,

$$\hat{e}_t = y_t - \hat{\beta}'_n x_t - \hat{\gamma}'_n v_t.$$

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

$$\hat{e}_t = \hat{\rho}_n \hat{e}_{t-1} + \hat{u}_t. \tag{8}$$

It can be shown that $n(\hat{\rho}_n - 1) = O_p(1)$. Conduct the following Cochrane–Orcutt transformation of the data:

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

Then consider OLS estimation of the regression

$$\tilde{y}_t = \beta' \tilde{x}_t + \gamma' \tilde{v}_t + error = \theta' \tilde{z}_t + error, \tag{10}$$

where $\tilde{z}_t = (\tilde{x}_t', \tilde{v}_t)'$. The OLS estimator of θ in (10) is computed as

$$\tilde{\theta}_{\text{fglS}} = \left[\sum_{t=1}^n \tilde{z}_t \tilde{z}_t' \right]^{-1} \left[\sum_{t=1}^n \tilde{z}_t \tilde{y}_t \right]. \tag{11}$$

The limiting distribution of $\tilde{\theta}_{\text{fglS}}$ can be shown to be the same as in (7). Intuitively, even though the dynamic OLS estimator is inconsistent, the residual is unit-root nonstationary because no linear combination of y_t and x_t is stationary. Therefore, $\hat{\rho}_n$ approaches unity in the limit, and \tilde{z}_t behaves asymptotically equivalently to Δz_t . A detailed proof of results in this section is given in Appendix A.

⁴This is a conventional GLS procedure when u_t is *i.i.d.* When u_t is serially correlated as in our approach, we name this procedure GLS corrected dynamic estimation.

2.2. Regressions with $I(0)$ error

In this section, we consider the asymptotic distributions of the three estimators (the DOLS estimator, the GLS corrected estimator, and the FGLS corrected estimator) under the assumption of cointegration, i.e., $|\rho| < 1$ in (3).

2.2.1. The dynamic OLS estimation

Under the assumption of cointegration, the DGP of y_t is

$$y_t = \beta' x_t + \gamma' v_t + e_t, \quad e_t = \rho e_{t-1} + u_t, \quad |\rho| < 1. \quad (12)$$

Applying the invariance principle, for $r \in [0, 1]$, $n^{-1/2} \sum_{t=1}^{[nr]} e_t \rightarrow_d E(r)$, where $E(r)$ is a Brownian motion with variance $\sum_{j=-\infty}^{\infty} E(e_t e_{t-j})$. The limiting distribution of the OLS estimator of β , which is asymptotically independent of $\hat{\gamma}_n$, is known to be

$$n(\hat{\beta}_{\text{dols}} - \beta_0) \rightarrow_d \left(\int_0^1 V(r) V(r)' dr \right)^{-1} \left(\int_0^1 V(r) dE(r) \right). \quad (13)$$

2.2.2. GLS corrected estimation

We now take the full first difference as we did in the spurious regressions, the regression becomes

$$\Delta y_t = \beta' \Delta x_t + \gamma' \Delta v_t + e_t - e_{t-1} = \theta' \Delta z_t + e_t - e_{t-1}. \quad (14)$$

Note that this transformation leads to a loss in efficiency since the estimator $\tilde{\beta}_{\text{dglS}}$ is now \sqrt{n} convergent rather than n convergent as the DOLS estimator is. With some minor revisions to equation (7), the limiting distribution of the estimator in this case can be written as

$$\sqrt{n}(\tilde{\theta}_{\text{dglS}} - \theta_0) \rightarrow_d N(\mathbf{0}, \Omega^*), \quad (15)$$

where $\Omega^* = Q^{-1} A^* Q^{-1}$. Q is again defined as $Q = E(\Delta z_t \Delta z_t')$ and A^* is the long-run variance matrix of vector $\Delta z_t \Delta e_t$. In the special case when $m = 1$, $\{v_{1t}\}$ and $\{u_t\}$ are *i.i.d.* sequences, $\eta_t = e_t$, and $\Omega^* = 2\sigma_e^2(1 - \psi_e)/\sigma_{1v}^2$, where ψ_e is the first-order autocorrelation coefficient of $\{e_t\}$.

2.2.3. The FGLS estimation

Instead of taking the full first difference, if we estimate the autoregression coefficient in the error and use this estimator to filter all sequences, we will obtain an estimator that is asymptotically equivalent to the DOLS estimator. Intuitively, in the case when the error $e_t = u_t$ is serially uncorrelated, the AR(1) coefficient $\hat{\rho}_n$ will converge to zero, and hence the transformed regression will be asymptotically equivalent to the original regression. If, on the other hand, the error is stationary and serially correlated, then the AR(1) coefficient will be less than unity, and, as shown in Phillips and Park (1988), the GLS estimator and the OLS estimator in a cointegration regression are asymptotically equivalent.

If we conduct the Cochrane–Orcutt transformation (9) and estimate β in the regression

$$\tilde{y}_t = \tilde{\beta}'_n \tilde{x}_t + \tilde{\gamma}'_n \tilde{v}_t + \text{error}, \quad (16)$$

then Appendix B shows that the limiting distribution of $\tilde{\beta}_n$ is the same as the limit of the OLS estimator given in (13).

2.3. FGLS: a robust estimator with respect to the order of errors

From our discussions on the FGLS estimator in Sections 2.1.3 and 2.2.3, we can summarize the FGLS corrected estimator in the following proposition:

Proposition 1. *Suppose Assumption 1 holds. In spurious regressions, the FGLS corrected estimator is asymptotically equivalent to the GLS corrected estimator, and its limit distribution can be written as*

$$\sqrt{n}(\tilde{\theta}_{\text{fgls}} - \theta_0) \rightarrow_d N(\mathbf{0}, \Omega).$$

In cointegration regressions, the FGLS corrected estimator is asymptotically equivalent to the DOLS estimator, and its limit distribution can be written as

$$n(\hat{\beta}_{fpls} - \beta_0) \rightarrow_d \left(\int_0^1 V(r)V(r)' dr \right)^{-1} \left(\int_0^1 V(r) dE(r) \right).$$

So FGLS is not only valid when the regression is spurious but also asymptotically efficient when the regression is cointegration.

Remarks.

1. If a constant is added to (4), we can show that the GLS or FGLS corrected estimators are asymptotically equivalent to that given in (7) under the assumption of spurious regressions.
2. If a trend term is added to (4) (this is the case in which the deterministic cointegration restriction is not satisfied in the terminology of Ogaki and Park, 1998), then the GLS corrected estimation leads to a singular covariance matrix for the estimator when ρ is less than one in absolute value. This is because a trend term in (4) leads to a constant term in the first differenced regression (14) and because the long-run variance of the first difference of e_t multiplied by a constant is zero. Therefore, our methods do not apply to regressions with time trends.
3. Under some conditions, the methods proposed in this paper also apply to other model configurations, such as regressions where the regressors have drifts. These extensions will be studied in future work.

2.4. Finite sample performance of the three estimators

From the above analysis, we show that FGLS corrected estimation is a robust procedure with respect to error specifications. In this section, we use simulations to study its finite sample performances compared to the other two estimators. In the simulation we consider the case when x_t is a scalar variable and generate v_t and u_t from two independent standard normal distributions while letting $e_t = \rho e_{t-1} + u_t$. The structural parameter is set to $\beta = 2$, and $\gamma'v_t = 0.5v_t$. The number of iterations in each simulation is 5000, and in each replication $100 + n$ observations are generated, of which the first 100 observations are discarded.

Table 1 shows the bias and the mean square error (MSE) of all three estimators for $\rho = 0, 0.95$, and 1. When $\rho = 0$, the regression is cointegration with *i.i.d.* error. It is clear that the DOLS estimator is the best one when $n = 50$. When n reaches 100, however, the FGLS estimator becomes almost as good as the DOLS estimator. When $\rho = 0.95$, the regression is cointegration with serially correlated error. In this case, the GLS and FGLS estimators are much better than the DOLS estimator. When the sample size increases, the FGLS estimator

Table 1
The bias and square root of the mean square error of three estimators

ρ	n	DOLS estimator		GLS corrected estimator		FGLS corrected estimator	
		Bias	Square root of MSE	Bias	Square root of MSE	Bias	Square root of MSE
$\rho = 0$	$n = 50$	-0.0000	0.0290	0.0031	0.2156	-0.0000	0.0296
	$n = 100$	-0.0002	0.0169	0.0004	0.1448	-0.0002	0.0170
	$n = 500$	-0.0000	0.0040	0.0010	0.0634	-0.0000	0.0040
$\rho = 0.95$	$n = 50$	-0.0007	0.3555	0.0019	0.2104	0.0000	0.2162
	$n = 100$	0.0070	0.2415	-0.0010	0.1472	0.0036	0.1375
	$n = 500$	0.0006	0.0773	-0.0007	0.0637	-0.0000	0.0475
$\rho = 1$	$n = 50$	-0.0347	1.4769	0.0024	0.2103	-0.0095	0.5968
	$n = 100$	-0.0113	1.2809	0.0028	0.1439	0.0031	0.3635
	$n = 500$	-0.0086	0.9895	0.0003	0.0643	-0.0011	0.1167

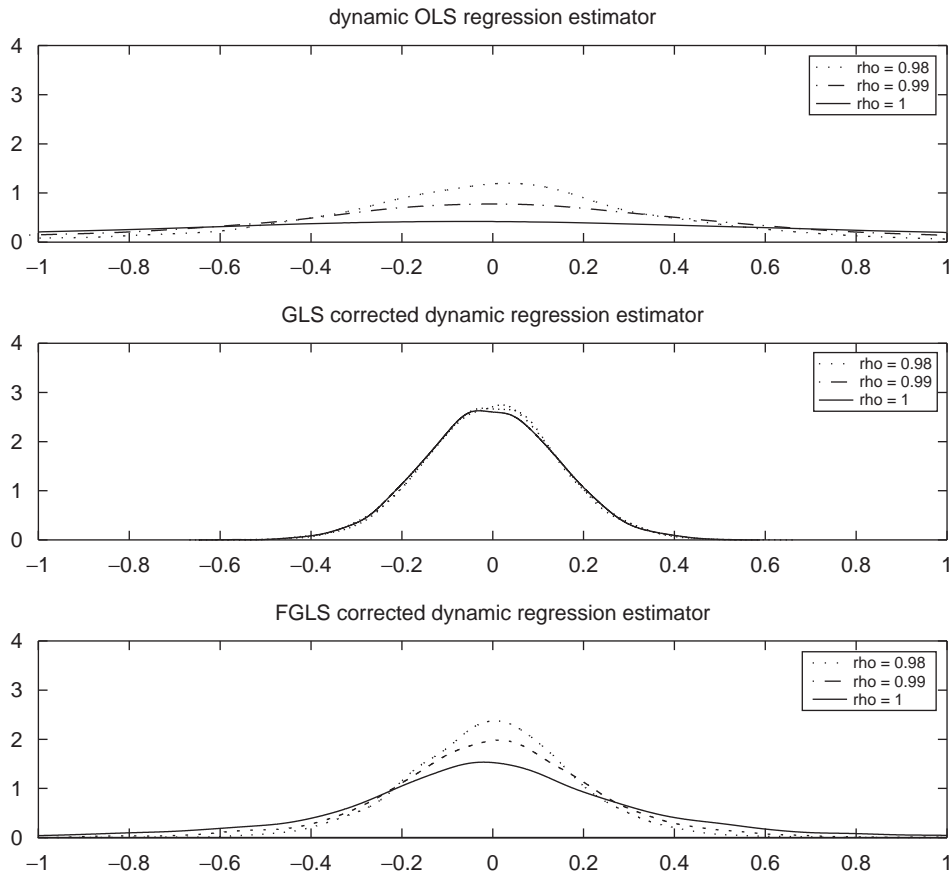


Fig. 1. Comparison of three estimators when $n = 100$ and $\rho \rightarrow 1$.

becomes the best one. Finally, when $\rho = 1$, the regression is spurious, and, as expected, the GLS corrected estimator performs best.

Fig. 1 plots the empirical distribution of these three estimators (minus the true value) when $n = 100$ and as ρ approaches 1. The figures show that the DOLS estimator becomes flatter and flatter as $\rho \rightarrow 1$. The GLS estimator remains largely the same for ρ close to unity. The FGLS estimator becomes a bit flatter when ρ reaches 1, but it still shows a clear peak around zero.

From the finite sample performance, it can be seen that the FGLS estimator is almost as good as the DOLS estimator in cointegration and significantly outperforms the DOLS estimator in spurious regressions. The GLS estimator is the best when ρ approaches 1, but it suffers from a significant loss in efficiency when ρ is small. So we may want to take the full difference only when we are very sure that the error is unit root nonstationary. Otherwise, the FGLS estimator is a good choice.

2.5. Hausman specification test for cointegration

2.5.1. The test statistic and its asymptotic properties

In this section, we construct a Hausman-type cointegration test based on the difference of two estimators: an OLS estimator ($\hat{\beta}_{\text{dols}}$) and a GLS corrected estimator ($\tilde{\beta}_{\text{dglS}}$). This is equivalent to comparing estimators in a level regression and in a differenced regression. The test is for the null of cointegrating relationships against the alternative of a spurious regression:

$$H_0: |\rho| < 1 \quad \text{against} \quad H_A: \rho = 1.$$

Our discussions so far suggest that under the null of cointegration, both OLS and GLS are consistent but the OLS estimator is more efficient. Under the alternative of a spurious regression, however, only the GLS corrected estimator is consistent.

Let \hat{V}_β denote a consistent estimator for the asymptotic variance of $\sqrt{n}(\tilde{\beta}_{\text{dglS}} - \beta)$. Under our assumptions, it converges to the corresponding submatrix of Ω^* under the null hypothesis and to the corresponding submatrix of Ω under the alternative. For example, when $m = 1$, $\{v_{1t}\}$ and $\{u_t\}$ are independent *i.i.d.*, and $\eta_t = e_t$, take $\hat{V}_\beta = (\frac{1}{n}\sum_{t=1}^n \hat{w}_t^2) / (\frac{1}{n}\sum_{t=1}^n \Delta x_t^2)$, where \hat{w}_t denotes the residuals from OLS estimation of the differenced regression. Under the null of cointegration, $\hat{V}_\beta \rightarrow_p 2\sigma_e^2(1 - \psi_e) / \sigma_{1v}^2$. Under the alternative of spurious regression, $\hat{V}_\beta \rightarrow_p \sigma_u^2 / \sigma_{1v}^2$.

We define the Hausman-type test statistic as:

$$h_n = n(\tilde{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}})' \hat{V}_\beta^{-1} (\tilde{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}}). \tag{17}$$

Proposition 2. *Suppose Assumption 1 holds. Under the null hypothesis of cointegration, $h_n \rightarrow_d \chi^2(m)$. Under the alternative of spurious regressions, $h_n = O_p(n)$.*

Proof. Under the null of cointegration,

$$\begin{aligned} \sqrt{n}(\tilde{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}}) &= \sqrt{n}(\tilde{\beta}_{\text{dglS}} - \beta_0) - \sqrt{n}(\hat{\beta}_{\text{dols}} - \beta_0) \\ &= \sqrt{n}(\tilde{\beta}_{\text{dglS}} - \beta_0) + o_p(1) \\ &\rightarrow_d N(0, V_\beta), \end{aligned}$$

where V_β is the asymptotic variance of $\tilde{\beta}_{\text{dglS}}$ under the assumption of cointegration. Therefore, if \hat{V}_β is a consistent estimator for V_β ,

$$h_n = n(\tilde{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}})' (\hat{V}_\beta)^{-1} (\tilde{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}}) \rightarrow_d \chi^2(m).$$

Under the alternative of spurious regressions,

$$\begin{aligned} \sqrt{n}(\tilde{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}}) &= \sqrt{n}(\tilde{\beta}_{\text{dglS}} - \beta_0) - \sqrt{n}(\hat{\beta}_{\text{dols}} - \beta_0) \\ &= O_p(1) + O_p(\sqrt{n}) \\ &= O_p(\sqrt{n}). \end{aligned}$$

Hence, $h_n = O_p(n)$ under the alternative.

We can extend the test to allow endogeneity under the alternative. Consider the following DGP:

$$\begin{aligned} y_t &= \beta' x_t + \gamma' v_t + \phi s_t + e_t, \\ e_t &= \rho e_{t-1} + u_t, \end{aligned}$$

where $\{s_t\}$ satisfies the same conditions as u_t and v_t but is correlated with $\{v_t\}$. The statistic defined in (17) can be applied to test the hypotheses:

$$\begin{aligned} H'_0: & |\rho| < 1 \quad \text{and} \quad \phi = 0 \\ \text{against } H'_A: & \rho = 1 \quad \text{and} \quad \phi \neq 0. \end{aligned}$$

The asymptotics of h_n under the null H'_0 are the same as that under H_0 . Under the alternative H'_A , we show in Appendix C that the DOLS estimator has the same asymptotic distribution as that under H_A and $h_n = O_p(n)$. Therefore, this Hausman-type test is consistent for the null hypothesis of cointegration against the alternative of spurious regressions, regardless of whether the exogeneity assumption holds under the alternative.

2.5.2. Finite sample properties of the Hausman-type cointegration test

Before applying the Hausman-type cointegration test empirically, it will be instructive to examine its finite sample properties in comparison with other comparable tests under the same null hypothesis. To this end, we

Table 2
Finite sample performance of the Hausman-type cointegration test

T	Hausman-type test		Shin's test	
	Power	Size (5%)	Power	Size (5%)
50	0.621	0.114	0.249	0.141
100	0.688	0.072	0.402	0.171
200	0.754	0.050	0.652	0.199
300	0.783	0.039	0.775	0.184
500	0.816	0.040	0.882	0.181

Note: The Hausman-type cointegration test is stipulated in Section 2.5. Nonparametric estimator of long run variance used is based on the QS kernel with the bandwidth of 'integer $\lceil 8(T/100)^{1/4} \rceil$ '.

conduct a small simulation experiment based on the following dynamic regression model:

$$y_t = \gamma_1 \Delta x_{t+1} + \beta x_t + \gamma_2 \Delta x_{t-1} + e_t, \quad (18)$$

$$e_t = \rho e_{t-1} + u_t, \quad (19)$$

where $\gamma_1 = 0.3$, $\beta = 2$, $\gamma_2 = -0.5$, and setting $\rho = 0.9$ for the size performance and $\rho = 1$ for the power performance. We consider sample sizes of $n \in \{50, 100, 200, 300, 500\}$ that are commonly encountered in empirical analysis. In the simulations, pseudo-random numbers are generated using the GAUSS (version 6.0) RNDNS procedures. Each simulation run is carried out with 5000 replications. At each replication, $100 + n$ random numbers are generated, of which the first 100 observations are discarded to avoid a start-up effect.

Table 2 reports selected finite sample properties of the Hausman-type cointegration test together with a residual-based test under the null of cointegration due to Shin (1994, Shin's test), who extended the KPSS test in the parametrically corrected cointegrating regression. In the simulations, the lengths of the lead and lag terms for DOLS and DGLS are chosen by the BIC rule.⁵ A nonparametric estimation method for long-run variance estimation is employed using the QS kernel with the bandwidth of 'integer $\lceil 8(n/100)^{1/4} \rceil$ '. The results in Table 2 illustrate two points. First, the empirical size of the Hausman-type test is close to the nominal size, in particular, when the sample size is relatively large, whereas Shin's test suffers from a serious oversize problem. Second, in terms of power, the Hausman-type test dominates Shin's test for moderate sample sizes that are very likely to be encountered empirically. Shin's test seems more powerful when n is relatively large, but only at the cost of severe size distortions. Overall, our simulation results provide evidence in favor of the Hausman-type test.

3. Empirical applications

In this section we apply the GLS-type correction methods and the Hausman-type cointegration test to analyze three macroeconomic issues: (i) long run money demand in the U.S., (ii) output convergence among industrial and developing countries, and (iii) PPP for traded and non-traded goods. The main purpose of the first application is to illustrate the spurious regression approach to estimating unknown structural parameters. Identification of the structural parameters in this application is based on nonstationary measurement error or nonstationary omitted variables that are explained in the following two sections. The main purpose of the other two applications is to apply the Hausman-type cointegration test. The alternative hypothesis is not taken as structural spurious regressions in the last two applications.

3.1. A model of nonstationary measurement error

Consider a set of variables that are cointegrated. One model of a structural spurious regression is based on the case in which one of the variables is measured with nonstationary measurement error. Let y_t^0 be the true

⁵It is an interesting research topic to investigate the performance of various lag length selection rules, but would be beyond the scope of this paper.

value of y_t , and assume that

$$y_t^0 = \beta'x_t + \gamma^0v_t + e_t^0 \quad (20)$$

is a dynamic cointegrating regression that satisfies the strict exogeneity assumption.⁶ Let y_t be the measured value of y_t^0 , and assume that the measurement error satisfies

$$y_t - y_t^0 = \gamma^mv_t + e_t^m, \quad (21)$$

where e_t^m is $I(1)$ and its expectation conditional on x_s for all s is zero. Here, the crucial assumption for identification is that the measurement error is not cointegrated with x_t . Then

$$y_t = \beta'x_t + \gamma'v_t + e_t, \quad (22)$$

where $\gamma = \gamma^0 + \gamma^m$, and e_t is $I(1)$ and satisfies the strict exogeneity assumption.⁷

3.2. A model of nonstationary omitted variables

Another case that leads to a structural spurious regression is a model of nonstationary omitted variables.

$$y_t = \beta'x_t + \theta'x_t^0 + \gamma^lv_t + \gamma^0v_t^0 + e_t^0, \quad (23)$$

where x_t^0 is a vector of $I(1)$ variables and v_t^0 is a vector of leads and lags of the first differences of x_t^0 . We imagine that the econometrician omits x_t^0 from his regression. We assume that

$$\theta'x_t^0 + \gamma^0v_t^0 = \gamma^mv_t + e_t^m, \quad (24)$$

where e_t^m is $I(1)$ and its expectation conditional on x_s for all s is zero. Here, the crucial assumption for identification is that the x_t^0 is not cointegrated with x_t . Then

$$y_t = \beta'x_t + \gamma'v_t + e_t, \quad (25)$$

where $\gamma = \gamma^l + \gamma^m$, and e_t is $I(1)$ and satisfies the strict exogeneity assumption. This model is observationally equivalent to the model of nonstationary measurement error within our single equation approach. However, the assumptions made in both cases are conceptually different.

3.3. U.S. money demand

The long-run money demand function has often been estimated under a cointegrating restriction among real balances, real income, and the interest rate. The restriction is legitimate if the money demand function is stable in the long run and if all variables are measured without nonstationary error. Indeed, [Stock and Watson \(1993\)](#) found supportive evidence of stable long-run M1 demand by estimating cointegrating vectors. However, if either money is measured with a nonstationary measurement error or nonstationary omitted variables exist, then we have a spurious regression, and the estimation results based on a cointegration regression are questionable.

First, consider the model of a nonstationary measurement error described above. To be specific, we follow [Stock and Watson \(1993\)](#) and assume that the dynamic regression error is stationary and the strict exogeneity assumption holds for the dynamic regression error when money is correctly measured. We then assume that money is measured with a multiplicative measurement error. We assume that the log measurement error is unit-root nonstationary and that the residuals of the projection of the log measurement error on the leads and lags of the regressors in the dynamic regression satisfy the strict exogeneity assumption. Given that a large component of the measurement error is arguably currency held by foreign residents and black market participants, the log measurement error is likely to be very persistent. Therefore, the assumption that the log measurement error is unit-root nonstationary may be at least a good approximation. The assumption that the

⁶Note that any variable can be chosen as the regressand in a cointegrating regression. Therefore, we choose the variable with nonstationary measurement error as the regressand.

⁷Here, we assume that the dimensions of γ^0 and γ^m are the same without loss of generality because we can add zeros as elements of γ^0 and γ^m as needed.

measurement error is not cointegrated with the regressors is plausible if the error is mainly due to currency held by foreign residents.

Second, consider the model of nonstationary omitted variables. A possible omitted variable is a measure of the “shoe leather cost” that represents transaction costs. For example, in the literature of money demand estimation, the real wage rate has sometimes been used as a regressor for this reason. Because the real wage rate is $I(1)$ in standard dynamic stochastic general equilibrium models with an $I(1)$ technological shock, the omitted measure of the “shoe leather cost” is nonstationary. If the real wage rate is the omitted variable, the assumption that it is not cointegrated with the regressors that include log income is not very plausible. However, it is possible that the true omitted variable that represents the “shoe leather cost” is not the real wage rate and is not cointegrated with log income.

We apply our GLS correction method to estimate the long-run income and interest elasticities of M1 demand during the period of 1947–1997.⁸ To this end, the regression equations are set up with the real money balance (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 M/P , the real net national product for y and the six-month commercial paper rate in percentages for i . M/P and y are in logarithms. 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):

$$\ln \left(\frac{M}{P} \right)_t = \alpha + \beta \ln(y_t) + \gamma i_t + e_t \quad (\text{equation 1}),$$

$$\ln \left(\frac{M}{P} \right)_t = \alpha + \beta \ln(y_t) + \gamma \ln(i_t) + e_t \quad (\text{equation 2}),$$

$$\ln \left(\frac{M}{P} \right)_t = \alpha + \beta \ln(y_t) + \gamma \ln \left[\frac{1+i_t}{i_t} \right] + e_t \quad (\text{equation 3}).$$

It is worth noting that the liquidity trap is possible for the latter two functional forms as emphasized by Bae and de Jong (2007).⁹ When the data contain periods with very low nominal interest rates, the latter two functional forms may be more appropriate.

Table 3 presents the point estimates for β (income elasticity of money demand) and γ based on the three estimators under scrutiny: the dynamic OLS estimator, the GLS corrected dynamic regression estimator, and the FGLS corrected dynamic regression estimator.¹⁰ Several features emerge from the table. First, all the estimated coefficients have theoretically ‘correct’ signs: positive signs for income elasticities and negative signs for γ for the first two functional forms and positive signs for γ for the third functional form. Second, the GLS corrected estimates of the income elasticity are implausibly low for all three functional forms for low values of k and increase to more plausible values near one as k increases.¹¹ The fact that the results become more plausible as k increases suggests that the endogeneity correction of dynamic regressions works in this application for moderately large values of k such as three and four. The results for lower values of k are consistent with those of the low income elasticity estimates of first differenced regressions that were used in the literature before 1980. Therefore, the estimators in the old literature of first differenced regressions before cointegration became popular are likely to be downward biased because of the endogeneity problem. Third, all point estimates of the three estimators are very similar, and the Hausman-type test fails to reject the null hypothesis of cointegration for large enough values of k . Hence, there is little evidence against cointegration. However, it should be noted that a small random walk component is very hard to detect

⁸Readers are referred to Appendix D for the empirical guidelines on the use of estimation and testing techniques developed in this paper. We thank Youngsoo Bae for providing the data used in Bae and de Jong (2007) to us. This data set extends Stock and Watson’s data up to 1997 when the six-month commercial rate was discontinued.

⁹Bae and de Jong point out that nonlinear cointegration methods are needed if we are to evaluate these different functional forms with a common set of assumptions. It is beyond the scope of the present paper to develop spurious regression methods for nonlinear cointegration models.

¹⁰For the FGLS corrected dynamic regression estimator, the serial correlation coefficient of the error term is estimated before being applied to the Cochrane–Orcutt transformation. This coefficient is assumed to be unity in the GLS corrected dynamic regression estimator which is equivalent to regressing the first difference of variables without a constant term.

¹¹When k is increased beyond five (the maximum k in the table), point estimates for income elasticity estimates stabilize around unity.

Table 3
Application to long run U.S. money demand

Estimator	k	Equation 1		Equation 2		Equation 3	
		$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$
DOLS	0	0.891 (0.080)	−0.084 (0.025)	0.860 (0.061)	−0.307 (0.069)	0.858 (0.061)	0.316 (0.071)
	1	0.907 (0.079)	−0.092 (0.025)	0.861 (0.047)	−0.313 (0.053)	0.859 (0.047)	0.322 (0.054)
	2	0.924 (0.074)	−0.098 (0.024)	0.863 (0.040)	−0.321 (0.044)	0.861 (0.040)	0.329 (0.046)
	3	0.938 (0.059)	−0.102 (0.019)	0.865 (0.031)	−0.327 (0.034)	0.862 (0.031)	0.336 (0.036)
	4	0.945 (0.062)	−0.104 (0.020)	0.861 (0.029)	−0.330 (0.033)	0.858 (0.030)	0.339 (0.034)
	5	0.954 (0.056)	−0.108 (0.018)	0.863 (0.029)	−0.339 (0.033)	0.860 (0.029)	0.348 (0.034)
	BIC [lag]	[4]			[5]		[5]
GLS-corrected	0	0.405 (0.080)	−0.014 (0.004)	0.415 (0.079)	−0.086 (0.021)	0.415 (0.079)	0.089 (0.022)
	1	0.639 (0.113)	−0.029 (0.009)	0.664 (0.109)	−0.191 (0.042)	0.664 (0.109)	0.199 (0.044)
	2	0.799 (0.127)	−0.047 (0.012)	0.815 (0.123)	−0.241 (0.049)	0.815 (0.123)	0.249 (0.051)
	3	0.811 (0.137)	−0.058 (0.015)	0.850 (0.131)	−0.270 (0.057)	0.849 (0.131)	0.278 (0.058)
	4	0.877 (0.149)	−0.066 (0.018)	0.870 (0.142)	−0.262 (0.062)	0.868 (0.142)	0.267 (0.064)
	5	0.932 (0.160)	−0.070 (0.019)	0.908 (0.149)	−0.284 (0.065)	0.904 (0.148)	0.291 (0.067)
	BIC [lag]	[2]			[2]		[2]
FGLS-corrected AR(1)	0	0.887 (0.050)	−0.078 (0.023)	0.862 (0.044)	−0.288 (0.075)	0.861 (0.158)	0.296 (0.116)
	1	0.757 (0.054)	−0.041 (0.009)	0.826 (0.036)	−0.264 (0.032)	0.826 (0.035)	0.273 (0.032)
	2	0.834 (0.054)	−0.063 (0.011)	0.862 (0.031)	−0.303 (0.030)	0.860 (0.031)	0.311 (0.031)
	3	0.884 (0.047)	−0.082 (0.012)	0.874 (0.026)	−0.325 (0.026)	0.872 (0.026)	0.334 (0.027)
	4	0.910 (0.053)	−0.089 (0.013)	0.874 (0.027)	−0.327 (0.027)	0.871 (0.027)	0.335 (0.028)
	5	0.944 (0.046)	−0.100 (0.012)	0.879 (0.027)	−0.340 (0.027)	0.876 (0.027)	0.349 (0.028)
	BIC [lag]	[2]			[2]		[2]
FGLS-corrected AR(2)	0	0.887 (0.050)	−0.078 (0.023)	0.862 (0.044)	−0.288 (0.075)	0.861 (0.158)	0.296 (0.116)
	1	0.774 (0.052)	−0.043 (0.009)	0.825 (0.038)	−0.261 (0.032)	0.825 (0.037)	0.270 (0.033)
	2	0.847 (0.048)	−0.068 (0.011)	0.861 (0.028)	−0.307 (0.028)	0.859 (0.028)	0.316 (0.029)
	3	0.909 (0.041)	−0.090 (0.011)	0.877 (0.024)	−0.328 (0.025)	0.875 (0.025)	0.336 (0.026)
	4	0.937 (0.043)	−0.098 (0.012)	0.877 (0.025)	−0.330 (0.025)	0.874 (0.025)	0.338 (0.026)
	5	0.960 (0.039)	−0.105 (0.011)	0.878 (0.021)	−0.343 (0.022)	0.874 (0.022)	0.352 (0.023)
	BIC [lag]	[2]			[5]		[5]
HAUSMAN test	0	3.510		2.744		2.677	
	1	0.655		0.100		0.092	
	2	0.656		0.010		0.007	
	3	0.166		0.004		0.005	
	4	0.015		0.099		0.098	
	5	0.468		0.115		0.115	
ADF-BASED test	2	−3.288 [‡]		−3.891 [‡]		−3.850 [‡]	

Note:

$$\ln\left(\frac{M}{P}\right)_t = \alpha + \beta \ln(y_t) + \gamma i_t + e_t \quad (\text{equation 1}),$$

$$\ln\left(\frac{M}{P}\right)_t = \alpha + \beta \ln(y_t) + \gamma \ln(i_t) + e_t \quad (\text{equation 2}),$$

$$\ln\left(\frac{M}{P}\right)_t = \alpha + \beta \ln(y_t) + \gamma \ln\left[\frac{1+i_t}{i_t}\right] + e_t \quad (\text{equation 3}).$$

‘GLS-corrected (FGLS-corrected)’ denotes the GLS (FGLS) corrected dynamic regression estimator. Figures in the parenthesis represent standard errors. ‘k’ denotes the maximum length of leads and lags. In FGLS corrected estimation, the serial correlation coefficient in the error term is estimated before being applied to the Cochrane–Orcutt transformation, whereas it is assumed to be unity in GLS corrected estimation which is analogous to regressing the first difference of variables without a constant term. Hausman test represents the Hausman-type cointegration test as stipulated in Section 2.5. The test statistic is constructed as $(\hat{\Gamma}_{dglS} - \tilde{\Gamma}_{dols})\Sigma(\hat{\Gamma}_{dglS} - \tilde{\Gamma}_{dols})' \rightarrow d\chi^2(2)$

where $\Gamma = [\beta, \gamma]$ and $\Sigma = \begin{bmatrix} \text{var}(\hat{\beta}_{dglS}) & \text{cov}(\hat{\beta}_{dglS}, \hat{\gamma}_{dglS}) \\ \text{cov}(\hat{\beta}_{dglS}, \hat{\gamma}_{dglS}) & \text{var}(\hat{\gamma}_{dglS}) \end{bmatrix}$. The critical values of $\chi^2(2)$ are 4.61, 5.99 and 9.21 for 10%, 5%, and 1% significance levels. The critical values of the ADF-based tests are −2.88 and −2.57 for 5% and 10% significance levels.

[‡]represents that the null hypothesis can be rejected at 5%.

with any test for cointegration. Therefore, it is assuring to know that all three estimators are similar for large enough values of k and that the estimates are robust with respect to whether the regression error is $I(0)$ or $I(1)$.

We report the value of k chosen by the Bayesian information criterion (BIC) rule throughout our empirical applications in order to give some guidance in interpreting results.¹² A detailed analysis of how k should be chosen is beyond the scope of this paper because this issue has not been settled in the literature of dynamic cointegrating regressions.

Table 3 also reports the results when the ADF test is applied to the OLS residuals. The results show evidence against the null hypothesis of structural spurious regressions and thereby corroborate the results from the Hausman-type test under the opposite null hypothesis.

3.4. Output convergence across national economies

In this section, we apply the techniques to re-examine a long standing issue in macroeconomics, the hypothesis of output convergence. For this application and the next, our main purpose is not to estimate unknown structural parameters but to test the null hypothesis of cointegration with the Hausman-type test. For this purpose, we do not need the strict exogeneity assumption under the alternative hypothesis of no cointegration (or a spurious regression).

As a key proposition of the neoclassical growth model, the convergence hypothesis has been popular in macroeconomics and has attracted considerable attention in the empirical field, particularly during the last decade. Besides its important policy implications, the convergence hypothesis has been used as a criterion to discern between the two main growth theories, exogenous growth theory and endogenous growth theory. Despite this attention, it remains the subject of continuing debate mainly because the empirical evidence supporting the hypothesis is mixed. Nevertheless, the established literature based on popular international data sets such as the Summers–Heston (Summers and Heston, 2006) data set suggests as a stylized fact output convergence among industrialized countries but not among developing countries and not between industrialized and developing countries.

Given that a mean stationary stochastic process of output disparities between two economies is interpreted as supportive evidence of stochastic convergence, unit-root or cointegration testing procedures are often used by empirical researchers to evaluate the convergence hypothesis. In this vein, our techniques proposed here fit in the study of output convergence. We consider four developing countries (Columbia, Ecuador, Egypt, and Pakistan) along with four industrial countries (Denmark, New Zealand, South Africa, Switzerland). The raw data are extracted from the *Penn World Tables* of Summers and Heston (2006) and consist of annual real GDP per capita (RGDPCH) over the period of 1951–2003. The following two regression equations are considered with regard to the cointegration relation:

$$y_t^D = \alpha + \beta y_t^I + e_t, \quad (26)$$

$$y_t^I = \alpha + \beta y_t^I + e_t, \quad (27)$$

where y_t^D and y_t^I denote log real GDP per capita for developing and industrial countries, respectively.

Table 4 presents the results which exhibit a large variation in estimated coefficients. Recall that our interest in this application lies in the cointegration test based on the Hausman-type test. As can be seen from Table 4, irrespective of country combinations, the null hypothesis of cointegration can be rejected when developing countries are regressed onto industrial countries, indicating that there is little evidence of output convergence between developing countries and industrial countries. The picture changes dramatically when industrial countries are regressed onto industrial countries as in (27). Table 4 also reports that the Hausman-type test fails to reject the null of cointegration in all cases considered. Our finding is therefore consistent with the notion of convergence clubs which is taken as a stylized fact in the growth literature (e.g. Durlauf and Quah, 1999; Easterly, 2001).

¹²In fact, we are not sure whether the BIC is the right method. Given the sensitivity of estimation results, we report the results from various k 's in the money demand application. In the following two applications for the PPP and output convergence, we just report the results based on BIC mainly because the results are not very sensitive to k around the BIC choice.

Table 4
Application to output convergence

	Regressand	Regressor	k	DOLS	GLS-corrected	FGLS-corrected	Hausman test
Regression 1	COL	DEN	[1]	0.815 (0.168)	0.569 (0.088)	0.767 (0.127)	5.277 [‡]
		NZL	[0]	1.267 (0.337)	0.212 (0.086)	1.254 (0.049)	43.111 [‡]
		SWI	[0]	1.032 (0.225)	0.394 (0.077)	1.035 (0.026)	17.665 [‡]
		ZAF	[0]	1.336 (0.383)	0.368 (0.125)	1.337 (0.051)	30.998 [‡]
	ECU	DEN	[0]	0.789 (0.341)	0.394 (0.164)	0.819 (0.003)	3.608 [‡]
		NZL	[0]	1.238 (0.711)	0.262 (0.159)	1.240 (0.052)	3.082 [†]
		SWI	[0]	1.027 (0.430)	0.358 (0.162)	1.050 (0.030)	7.086 [‡]
		ZAF	[0]	1.381 (0.461)	0.321 (0.237)	1.416 (0.064)	8.203 [‡]
	EGT	DEN	[0]	1.144 (0.215)	0.664 (0.150)	1.120 (0.042)	5.148 [‡]
		NZL	[4]	2.277 (0.354)	1.351 (0.347)	2.504 (0.234)	9.570 [‡]
		SWI	[2]	1.554 (0.725)	0.943 (0.271)	1.038 (0.642)	5.867 [‡]
		ZAF	[0]	1.790 (0.685)	0.505 (0.236)	1.753 (0.115)	10.107 [‡]
	PAK	DEN	[0]	1.138 (0.181)	0.397 (0.132)	1.130 (0.049)	5.904 [‡]
		NZL	[0]	1.744 (0.409)	0.012 (0.135)	1.702 (0.115)	41.209 [‡]
		SWI	[2]	1.532 (0.304)	0.890 (0.207)	2.062 (0.235)	4.048 [‡]
		ZAF	[0]	1.813 (0.483)	0.578 (0.182)	1.802 (0.127)	15.634 [‡]
Regression 2	DEN	NZL	[4]	1.665 (0.042)	1.569 (0.201)	1.656 (0.041)	0.149
		SWI	[4]	1.303 (0.324)	1.353 (0.140)	1.285 (0.253)	0.011
		ZAF	[2]	1.683 (0.283)	1.235 (0.238)	1.599 (0.227)	2.104
	NZL	DEN	[0]	0.629 (0.033)	0.496 (0.133)	0.635 (0.028)	0.396
		SWI	[0]	0.785 (0.088)	0.455 (0.133)	0.794 (0.019)	0.813
		ZAF	[2]	1.051 (0.158)	0.880 (0.244)	1.032 (0.146)	0.122
	SWI	DEN	[1]	0.749 (0.048)	0.774 (0.137)	0.670 (0.045)	0.048
		NZL	[2]	1.215 (0.098)	0.813 (0.238)	1.111 (0.093)	1.779
		ZAF	[0]	1.273 (0.130)	0.690 (0.175)	1.289 (0.018)	0.727
	ZAF	DEN	[1]	0.543 (0.075)	0.524 (0.107)	0.482 (0.068)	0.078
		NZL	[2]	0.870 (0.129)	0.667 (0.143)	0.706 (0.119)	0.822
		SWI	[1]	0.735 (0.106)	0.606 (0.114)	0.647 (0.114)	0.289

Note: See the notes in Table 2. Annual data covering 1951–2003 are used for four developing countries (COL: Columbia; ECU: Ecuador; EGT: Egypt; PAK: Pakistan) and four industrial countries (DEN: Denmark; NZL: New Zealand; SWI: Switzerland; ZAF: South Africa). k denotes the length of lead and lag terms for DOLS and DGLS chosen by the BIC rule.

$$\text{regression 1: } \ln(y_{DEV}) = \alpha + \beta \ln(y_{IND}) + e,$$

$$\text{regression 2: } \ln(y_{IND}) = \alpha + \beta \ln(y_{IND}) + e.$$

3.5. PPP for traded and non-traded goods

As a major building block for many models of exchange rate determination, PPP has been one of the most heavily studied subjects in international macroeconomics. Despite extensive research, the empirical evidence on PPP remains inconclusive, largely due to the econometric challenges involved in determining its validity. As is generally agreed, most real exchange rates show very slow convergence which makes estimating long-run relationships difficult with existing statistical tools. The literature suggests a number of potential explanations for the very slow adjustment of relative prices: volatility of the nominal exchange-rate, market frictions such as trade barriers and transportation costs, imperfect competition in product markets, and the presence of non-traded goods in the price basket. According to the commodity-arbitrage view of PPP, the law of one price holds only for traded goods, and the departures from PPP are primarily attributed to the large weight placed on non-traded goods in the CPI. This view has obtained support from many empirical studies based on

Table 5
Application to PPP for traded and non-traded goods

Estimator	Traded goods					Non-traded goods					
	FRA	ITA	JPN	U.K.	U.S.	FRA	ITA	JPN	U.K.	U.S.	
DOLS	1.149 (0.312)	1.379 (0.165)	1.558 (0.326)	1.306 (0.201)	1.053 (0.198)	1.872 (0.165)	2.142 (0.241)	2.357 (0.299)	2.059 (0.278)	1.711 (0.443)	
BIC	[0]	[0]	[0]	[2]	[0]	[0]	[0]	[5]	[1]	[0]	
GLS-corrected	0.833 (0.393)	1.114 (0.381)	1.086 (0.411)	1.030 (0.365)	0.919 (0.140)	0.375 (0.178)	0.448 (0.176)	0.372 (0.198)	0.351 (0.171)	0.159 (0.080)	
BIC	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[1]	[0]	
FGLS-corrected	1.229 (0.329)	1.456 (0.149)	1.717 (0.396)	1.248 (0.162)	0.766 (0.242)	1.932 (0.152)	2.005 (0.201)	2.214 (0.313)	1.671 (0.226)	0.392 (0.216)	
BIC	[1]	[1]	[2]	[1]	[1]	[2]	[1]	[5]	[1]	[3]	
Hausman test	[1]	0.022	0.209	0.002	0.002	0.006	5.097 [‡]	10.603 [‡]	9.597 [‡]	8.118 [‡]	17.424 [‡]

Note: Results are for $f_t^T = \alpha + \beta p_t^T + e_t$ and $f_t^N = \alpha + \beta p_t^N + e_t$ using Canada as a base country. Figures in parenthesis represent standard errors. Entries inside square brackets represent the length of leads and lags chosen by BIC. Hausman test represents the Hausman-type cointegration test as stipulated in Section 2.5. The test statistic is constructed as $(\hat{\beta}_{dglS} - \tilde{\beta}_{dols})^2 / \text{Var}(\tilde{\beta}_{dglS}) \rightarrow d\chi^2(1)$. The critical values of $\chi^2(1)$ are 2.71, 3.84 and 6.63 for ten, five, and one percent significance level.

[‡]represents that the null hypothesis of $\hat{\beta}_{dglS} = \tilde{\beta}_{dols}$ can be rejected at 5% significance level.

disaggregated price indices. They tend to provide ample evidence that prices for non-traded goods are much more dispersed than for their traded counterparts and consequently non-traded goods exhibit far larger deviations from PPP than traded goods. Given that general price indices involve a mix of both traded and non-traded goods, highly persistent deviations of non-traded goods from PPP can lead to the lack of conclusive evidence on the long run PPP relationship. As in the previous application, our main purpose for this application is not to estimate unknown structural parameters but to test the null hypothesis of cointegration with the Hausman-type test.

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 the 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 $e_t \sim I(0)$ in

$$\begin{aligned} f_t^T &= \alpha + \beta p_t^T + e_t, \\ f_t^N &= \alpha + \beta p_t^N + e_t, \end{aligned}$$

where the superscripts T and N denote the price levels of traded goods and non-traded goods, respectively.

Following the method of Stockman and Tesar (1995), Kim (2005) recently analyzed the real exchange rate for total consumption using the general price deflator and the real exchange rate for traded and non-traded goods using implicit deflators for non-service consumption and service consumption, respectively.¹³ We use Kim's data set to apply our techniques to the linear combination of sectorally decomposed variables. Table 5 presents the results using quarterly price and exchange rate data for six countries: Canada, France, Italy, Japan, U.K., and U.S. for the period of 1974 Q1 through 1998 Q4. With the Canadian dollar used as numeraire, Table 5 presents the estimates for β which should be close to unity according to long-run PPP. For traded goods, estimates are above unity in most cases, but the variation across estimates does not seem substantial, resulting in non-rejection of the null of cointegration in all cases considered. By sharp contrast, the Hausman-type cointegration test rejects the null hypothesis in every country when the price for non-traded goods is used. It is noteworthy that there exists a considerable difference between the GLS-corrected estimates for β and their DOLS and FGLS counterparts which are far greater than unity. That is, supportive evidence of

¹³For details, see the Appendix for the description of the data. We thank J.B. Kim for sharing the data set.

PPP is found for traded goods but not for non-traded goods, congruent with the general intuition as well as the findings by other studies in the literature such as Kakkar and Ogaki (1999) and Kim (2005).¹⁴

4. Concluding remarks and future work

In this paper, we analyzed an approach to correcting spurious regressions involving unit-root nonstationary variables by generalized least squares using asymptotic theory. This analysis leads to a new robust estimator and a new test for dynamic regressions.

We considered two estimators to estimate structural parameters in spurious regressions: the GLS corrected dynamic regression estimator suggested by Choi and Ogaki (2001) and the FGLS corrected dynamic regression estimator. A GLS corrected dynamic regression estimator is a first differenced version of a dynamic OLS regression estimator. Asymptotic theory shows that, under some regularity conditions, the endogeneity correction of the dynamic regression works for the first differenced regressions for both cointegrating and spurious regressions. This result is useful because it is not intuitively clear that the endogeneity correction works in regressions with stationary first differenced variables even though it has been used for cointegrating regressions.

For the purpose of the estimating structural parameters when the possibility of nonstationary measurement error or nonstationary omitted variables cannot be ruled out, we recommend the FGLS corrected dynamic regression estimators because they are robust. They are consistent both when the error is $I(0)$ and $I(1)$. They are asymptotically as efficient as dynamic OLS when the error is $I(0)$ and as efficient as GLS corrected dynamic regression when the error is $I(1)$. This feature may be especially attractive when the FGLS corrected dynamic estimator is extended to a panel data setting. Past works on nonstationary time series panels assume that all regressions in a panel are either cointegrating or spurious. However, when the number of cross-sectional observations increases, it is very likely that we may observe an $I(0)/I(1)$ mixed panel, i.e., the regression errors are $I(0)$ in some regressions and $I(1)$ in others. One example is that we may reject PPP in some countries while not rejecting it in others. To estimate the structural parameter in an $I(0)/I(1)$ mixed panel, we can first run an FGLS correction of each individual equation, so that the pooled panel estimator always takes the fastest convergence rate. Hu (2005) studies this extension.

We also developed a Hausman-type cointegration test by comparing the dynamic OLS regression estimator and the GLS corrected dynamic regression estimator. As noted in the introduction, this task is important not merely because few tests for cointegration have been developed for dynamic OLS, but also because tests for the null hypothesis of cointegration are useful in many applications. For this test, the spurious regression obtained under the alternative hypothesis does not have to be structural.

We demonstrated our estimation and testing methods in three applications: (i) long-run money demand in the U.S., (ii) output convergence among industrial and developing countries, and (iii) PPP for traded and non-traded goods.

In the first application of estimating the money demand function, the results suggest that the endogeneity correction of the dynamic regression works with a moderately large number of leads and lags for the GLS corrected dynamic regression estimator. The GLS corrected dynamic regression estimates of the income elasticity of money demand are very low with low orders of leads and lags, and then increase to more plausible values as the order of leads and lags increases. Dynamic OLS estimates are close to the GLS corrected dynamic regression estimates for a large enough order of leads and lags, and we find little evidence against cointegration with the Hausman-type cointegration test. The FGLS corrected dynamic regression estimates are very close to the GLS corrected dynamic regression estimates and the dynamic OLS estimates for sufficiently large orders of leads and lags. Hence, in the first application, the FGLS corrected dynamic regression estimator works well in the sense that it yields estimates that are close to those of the estimator that seems to be correctly specified. This is confirmed by our simulation results in Section 2, indicating that the

¹⁴Engel (1999) finds little evidence for long-run PPP for traded goods with his variance decomposition method. However, it should be noted that his method is designed to study variations of real exchange rates over relatively shorter periods. Park and Ogaki (2007) show that this variance decomposition has an unexpected property when the variance of long-run differences is used and provides little information about long-run variations of real exchange rates.

small sample efficiency loss from using the FGLS corrected dynamic regression estimator is negligible for reasonable sample sizes. Therefore, we recommend the robust FGLS corrected dynamic regression estimator when the researcher is unsure about whether or not the regression error is $I(0)$ or $I(1)$. This is important because it is difficult to detect a small random walk component in the error term when the error is actually $I(1)$, and it is difficult to detect a small deviation from a unit-root when the error is actually $I(0)$ but the dominant autoregressive root is very close to unity.

In the second application, we applied the Hausman-type cointegration test to the log real output of pairs of countries to study output convergence across national economies. Our test results are consistent with the stylized fact of convergence clubs in that we reject the null hypothesis of cointegration between developing and developed countries while failing to reject the null hypothesis of cointegration between two developed countries. Finally, we apply the Hausman-type cointegration test to study long-run PPP. Our test results support the commodity-arbitrage view that long-run PPP holds for traded goods but not for non-traded goods.

In future work, it will be important to study the choice of k , the number of leads and lags in the endogeneity correction. Another aspect that will be useful in empirical work is the study of possible deterministic time trends and seasonal effects in the model.

Acknowledgments

We thank Don Andrews, Youngsoo Bae, Yanqin Fan, Clive Granger, Lars Hansen, Kotaro Hitomi, Hide Ichimura, Hyeok Jeong, Lutz Kilian, Yuichi Kitamura, Yoshihiko Nishiyama, Sungwook Park, Peter Phillips, Ping Wang, Arnold Zellner, and seminar participants at the Bank of Japan, Kyoto University, National Chung Cheng University, New York University, Ohio State University, University of Southern California, University of Tokyo, Vanderbilt University, Yale University, the 2004 Far Eastern Meeting of the Econometric Society, and the 2004 Annual Meeting of the Midwest Econometrics Group for helpful comments. Any remaining errors are the authors'.

Appendix A. Proof of results in Section 2.1

To show the distribution of the OLS estimator in regression (4) with $\rho = 1$, define $\mathbf{X} = [x_1, \dots, x_n]'$, $\mathbf{V} = [v_1, \dots, v_n]'$, and $\mathbf{e} = [e_1, \dots, e_n]'$. Also define $M_v = I_n - \mathbf{V}(\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'$ and $M_x = I_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Then the OLS estimator for β and γ can be written as

$$\begin{aligned} \begin{bmatrix} \hat{\beta}_n - \beta_0 \\ \hat{\gamma}_n - \gamma_0 \end{bmatrix} &= \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{V} \\ \mathbf{V}'\mathbf{X} & \mathbf{V}'\mathbf{V} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{e} \\ \mathbf{V}'\mathbf{e} \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{X}'M_v\mathbf{X})^{-1} & -(\mathbf{X}'M_v\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}(\mathbf{V}'\mathbf{V})^{-1} \\ -(\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'\mathbf{X}(\mathbf{X}'M_v\mathbf{X})^{-1} & (\mathbf{V}'M_x\mathbf{V})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{X}'\mathbf{e} \\ \mathbf{V}'\mathbf{e} \end{bmatrix}. \end{aligned}$$

We are mostly interested in the structural parameter $\hat{\beta}$. Write its limit distribution as

$$\begin{aligned} \hat{\beta}_n - \beta_0 &= (\mathbf{X}'M_v\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} - (\mathbf{X}'M_v\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}(\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'\mathbf{e} \\ &= \left(\frac{\mathbf{X}'M_v\mathbf{X}}{n^2}\right)^{-1} \left(\frac{\mathbf{X}'\mathbf{e}}{n^2}\right) - \frac{1}{n} \left(\frac{\mathbf{X}'M_v\mathbf{X}}{n^2}\right)^{-1} \left(\frac{\mathbf{X}'\mathbf{V}}{n}\right) \left(\frac{\mathbf{V}'\mathbf{V}}{n}\right)^{-1} \left(\frac{\mathbf{V}'\mathbf{e}}{n}\right) \\ &= \left(\frac{\mathbf{X}'M_v\mathbf{X}}{n^2}\right)^{-1} \left(\frac{\mathbf{X}'\mathbf{e}}{n^2}\right) + o_p(1) \\ &\rightarrow_d \left[\int_0^1 V(r)V(r)' dr \right]^{-1} \left[\int_0^1 V(r)U(r) dr \right] \equiv h_1, \end{aligned}$$

which gives Eq. (5).

To show the limit distribution of the GLS corrected estimator in regression (6), write

$$\sqrt{n}(\tilde{\theta}_{\text{dglS}} - \theta_0) = \left[n^{-1} \sum_{t=1}^n \Delta z_t \Delta z_t' \right]^{-1} \left[n^{-1/2} \sum_{t=1}^n r \Delta z_t u_t \right]. \tag{28}$$

For the denominator,

$$n^{-1} \sum_{t=1}^n \Delta z_t \Delta z_t' = \begin{bmatrix} n^{-1} \sum_{t=1}^n v_t v_t' & n^{-1} \sum_{t=1}^n v_t \Delta \mathbf{v}_t' \\ n^{-1} \sum_{t=1}^n \Delta \mathbf{v}_t v_t' & n^{-1} \sum_{t=1}^n \Delta \mathbf{v}_t \Delta \mathbf{v}_t' \end{bmatrix} \rightarrow_d \begin{bmatrix} \Sigma_v & \Gamma'_{v, \Delta v} \\ \Gamma_{v, \Delta v} & \Gamma_{\Delta v, \Delta v} \end{bmatrix} \equiv Q, \tag{29}$$

where Σ_v is the variance matrix of $\{v_t\}$ and Γ is a matrix with elements computed from the autocovariances of $\{v_t\}$.

For the numerator, the assumptions on the innovation processes ensure that the CLT holds:

$$n^{-1/2} \sum_{t=1}^n \Delta z_t u_t = \begin{bmatrix} n^{-1/2} \sum_{t=1}^n v_t u_t \\ n^{-1/2} \sum_{t=1}^n \Delta \mathbf{v}_t u_t \end{bmatrix} \rightarrow_d N(0, A), \tag{30}$$

where A is the long-run covariance matrix of the vector $\Delta z_t u_t$:

$$A = \begin{bmatrix} \sum_{j=-\infty}^{\infty} h^{\infty} E(v_t v_{t-j}' u_t u_{t-j}) & \sum_{j=-\infty}^{\infty} E(v_t \Delta \mathbf{v}_{t-j}' u_t u_{t-j}) \\ \sum_{j=-\infty}^{\infty} E(\Delta \mathbf{v}_t v_{t-j}' u_t u_{t-j}) & \sum_{j=-\infty}^{\infty} E(\Delta \mathbf{v}_t \Delta \mathbf{v}_{t-j}' u_t u_{t-j}) \end{bmatrix}.$$

Hence, for the quantity defined in (28), we have the limit distribution given in (7):

$$\sqrt{n}(\tilde{\theta}_{\text{dglS}} - \theta_0) \rightarrow_d N(\mathbf{0}, \Omega),$$

where $\Omega = Q^{-1} A Q^{-1}$.

To derive the limit distribution for the FGLS estimator, we can first show that

$$n(\hat{\rho}_n - 1) = O_p(1).$$

The proof is similar as that in Phillips and Hodgson (1994). Then for the sequence of \tilde{y}_t , we can write it as

$$\tilde{y}_t = \theta'_0 \tilde{z}_t + u_t + (1 - \hat{\rho}_n) e_{t-1}.$$

Now, we can write

$$\hat{\theta}_{\text{fglS}} - \theta_0 = \left[\sum_{t=1}^n \tilde{z}_t \tilde{z}_t' \right]^{-1} \left[\sum_{t=1}^n \tilde{z}_t [u_t + (1 - \hat{\rho}_n) e_{t-1}] \right]. \tag{31}$$

Write the denominator as

$$\sum_{t=1}^n \tilde{z}_t \tilde{z}_t' = \begin{bmatrix} \sum_{t=1}^n \tilde{x}_t \tilde{x}_t' & \sum_{t=1}^n \tilde{x}_t \tilde{\mathbf{v}}_t' \\ \sum_{t=1}^n \tilde{\mathbf{v}}_t \tilde{x}_t' & \sum_{t=1}^n \tilde{\mathbf{v}}_t \tilde{\mathbf{v}}_t' \end{bmatrix}.$$

First,

$$\begin{aligned} \sum_{t=1}^n \tilde{x}_t \tilde{x}'_t &= \sum_{t=1}^n (x_t - \hat{\rho}_n x_{t-1})(x_t - \hat{\rho}_n x_{t-1})' \\ &= (1 - \hat{\rho}_n)^2 \sum_{t=1}^n x_{t-1} x'_{t-1} + (1 - \hat{\rho}_n) \sum_{t=1}^n [x_{t-1} v'_t + v_t x'_{t-1}] + \sum_{t=1}^n v_t v'_t. \end{aligned}$$

Hence,

$$\begin{aligned} n^{-1} \sum_{t=1}^n \tilde{x}_t \tilde{x}'_t &= n(1 - \hat{\rho}_n)^2 \left(n^{-2} \sum_{t=1}^n x_{t-1} x'_{t-1} \right) \\ &\quad + (1 - \hat{\rho}_n) \left(n^{-1} \sum_{t=1}^n x_{t-1} v'_t + n^{-1} \sum_{t=1}^n v_t x'_{t-1} \right) + n^{-1} \sum_{t=1}^n v_t v'_t \\ &= n^{-1} \sum_{t=1}^n v_t v'_t + o_p(1) \rightarrow_p \Sigma_v. \end{aligned}$$

Similarly, we can show that

$$\begin{aligned} n^{-1} \sum_{t=1}^n \tilde{x}_t \tilde{v}'_t &= n^{-1} \sum_{t=1}^n v_t (v_t - v_{t-1})' + o_p(1) \rightarrow_p \Gamma_{v, \Delta v}, \\ n^{-1} \sum_{t=1}^n \tilde{v}_t \tilde{v}'_t &= \frac{1}{n} \sum_{t=1}^n \Delta v_t \Delta v'_t + o_p(1) \rightarrow_p \Gamma_{\Delta v, \Delta v}. \end{aligned}$$

Hence,

$$n^{-1} \sum_{t=1}^n \tilde{z}_t \tilde{z}'_t = n^{-1} \sum_{t=1}^n \Delta z_t \Delta z'_t + o_p(1) \rightarrow_p Q. \tag{32}$$

Next, consider the numerator in (31)

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

It is not hard to see that $n^{-1} \sum_{t=1}^n \tilde{z}_t [u_t + (1 - \hat{\rho}_n) e_{t-1}] \rightarrow_p 0$. Intuitively, \tilde{z}_t behaves asymptotically like the differenced regressors $(v'_t, \Delta v'_t)'$, and u and v are uncorrelated by assumption. Our remaining task is to show that

$$n^{-1/2} \sum_{t=1}^n \tilde{z}_t [u_t + (1 - \hat{\rho}_n) e_{t-1}] = n^{-1/2} \sum_{t=1}^n \Delta z_t u_t + o_p(1) \rightarrow_d N(\mathbf{0}, A). \tag{33}$$

This can be shown using arguments similar to those used in proving (32). Combining (33) with (32), we obtain the limit distribution for $\hat{\theta}_{d\text{gls}}$ as given in (11). \square

Appendix B. Proof of results in Section 2.2

To show the limit distribution of the dynamic OLS estimator in the cointegration, define

$$H_n = \begin{bmatrix} I_m n & 0 \\ 0 & I_{m(2k+1)} n^{1/2} \end{bmatrix}. \tag{34}$$

We can write

$$H_n(\hat{\theta}_{\text{dols}} - \theta_0) = \begin{bmatrix} n(\hat{\beta}_n - \beta_0) \\ n^{1/2}(\hat{\gamma}_n - \gamma_0) \end{bmatrix} = \begin{bmatrix} n^{-2} \sum_{t=1}^n x_t x_t' & n^{-3/2} \sum_{t=1}^n x_t \mathbf{v}_t' \\ n^{-3/2} \sum_{t=1}^n \mathbf{v}_t x_t' & n^{-1} \sum_{t=1}^n \mathbf{v}_t \mathbf{v}_t' \end{bmatrix}^{-1} \begin{bmatrix} n^{-1} \sum_{t=1}^n x_t e_t \\ n^{-1/2} \sum_{t=1}^n \mathbf{v}_t e_t \end{bmatrix}.$$

For the denominator,

$$\begin{bmatrix} n^{-2} \sum_{t=1}^n x_t x_t' & n^{-3/2} \sum_{t=1}^n x_t \mathbf{v}_t' \\ n^{-3/2} \sum_{t=1}^n \mathbf{v}_t x_t' & n^{-1} \sum_{t=1}^n \mathbf{v}_t \mathbf{v}_t' \end{bmatrix} \rightarrow_d \begin{bmatrix} \int_0^1 V(r)V(r)' dr & 0 \\ 0 & \Gamma_{\mathbf{v},\mathbf{v}} \end{bmatrix}. \tag{35}$$

Thus, the estimator of the $I(1)$ and $I(0)$ components are asymptotically independent. For the numerator,

$$\begin{bmatrix} n^{-1} \sum_{t=1}^n x_t e_t \\ n^{-1/2} \sum_{t=1}^n \mathbf{v}_t e_t \end{bmatrix} \rightarrow_d \begin{bmatrix} \int_0^1 V(r) dE(r) \\ N(0, A_{\mathbf{v},e}) \end{bmatrix}, \tag{36}$$

where $A_{\mathbf{v},e}$ is the long-run variance of $\mathbf{v}_t e_t$. Eq. (13) then follows.

To show the limit distribution for FGLS estimator in regression (16), write

$$\begin{aligned} n^{-1} \sum_{t=1}^n \hat{e}_t^2 &= n^{-1} \sum_{t=1}^n e_t^2 + o_p(1) \rightarrow_p \sigma_e^2, \\ n^{-1} \sum_{t=1}^n \hat{e}_t \hat{e}_{t-1} &= n^{-1} \sum_{t=1}^n e_t e_{t-1} + o_p(1) \rightarrow_p \psi_e \sigma_e^2, \end{aligned}$$

where ψ_e is the first-order autocorrelation coefficient of $\{e_t\}$. Then the OLS estimator is

$$\hat{\rho}_n = \frac{n^{-1} \sum_{t=1}^n \hat{e}_t e_{t-1}}{n^{-1} \sum_{t=1}^n \hat{e}_t^2} \rightarrow_p \psi_e.$$

Conduct the Cochrane–Orcutt transformation (9) and estimate

$$\tilde{y}_t = \beta' \tilde{x}_t + \gamma' \tilde{\mathbf{v}}_t + \text{error}.$$

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

$$\tilde{y}_t = \beta'_0 \tilde{x}_t + \gamma'_0 \tilde{\mathbf{v}}_t + \tilde{e}_t,$$

where $\tilde{e}_t = e_t - \hat{\rho}_n e_{t-1}$. Using the same weight matrix H_n , write

$$\begin{bmatrix} n(\tilde{\beta}_n - \beta_0) \\ n^{1/2}(\tilde{\gamma}_n - \gamma_0) \end{bmatrix} = \begin{bmatrix} H_n^{-1} \begin{bmatrix} \sum_{t=1}^n \tilde{x}_t \tilde{x}_t' & \sum_{t=1}^n \tilde{x}_t \tilde{\mathbf{v}}_t' \\ \sum_{t=1}^n \tilde{\mathbf{v}}_t \tilde{x}_t' & \sum_{t=1}^n \tilde{\mathbf{v}}_t \tilde{\mathbf{v}}_t' \end{bmatrix} H_n^{-1} \end{bmatrix}^{-1} \begin{bmatrix} n^{-1} \sum_{t=1}^n \tilde{x}_t \tilde{e}_t \\ n^{-1/2} \sum_{t=1}^n \tilde{\mathbf{v}}_t \tilde{e}_t \end{bmatrix}. \tag{37}$$

Define $\tilde{E}(r) = (1 - \psi_e)E(r)$ and $\tilde{V}(r) = (1 - \psi_e)V(r)$. By Lemma 2.1 in Phillips and Ouliaris (1990), $n^{-1/2}\tilde{x}_{[nr]} \rightarrow_d \tilde{V}(r)$ and $n^{-1/2}\sum_{t=1}^{[nr]}\tilde{e}_t \rightarrow_d \tilde{E}(r)$. Therefore we can show that

$$\begin{aligned} n^{-2} \sum_{t=1}^n \tilde{x}_t \tilde{x}'_t &\rightarrow_d \int_0^1 \tilde{V}(r) \tilde{V}(r)' dr, \\ n^{-3/2} \sum_{t=1}^n \tilde{x}_t \tilde{v}'_t &\rightarrow_p 0, \\ n^{-1} \sum_{t=1}^n \tilde{v}_t \tilde{v}'_t &\rightarrow_p P, \quad \text{say.} \end{aligned}$$

Hence, the limit of the denominator in (37) is

$$H_n^{-1} \left[\sum_{t=1}^n z_t z'_t \right] H_n^{-1} \rightarrow_d \begin{bmatrix} (1 - \psi_e)^2 \int_0^1 V(r) V(r)' dr & 0 \\ 0 & P \end{bmatrix}.$$

Next, consider the numerator in (37). In fact, we are only interested in the first element,

$$n^{-1} \sum_{t=1}^n \tilde{x}_t \tilde{e}_t \rightarrow_d \int_0^1 \tilde{V}(r) d\tilde{E}(r) = (1 - \psi_e)^2 \int_0^1 V(r) dE(r).$$

Therefore, we obtain the limit distribution for $\tilde{\beta}_{\text{fgls}}$,

$$\begin{aligned} n(\tilde{\beta}_{\text{fgls}} - \beta_0) &\rightarrow_d \left(\int_0^1 \tilde{V}(r) \tilde{V}(r)' dr \right)^{-1} \left(\int_0^1 \tilde{V}(r) d\tilde{E}(r) \right) \\ &= \left(\int_0^1 V(r) V(r)' dr \right)^{-1} \left(\int_0^1 V(r) dE(r) \right), \end{aligned}$$

which is the same as the limit of $\hat{\beta}_{\text{dols}}$. \square

Appendix C. Proof of results in Section 2.5

In the extended test where we allow endogeneity under the alternative, the regression can be written as

$$y_t = \beta' x_t + \gamma' v_t + (\phi s_t + e_t).$$

Define $\mathbf{s} = [\phi s_1 + e_1, \dots, \phi s_n + e_n]'$. Note that $n^{-1/2}\mathbf{s} = n^{-1/2}\mathbf{e} + o_p(1)$. Similar to Appendix A, the OLS estimators for β under the alternative of a spurious regression can be written as

$$\begin{aligned} \hat{\beta}_n - \beta_0 &= (\mathbf{X}' M_v \mathbf{X})^{-1} \mathbf{X}' \mathbf{s} - (\mathbf{X}' M_v \mathbf{X})^{-1} \mathbf{X}' \mathbf{V} (\mathbf{V}' \mathbf{V})^{-1} \mathbf{V}' \mathbf{s} \\ &= \left(\frac{\mathbf{X}' M_v \mathbf{X}}{n^2} \right)^{-1} \left(\frac{\mathbf{X}' \mathbf{e}}{n^2} \right) - \frac{1}{n} \left(\frac{\mathbf{X}' M_v \mathbf{X}}{n^2} \right)^{-1} \left(\frac{\mathbf{X}' \mathbf{V}}{n} \right) \left(\frac{\mathbf{V}' \mathbf{V}}{n^2} \right)^{-1} \left(\frac{\mathbf{V}' \mathbf{e}}{n} \right) + o_p(1) \\ &= \left(\frac{\mathbf{X}' M_v \mathbf{X}}{n^2} \right)^{-1} \left(\frac{\mathbf{X}' \mathbf{e}}{n^2} \right) + o_p(1) \\ &\rightarrow_d \left[\int_0^1 V(r) V(r)' dr \right]^{-1} \left[\int_0^1 V(r) U(r) dr \right]. \end{aligned}$$

Due to endogeneity, the estimator in the differenced regression is not consistent either. The estimators $(\tilde{\beta}'_{\text{dglS}} - \beta'_0, \tilde{\gamma}'_{\text{dglS}} - \gamma'_0) \rightarrow_p Q^{-1} \phi (E(v_t \Delta s_t)', E(\Delta v_t \Delta s_t)')$. Let $\tilde{\beta}$ denote the limit of $\tilde{\beta}_{\text{dglS}}$, then

$$\begin{aligned} \sqrt{n}(\tilde{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}}) &= \sqrt{n}(\tilde{\beta}_{\text{dglS}} - \tilde{\beta}) - \sqrt{n}(\hat{\beta}_{\text{dols}} - \tilde{\beta}) \\ &= O_p(1) + O_p(\sqrt{n}) \\ &= O_p(\sqrt{n}). \end{aligned}$$

Finally, in the differenced regression, the variance estimate still converges. Therefore, the Hausman-type test statistic is of order n under the alternative of spurious regressions no matter whether exogeneity holds or not. \square

Appendix D. Data descriptions

In the application of the U.S. money demand, we use the same data set as in Stock and Watson (1993, p. 817) for M1, GNP, price deflator, and 6-month commercial CP rates during 1947–1997. Readers are referred to the original work for further details on data. In the second application, we retrieved the per capita output series from the *Penn World Tables: version 6.2* of Summers and Heston (2006). This consists of annual data on real GDP per capita (RGDPCH) for four developing countries (Columbia, Ecuador, Egypt, and Pakistan) along with four industrial countries (Denmark, New Zealand, South Africa, and Switzerland) over the period of 1951–2003. In the PPP application we borrow the data set from Kim (2005) who constructed the real exchange rate for total consumption using the general price deflator and the real exchange rate for traded and non-traded goods using implicit deflators for non-service consumption and service consumption, respectively. Data are quarterly observations spanning from 1974 Q1 to 1998 Q4. The exchange rates for Canada, France, Italy, Japan, the United Kingdom, and the United States are taken from the International Financial Statistics (IFS) CD-ROM, and bilateral real exchange rates of traded and non-traded goods classified by type and total consumption deflators are from the Quarterly National Accounts and Data Stream.

Appendix E. Guidelines on empirical application

Procedures for GLS- and FGLS-corrected estimations

Step 1: Choose a length (p) of lead and lag terms using popular lag selection rules such as AIC, BIC, or their modified versions due to Ng and Perron (2001). Given that the lead and lag length selection issue has not been settled in the dynamic OLS literature, we recommend to report results from different orders together with BIC as a rough guideline. To correct for the endogeneity problem, the instrumental variable (IV) approach can also be applied. The IV approach is appealing as it does not involve choosing the proper length of leads and lags, but the downside is that it is not easy to find good instruments in practice.

$$y_t = \sum_{k=1}^p \gamma_k \Delta x_{t+k} + \beta x_t + \sum_{k=0}^p \phi_k \Delta x_{t-k} + e_t.$$

Step 2: (GLS-corrected estimation) Filter all variables in the above equation by taking the full difference

$$\Delta y_t = \sum_{k=1}^{p-1} \gamma_k \Delta^2 x_{t+k} + \beta \Delta x_t + \sum_{k=0}^{p-1} \phi_k \Delta^2 x_{t-k} + \Delta e_t,$$

$$\Delta y_t = \theta' \Delta z_t + \Delta e_t.$$

Step 2': (FGLS-corrected estimation) Retrieve the OLS residuals such that $\hat{e}_t = y_t - \sum_{k=1}^p \hat{\gamma}_k \Delta x_{t+k} + \hat{\beta} x_t + \sum_{k=0}^p \hat{\phi}_k \Delta x_{t-k}$, and obtain $\hat{\rho}$ from regressing \hat{e}_t onto \hat{e}_{t-1} . After n' iterations $\hat{\rho}_n$ can be obtained. The variables are transformed such that $\tilde{y}_t = y_t - \hat{\rho}_n y_{t-1}$, $\tilde{x}_t = x_t - \hat{\rho}_n x_{t-1}$, and $\Delta \tilde{x}_{t+k} = \Delta x_{t+k} - \hat{\rho}_n \Delta x_{t+k-1}$.

Step 3: Apply OLS to estimate $\theta = \{\gamma_1, \dots, \gamma_{p-1}, \beta, \phi_1, \dots, \phi_{p-1}\}$. The obtained estimates are the (F)GLS corrected estimates of θ .

The Hausman-type cointegration test

Step 1: Obtain the DOLS and GLS-corrected estimates for the parameters. We recommend to report the results from different orders of lead and lag terms together with the one chosen by the BIC rule as a guideline. When selecting lead and lag lengths through the BIC rule, it is recommended to choose the lengths for DOLS

and DGLS separately. That is, $\hat{\theta}_{\text{dols}}$ is obtained using the BIC lag length from DOLS regression equation while $\hat{\theta}_{\text{dglS}}$ is obtained using the BIC lag length selected from DGLS regression equation as described above.

Step 2: Compute \hat{V}_{β} , a consistent estimate for the long run variance matrix of $\sqrt{n}(\hat{\theta}_{\text{dglS}} - \theta)$ using the heteroskedasticity and autocorrelation consistent (HAC) estimator. In the empirical part of this paper we adopted the long-run variance estimator from Andrews and Monahan (1992) with a quadratic spectral (QS) kernel using prewhitening. Readers are also referred to the recent study by Sul et al. (2005) who propose a recursive demeaning and recursive Cauchy estimation to reduce the small sample bias in prewhitening coefficient estimates as well as a sample-size-dependent boundary condition rule that substantially enhances power without compromising size.

Step 3: Construct the test statistic

$$h_n = n(\hat{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}})' \hat{V}_{\beta}^{-1} (\hat{\beta}_{\text{dglS}} - \hat{\beta}_{\text{dols}}) \rightarrow_d \chi^2(m).$$

References

- Andrews, D.W.K., Monahan, J.C., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60, 953–966.
- Bae, Y., de Jong, R., 2007. Money demand function estimation by nonlinear cointegration. *Journal of Applied Econometrics*, forthcoming.
- de Jong, R., Davidson, J., 2000. The functional central limit theorem and weak convergence to stochastic integrals I: weakly dependent processes. *Econometric Theory* 16, 621–642.
- Durbin, J., 1954. Errors in variables. *Review of the International Statistical Institute* 22, 23–32.
- Durlauf, S.N., Phillips, P.C.B., 1988. Trends versus random walks in time series analysis. *Econometrica* 56, 1333–1335.
- Durlauf, S.N., Quah, D., 1999. The new empirics of economic growth. In: Taylor, J.B., Woodford, M. (Eds.), *Handbook of Macroeconomics*, vol. 1A. North-Holland, Amsterdam, pp. 235–308.
- Easterly, W., 2001. The middle class consensus and economic development. *Journal of Economic Growth* 6, 317–336.
- Engel, C., 1999. Accounting for US real exchange rate changes. *Journal of Political Economy* 107, 507–538.
- Fernández-Macho, J., Mariel, P., 1994. Hausman-like and variance-ratio tests for cointegrated regressions. *Vysoké Školy Ekonomické*, 27–42.
- Granger, C.W.J., Newbold, P., 1974. Spurious regressions in econometrics. *Journal of Econometrics* 74, 111–120.
- Hausman, J., 1978. Specification tests in econometrics. *Econometrica* 46, 1251–1271.
- Hu, L., 2005. A simple panel cointegration test and estimation in $I(0)/I(1)$ mixed panel data. Unpublished manuscript, Ohio State University.
- Hu, L., 2006. Residual-based cointegration test for the null of structural spurious regressions. Unpublished manuscript, Ohio State University, Yale University.
- Kakkar, V., Ogaki, M., 1999. Real exchange rates and nontradables: a relative price approach. *Journal of Empirical Finance* 6, 193–215.
- Kim, J., 2005. Convergence rates to PPP for traded and non-traded goods: a structural error correction model approach. *Journal of Business and Economic Statistics* 23, 76–86.
- Lewbel, A., Ng, S., 2005. Demand system with nonstationary prices. *Review of Economics and Statistics* 87, 479–494.
- Nelson, C.R., Kang, H., 1981. Spurious periodicity in inappropriately detrended time series. *Econometrica* 49, 741–751.
- Nelson, C.R., Kang, H., 1983. Pitfalls in the use of time as an explanatory variable in regression. *Journal of Business and Economic Statistics* 2, 73–82.
- Ng, S., Perron, P., 2001. Lag length selection and the construction of unit-root tests with good size and power. *Econometrica* 69, 1519–1554.
- Ogaki, M., Choi, C.Y., 2001. The Gauss–Markov theorem and spurious regressions. WP 01-13, Department of Economics, The Ohio State University.
- Ogaki, M., Park, Y.J., 1998. A cointegration approach to estimating preference parameters. *Journal of Econometrics* 82, 107–134.
- Park, S., Ogaki, M., 2007. Long-run real exchange rate changes and the properties of the variance of k-differences. WP 07-01. Department of Economics, The Ohio State University.
- Phillips, P.C.B., 1986. Understanding spurious regressions in econometrics. *Journal of Econometrics* 33, 311–340.
- Phillips, P.C.B., 1989. Partially identified econometric models. *Econometric Theory* 5, 95–132.
- Phillips, P.C.B., 1998. New tools for understanding spurious regression. *Econometrica* 66, 1299–1325.
- Phillips, P.C.B., Hodgson, D.J., 1994. Spurious regression and generalized least squares. *Econometric Theory* 10, 957–958.
- Phillips, P.C.B., Ouliaris, S., 1990. Asymptotic properties of residual based tests for cointegration. *Econometrica* 58, 165–193.
- Phillips, P.C.B., Park, J.Y., 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.

- Shin, Y., 1994. A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory* 10, 91–115.
- Stock, J.H., Watson, M.W., 1993. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820.
- Stockman, A.C., Tesar, L., 1995. Tastes and technology in a two-country model of the business cycle: explaining international comovements. *American Economic Review* 85, 168–185.
- Sul, D., Phillips, P.C.B., Choi, C.Y., 2005. Prewhitening bias in HAC estimation. *Oxford Bulletin of Economics and Statistics* 67, 517–546.
- Summers, R., Heston, A., 2006. Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- Wu, D., 1973. Alternative tests of independence between stochastic regressors and disturbances. *Econometrica* 41, 733–750.