

NBER WORKING PAPER SERIES

RELATIVE LABOR PRODUCTIVITY  
AND THE REAL EXCHANGE RATE IN  
THE LONG RUN: EVIDENCE FOR A PANEL  
OF OECD COUNTRIES

Matthew B. Canzoneri  
Robert E. Cumby  
Behzad Diba

NBER Working Paper 5676

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
July 1996

Helpful discussions with José Vinals and generous assistance and numerous suggestions from Peter Pedroni are acknowledged with thanks. This paper is part of NBER's research program in International Finance and Macroeconomics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1996 by Matthew B. Canzoneri, Robert E. Cumby and Behzad Diba. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

RELATIVE LABOR PRODUCTIVITY  
AND THE REAL EXCHANGE RATE IN  
THE LONG RUN: EVIDENCE FOR A PANEL  
OF OECD COUNTRIES

**ABSTRACT**

The Balassa-Samuelson model, which explains real exchange rate movements in terms of sectoral productivities, rests on two components. First, for a class of technologies including Cobb-Douglas, the model implies that the relative price of nontraded goods in each country should reflect the relative productivity of labor in the traded and nontraded goods sectors. Second, the model assumes that purchasing power parity holds for traded goods in the long-run. We test each of these implications using data from a panel of OECD countries. Our results suggest that the first of these two fits the data quite well. In the long run, relative prices generally reflect relative labor productivities. The evidence on purchasing power parity in traded goods is considerably less favorable. When we look at US dollar exchange rates, PPP does not appear to hold for traded goods, even in the long run. On the other hand, when we look at DM exchange rates purchasing power parity appears to be a somewhat better characterization of traded goods prices.

Matthew B. Canzoneri  
Department of Economics  
Georgetown University  
Washington, DC 20057

Robert E. Cumby  
Department of Economics  
Georgetown University  
Washington, DC 20057  
and NBER

Behzad Diba  
Department of Economics  
Georgetown University  
Washington, DC 20057

Changes in real exchange rates—defined as the relative price of national outputs—have been so persistent that they question the very notion of purchasing power parity (PPP). Even the evidence suggesting that deviations from PPP are temporary points to a half life of four to five years.<sup>1</sup>

Explanations of persistent real exchange rate changes have often followed the lead of Balassa (1964) and Samuelson (1964), who divide national output into traded goods and nontraded goods and explain real exchange rates in terms of sectoral productivity. The “Balassa-Samuelson” hypothesis divides real exchange rate movements into two components. Competitive behavior implies that the relative price of nontraded goods depends on the ratio of the marginal costs in the two sectors, and we will show that for a wide class of production technologies the ratio of marginal costs is proportional to the ratio of average labor products in the two sectors. So, the first component of the hypothesis is the assumption that the relative price of tradeables is proportional to the ratio of average labor products. The second component is the assumption of PPP for traded goods.

The two components combine to produce a simple model of real exchange rate movements. If, for example, the ratio of traded goods productivity to nontraded goods productivity is growing faster at home than abroad, then the relative price of nontraded goods must be growing faster at home than abroad, and the price of home national output must be rising relative to the price of foreign national output (since by assumption the prices of traded goods equalize). In other words, if traded good productivity relative to nontraded good productivity is

---

<sup>1</sup> Frankel and Rose (1996) and Lothian (1994) report half lives of this magnitude using aggregate price levels as do Wei and Parsley (1994) using sectoral prices.

growing faster at home than abroad, then the home country should experience a real appreciation.

How well does the Balassa-Samuelson model explain real exchange rate movements? In figure 1 we plot four bilateral real exchange rates along with the ratio of labor productivity in the traded and nontraded goods sectors for each country. The Balassa-Samuelson hypothesis clearly fails to explain the short run movements in real exchange rates, and figure 1 suggests that there may also be problems in explaining the long run movements, especially where the U.S. dollar is concerned. Relative traded goods productivity has risen much faster in the United States than in Germany, especially since the mid-1970s. This should have led to a real appreciation of the dollar (a fall in the real exchange rate as plotted in figure 1); instead, the dollar has depreciated. And the negligible difference between U.S. and Japanese productivity trends has little hope of explaining the real depreciation of the dollar against the yen. The Balassa-Samuelson hypothesis does seem to fare better with the DM. Since the mid-1970s, relative traded goods productivity has been growing more rapidly in both Italy and Japan than it has in Germany, and both the lira and the yen have appreciated in real terms against the DM.

Figure 1 points to some problems with the Balassa-Samuelson hypothesis, but it does not reveal the source of the problems. Nor does it explain why there might be a problem with the dollar real exchange rates. Either component of the hypothesis can be challenged, and the special problem with the dollar—if one exists—could reside in either place.

In this paper we examine the two components of the Balassa-Samuelson hypothesis using a panel of 13 OECD countries. We do not require that either part of the hypothesis holds in the short run. In fact, both theory and a substantial body of empirical evidence suggest that short run deviations from both can be substantial. Instead, we focus our tests on the long-run or trend

behavior of the series.

Our work differs from the work of others in three important ways:<sup>2</sup> we use the average product of labor to measure productivity instead of Solow residuals, or total factor productivity (TFP);<sup>3</sup> we employ recently developed time series techniques that allow us to use panel data; and finally we develop Monte Carlo experiments to explain certain small sample biases that remain. Each of these innovations may need some elaboration.

We use the average product of labor instead of TFP for four reasons: (1) The almost exclusive focus on TFP in the recent literature reflects (at least in part) an interest in assessing the relative importance of supply shocks (as proxied by TFP) and demand shocks (as proxied by government expenditures, etc.) in explaining real exchange rate movements. Interpreting movements in Solow residuals as exogenous supply shocks is problematic in and of itself,<sup>4</sup> and in any event we think that this is a case of putting the cart before the horse. We do not need to take a stand on the relative importance of supply and demand shocks to test the first component of the Balassa-Samuelson hypothesis; we will show that in using average labor productivity we are implicitly allowing both supply and demand shocks to affect real exchange rates. We think that a

---

<sup>2</sup> Asea and Corden (1994) and Froot and Rogoff (1995) provide surveys of the literature.

<sup>3</sup> Although Marston (1987) and others have used average labor productivity, most of the recent literature has focused on total factor productivity computed from a Cobb-Douglas production function. See, in particular, Chinn (1995), Chinn and Johnston (1996), De Gregorio, Giovannini and Krueger (1993), De Gregorio, Giovannini and Wolf (1993), and Froot and Rogoff (1991).

<sup>4</sup> Evans (1992), for example, shows that measured Solow residuals are Granger caused by money, interest rates, and government spending and finds that one fourth to one half of their variation is attributable to variations in aggregate demand. Summers (1986) and Mankiw (1989) present earlier skeptical views on the exogeneity of total factor productivity.

test of the basic Balassa-Samuelson assumption of marginal cost pricing should precede any attempt to assess the relative importance of supply and demand shocks. (2) We don't need data on sectoral capital stocks, which are likely to be less reliable than data on sectoral employment and value added. (3) Our development of the first component of the Balassa-Samuelson hypothesis holds for a broader class of technologies than the Cobb-Douglas production function, which is used to compute Solow residuals. (4) Our specification allows a sharper test of the first component of the Balassa-Samuelson hypothesis; in particular, we do not need to rely on outside estimates of labor's share in production.

Two unfortunate facts have combined to plague the time series literature on real exchange rates: we typically only have 20 to 30 years of data on any pair of countries, and unit root tests have notoriously low power in small samples to distinguish between series that are nonstationary and series that are stationary but have highly persistent dynamics. Recently developed techniques allow us to deal with nonstationary data in heterogeneous panels,<sup>5</sup> and in fact combining the data from 13 countries yields substantial benefits. We are able to confirm the existence of a long run (or cointegrating) relationship between the relative price of nontradeables and the ratio of average labor products, and we are able to estimate its parameters with a surprising degree of precision.

The precision is so great that we are seemingly able to reject the first component of Balassa-Samuelson model. This leads us to perform some Monte Carlo experiments to show that the remaining small sample bias is capable of explaining our apparent rejection of the model. We

---

<sup>5</sup> Chinn and Johnston (1996) use one of the tests that we use for dealing with nonstationary panels. The focus of their work is, however, quite different. They use the Balassa-Samuelson model as a broad motivation for specifying a number of forecasting equations for the trend in the real exchange rate. In contrast we test the two component hypotheses that make up the Balassa-Samuelson model

will argue that this firmly establishes the first component of the Balassa-Samuelson hypothesis.

We will argue instead that the problems with the Balassa-Samuelson hypothesis lie in the failure of PPP to explain traded goods prices, especially for the U.S. dollar. Once again, using panel data we are able to establish that the appropriate long run (or cointegrating) relationship exists, but our estimates of its parameters are much less precise, and none of the parameter estimates are close to the values implied by PPP. And, as figure 1 would seem to suggest, we find that the failure of PPP to hold for traded goods may be largely a U.S. dollar phenomenon. When we use the DM as numeraire, the results are much more favorable.

The plan of the paper is as follows. In section 2 we describe our notation and set out the two components of the Balassa-Samuelson hypothesis. We review the econometric methods we use in section 3, we present our empirical results in section 4. In section 5 we offer some concluding remarks.

## 2. The Analytical Framework

We begin with the link between the relative price of nontraded goods and the relative productivities in the traded and nontraded goods sectors. The analytical framework we use is quite general; it could be embedded in a wide class of models. In each country, capital and labor are employed in the production of traded goods, T, and non-traded goods, N. Competition implies that labor is paid the value of its marginal product, and labor mobility implies that the nominal wage rate, W, is equal in the two sectors.

$$\frac{\partial T / \partial L^T}{\partial N / \partial L^N} = \frac{W / P^T}{W / P^N} = \frac{P^N}{P^T} = q \quad (1)$$

Equation (1) states the familiar condition that the relative price of nontraded goods, which we denote as  $q$ , is equal to the slope of the production possibility curve.

To measure the marginal products in equation (1), we assume that the marginal product of labor is proportional to the average product of labor in each sector.

$$\frac{\partial T / \partial L^T}{\partial N / \partial L^N} = \frac{\varphi(T / L^T)}{\psi(N / L^N)} \quad (2)$$

With Cobb Douglas technologies,  $\varphi$  and  $\psi$  are the labor shares in value added in the traded and nontraded goods sectors. But equation (2) will hold under assumptions that are much less restrictive than Cobb Douglas. Average and marginal products will be proportional if the production functions can be expressed as,

$$T = F(K^T)(L^T)^\varphi, \quad N = G(K^N)(L^N)^\psi \quad (3)$$

where  $F(\cdot)$  and  $G(\cdot)$  are arbitrary functions that may, for example, depend on inputs other than the firms' choices of labor and capital as in the endogenous growth literature.<sup>6</sup> Moreover, if labor and capital are both mobile across sectors, then (2) holds in equilibrium even for some technologies that cannot be represented in the form of (3); it is, for example, straightforward to check that CES production functions with constant returns to scale satisfy (3) in equilibrium if both factors are mobile.

When average and marginal products are proportional, the relative price of non-traded

---

<sup>6</sup> The production functions specified in (3) are the general solutions to the differential equations:  $\partial T / \partial L^T = \varphi(T / L^T)$  and  $\partial N / \partial L^N = \psi(N / L^N)$ . We require only that they satisfy the standard properties of production functions.



goods,  $q$ , is proportional to the ratio of the average products of labor,  $x/n$ , in the two sectors:

$$q = \frac{P^N}{P^T} = \frac{\varphi}{\psi} \frac{T/L^T}{N/L^N} = \frac{\varphi}{\psi} \frac{x}{n} \quad (4)$$

Adding subscripts to denote country  $i$  at date  $t$  and taking natural logarithms of both sides of equation (4), we get

$$\ln(q_{i,t}) = \ln\left(\frac{\varphi_i}{\psi_i}\right) + \ln\left(\frac{x_{i,t}}{n_{i,t}}\right) \quad (5)$$

Although the model described above does not distinguish short-run and long-run fluctuations, in our empirical tests we interpret equation (5) as a restriction on the long-run trends in the relative price of non-traded goods and relative labor productivity in the two sectors. Thus, in Section 4 we test whether  $\ln(q)$  and  $\ln(x/n)$  are cointegrated and whether the cointegrating slope is one.

We now turn the second part of the Balassa-Samuelson model, the assumption that traded goods prices are characterized by purchasing power parity. Let  $E_{i,t}$  be the nominal exchange rate of currency  $i$  relative to currency 1, the numeraire currency at time  $t$ .<sup>7</sup> If  $E_{i,t}$  is expressed in units of the numeraire currency per unit of currency  $i$ , purchasing power parity implies that

$E_{i,t} P_{i,t}^T = P_{1,t}^T$ . If we define the PPP exchange rate for country  $i$  at date  $t$  as

$$r_{i,t} = \frac{P_{1,t}^T}{P_{i,t}^T} \quad (6)$$

---

<sup>7</sup> We use two numeraire currencies, the U.S. dollar and the DM, in our tests.

purchasing power parity implies that the nominal exchange rate is equal to the PPP exchange rate. The interesting question is whether purchasing power parity holds in the long run for traded goods. As is well known, it fails dramatically in the short run. We therefore test whether  $\ln(E_{i,t})$  is cointegrated with  $\ln(r_{i,t})$ , and whether the cointegrating slope is unity.

### 3. A Brief Review of the Econometric Methods

Because the tests that we consider focus on the long-run or trend behavior of the relative prices, relative productivities and nominal and PPP exchange rates, we begin by examining the long-run or trend behavior of each of the series and then test the restrictions that the hypotheses imply on that long-run behavior. In all of the tests that we implement, we allow the short-run dynamics to be relatively unconstrained and focus only on the long-run behavior of the data.

Before proceeding, we ask whether each of the series is better characterized by stationary deviations from a deterministic trend or by stochastic trends, or possibly by both a deterministic and stochastic trend. We therefore consider a univariate autoregression for each series,

$$\Delta z_{i,t} = \theta_i + (\rho_i - 1) z_{i,t-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta z_{i,t-j} + v_{i,t} \quad (7)$$

and carry out three tests. The first is an augmented Dickey Fuller (1979) t-test of the hypothesis that  $\rho_i=1$ . We allow the data to determine the number of lags,  $k_i$ , using the general-to-specific procedures suggested by Hall (1994) and Ng and Perron (1995). The other two tests, proposed by Phillips and Perron (1988), set  $k_i = 0$  when estimating  $\rho$  and adjust the test statistics using  $k_i$  autocovariances of the residual to account for autocorrelation. We also carry out the three tests

adding a trend to (7).

Im, Pesaran and Shin (1995) propose testing the null hypothesis that  $\rho_i = 1$  for  $i = 1, \dots, N$  by computing the average t-ratio,  $\bar{t}_N = N^{-1} \sum_{i=1}^N (\hat{\rho}_i - 1) / \hat{\sigma}_p$ . They show that the statistic,

$$z_{N,T} = \sqrt{N} \left( \frac{\bar{t}_N - a_T}{\sqrt{b_T}} \right) \quad (8)$$

is distributed asymptotically as a standard normal. The mean and variance of the average t-ratio,  $a_T$  and  $b_T$  depend only on  $T$  and  $\bar{k}$  (the average of the  $k_i$ ) and are tabulated by Im, Pesaran, and Shin.<sup>8</sup>

If the data contain stochastic trends, the Balassa-Samuelson model implies that pairs of series must share the same stochastic trend, that is they must be cointegrated. We test for cointegration using both standard, residual-based tests for each country and the panel tests proposed by Pedroni (1995). These tests examine the residual from the regression,  $y_{i,t} = \alpha_i + \beta_i x_{i,t} + \epsilon_{i,t}$ .<sup>9</sup> We then use the estimated residuals,  $\hat{\epsilon}_{i,t}$ , in the autoregression,

---

<sup>8</sup> The asymptotic distribution is obtained by letting  $N \rightarrow \infty$ ,  $T \rightarrow \infty$ , and  $N/T \rightarrow \infty$  and assuming that the data are generated independently across  $i$ . They suggest taking out common time effects by using time dummies or (equivalently) computing the test statistics using deviations from cross section means to account for correlation across cross section units. Separate values for the constants are tabulated for t-ratios computed from regression with and without a trend. Our cross sections differ in the number of observations that we have available so we use all of the available data to compute the t-ratios for each cross section.

<sup>9</sup> In one set of tests  $y_{i,t}$  is the log relative price of nontraded goods and  $x_{i,t}$  is the log of the relative average products of labor. In the second set of tests,  $y_{i,t}$  is the nominal exchange rate of currency  $i$  with the numeraire currency and  $x_{i,t}$  is the corresponding PPP exchange rate.

$$\Delta \hat{\epsilon}_{i,t} = (\rho_i - 1) \hat{\epsilon}_{i,t-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \hat{\epsilon}_{i,t-j} + v_{i,t} \quad (9)$$

and test the null hypothesis that  $\rho_i = 1$  using the same three unit root tests that we used on the raw data as suggested by Engle and Granger (1987) and Phillips and Ouliaris (1990). Pedroni (1995) proposes a series of tests of cointegration in heterogeneous panels that can be viewed as extensions of the single-equation tests that we use.<sup>10</sup> Before describing the test statistics, we define some notation. Let:

$\hat{\mu}_{i,t}$  be the residual from the autoregression,  $\hat{\epsilon}_{i,t} = \rho_i \hat{\epsilon}_{i,t-1} + \mu_{i,t}$ ,

$\hat{\sigma}_i^2$  be the estimated long-run variance of  $\hat{\mu}_{i,t}$  ( $2\pi f_\mu(0)$ , where  $f_\mu(\omega)$  is the spectral density of  $\mu$  at frequency  $\omega$ ),

$$\hat{s}_i^2 = N^{-1} \sum_{t=1}^T \hat{\mu}_{i,t}^2,$$

$$\lambda_i = .5 (\sigma_i^2 - s_i^2),$$

$\xi_{i,t}$  be the vector of innovations to the process  $\{y_{i,t}, x_{i,t}\}$ ,

$\Omega_i$  be the long-run covariance matrix of  $\xi_{i,t}$ ,

$$L_i = \left( \Omega_i(1,1) - \frac{\Omega_i^2(2,1)}{\Omega_i(2,2)} \right)^{1/2},$$

$$\text{and } \hat{\sigma}_{NT}^2 = N^{-1} \sum_{i=1}^N L_i^{-2} \hat{\sigma}_i^2.$$

The first of Pedroni's panel cointegration tests is an extension of the semiparametric (Phillips-Perron and Phillips-Ouliaris)  $T(\hat{\rho}-1)$  test. Pedroni (1995) shows that,

---

<sup>10</sup> The parameters  $\alpha_i$  and  $\beta_i$  in the cointegrating regression for each cross section are allowed to differ as are the dynamics of  $y_{i,t}$  and  $x_{i,t}$ .

$$\mathbf{T} \sqrt{N} (\mathbf{Z}_{\hat{\rho}} - 1) = \sqrt{N} \frac{\sum_{i=1}^N \mathbf{T}_i^{-1} \hat{\mathbf{L}}_i^{-2} \sum_{t=1}^{T_i} (\hat{\epsilon}_{i,t-1} \Delta \hat{\epsilon}_{i,t} - \hat{\lambda}_i)}{\sum_{i=1}^N \mathbf{T}_i^{-2} \hat{\mathbf{L}}_i^{-2} \sum_{t=1}^{T_i} \hat{\epsilon}_{i,t-1}^2} \quad (10)$$

is asymptotically normal and presents critical values of the distribution for several combinations of N and T.<sup>11</sup> The second test extends the semiparametric t-test. Pedroni (1995) shows that,

$$Z_{\hat{\rho}} = \frac{\sum_{i=1}^N \mathbf{T}_i^{-1} \hat{\mathbf{L}}_i^{-2} \sum_{t=1}^{T_i} (\hat{\epsilon}_{i,t-1} \Delta \hat{\epsilon}_{i,t} - \hat{\lambda}_i)}{\left( \hat{\sigma}_{NT}^2 \sum_{i=1}^N \mathbf{T}_i^{-2} \hat{\mathbf{L}}_i^{-2} \sum_{t=1}^{T_i} \hat{\epsilon}_{i,t-1}^2 \right)^{1/2}} \quad (11)$$

is asymptotically normal and presents critical values for the distribution for several combinations of N and T. The third test, which extends the ADF test of cointegration, is similar to  $Z_{\hat{\rho}}$  but uses the cross moments computed from (9).

The Balassa-Samuelson model not only implies that pairs of variables should be cointegrated, it implies that the slopes,  $\beta_i$ , in the cointegrating relationships should have the common value of 1.0. We test this hypothesis in two ways. First, we impose  $\beta_i = 1.0$  and test for unit roots in the difference  $y_{i,t} - x_{i,t}$  using the same set of tests (the Phillips-Perron and ADF tests for each cross section along with the Im-Pesaran-Shin tests for the panel) that we use to test for unit roots in the individual series. Next, we compute the Phillips-Hansen fully modified OLS

---

<sup>11</sup> This statistic is a slight modification of Pedroni's that allows for a different number of time series observations in each cross section. Pedroni's panel tests, like the Im, Pesaran, Shin tests, require cross sectional independence after removing common time effects. We account for cross section correlation by using common time dummies to remove shocks common to all cross sections for each year.

estimates of  $\beta$  and the corresponding standard error and test the null hypothesis that  $\beta_i = 1.0$ .

Pedroni (1996) proposes two tests of the hypothesis that the (common) cointegrating slope for heterogenous panels,  $\beta$ , is equal to some value,  $\beta_0$ , that we also use to test the null that the  $\beta_i$  are jointly equal to 1.0.

Again, before describing the tests, we define some notation. The model is,

$$\begin{aligned} y_{i,t} &= \alpha_{1,i} + \beta_i x_{i,t} + \epsilon_{i,t} \\ \Delta x_{i,t} &= \alpha_{2,i} + \zeta_{i,t} \end{aligned} \tag{12}$$

where we make no assumptions about the exogeneity of the regressors. Let:

$\Omega_i = \Omega_{0,i} + \Gamma_i + \Gamma_i'$  be the long-run covariance matrix of  $\eta_{i,t} = (\epsilon_{i,t}, \zeta_{i,t})'$ ,

$C_i$  be the Cholesky factor of  $\Omega_i$  with  $C_i(1,1) = \left( \Omega_i(1,1) - \frac{\Omega_i^2(2,1)}{\Omega_i(2,2)} \right)^{1/2}$ ,

$y_{i,t}^* = y_{i,t} - \bar{y}_i - \frac{\hat{C}_i(2,1)}{\hat{C}_i(2,2)} (\Delta x_{i,t} - \overline{\Delta x_i})$ ,

and  $\gamma_i = \Gamma_i(2,1) + \Omega_{0,i}(2,1) - (\Gamma_i(2,2) + \Omega_{0,i}(2,2)) C_i(2,1)/C_i(2,2)$ .

The fully modified OLS estimate of  $\beta$  is,

$$\hat{\beta}_{i,FMOLS} = \frac{\sum_{t=1}^{T_i} (x_{i,t} - \bar{x}_i) y_{i,t}^* - T_i \hat{\gamma}_i}{\sum_{t=1}^{T_i} (x_{i,t} - \bar{x}_i)^2}$$

and the corresponding standard error is  $\hat{C}_i(1,1) / \sum_{t=1}^{T_i} (x_{i,t} - \bar{x}_i)^2$ .

Both tests suggested by Pedroni (1996) can be thought of as tests that extend fully modified OLS to heterogeneous panels in which the intercepts  $\alpha_{1,i}$  and  $\alpha_{2,i}$  can differ across  $i$  as can the dynamics of  $\eta_{i,t}$ . The first is the group mean t-ratio (as in Im, Pesaran, and Shin (1995)),

which Pedroni (1996) shows is distributed as standard normal under the null hypothesis that  $\beta_i = \beta_0$  for all  $i$ .

$$\sqrt{N} \bar{t}_{\beta} = \frac{1}{\sqrt{N}} \sum_{i=1}^N \hat{C}_i^{-1}(1,1) \sqrt{\sum_{t=1}^{T_i} (x_{i,t} - \bar{x}_i)^2} \left( \sum_{t=1}^{T_i} (x_{i,t} - \bar{x}_i) y_{i,t}^* - T_i \gamma_i - \beta_0 \right) \quad (13)$$

The second is the t-ratio,  $\hat{t}_{\beta}$ ,

$$\left( \sum_{i=1}^N \hat{C}_i^{-1}(1,1) \hat{C}_i^{-1}(2,2) \left( \sum_{t=1}^{T_i} (x_{i,t} - \bar{x}_i) \mu_{i,t}^* - T_i \gamma_i \right) \right) \sqrt{\sum_{i=1}^N \hat{C}_i^{-2}(2,2) \sum_{t=1}^{T_i} (x_{i,t} - \bar{x}_i)^2}$$

where  $\mu_{i,t}^* = y_{i,t}^* + \frac{\hat{C}_i(1,1) - \hat{C}_i(2,2)}{\hat{C}_i(2,2)} (x_{i,t} - \bar{x}_i) \beta_0$ .

#### 4. Empirical Results

Before examining either of the hypotheses of interest, we begin by examining the trend behavior of each of these series. Table 1 contains the results of tests of the hypothesis that the relative prices of nontraded goods in the countries in our panel contain a unit root. There is little evidence against the null hypothesis of a unit root. When we look at each country individually, the null hypothesis is rejected for Germany with one of the three tests and for Denmark with two of the three tests. The panel test yields no evidence against the null—the mean t-ratio is negative, but its expected value is negative and larger in absolute value. And once we include time the average t-ratio is essentially equal to its expected value.

Because the data may contain deterministic as well as stochastic trends, we also carry out the unit root tests including a time trend. The panel tests of a unit root again yield no evidence against the null. We also carry out an F-test of the null that  $\rho_i = 1$  and the trend coefficient is zero for each country and find evidence of a deterministic trend for Denmark and possibly for Canada. Interestingly, once a time trend is included in the regression for Denmark, the evidence against a unit root disappears. On the other hand, when a time trend is included in the regression for Canada, we reject the null of a unit root.

Table 2 contains the unit root tests for relative productivities. Again, we find little evidence against the null hypothesis of a unit root. Regardless of whether we include common time dummies, the panel tests do not come close to rejecting a unit root at conventional significance levels. Among the individual regressions, only for Japan do we reject a unit root and this is no longer the case once we include a deterministic trend in the regressions. The country-by-country F-tests point to evidence of deterministic trends in the regressions for Japan, Great Britain, and Austria. And when we include a deterministic trend, we can reject a unit root in the regressions for Great Britain and Austria.

The evidence points overwhelmingly to unit roots in both relative productivities and the relative price of nontraded goods. We therefore proceed with our tests looking at cointegrating relationships. But because the unit root assumption might be questionable for Canada, Great Britain, and Austria, we will carry out all joint tests of this first part of the Balassa-Samuelson hypothesis both including and excluding these three countries.

Tables 3 and 4 contain the results of the unit root tests for nominal exchange rates and the PPP exchange rates, respectively. There is no evidence against the null of a unit root for either of



the series, both when we consider each country individually and when we use the panel to carry our joint tests. In addition, in no instance do we find evidence of deterministic trends at the 5 percent level and in only one instance for each series do we find evidence at the 10 percent level. We therefore proceed assuming that both the nominal exchange rate and the PPP exchange rate for all countries in our panel contain a stochastic trend but no deterministic trend. Now that the preliminary testing is completed, we turn to the hypotheses of interest.

#### 4.1 Relative Prices and Relative Productivities

We begin with the hypothesis that, in the long run, the relative price of nontraded goods reflects relative productivities in the traded and nontraded goods sectors. Table 5 contains the results of tests of the null hypothesis that the relative prices and the relative productivities are not cointegrated. The tests using data from an individual country yield mixed evidence. The augmented Dickey-Fuller tests reject the null hypothesis of no cointegration for 7 of the 13 countries at the 10 percent level, and 6 of these 7 test statistics are also significant at the 5 percent level. The other two tests provide less evidence against the null. The results of the panel tests provide strong evidence that relative prices and relative productivities are, in fact, cointegrated. We reject the null of no cointegration at the 5 percent level with all three tests, but when common time dummies are included only the t-tests yield rejections.

Because there is some evidence that deterministic trends might be in the data for some of the countries, we also carry out the panel cointegration tests including a time trend in the cointegrating regression. The evidence from these tests is not as strong, but, because only some of the series may have deterministic trends and because the critical values with deterministic trends are higher than those without deterministic trends, the tests are probably overly

conservative. The evidence is weakest when we include common time dummies. When we exclude the 3 countries (Canada, Great Britain, and Austria) for which the data may not fit the stochastic assumptions that the test rely on, we reject at the 5% level for one of the three tests.

We conclude from table 5 that the relative prices of nontraded goods and the relative productivities in the traded and nontraded goods sectors are cointegrated as the Balassa-Samuelson model predicts. In tables 6 and 7 we turn to the stronger predictions of the Balassa-Samuelson model that the slope in the cointegrating relationship is 1.0. In table 6 we test this implication of the model by testing the null hypothesis that there is a unit root in the difference between the (log) relative price and the (log) relative productivities. Once again the evidence is mixed when we look at the tests carried out on the data for each country individually. We reject the null hypothesis of a unit root in the difference at the 5 percent level for 5 of the 13 countries and at the 10 percent level for one additional country. Once again, however, the gain in power from using joint tests in a panel is clear. We reject the null hypothesis of a unit root at the 5 percent level for all three groups of countries that we consider. The evidence from the tests without a time trend is clearly consistent with the hypothesis that the slope of the cointegrating relationship is 1.0, or, alternatively, that the relative price of nontraded goods is proportional to the relative productivities in the traded and nontraded goods sectors.

Including a time trend again weakens the evidence against the null hypothesis of a unit root. When time common dummies are included, we reject the null hypothesis of a unit root in the difference only when we consider all countries together. There are several interpretations for this finding. It may be that the tests are overly conservative because only some of the series may have deterministic trends. It may be that relative prices and relative productivities are

proportional in the long run for only some of the countries. Or it may be that the added power from using the full sample is responsible for the rejection when the full set of countries is used.

The evidence in table 6 is generally consistent with the long run proportionality of relative prices and relative productivities. An alternative means of testing whether the slope of the cointegrating relationship is equal to 1.0 is to estimate the slope directly and do a t-test on the estimated slope. We report the results of these tests in table 7. With three exceptions, Canada, Denmark, and Sweden, the slope coefficients are generally close to 1.0. The average fully modified OLS slope estimate is roughly 0.8, and excluding the three countries with the slope around 0.5, the average is nearly 0.9. The slopes are fairly precisely estimated, however, and we can reject the null hypothesis that the slope is 1.0 at the 5 percent level for 9 of the 13 countries. The panel tests confirm these results. When we include common time dummies we obtain a t-ratio for the null that  $\beta = 1.0$  of around -8.0 when we use all of the countries in the sample or when we exclude the three countries for which the series are most likely to violate the assumptions about the stochastic trends.

The results of the formal tests are reflected in figure 2 where we plot the (log of the) relative price of nontraded goods and the (log of the) relative labor productivities in traded and nontraded goods for four countries, Germany, Japan, Italy, and the United States. The two series are normalized so that 1970 = 0 for each country. The differences between the two series generally appear to be transitory although the data for the United States and Italy exhibit a tendency common to many of the countries we examine for the relative productivities to grow by more than the relative price. This tendency is reflected in the slope estimates, which are generally less than one.

## 4.2 Monte Carlo Experiments

Tables 6 and 7 contain contradictory results. The results in table 6 are consistent with a slope of 1.0 while the result from table 7 are not. One possible explanation for the difference is the small sample properties of the fully modified OLS estimators. Phillips and Hansen (1990) find that the FMOLS estimates can exhibit considerable small sample bias when  $\epsilon$  and  $\zeta$  (from (12)) are persistent and when they are highly correlated. In order to explore whether small sample problems might be responsible for the differences between the results in tables 6 and 7, we conduct a series of Monte Carlo experiments. First we verify that the FMOLS estimates perform well even in samples of 25 observations when  $\epsilon$  and  $\zeta$  are independent and identically distributed. Then we choose three countries: one (the United States) with a slope estimate of around 0.9, one (Belgium) with a slope estimate of around 0.8, and one (Denmark) with a slope estimate around 0.6. For each of the countries we estimate a VAR for  $\eta=(\epsilon, \zeta)'$  and use the parameters and error covariance matrix to generate 15,000 samples of 25 observations in which the true value of  $\beta$  is set to 1.0. We also examine the small sample properties of the pooled tests statistics from the corresponding 1,000 samples of 15 countries and 25 observations.

In all three cases we find that the FMOLS estimate of the slope is biased and that using the asymptotic distribution would lead to rejecting the null hypothesis of  $\beta=1.0$  too frequently. The mean FMOLS estimates of  $\beta$  are: 0.90 when we use the U.S. parameters, 0.77 when we use the Belgian parameters, and 0.86 when we use the Danish parameters. Thus for two of the three we find that our point estimate from table 7 is roughly equal to the mean value estimated from the Monte Carlos in which the true value of  $\beta$  is unity. The point estimate in table 7 computed from the Danish data (0.54) is below the average for the third and is significantly different from 1.0 at

the 5 percent level even using the empirical distribution from the Monte Carlo experiments.

The tendency for the FMOLS estimates to reject the null of  $\beta = 1.0$  too frequently and the downward bias in the estimate of  $\beta$  is also reflected in the panel test statistics. As is the case with the individual country tests, the panel tests (especially the group mean test) reject too frequently. When the U.S. parameters are used to generate the data and all countries are used, the 95 percent critical value from the Monte Carlos is below the value reported in table 7 for all countries. The critical values computed from the Monte Carlos using the other two sets of parameters all exceed the test statistics reported in table 7.

The Monte Carlo evidence suggest that one would tend to find slope estimates below 1.0 (relative productivity changes outstripping relative price changes) even if the true slope was, in fact, 1.0. Thus the test statistics reported in table 7 should probably not be taken as evidence against the null hypothesis that  $\beta$  is one.<sup>12</sup>

Figure 2, the results in tables 5, 6 and 7, and the Monte Carlo experiments taken together lend support to the first part of the Balassa-Samuelson hypothesis. The relative prices of nontraded goods and the relative productivities in the traded and nontraded goods sectors appear to be cointegrated and the slope of the cointegrating relationship is close to 1.0 as the hypothesis predicts. Thus relative prices and relative productivities appear to be proportional in the long run.

#### 4.3 Purchasing Power in Traded Goods: U.S. Dollar as the Numeraire Currency

Next we turn to testing long run purchasing power parity in traded goods using the U.S. dollar as the numeraire currency. Table 8 presents the results of tests of the hypothesis that the

---

<sup>12</sup> Pedroni (1996) presents more extensive evidence on the small sample properties of the panel test statistics.

nominal exchange rate and the PPP exchange rate are not cointegrated. The tests carried out on the data from each individual country yield little evidence against the null of no cointegration. The ADF tests reject the null at the 5 percent level for only two countries, Great Britain and Finland. We find no rejections with either of the other two tests. The benefit of pooling the data in the panel tests is, once again, clearly apparent. The panel tests provide evidence that the nominal and PPP exchange rates are cointegrated. When we include common time dummies, all three test rejects the null of no cointegration at the 5 percent level.

The results in table 8 suggest that nominal exchange rates and PPP exchange rates are probably cointegrated when the U.S. dollar is the numeraire currency. If purchasing power parity holds for traded goods, they should be cointegrated with a cointegrating slope of 1.0. Next we test this stronger restriction of purchasing power parity and report the results in tables 9 and 10. If purchasing power parity holds in the long run for traded goods, the difference between the nominal exchange rate and the PPP exchange rate should be stationary. In table 9 we report the results of tests of the null hypothesis that there is a unit root in the difference. The tests using the data from each country individually provide no evidence against the null hypothesis of a unit root. The panel test points to a rejection of the null hypothesis when common time dummies are used.<sup>13</sup> Otherwise the panel test yield no evidence against the null. Although nominal exchange rates and PPP exchange rates are cointegrated, the difference between them appears to be nonstationary.

---

<sup>13</sup> This rejection might appear surprising given the failure to reject the null of no cointegration on the nominal and PPP exchange rates with the other tests. The reason for this rejection becomes clearer in the next section when we change numeraire currencies. Because changing the numeraire amounts to adding the same log exchange rate to each cross-sectional unit, it will have no impact on tests computed using deviations from cross-section means, at least for large N. As can be seen by comparing the results in tables 8 and 12, with N=12, a change of numeraire results in only minimal changes to the test statistics. See O'Connell (1996).

Contrary to the predictions of purchasing power parity in traded goods, the slope of the cointegrating relationship appears not to be 1.0.

In table 10 we present estimates of the slopes of the cointegrating relationships along with tests of the hypothesis that the slope is 1.0. Consistent with the results in table 9, we find that we reject the null hypothesis that the slope is 1.0 for 10 of the 12 countries and the panel test strongly reject the null hypothesis that the slopes are jointly equal to one. We also found a large number of rejections of the null when we looked at the cointegrating slopes for relative prices and relative productivities, but the reason for the large number of rejections is quite different here. There the rejections arise because the slope estimates, which are fairly close to 1.0, are precisely estimated. In contrast, here the point estimates are all far from 1.0 and we reject the null despite fairly large standard errors. The wide range of point estimates that we report in table 10 is roughly consistent with previous work that has looked at purchasing power parity. In their survey, Froot and Rogoff (1995) report that it is common in the literature to find slopes that vary widely across countries and are frequently far from the value of 1.0 that is implied by purchasing power parity.

#### 4.4 Purchasing Power Parity in Traded Goods: DM as the Numeraire Currency

Do the results on purchasing power parity in traded goods depend on the choice of the U.S. dollar as the numeraire currency?<sup>14</sup> In order to shed some light on this question, we now reconsider our tests using the DM as the numeraire currency. In table 11 we present the results of

---

<sup>14</sup> Two pieces of existing evidence suggest that it might. Papell (1995) tests the null hypothesis of a unit root in real exchange rates computed with consumer prices and finds stronger rejections when the DM is used as the reference currency than when the dollar is used as the reference currency. Similarly, Jorion and Sweeney (1994) find stronger evidence that real exchange rates are mean reverting when the DM is used as the numeraire. On the other hand, Wei and Parsley (1995) use prices in 12 tradable sectors and find that rates of convergence to purchasing power parity are similar when the DM and the U.S. dollar are used as the numeraire.

tests of the null hypothesis that nominal exchange rates and PPP exchange rates are not cointegrated.<sup>15</sup> The tests carried out on individual country data reject the null hypothesis only for the three Scandinavian countries. The results of the joint tests are somewhat similar to those found using the U.S. dollar as the numeraire. The added power derived from using a panel is again apparent and the null of no cointegration is strongly rejected.

The results in table 11 suggest that nominal and PPP exchange rates relative to the DM (like those relative to the dollar) are cointegrated. Next we turn to tests of the prediction that the cointegrating slope coefficient is one. Again, the first test imposes the null hypothesis that the slope is 1.0 and examines whether the difference between nominal and PPP exchange rates is stationary. We present the results of these unit root tests in Table 12 and find fairly strong evidence against the null hypothesis of a unit root in the difference. We reject the null hypothesis with at least two of the three tests for five of the twelve countries taken individually and with one of the tests for a sixth country. The joint tests yield rejections at the 95% level or greater.

Next we estimate the cointegrating slope and report the results in table 13. These results provide an interesting contrast to those obtained using the U.S. dollar as the numeraire currency. Although we reject the null hypothesis that the slope is 1.0 for five of the twelve currencies, the point estimates are fairly close to 1.0, and are much closer to 1.0 than are those obtained with a

---

<sup>15</sup> We also test for unit roots in nominal exchange rates and PPP exchange rates (both relative to the DM) but do not report the results. We find no evidence against the null of a unit root in nominal exchange rates and only for Belgium do we find strong evidence against the null of a unit root in PPP exchange rates. We also reject the null of a unit root in the PPP exchange rates for Sweden and Austria using one of the three tests. Finally, we do not find deterministic trends in either the nominal or PPP exchange rates.



dollar numeraire.<sup>16</sup> Even for the currencies for which we reject the null hypothesis of a unit slope, the point estimates are closer to one than are any of the slopes we obtain using the dollar as the numeraire.

The results we report in tables 11 - 13 lend somewhat favorable support for the hypothesis of purchasing power parity in traded goods when the DM is the numeraire currency. Nominal exchange rates and PPP exchange rates appear to be cointegrated and, although we can frequently reject that the two are proportional in the long run, our estimates suggest that they are nearly proportional for the currencies we consider. Interestingly, the evidence does not indicate that purchasing power parity in traded goods is any more likely to hold for European currencies than for non-European currencies relative to the DM.

In figure 3 we plot the (logs of the) nominal and PPP exchange rates for the DM and yen relative to the U.S. dollar and the lira and yen relative to the DM. Again, each series is normalized so that 1970 = 0. Large and long-lasting deviations from PPP in traded goods are evident for both exchange rates relative to the U.S. dollar. As is consistent with the formal tests, these deviations are smaller and less persistent when we examine the two exchange rates relative to the DM.

## 5. Concluding Remarks

How well does the Balassa-Samuelson model explain the behavior of real exchange rates?

---

<sup>16</sup> The one notable exception is the Belgian franc regression where we obtain a negative point estimate and a large standard error. Recall that we rejected the null hypothesis of a unit root in the Belgian franc - DM PPP exchange rate but not for the nominal exchange rate. The two therefore cannot be cointegrated so that the estimate is not particularly bothersome. The slope for the U.S. dollar - DM exchange rate is, of course, identical to that reported in table 10.

The evidence from a panel of 13 OECD countries supports the hypothesis that the relative price of nontraded goods reflects the relative labor productivities in the traded and nontraded goods sectors. The results suggest that the relative prices of nontraded goods and the relative productivities in the traded and nontraded goods sectors are cointegrated and that the slope of the cointegrating relationship is generally close to 1.0. For some countries, growth in relative productivities appears to outstrip growth in relative prices (or equivalently the slope is less than 1.0), but Monte Carlo evidence suggests that this difference is within the range of sampling variation in a sample like ours. Thus relative prices and relative productivities appear to be proportional in the long run.

But the Balassa-Samuelson model also assumes that traded goods prices are characterized by purchasing power parity. As is consistent with the evidence presented in Engel (1995), we find large and long-lived deviations from PPP in traded goods when we look at U.S. dollar exchange rates. Although nominal exchange rates and PPP exchange rates appear to be cointegrated we find that the slopes of the cointegrating regressions differ vary widely and differ substantially from one. But when we examine DM exchange rates the evidence is considerably more favorable to purchasing power parity in traded goods. Nominal and PPP exchange rates appear to be cointegrated and the slopes of the cointegrating regressions are much closer to one than is the case with U.S. dollar exchange rates. This evidence suggests that Engel's results may well be sensitive to the choice of numeraire currency. Interestingly, apart from the U.S. dollar - DM exchange rates, purchasing power parity does not appear to be a better characterization of DM exchange rates relative to European currencies than relative to non-European currencies.

## References

- Asea, Patrick K. and Enrique G. Mendoza (1994), "The Balassa-Samuelson Model: A General Equilibrium Appraisal," Review of International Economics 2, pp. 244-267.
- Asea, Patrick K and Max Corden (1994), "The Balassa-Samuelson Model: An Overview," Review of International Economics 2, pp.
- Balassa, Bela (1964), "The Purchasing Power Parity Doctrine: A Reappraisal," Journal of Political Economy 72, pp. 584-596.
- Chinn, Menzie David (1995), "Whither the Yen? Implications of an Intertemporal Model of the Yen/Dollar Rate," working paper, Department of Economics, University of California, Santa Cruz, October 1995.
- Chinn, Menzie David, and Louis Johnston (1996), "Real Exchange Rate levels, Productivity and Demand Shocks: Evidence from a Panel of 14 Countries," working paper, Department of Economics, University of California, Santa Cruz, January 1996.
- De Gregorio, Jose, Alberto Giovannini, and Thomas Krueger (1993), "The Behavior of Nontradable Goods Prices in Europe: Evidence and Interpretation," working paper, International Monetary Fund, April 1993.
- De Gregorio, Jose, Alberto Giovannini, and Holger Wolf (1993), "International Evidence on Tradables and Nontradables Inflation," working paper, International Monetary Fund, September 1993.
- Dickey, David A. And Wayne A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Association 74, pp. 427-431.
- Engel, Robert F. and Clive W. J. Granger (1987), "Co-integration and Error Correction: Representation, Estimation, and Testing," Econometrica 55, pp. 251-276.
- Engle, Charles (1995), "Accounting for U.S. Real Exchange Rate Changes," NBER working paper no. 5394, December 1995.
- Evans, Charles (1992), "Productivity Shocks and Real Business Cycles," Journal of Monetary Economics 29, pp. 191-208.
- Froot, Kenneth, and Kenneth Rogoff (1991), "The EMS, the EMU and the Transition to a

Common Currency," NBER Macroeconomics Annual 6, pp. 269-317.

Froot, Kenneth, and Kenneth Rogoff (1995), "Perspectives on PPP and Long-Run Real Exchange Rates," in Gene Grossman and Kenneth Rogoff (eds.) Handbook of International Economics, Volume 3, North Holland.

Hall, Alastair (1994), "Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection," Journal of Business and Economic Statistics 12, pp. 461-470.

Im, Kyungso, M. Hashem Pesaran, and Yongcheol Shin (1995), "Testing for Unit Roots in Heterogeneous Panels," working paper, Department of Applied Economics, University of Cambridge, June 1995.

Jorion, Philippe and Richard D. Sweeney (1994), "Mean Reversion in Real Exchange Rates: Evidence and Implications for Forecasting," working paper, Georgetown University.

Marston, Richard (1987), "Real Exchange Rates and Productivity Growth in the United States and Japan," in Arndt, Sven and J. David Richardson (eds), Real-Financial Linkages among Open Economies, MIT Press, 1987.

Mankiw, Gregory (1989), "Real Business Cycles: A New Keynesian Perspective," Journal of Economic Perspectives 3 pp. 79-90.

Newey, Whitney K. And Kenneth D. West (1994), "Automatic Lag Selection in Covariance Matrix Estimation," Review of Economic Studies 61, pp. 631-653.

Ng, Serena and Pierre Perron (1995), "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag," Journal of the American Statistical Association 90, pp.268-281.

O'Connell, Paul (1996), "The Overvaluation of Purchasing Power Parity," working paper, Harvard University, April 1996.

Papell, David H. (1995), "Searching for Stationarity: Purchasing Power Parity Under the Current Float," working paper, Department of Economics, University of Houston, October 1995.

Pedroni, Peter (1995), "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with and Application to the PPP Hypothesis," working paper, Department of Economics, Indiana University, June 1995.

Pedroni, Peter (1996), "Fully Modified OLS for Heterogeneous Panels," working paper, Department of Economics, Indiana University, March 1996.

Phillips, Peter C.B. and Bruce E. Hansen (1990), "Statistical Inference in Instrumental Variables Regressions with I(1) Processes," Review of Economic Studies 57, pp. 99-125.

Phillips, Peter C.B. and Sam Ouliaris (1990), "Asymptotic Properties of Residual Based Tests for Cointegration," Econometrica 58, pp. 165-193.

Phillips, Peter C. B. and Pierre Perron (1988), "Testing for a Unit Root in Time Series Regression," Biometrika 75, pp. 335-346.

Samuelson, Paul (1964), "Theoretical Notes on Trade Problems," Review of Economics and Statistics 46, pp. 145-154.

Summers, Lawrence (1986), "Some Skeptical Observations on Real Business Cycle Theory," Federal Reserve Bank of Minneapolis Quarterly Review 10, pp. 23-27.

Wei, Shang-Jin and David C. Parsley (1995), "Purchasing Power Dis-Parity During the Floating Rate Period: Exchange Rate Volatility, Trade Barriers, and Other Culprits," working paper, Kennedy School of Government, March 1995.

## **Appendix: THE DATA**

Nominal exchange rates are from International Financial Statistics. Sectoral price and productivity data for Belgium, Canada, Denmark, Great Britain, Finland, France, Germany, Italy, Japan, Sweden, and the US come from the OECD National Accounts. Price and productivity data for Austria and Spain come from national statistics. (Francisco de Castro of the Bank of Spain collected and documented the data.) These sources provide annual sectoral data on nominal value added, real value added and number of employees. The traded sector is made up of the "manufacturing" sector and the "agriculture, hunting forestry and fishing" sector. The non-traded sector is made up of the "wholesale and retail trade, restaurants and hotels" sector, the "transport, storage and communication" sector, the "finance, insurance, real estate and business services" sector, the "community social and personal services" sector, and the "non-market services" sector; the "non-market services" sector is made up of the "producers of government services" subsector and the "other producers" subsector.

Data consistency is always an issue. We are aware of the following anomalies in the OECD data: (1) The German market services employment figures do not include the "real estate and business services" sector; value added figures do. (2) The Italian, British and Belgian value added and employment figures do not include "real estate and business services" sector. (3) British value added and employment figures consist of the "producers of government services" sector; they do not include the "other producers" sector.

Table 1: Tests for Unit Roots in  $\ln(q_{i,t})$

$$\Delta \ln(q_{i,t}) = \theta_i + (\rho_i - 1) \ln(q_{i,t-1}) + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \ln(q_{i,t-j}) + v_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_{\hat{\rho}}(PP)$	$t_{\hat{\rho}}(ADF)$	$k_i$
United States	0.558	0.829	0.801	0
Canada	-7.023	-1.669	-0.942	2
Japan	-0.785	-1.478	-1.410	0
Germany	-1.931	** -3.116	-2.515	1
France	-1.180	-1.408	-1.345	4
Italy	0.587	1.220	1.166	0
Great Britain	-2.234	-0.949	-1.757	1
Belgium	-1.326	-2.406	-2.283	0
Denmark	-2.760	** -3.992	** -4.178	3
Sweden	-0.096	-0.043	-0.041	0
Finland	1.919	0.847	0.986	2
Austria	-0.683	-1.321	-1.259	0
Spain	-1.897	-1.264	-0.563	4
All Countries			1.858	
G-7 Countries			1.836	

Panel Unit Root Test with Common Time Dummies

All Countries	0.060
G-7 Countries	0.064

Joint Tests of Unit Roots Allowing for Deterministic Trends

All Countries	-0.808
G-7 Countries	-0.746

Panel Unit Root Test with Trends and Common Time Dummies

All Countries	1.587
G-7 Countries	1.020

Notes: Statistical significance at the 95% level or greater is signified by \*\*. Significance at the 90% level or greater is signified by \*. The  $T(\hat{\rho}-1)$  and  $t_{\hat{\rho}}(PP)$  statistics are tests proposed by Phillips and Perron and are computed from regressions setting  $k_i=0$ . The  $t_{\hat{\rho}}(ADF)$  is the augmented Dickey-Fuller statistic. The joint tests are the panel unit root tests proposed by Im, Peseran, and Shin (1995) and are distributed as standard normal.

Table 2: Tests for Unit Roots in  $\ln(x_{i,t}/n_{i,t})$

$$\Delta \ln(x_{i,t}/n_{i,t}) = \theta_i + (\rho_i - 1) \ln(x_{i,t}/n_{i,t}) + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \ln(x_{i,t}/n_{i,t}) + v_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_{\rho}^{\wedge}(PP)$	$t_{\rho}^{\wedge}(ADF)$	$k_i$
United States	0.476	0.717	0.692	0
Canada	-1.499	-0.829	0.248	3
Japan	** -13.222	** -3.360	** -3.161	3
Germany	-2.017	-2.482	-2.400	0
France	-0.469	-0.922	-1.061	2
Italy	-0.061	-0.123	-0.883	3
Great Britain	-2.263	-1.555	-1.476	0
Belgium	-0.600	-1.327	-1.259	0
Denmark	-2.705	-2.500	-2.402	0
Sweden	-0.290	-0.274	0.283	4
Finland	1.496	0.760	0.735	0
Austria	0.327	1.045	1.867	4
Spain	-0.569	-1.184	-1.647	4
All Countries			3.574	
G-7 Countries			2.441	

Panel Unit Root Test with Common Time Dummies

All Countries	-1.304
G-7 Countries	-0.430

Joint Tests of Unit Roots Allowing for Deterministic Trends

All Countries	0.484
G-7 Countries	0.131

Panel Unit Root Test with Trends and Common Time Dummies

All Countries	-0.841
G-7 Countries	-1.603

Notes: Statistical significance at the 95% level or greater is signified by \*\*. Significance at the 90% level or greater is signified by \*. The  $T(\hat{\rho}-1)$  and  $t_{\rho}^{\wedge}(PP)$  statistics are tests proposed by Phillips and Perron and are computed from regressions setting  $k_i=0$ . The  $t_{\rho}^{\wedge}(ADF)$  is the augmented Dickey-Fuller statistic. The joint tests are the panel unit root tests proposed by Im, Peseran, and Shin (1995) and are distributed as standard normal.



Table 3: Tests for Unit Roots in  $\ln(E_{i,t})$

$$\Delta \ln(E_{i,t}) = \theta_i + (\rho_i - 1) \ln(E_{i,t-1}) + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \ln(E_{i,t-j}) + v_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_{\hat{\rho}}(PP)$	$t_{\hat{\rho}}(ADF)$	$k_i$
Canada	-3.093	-1.316	-1.529	1
Japan	-1.574	-0.853	-0.811	0
Germany	-1.500	-0.738	-0.993	1
France	-7.290	-1.290	-2.690	1
Italy	-1.859	-0.990	-1.303	1
Great Britain	-3.865	-1.590	-1.313	4
Belgium	-5.873	-1.618	-2.352	1
Denmark	-6.344	-1.793	-2.468	1
Sweden	-3.504	-1.359	-1.702	1
Finland	-5.492	-1.895	-1.833	0
Austria	-1.329	-0.610	-1.009	1
Spain	-2.999	-1.342	-1.830	1
All Countries			-0.531	
G-7 Countries			0.152	

Panel Unit Root Test with Common Time Dummies

All Countries	0.284
G-7 Countries	1.000

Joint Tests of Unit Roots Allowing for Deterministic Trends

All Countries	-2.335
G-7 Countries	-1.628

Panel Unit Root Test with Trends and Common Time Dummies

All Countries	0.635
G-7 Countries	-1.187

Notes: Statistical significance at the 95% level or greater is signified by \*\*. Significance at the 90% level or greater is signified by \*. The  $T(\hat{\rho}-1)$  and  $t_{\hat{\rho}}(PP)$  statistics are tests proposed by Phillips and Perron and are computed from regressions setting  $k_i=0$ . The  $t_{\hat{\rho}}(ADF)$  is the augmented Dickey-Fuller statistic. The joint tests are the panel unit root tests proposed by Im, Peseran, and Shin (1995) and are distributed as standard normal.

Table 4: Tests for Unit Roots in  $\ln(r_{i,t})$

$$\Delta \ln(r_{i,t}) = \theta_i + (\rho_i - 1) \ln(r_{i,t}) + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \ln(r_{i,t}) + v_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_{\hat{\rho}}(\text{PP})$	$t_{\hat{\rho}}(\text{ADF})$	$k_i$
Canada	-0.310	-0.215	0.091	2
Japan	-0.061	-0.037	0.42	0
Germany	-2.210	-1.022	-1.325	1
France	0.192	0.159	-0.423	0
Italy	0.193	0.275	-0.175	2
Great Britain	0.102	0.110	-0.201	1
Belgium	-2.140	-0.971	-0.797	0
Denmark	-0.497	-0.542	-0.520	0
Sweden	-0.342	-0.382	-0.274	0
Finland	-1.094	-1.503	-1.353	2
Austria	-2.227	-0.988	-1.583	2
Spain	0.280	0.438	-0.265	1
All Countries			3.566	
G-7 Countries			3.177	

Panel Unit Root Test with Common Time Dummies

All Countries	0.475
G-7 Countries	1.939

Joint Tests of Unit Roots Allowing for Deterministic Trends

All Countries	-0.425
G-7 Countries	-0.010

Panel Unit Root Test with Trends and Common Time Dummies

All Countries	-1.065
G-7 Countries	1.037

Notes: Statistical significance at the 95% level or greater is signified by \*\*. Significance at the 90% level or greater is signified by \*. The  $T(\hat{\rho}-1)$  and  $t_{\hat{\rho}}(\text{PP})$  statistics are tests proposed by Phillips and Perron and are computed from regressions setting  $k_i=0$ . The  $t_{\hat{\rho}}(\text{ADF})$  is the augmented Dickey-Fuller statistic. The joint tests are the panel unit root tests proposed by Im, Peseran, and Shim (1995) and are distributed as standard normal.

Table 5: Tests for Cointegration of  $\ln(q_{i,t})$  and  $\ln(x_{i,t}/n_{i,t})$

$$\ln(q_{i,t}) = \alpha_i + \beta_i \ln(x_{i,t}/n_{i,t}) + \epsilon_{i,t}$$

$$\Delta \hat{\epsilon}_{i,t} = (\rho_i - 1) \hat{\epsilon}_{i,t-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \hat{\epsilon}_{i,t-j} + v_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_{\hat{\rho}}(PP)$	$t_{\hat{\rho}}(ADF)$	$k_i$
United States	*-18.106	** -3.483	** -3.367	4
Canada	-11.921	-2.856	** -5.446	2
Japan	-8.643	-3.052	** -3.521	3
Germany	-16.084	* -3.200	** -3.946	3
France	-8.976	-2.425	-2.642	3
Italy	-5.798	-1.650	-2.949	3
Great Britain	-8.277	-2.029	-1.732	0
Belgium	-12.641	* -3.184	-3.025	0
Denmark	-6.546	-1.894	-1.762	2
Sweden	-6.851	-1.867	** -4.539	5
Finland	* -18.637	** -3.590	** -5.593	1
Austria	-11.208	* -3.172	-3.021	0
Spain	-9.809	-2.494	* -3.356	3
All Countries	** -33.906	** -9.421	** -14.606	
No CA, UK, AU	** -29.077	** -8.172	** -13.636	
G-7 Countries	** -24.841	** -7.220	** -11.621	

Panel Tests of Cointegration with Time Dummies

All Countries	-23.707	-7.494	** -8.889
No CA, UK, AU	-18.177	** -7.811	** -9.881
G-7 Countries	-14.457	** -7.074	** -7.586

Joint Tests of Cointegration Allowing for Deterministic Trends

All Countries	-36.935	-9.235	** -12.132
No CA, UK, AU	-33.096	-8.261	** -10.981
G-7 Countries	-25.684	-6.614	** -8.488

Panel Tests of Cointegration with Trends and Time Dummies

All Countries	-31.293	-7.820	-8.996
No CA, UK, AU	-32.994	-8.306	** -10.878
G-7 Countries	-25.380	-5.914	-7.214

Notes: Significance at the 95% and 90% levels are noted by \*\* and \*, respectively. The  $T(\hat{\rho}-1)$  and  $t_{\hat{\rho}}(PP)$  statistics are Phillips-Perron tests applied to the residuals from the first regression and are computed with  $k_i=0$ . The  $t_{\hat{\rho}}(ADF)$  is the augmented Dickey-Fuller statistic. Critical values are taken from the tables compiled by Phillips and Ouliaris (1990). The panel cointegration tests are those proposed by Pedroni (1995), which is the source of the critical values.

Table 6: Tests for Unit Roots in  $\delta_{i,t} = \ln(q_{i,t}) - \ln(x_{i,t} / n_{i,t})$

$$\Delta \delta_{i,t} = \theta_i + (\rho_i - 1) \delta_{i,t-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \delta_{i,t-j} + \eta_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_p^{\wedge}(\text{PP})$	$t_p^{\wedge}(\text{ADF})$	$k_i$
United States	-9.458	-2.367	-2.287	0
Canada	-6.858	-2.437	-2.312	0
Japan	-7.996	** -3.307	** -4.210	3
Germany	** -15.327	** -3.075	** -3.208	1
France	-4.100	-1.411	-0.451	2
Italy	-4.938	-1.775	** -3.060	3
Great Britain	-6.467	-2.139	-2.030	0
Belgium	0.046	0.026	1.238	5
Denmark	-2.850	-1.650	-1.540	0
Sweden	-6.361	-2.079	-2.292	1
Finland	** -21.268	** -3.811	** -5.322	1
Austria	* -10.391	* -2.841	* -2.709	0
Spain	-8.053	-2.100	** -3.819	3
All Countries			** -3.721	
Excluding Canada, Great Britain, Austria			** -3.473	
G-7 Countries			** -2.638	

Panel Unit Root Test with Common Time Dummies

All Countries	** -4.114
Excluding Canada, Great Britain, Austria	** -3.703
G-7 Countries	** -2.040

Joint Tests of Unit Roots Allowing for Deterministic Trends

All Countries	** -4.164
Excluding Canada, Great Britain, Austria	** -4.261
G-7 Countries	** -3.281

Panel Unit Root Test with Trends and Common Time Dummies

All Countries	** -1.977
Excluding Canada, Great Britain, Austria	-0.403
G-7 Countries	-0.512

Notes: Significance at the 95% and 90% levels is noted by \*\* and\*, respectively. The  $T(\hat{\rho}-1)$  and  $t_p^{\wedge}(\text{PP})$  statistics are tests proposed by Phillips and Perron and are computed setting  $k_i=0$ . The  $t_p^{\wedge}(\text{ADF})$  is the augmented Dickey-Fuller statistic. The joint tests are the panel unit root tests proposed by Im, Peseran, and Shin (1995) and are distributed as standard normal.

Table 7: Estimates of the Cointegrating Slope Coefficient of  $\ln(q_{i,t})$  and  $\ln(x_{i,t}/n_{i,t})$

$$\ln(q_{i,t}) = \alpha_i + \beta_i \ln(x_{i,t}/n_{i,t}) + \epsilon_{i,t}$$

Country	$\hat{\beta}_{i,OLS}$	$\hat{\beta}_{i,FMOLS}$	$t(\hat{\beta}_{i,FMOLS}=1)$	No. Lags
United States	0.887	0.887	** -3.852	3
Canada	0.455	0.442	** -6.243	2
Japan	1.045	1.006	0.092	2
Germany	1.050	1.037	1.013	2
France	0.771	0.755	** -3.791	2
Italy	0.883	0.892	** -2.007	2
Great Britain	0.700	0.728	** -2.082	2
Belgium	0.791	0.785	** -8.369	2
Denmark	0.544	0.512	** -10.029	1
Sweden	0.574	0.537	** -3.929	2
Finland	0.956	0.909	-1.455	5
Austria	0.959	0.929	-1.398	2
Spain	0.895	0.859	** -2.141	3

Panel Tests of  $\hat{\beta} = 1$

	$\hat{t}_{\beta}$	$\sqrt{N} \bar{t}$
All Countries	** -10.295	** -12.256
Excluding CA, UK, AU	** -9.865	** -10.899
G-7 Countries	** -5.203	** -6.376

Panel Tests of  $\hat{\beta} = 1$  with Common Time Dummies

	$\hat{t}_{\beta}$	$\sqrt{N} \bar{t}$
All Countries	** -8.318	** -7.866
Excluding CA, UK, AU	** -7.107	** -8.917
G-7 Countries	** -3.086	** -2.296

Notes: The second column contains the ordinary least squares estimates of the slope coefficient. Column 3 contains the Phillips-Hansen (1990) fully modified OLS estimates of the slope and column 4 contains the t-ratio formed by subtracting one from the fully modified OLS estimate of the slope and dividing by the corresponding standard error. The number of lags used in computing the fully modified OLS estimate and its standard error is reported in column 5 and are chosen using the data dependent procedure proposed by Newey and West (1994). The panel tests are those proposed by Pedroni (1996).

Table 8: Tests for Cointegration of  $\ln(E_{i,t})$  and  $\ln(r_{i,t})$

$$\ln(E_{i,t}) = \alpha_i + \beta_i \ln(r_{i,t}) + \epsilon_{i,t}$$

$$\Delta \hat{\epsilon}_{i,t} = (\rho_i - 1) \hat{\epsilon}_{i,t-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \hat{\epsilon}_{i,t-j} + v_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_{\hat{\rho}}(PP)$	$t_{\hat{\rho}}(ADF)$	$k_i$
Canada	-8.067	-1.927	-2.920	1
Japan	-7.149	-2.401	-3.014	5
Germany	-8.135	-2.043	-2.395	1
France	-9.676	-2.211	-3.319	1
Italy	-9.051	-2.089	-3.120	1
Great Britain	-13.599	-2.757	** -3.488	1
Belgium	-7.328	-1.880	-2.612	1
Denmark	-6.351	-1.790	-2.475	1
Sweden	-6.084	-1.777	-2.505	5
Finland	-9.193	-2.251	** -3.740	5
Austria	-7.189	-1.937	-1.931	0
Spain	-6.858	-1.755	-3.057	1
All Countries	-24.030	-6.268	** -9.415	
G-7 Countries	-18.984	-4.930	** -7.633	

Panel Tests of Cointegration with Time Dummies

All Countries	** -33.515	** -8.495	** -9.696
G-7 Countries	** -26.054	* -5.956	** -6.759

Notes: Statistical significance at the 95% level or greater is signified by \*\*. Significance at the 90% level or greater is signified by \*. The  $T(\hat{\rho}-1)$  and  $t_{\hat{\rho}}(PP)$  statistics are Phillips and Perron tests applied to the residuals from the first regression and are computed from regressions setting  $k_i=0$  in the second regression. The  $t_{\hat{\rho}}(ADF)$  is the augmented Dickey-Fuller statistic. Significance levels for these tests are taken from the tables compiled by Phillips and Ouliaris (1990). The joint tests are the panel cointegration tests proposed by Pedroni (1995). Significance levels for these tests are computed from tables compiled by Pedroni.

Table 9: Tests for Unit Roots in  $\mu_{i,t} = \ln(E_{i,t}) - \ln(r_{i,t})$

$$\Delta \mu_{i,t} = \theta_i + (\rho_i - 1) \mu_{i,t-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \mu_{i,t-j} + \eta_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_p(\hat{\rho})$ (PP)	$t_p(\hat{\rho})$ (ADF)	$k_i$
Canada	-3.163	-0.960	-1.870	1
Japan	-4.754	-1.780	-1.693	0
Germany	-3.427	-1.220	-1.433	1
France	-6.623	-1.653	-2.617	1
Italy	-3.524	-1.105	-1.722	1
Great Britain	-4.046	-1.200	-0.698	4
Belgium	7.866	-1.996	*-2.644	1
Denmark	-3.528	-1.224	-1.825	1
Sweden	-4.965	-1.575	-1.872	1
Finland	-1.261	-1.112	-1.076	0
Austria	-2.292	-0.897	-0.867	0
Spain	-2.442	-0.800	-1.761	1
All Countries			-0.171	
G-7 Countries			0.207	

Panel Unit Root Test with Common Time Dummies

All Countries	** -2.754
G-7 Countries	-0.651

Notes: Statistical significance at the 95% level or greater is signified by \*\*. Significance at the 90% level or greater is signified by \*. The  $T(\hat{\rho}-1)$  and  $t_p(\hat{\rho})$ (PP) statistics are tests proposed by Phillips and Perron and are computed setting  $k_i=0$ . The  $t_p(\hat{\rho})$ (ADF) is the augmented Dickey-Fuller statistic, the  $t$  ratio. The joint tests are the panel unit root tests proposed by Im, Peseran, and Shin (1995) and are distributed as standard normal.

Table 10: Estimates of the Cointegrating Slope Coefficient of  $\ln(E_{i,t})$  and  $\ln(r_{i,t})$

$$\ln(E_{i,t}) = \alpha_i + \beta_i \ln(r_{i,t}) + \epsilon_{i,t}$$

Country	$\hat{\beta}_{i,OLS}$	$\hat{\beta}_{i,FMOLS}$	$t(\hat{\beta}_{i,FMOLS}=1)$	No. Lags
Canada	0.516	0.483	** -4.853	2
Japan	1.934	1.603	1.450	2
Germany	2.463	2.035	1.566	3
France	0.448	0.302	** -3.734	3
Italy	0.688	0.682	** -3.852	3
Great Britain	0.525	0.496	** -5.243	1
Belgium	0.543	0.103	** -2.247	3
Denmark	0.025	0.019	** -4.164	3
Sweden	0.557	0.510	** -2.221	3
Finland	0.264	0.264	** -6.994	3
Austria	2.670	2.235	** 1.999	3
Spain	0.596	0.562	** -3.395	3

Panel Tests of  $\hat{\beta} = 1$

	$\hat{t}_{\beta}$	$\sqrt{N} \bar{t}$
All Countries	** -11.177	** -9.184
G-7 Countries	** -7.459	** -5.987

Panel Tests of  $\hat{\beta} = 1$  with Common Time Dummies

	$\hat{t}_{\beta}$	$\sqrt{N} \bar{t}$
All Countries	** -32.277	** -22.448
G-7 Countries	** 6.458	** 6.148

Notes: The second column contains the ordinary least squares estimates of the slope coefficient. Column 3 contains the Phillips-Hansen (1990) fully modified OLS estimates of the slope and column 4 contains the t-ratio formed by subtracting one from the fully modified OLS estimate of the slope and dividing by the corresponding standard error. The number of lags used in computing the fully modified OLS estimate and its standard error is reported in column 5 and are chosen using the data dependent procedure proposed by Newey and West (1994). The panel tests are those proposed by Pedroni (1996).



Table 11: Tests for Cointegration of  $\ln(E_{i,t})$  and  $\ln(r_{i,t})$ , DM Numeraire

$$\ln(E_{i,t}) = \alpha_i + \beta_i \ln(r_{i,t}) + \epsilon_{i,t}$$

$$\Delta \hat{\epsilon}_{i,t} = (\rho_i - 1) \hat{\epsilon}_{i,t-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \hat{\epsilon}_{i,t-j} + v_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_{\hat{\rho}}(PP)$	$t_{\hat{\rho}}(ADF)$	$k_i$
United States	- 8.135	-2.043	-2.462	1
Canada	-9.193	-2.303	-2.663	4
Japan	-12.932	-2.886	-2.327	5
France	-7.204	-2.285	-2.115	5
Italy	-12.547	-2.768	-2.542	5
Great Britain	-6.471	-2.073	-2.034	0
Belgium	-0.474	-0.462	-0.789	3
Denmark	** -21.346	** -5.183	** -4.296	4
Sweden	* -18.432	** -3.871	** -5.933	1
Finland	* -19.841	** -3.466	** -4.084	1
Austria	-11.256	-2.902	-2.849	0
Spain	-14.197	-2.983	-2.930	0
All Countries	-21.146	-7.146	** -8.663	
European Countries	** -31.187	** -7.962	** -8.874	

Panel Tests of Cointegration with Time Dummies

All Countries	** -32.403	** -8.272	** -10.018
European Countries	** -33.130	** -8.455	** -10.494

Notes: Statistical significance at the 95% level or greater is signified by \*\*. Significance at the 90% level or greater is signified by \*. The  $T(\hat{\rho}-1)$  and  $t_{\hat{\rho}}(PP)$  statistics are Phillips and Perron tests applied to the residuals from the first regression and are computed from regressions setting  $k_i=0$  in the second regression. The  $t_{\hat{\rho}}(ADF)$  is the augmented Dickey-Fuller statistic. Significance levels for these tests are taken from the tables compiled by Phillips and Ouliaris (1990). The joint tests are the panel cointegration tests proposed by Pedroni (1995). Significance levels for these tests are computed from tables compiled by Pedroni.

Table 12: Tests for Unit Roots in  $\mu_{i,t} = \ln(E_{i,t}) - \ln(r_{i,t})$ , DM Numeraire

$$\Delta \mu_{i,t} = \theta_i + (\rho_i - 1) \mu_{i,t-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta \mu_{i,t-j} + \eta_{i,t}$$

Country	$T(\hat{\rho}-1)$	$t_{\hat{\rho}}(PP)$	$t_{\hat{\rho}}(ADF)$	$k_i$
United States	-3.427	-1.220	-1.433	1
Canada	-7.052	-1.890	-2.190	1
Japan	*-10.708	*-2.676	-2.545	0
France	-4.027	-1.826	-1.766	0
Italy	-7.718	-2.221	-2.148	0
Great Britain	-6.255	-2.076	-2.008	0
Belgium	-0.595	-0.475	-0.459	0
Denmark	** -20.501	** -4.304	** -4.128	0
Sweden	-10.110	** -3.043	-0.809	3
Finland	*-11.470	*-2.669	-1.391	3
Austria	*-10.843	*-2.862	*-2.769	0
Spain	** -14.185	** -3.016	*-2.902	0
All Countries			** -2.000	
European Countries			*-1.720	

Panel Unit Root Test with Common Time Dummies

All Countries	** -2.194
European	** -2.872

Notes: Statistical significance at the 95% level or greater is signified by \*\*. Significance at the 90% level or greater is signified by \*. The  $T(\hat{\rho}-1)$  and  $t_{\hat{\rho}}(PP)$  statistics are tests proposed by Phillips and Perron and are computed setting  $k_i=0$ . The  $t_{\hat{\rho}}(ADF)$  is the augmented Dickey-Fuller statistic, the t ratio. The joint tests are the panel unit root tests proposed by Im, Peseran, and Shin (1995) and are distributed as standard normal.

Table 13: Estimates of the Cointegrating Slope Coefficient of  $\ln(E_{i,t})$  and  $\ln(r_{i,t})$ , DM Numeraire

$$\ln(E_{i,t}) = \alpha_i + \beta_i \ln(r_{i,t}) + \epsilon_{i,t}$$

Country	$\hat{\beta}_{i,OLS}$	$\hat{\beta}_{i,FMOLS}$	$t(\hat{\beta}_{i,FMOLS}=1)$	No. Lags
United States	2.463	2.030	1.559	3
Canada	1.454	1.313	1.831	3
Japan	1.311	1.088	0.376	2
France	1.351	1.264	2.928	4
Italy	1.101	1.072	2.907	4
Great Britain	1.075	0.984	-0.155	4
Belgium	-1.064	-0.597	-0.689	4
Denmark	1.058	1.062	2.515	2
Sweden	1.154	1.168	5.107	4
Finland	0.873	0.857	-3.637	2
Austria	1.059	1.021	0.046	4
Spain	1.010	1.000	-0.015	2

Panel Tests of  $\hat{\beta} = 1$

	$\hat{t}_{\beta}$	$\sqrt{N} \bar{t}$
All Countries	**3.360	**2.983
European Countries	**2.780	**2.174

Panel Tests of  $\hat{\beta} = 1$  with Common Time Dummies

	$\hat{t}_{\beta}$	$\sqrt{N} \bar{t}$
All Countries	** -7.625	** -8.533
European Countries	** -6.152	** -7.676

Notes: The second column contains the ordinary least squares estimates of the slope coefficient. Column 3 contains the Phillips-Hansen (1990) fully modified OLS estimates of the slope and column 4 contains the t-ratio formed by subtracting one from the fully modified OLS estimate of the slope and dividing by the corresponding standard error. The number of lags used in computing the fully modified OLS estimate and its standard error is reported in column 5 and are chosen using the data dependent procedure proposed by Newey and West (1994). The panel tests are those proposed by Pedroni (1996).

Figure 1: Real Exchange Rates and Relative Labor Productivities

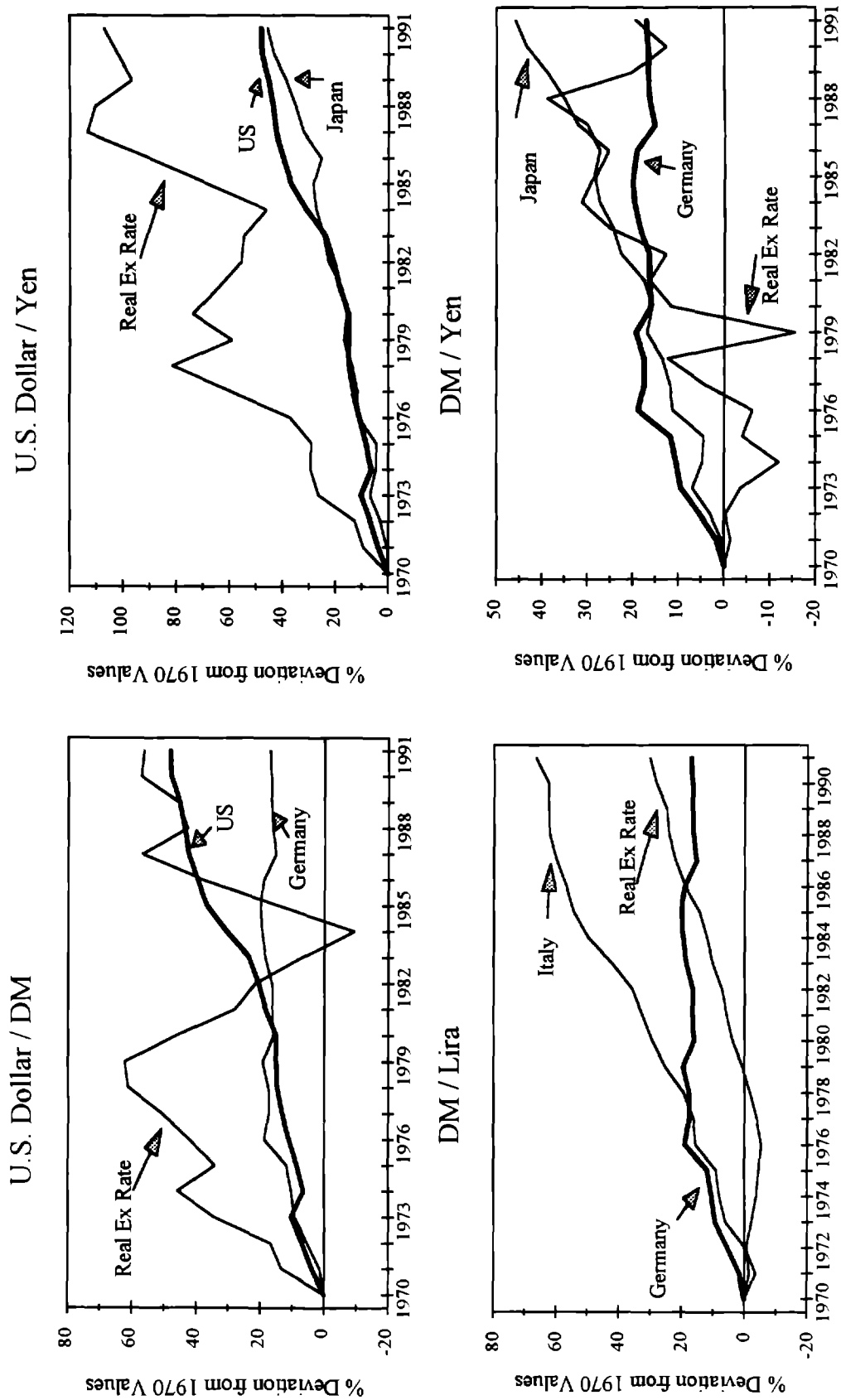


Figure 2: Relative Productivities and Relative Prices

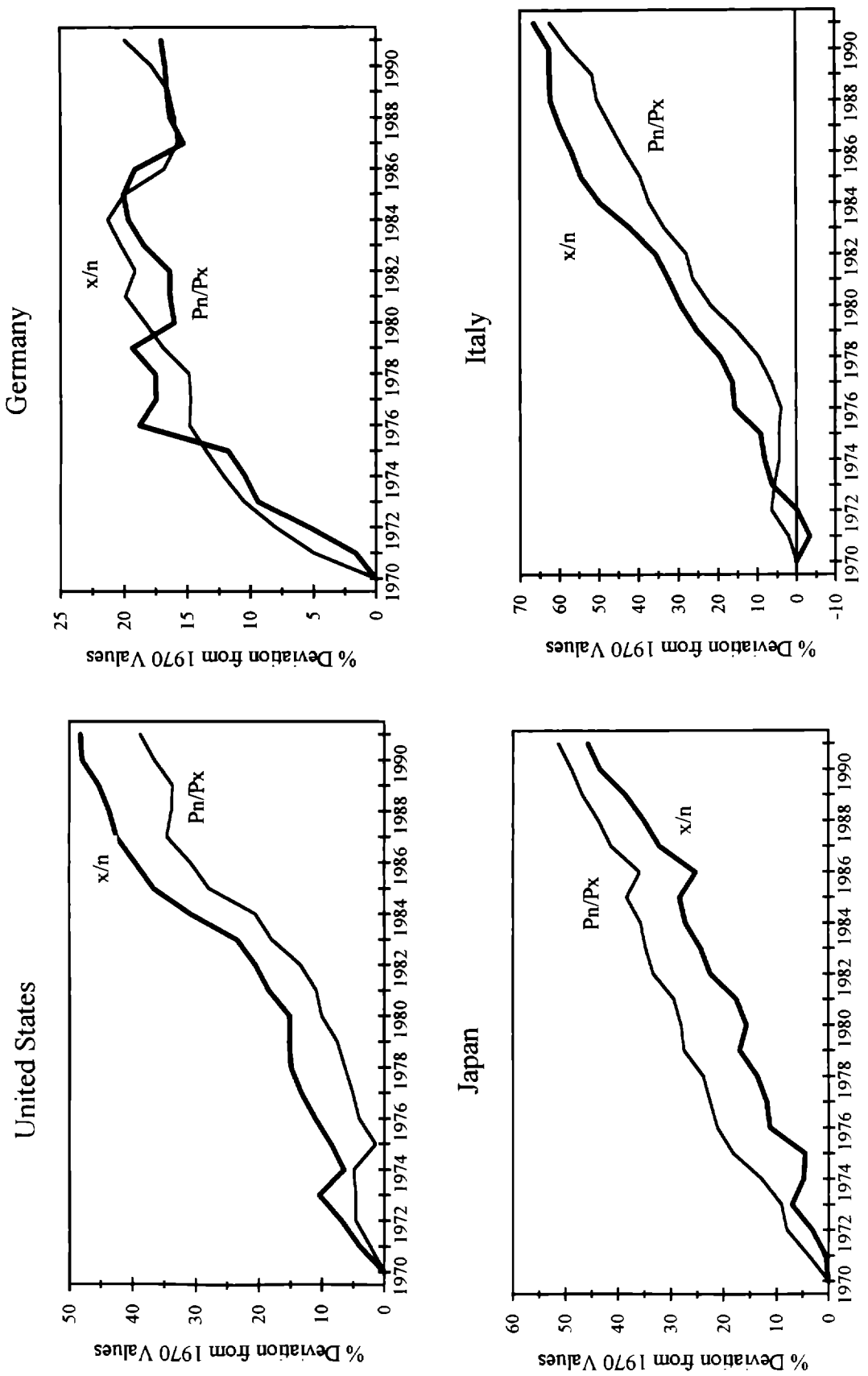


Figure 3: Nominal and Purchasing Power Parity Exchange Rates

