

Panel Evidence of Purchasing Power Parity Using Intranational and International Data*

Sarah E. Culver
Department of Economics
University of Alabama at Birmingham
UAB Station
Birmingham, AL 35294
(205) 934-8879
sculver@uab.edu

and

David H. Papell
Department of Economics
University of Houston
Houston, TX 77204-5882
(713) 743-3807
dpapell@uh.edu

June 1999

We investigate purchasing power parity (PPP) with CPI data for Canadian and United States cities, as well as for European countries. Using panel methods to test for the presence of a unit root, we find much less evidence of PPP with relative prices between cities within the same nation than with real exchange rates between European countries. The rates of price convergence are slower for United States cities than for Canadian cities or for European countries. We conduct a power analysis of the tests, and show that the results are consistent with differences in panel sizes and speeds of adjustment.

* We are grateful to John Rogers for providing us with the Canadian data.

I. INTRODUCTION

Purchasing power parity (PPP) remains an important research topic not only because it forms the basis of many international macroeconomic models but also because researchers have had a difficult time convincingly proving its existence and, when applicable, explaining the slower than expected rates of convergence. This is especially true for the post-Bretton-Woods era of flexible nominal exchange rates. After the first few years of generalized floating, it was obvious that PPP did not hold in the short-run. Price levels do not quickly counteract nominal exchange rate movements. However, whether *long-run* parity holds remains an actively researched topic.¹

Econometric analysis of long-run PPP typically involves conducting unit root tests for the real exchange rate. If the unit root null hypothesis is rejected, then the real exchange rate is mean stationary and any deviations from parity should diminish (even if slowly) over time.² Early investigations of PPP involved conducting Augmented-Dickey-Fuller (ADF) tests on univariate real exchange rates for industrialized countries. When these tests rarely rejected the unit root null, the validity of PPP was brought into question; today, however, those results are attributed mainly to the low power of the ADF tests with short time spans of data.³

One response to these findings is that to better test long-run reversion, longer spans of lower frequency data are needed.⁴ However, those long time series encompass periods in which nominal exchange rate regimes shifted from floating to fixed and back again. Since real exchange rate behavior varies over exchange rate regimes, these studies cannot

¹ Surveys of the PPP literature include Froot and Rogoff (1995) and Rogoff (1996).

² Alternatively, researchers have tested for cointegration between the nominal exchange rate, the domestic price level and the foreign price level. However, rejection of the no cointegration null provides necessary but not sufficient evidence of PPP; symmetry between the respected price levels and proportionality between relative prices and the nominal exchange rate must also hold. Otherwise, cointegration results offer evidence of “weak” PPP.

³ Cheung and Lai (1998) use the DF-GLS test of Elliot, Rothenberg, and Stock (1996) on post-73 real exchange rates for industrialized countries to provide some additional rejections of the unit root null hypothesis. Culver and Papell (1999), using tests where the null hypothesis is stationarity, also provide some additional evidence of PPP.

⁴ Frankel (1986), using 116 years of dollar/pound exchange rate data, was among the first to do this. More recent studies include Lothian and Taylor (1996). Froot and Rogoff (1995) show that, using a 5% Dickey Fuller critical value, 72 years of data would be needed to reject the unit root null hypothesis for a stationary AR(1) process with a PPP deviation half life of 3 years. Lothian and Taylor (1997) conclude

answer the question of whether the unit root null would be rejected with a century of flexible exchange rate data. Further, Engel (1999) shows that with longer spans of data, unit root (and cointegration) tests suffer from a serious size bias, whereby a large and economically significant unit root component in the real exchange rate can go undetected. Hegwood and Papell (1998), by demonstrating the importance of testing for possible structural breaks with longer spans of data, also caution against inferring that rejections of the unit root null constitute evidence of long-run PPP.⁵

An alternative approach has been to pool real exchange rates across countries and use panel econometric procedures. Inspired by Levin, Lin, and Chu (1997) and Im, Pesaran and Shin (1997), the advantage here is that in situations like the recent float, where there is not enough time series variation to produce good power in univariate unit root tests, a relatively small amount of cross-section variation can substantially improve power. Levin, Lin, and Chu (1997) and Bowman (1997) both report very high size adjusted power for panels of the size, time span and half-lives of the post-1973 period.

Despite the improved power, empirical studies of the current float have not provided persuasive evidence of long-run PPP with the U.S. dollar as the numeraire currency. Papell (1997) and Papell and Theodoridis (1998a), with a panel of quarterly data for 21 industrialized countries, cannot reject the unit root null at even the ten percent level of significance. These results are sensitive to the choice of numeraire. Jorion and Sweeney (1996), Papell (1997), and Papell and Theodoridis (1998a), using CPIs, and Wei and Parsley (1995) and Canzoneri, Cumby, and Diba (1999), using tradable goods prices, all report stronger rejections of unit roots in real exchange rates with the German mark, instead of the U.S. dollar, as the numeraire currency. Papell and Theodoridis (1998b) extend the analysis of the numeraire to European versus non-European currencies.

Perhaps another reason for the lack of conclusive evidence relates to international price level data limitations and real-world factors that interfere with relative price convergence. To rule out this possibility, many focus on their attention on disaggregated

from simulation studies that close to 200 years of real exchange rate data is needed to reject the unit root hypothesis for the dollar/pound exchange rate.

⁵ In addition, Hegwood and Papell (1998) report much faster rates of convergence when structural change is taken into account. Reported half lives range between 0.44 and 2.32 years.

and/or *intranational* price data. Disaggregated data avoids potential price stickiness problems that nontradable goods in the price index pose, while intranational price data avoids factors such as trade barriers, differences in market baskets, and other frictions that obstruct goods market arbitrage across borders.⁶ Within country comparisons also have better integrated markets, identical monetary policies, and, by definition, a fixed nominal exchange rate. Therefore, if PPP is a good model of long run international price movements, then it should certainly hold within countries and maintain faster rates of convergence.

Although the above framework for answering the PPP question is appealing, studies using these types of data have not been able to provide strong evidence of PPP. The following studies, all using annual (roughly 1918-1996) U.S. city CPI data from the Bureau of Labor Statistics (BLS), report some interesting results. Chen and Devereux (1997), using ADF and cointegration tests, find little evidence for PPP or weak PPP. Their univariate results are particularly sensitive to the choice of numeraire city. For example, with Atlanta as the base city, they report the greatest number (13 out of 18) of unit root null rejections at the 5% level of significance. At the other extreme is San Francisco as the base city, which never rejects the unit root null at the 5% level. Their reported convergence rates (for those city pairs that reject the unit root null) are similar to international comparisons.⁷ Sonora (1997) and Cecchetti, Mark and Sonora (1998), using Chicago as the numeraire, are able to reject the unit root null with panel methods, but report rates of convergence slower than those found across countries.⁸

We also employ panel econometric methods on intranational price data, but our study differs from those mentioned above in several ways. First, our period of interest is the post-1973 float. As such we use monthly data over the post-Bretton-Woods period to

⁶ The extreme case is testing the law of one price (LOOP) either across borders, such as Engel and Rogers (1996), or with-in borders, such as Parsley and Wei (1996) and O'Connell and Wei (1997). CPI analysis, though, is useful because of its role in formulating expectations and in calculating cost of living adjustments, inflation, and real output.

⁷ For those city pairs that reject the unit root null at the 5%, the half life is 3.91 years; at the 10%, the half life is 3.37 years. As a basis for comparison, studies using longer spans of data report very slow convergence rates back to PPP, with half lives of between 3-5 years. Panel studies using the current float report somewhat shorter half lives.

⁸ For example, half lives tabulated using Levin-Lin panel tests on the full sample range between 7.141-8.153 years for Sonora (1997) and 8.9 years for Cecchetti, Mark and Sonora (1998).

examine long-run PPP. Second, not only do we make use of U. S. city CPI data collected by the Bureau of Labor Statistics (BLS) but also Canadian city CPI data collected by Statistics Canada. As a benchmark we also use European Union real exchange rate data. Since trade barriers between European Union countries have mostly been eliminated and, through the EMS, currency fluctuations reduced for a subset of the countries, it provides an international data set that is as similar as possible to the intranational data sets. We construct two sets of intranational panels, U.S. and Canada, and compare these results against one another and against an international panel for the European Union.

We find much stronger evidence of PPP for the European Union panel than for the U.S. and Canada panels. These results are striking because, *a priori*, we expected to find more evidence of stationarity with intranational than with international data. The U.S. results are not sensitive to the numeraire choice: we are unable to reject the unit root null at even the 10% level for all panels but one, that panel with Detroit as the numeraire city. Choice of the numeraire matters more for the Canadian and European Union panels. We reject the unit root null for three out of nine panels at the 10% level only for the Canadian cities and for eight out of fifteen at the 5% level for the European panel of 14 countries.

Finally, we compare the rates of price convergence in the U.S. relative to Canada and the European Union. We find rates of price convergence for U.S. intranational city pairs to be slower than for Canada and the European Union. The comparison between the U.S. and Europe accords with the evidence from the unit root tests: the stronger rejections of the unit root null for the European Union countries are associated with faster convergence towards PPP. The evidence for Canada is puzzling. While the strength of the unit root rejections is comparable to the U.S., the speed of convergence towards PPP is comparable (and even slightly faster) than within the European Union. We provide some simulation evidence that the different results for Canada and Europe can be explained by the different size of the panels.

II. DATA DESCRIPTION

This paper sets out to test the purchasing power parity hypothesis over the post-Bretton-Woods flexible exchange rate period by exploiting the differences between

international and intranational data. We use three data sets in this paper. The first is United States city CPI data from the BLS. It contains monthly seasonally unadjusted all-items CPI observations for fourteen metropolitan areas, spanning from 1978:04 to 1997:04.⁹ The BLS reports monthly price data for four “core” metropolitan areas (Chicago, Los Angeles, New York City, and Philadelphia) and bimonthly data for ten others.¹⁰ Of these ten metropolitan areas, five report data on an odd-month basis (Baltimore, Boston, Miami, St. Louis, and Washington D.C.) and five report data on an even-month basis (Dallas/Ft. Worth, Detroit, Houston, Pittsburgh, and San Francisco).¹¹ The even-month series span from 1978:04 to 1997:04, and the odd-month series span from 1978:05 to 1997:05. Because we have monthly data for only four U.S. metropolitan areas, we construct monthly data from the bimonthly series.¹² This provides us with monthly observations for fourteen U.S. cities.

The second data set is monthly seasonally unadjusted all-items CPI observations for nine Canadian metropolitan areas. Obtained from Statistics Canada, the cities are Calgary, Edmonton, Montreal, Ontario, Quebec City, Regina, Toronto, Vancouver, and Winnipeg. To make the Canadian data comparable to the U.S. data, we restrict the data set to span from 1978:09 to 1997:06.

Although we would like to construct a comparable panel of European Union city pairs, the only city consumer price data available (that we are able to access) are Paris and Istanbul. When we calculate correlation coefficients between these cities and their respective national CPI, the values are 0.996 and 0.998, respectively. Given that the city CPI series are highly correlated with the respective country CPI series, we use European Union country data as a proxy for EU city CPI data. Obtained from *International Financial Statistics*, the European data set also includes seasonally adjusted monthly

⁹ Only seasonally unadjusted data is available on a monthly basis for the city price data.

¹⁰ They also report bimonthly data for an eleventh city, Cleveland, but we delete this from our analysis because it switched in 12:86 from odd-month reporting to even-month reporting.

¹¹ Although monthly data is available for part of the Detroit and San Francisco series, we treat both cities in the even-month series. In 12:86 Detroit changed from monthly to bi-monthly reporting and San Francisco changed from bimonthly to monthly reporting.

¹² We use an interpolate source program available in RATS.

exchange rate data for fifteen countries, spanning from 1978:01 to 1997:02. We restrict the time span to correspond with the U.S. and Canadian city data sets.¹³

III. UNIT ROOT TESTS

Let the relative price level (or real exchange rate) between cities (or countries) be calculated as

$$q = e + p^* - p \quad (1)$$

where q is the logarithm of the real exchange rate, e is the logarithm of the nominal (numeraire) exchange rate, p is the logarithm of the domestic Consumer Price Index, and p^* is the logarithm of the Consumer Price Index of the city (or country) whose currency we use as the numeraire currency. Note that because the within-country exchange rate is 1, e drops out of the equation so that the real exchange rate is just the relative price level, $p^* - p$. Also, because the IFS reports bilateral dollar exchange rates, for the European panels e is the difference between the logarithm of the nominal (dollar) exchange rate of the domestic country and the logarithm of the nominal (dollar) exchange rate of the country whose currency we use as the numeraire currency.

Univariate ADF tests regress the first difference of a variable (in this case the logarithm of the real exchange rate) on a constant, its lagged level and k lagged first differences using the following equation:

$$\Delta q_t = \mathbf{m} + \mathbf{a}q_{t-1} + \sum_{i=1}^k c_i \Delta q_{t-i} + \mathbf{e}_t \quad (2)$$

We omit a time trend in equation (2) to be theoretically consistent with long-run PPP. The null hypothesis of a unit root is rejected in favor of the alternative of level stationarity if α is significantly different from zero.

We use a recursive t-statistic procedure to select the value of k . We first set $k_{\max} = 36$ (months), which is the upper bound on k , and estimate the model with all lags.¹⁴ If the last included lag is significant then $k = k_{\max}$. If the last included lag is not significant, then we

¹³ The countries include Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. Ireland is excluded because it does not have monthly CPI data. Luxembourg is excluded because it has a currency union with Belgium.

reduce k by one until the last lag becomes significant. If no lags are significant, then $k = 0$. We use the ten percent value of the asymptotic normal distribution, 1.645, to determine significance. Campbell and Perron (1991) and Ng and Perron (1995) show that this recursive t-statistic procedure has better size and power properties than alternative procedures, such as selecting k based on AIC or BIC methods.

Very rarely do the ADF tests reject the unit root null for the within country relative price levels or for the international real exchange rates, and then only at the ten percent level of significance. ADF tests provide little if any evidence of long-run PPP for the post-73 float.¹⁵ This is not surprising, however, given the unit root test's low power against a highly persistent alternative.

Panel unit root tests, by exploiting both cross section and time series variation, have power and size advantages over univariate unit root tests in small spans of data like the recent float. Among those to apply panel test procedures to PPP during the recent float include Abuaf and Jorion (1990), Frankel and Rose (1996), Jorion and Sweeney (1996), O'Connell (1998), Papell (1997), and Papell and Theodoridis (1998a,b).

The ADF test in Equation (2) can be extended to a panel by estimating the following equations:

$$\Delta q_{jt} = \mathbf{m}_j + \mathbf{a}q_{jt-1} + \sum_{i=1}^k c_{ij} \Delta q_{jt-i} + \mathbf{e}_{jt} \quad (3)$$

where the subscript j indexes the cities (or countries) and \mathbf{m}_j denotes the heterogeneous intercept. We estimate Equation (3) by feasible GLS to account for contemporaneous correlation, with the coefficient \mathbf{a} equated across cities and the values for k taken from the results of a univariate Augmented Dickey-Fuller test.¹⁶ As in the univariate tests, these tests omit a time trend for theoretical consistency with PPP. The test statistic is the t-statistic on the coefficient \mathbf{a} where, like the ADF test, the null hypothesis of a unit root is

¹⁴ $K_{\max} = 24$ months was insufficient to account for serially correlation in the data.

¹⁵ These results are available upon request.

¹⁶ O'Connell (1998) has shown that, if the lag length k and the value of the c 's are equated across countries, panel tests of PPP using GLS are invariant to the choice of numeraire. Since we do not impose these restrictions, numeraire invariance is not imposed by our methods.

rejected in favor of the level stationarity alternative if \mathbf{a} is significantly different from zero.

17

With fourteen cities in the United States data set, we construct fourteen different panels (each with a different city as the numeraire city) of thirteen relative price levels each. For the Canadian data set we construct nine different panels of eight relative price levels each.¹⁸ For the European Union data we construct fifteen panels of fourteen real exchange rates each.

We calculate critical values using Monte Carlo methods with randomly generated data. First we fit autoregressive (AR) models to the first differences of each series, using the Schwartz criterion to choose the optimal AR model. Then we use the optimal AR model in order to generate the errors for each series. For each panel of n relative price levels (or real exchange rates) we use the optimal AR model with iid $N(0, \sigma^2)$ innovations to construct pseudo samples of size equal to the actual size of the n series (228 for U.S. monthly series, 226 for the Canadian series, and 230 for the European series). The critical values for the finite sample distributions are taken from the sorted vector of 5000 replicated statistics and reported along with the panel tests results. Our critical values are greater, in absolute value, than those in Levin, Lin, and Chu (1997) because we are accounting for serial correlation.¹⁹

We begin by conducting panel tests on the U.S. city relative price level series. Table 1 reports the results of the panel tests of the monthly city pairs for the United States. The panel tests fail to reject the unit root null hypothesis. Only one numeraire, Detroit, rejects the unit root null, and even then only at the ten percent level. This is much less evidence of PPP than is found using post-1973 monthly data with the U.S. dollar as numeraire. Papell (1997), testing all combinations of 13 (the size of the intranational panel of U.S.

¹⁷ These tests follow Levin, Lin, and Chu (1997) by restricting α to be equal across countries. Im, Peseran, and Shin (1997) develop tests where α can vary across countries. Based on the results of univariate ADF tests, this does not appear to be important for the real exchange rates or relative price levels investigated here. Bowman (1997) shows that, since the alternative hypothesis is that one element of the panel is stationary, the LLC tests are more conservative than the IPS tests.

¹⁸ Nine Canadian cities was the most available. Fourteen would have made a better cross panel comparison.

cities) out of 17 (the size of his international monthly data set) real exchange rates, rejects the unit root null at the 10% level in about 70% of the cases.

A common method for measuring persistence is to calculate the half lives of PPP deviations, the amount of time that it takes a shock to the series to revert halfway back to its mean value, with the formula defined as $\ln(.5)/\ln(1+\alpha)$.²⁰ The average half-life for U.S. city panels is 45.86 months, or 3.82 years. While this result is consistent with Rogoff's (1996) characterization of 3 - 5 year half-lives for PPP deviations, it represents a speed of reversion which is considerably slower than the 2.5 year half-life found in Papell (1997) for a panel of industrialized countries with the U.S. dollar as the numeraire currency. Even for the panel with Detroit as numeraire, which is the only panel that reports any evidence of stationarity, the calculated half-life is 36.13 months or 3.01 years.

We next examine how the U.S. within country findings compare with the Canadian relative price panels. Table 2 reports the findings from panel tests conducted on Canadian price level city pairs. For three out of the nine panels, the unit root hypothesis is rejected at the ten percent level. Although the Canadian panels are smaller (with 8 relative price levels) than the U.S. monthly panels (which have 13 relative price levels), with consequent reduction in power of the panel unit root tests, Canadian panels give more evidence against the unit root null. Using another metric, the average p-value for the Canadian panels is .166. This indicates much stronger evidence against the unit root null than the average p-value, .365, for the U.S. panels.

There are considerable differences in convergence rates between the U.S. panels and the Canadian panels. The average half-life for Canadian city panels is 22 months or 1.83 years. This speed of convergence is not only much faster than we found for U.S. cities, it is faster than the convergence speeds that are found for industrialized countries. For those

¹⁹ We do not incorporate contemporaneous correlation in the data generating process because, as shown by O'Connell (1998), the distribution of the FGLS estimate of α is invariant to the degree of correlation between real exchange rate innovations.

²⁰ Andrews and Chen (1994) show that the half-life is a good scalar measure of persistence.

Canadian numeraire cities where relative price levels are stationary, Quebec City, Toronto, and Winnipeg, the calculated half-lives are between 1.67 and 1.83 years.²¹

As a benchmark for comparison, we use European Union country real exchange rates as a proxy for European Union city pairs.²² These results are reported in Table 3. Like other international PPP studies using European data, we find strong rejections of the unit root null over the current floating exchange rate period. Out of fifteen panels, we are able to reject the unit root null hypothesis for 10 panels at the 10% level, 7 panels at the 5%, and 1 at the 1% level of significance. The average half-life for the European panels is 26.31 months or 2.19 years. This is very similar to the half-lives reported in Papell (1997) with the German mark as the numeraire currency. The calculated half-lives for the European real exchange rates that exhibit stationarity range between 1.83 to 2.28 years, slightly greater than those for Canada but less than those for the United States.

The comparison of results between the United States and the European Union panels are consistent with the findings in Sonora (1997) and Cecchetti, Mark, and Sonora (1998). Using long-term annual U.S. CPI data, they report half-lives which are considerable longer than are found in studies using comparable length international real exchange rate data. They are not, however, consistent with the results of Parsley and Wei (1996), who find that, using disaggregated data, the speed of convergence to the law of one price is substantially higher for intranational data than is found in cross-country data.

The findings for Canada are puzzling. Although the speed of convergence towards PPP is comparable (and even slightly faster) than within the European Union, there is little evidence against unit roots. An obvious difference is that the cross-section dimension is 8 for the Canadian panels versus 14 for the European panels. While the size of the panels of U.S. cities, 13, is similar to the European panels, the speed of convergence is much slower for U.S. cities. In order to explore both the hypothesis that the non-rejections of unit roots for the Canadian panels can be explained by the smaller size of the panels and the

²¹ Engel and Rogers (1996) report that average price volatility is higher between U.S. city pairs than between Canadian city pairs, and postulate that this might be because the U.S. is a more heterogeneous country.

²² We were only able to access two European city CPIs, those of Paris and Istanbul. Because two series makes for a small panel and because the city data was highly correlated with the corresponding country data, we use country CPI data as a proxy.

hypothesis that the non-rejections of unit roots for the U.S. panels can be explained by the slower speeds of convergence, we conduct a power analysis of the tests.

The results of the power analysis are presented in Table 4. Using Monte Carlo methods similar to those described above for the calculation of critical values, we generate panels of 8, 13, and 14 series. Each series is an AR(1), with the autoregressive coefficient, $1-\alpha$, calculated from the average value of α in the Canadian, U.S., and European panels, respectively.²³ These series are stationary by construction. The panel unit root tests are then conducted 5000 times. The entries in Table 4 describe the fraction of times that the unit root null can be rejected at the 1%, 5%, and 10% levels.

The power of the unit root tests is considerably lower for the panel replicating the characteristics of 8 Canadian cities than for the panel of 14 European countries, even though the speed of convergence is faster ($1-\alpha$ is smaller) for the Canadian panel. This can explain why the rejections of unit roots are so much stronger for the European countries than for the Canadian cities. The power of the tests is even lower for the panel replicating the U.S. cities. Since the number of elements is almost the same between the U.S. and the European panels, the fall in power is clearly due to the much slower speed of convergence for the U.S. cities. This can explain the almost total inability to reject the unit root null for the U.S. panels.

For all three panels, the power analysis predicts more rejections of the unit root null than occur with the actual data. This can be explained by contemporaneous correlation. For the power simulations, the disturbances are independent. With the actual data, because shocks to the "domestic" price level and, for the European panels, the nominal exchange rate are common across elements, there is considerable contemporaneous correlation. Since correlated disturbances reduce the additional information gained from increasing the size of the panels, the number of rejections of the unit root null with actual data would not be expected to match the power results with simulated data.

V. CONCLUSIONS

²³ The number of observations also match the actual series.

This paper sets out to provide evidence on the purchasing power parity hypothesis over the post-Bretton-Woods period by exploiting the differences between international and intranational data. Since intranational data avoids problems such as trade barriers, exchange rate volatility, differential monetary policies, nonintegrated markets, and other factors that restrict goods market arbitrage, we expected to find that the evidence of PPP was stronger among cities in Canada and the United States than across countries in Europe. However, we find just the opposite. Compared to the European Union panels, our results from the within country panels find at best weak evidence of PPP, and then only for the Canadian panels.

We calculate the half lives of PPP deviations as a metric to calibrate the rate of mean reversion. Here we find that the rate of price convergence is slower for U.S. cities than for either Canadian cities or for European countries. While this explains the stronger rejections of unit roots in real exchange rates for the European Union countries compared with the U.S. cities, the Canadian results are puzzling. While the speed of mean reversion is actually faster for Canadian cities than for European countries, the evidence of PPP is much weaker. We conduct a power analysis of the panel unit root tests, and show that the results can be explained by the different sizes of the panels.

Intranational data for U.S. cities has been used by Parsley and Wei (1996) to provide an upper bound of the rate of convergence to purchasing power parity. With the end of nominal exchange rate fluctuations among those European Union countries that have adopted the Euro, Cecchetti, Mark, and Sonora (1998) use evidence of price level convergence among U.S. cities to provide lessons for the European Central Bank. These studies reflect a presumption that the evidence of PPP should be stronger with intranational than with international data. Our major result, that the rejections of unit roots in relative prices are much stronger among European countries than among either U.S. or Canadian cities, is not in accord with that presumption. Further investigation of the power of the tests, however, shows that our results are in accord with what would be expected considering the differences in panel sizes and speeds of adjustment to PPP among the panels.

REFERENCES

- Abuaf, N. and P. Jorion (1990), "Purchasing Power Parity in the Long Run," *Journal of Finance*, 45, 157-174.
- Andrews, D., and H.-Y. Chen, (1994), "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics* 12, 187-204.
- Bowman, D., (1997), "Efficient Tests for Autoregressive Unit Roots in Panel Data," manuscript, Federal Reserve Board.
- Campbell, J. and P. Perron (1991), "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots," *NBER Macroeconomics Annual*, 141-201.
- Canzoneri, M., Cumby, R., and B. Diba, (1999), "Relative Labor Productivity and the Real Exchange Rate in the Long Run: Evidence for a Panel of OECD Countries," *Journal of International Economics*, 47, 245-266.
- Cecchetti, S., Mark, N., and R. Sonora, (1998), "Price level Convergence Among United States Cities: Lessons for the European Central Bank," manuscript, The Ohio State University.
- Chen, L. and J. Devereux (1997), "U.S. City Price levels in the Long Run," manuscript, University of Miami and University of Nevada.
- Cheung, Y.-W. and K. Lai, (1998), "On Cross-Country Differences in the Persistence of Real Exchange Rates, manuscript, University of California, Santa Cruz.
- Culver, S. and D. Papell (1999), "Long-Run Purchasing Power Parity with Short-Run Data: Evidence with a Null Hypothesis of Stationarity," forthcoming, *Journal of International Money and Finance*.
- Elliot, Rothenberg, and Stock (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813-836.
- Engel, C., (1999), "Long-Run PPP May Not Hold After All," forthcoming, *Journal of International Economics*.
- Engel, C. and J.H. Rogers, (1996) "How Wide Is the Border?," *American Economic Review*, 86, No.5, 1113-1125.
- Frankel, J. (1986), "International Capital Mobility and Crowding Out in the U.S. Economy: Imperfect Integration of Financial Markets or of Goods Markets?" in *How Open is the U.S. Economy?*, R. Hafer, ed., Lexington Books.

Frankel, J. and A. Rose (1996), "A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries," *Journal of International Economics*, 40, 209-224.

Froot, K. and K. Rogoff (1995), "Perspectives on PPP and Long-Run Real Exchange Rates" in *Handbook of International Economics*, G. Grossman and K. Rogoff, eds., North-Holland, Amsterdam.

Hegwood, N. and D. Papell (1998), "Quasi Purchasing Power Parity," *International Journal of Finance and Economics*, 3, 279-289.

Im, S., Pesaran, H. and Y. Shin (1997), "Testing for Unit Roots in Heterogeneous Panels," manuscript, University of Cambridge.

Jorion, P. and R. Sweeney (1996), "Mean Reversion in Real Exchange Rates: Evidence and Implications for Forecasting," *Journal of International Money and Finance*, 15, 535-550.

Levin, A., Lin, C. F. and J. Chu, (1997), "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," manuscript, Federal Reserve Board.
6, 19-35.

Lothian, J. and M. Taylor (1996), "Real Exchange Rate Behavior: The Recent Float From the Perspective of the Past Two Centuries," *Journal of Political Economy*, 104, 488-509.

Lothian, J. and M. Taylor (1997), "Real Exchange Rate Behavior," *Journal of International Money and Finance*, 16, 945-954.

Ng, S. and P. 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, 268-281.

O'Connell, P. (1998), "The Overvaluation of Purchasing Power Parity," *Journal of International Economics* 44, 1-19.

O'Connell, P. and S.J. Wei (1997), "The Bigger They Are, The Harder They Fall: How Price Differences between U.S. Cities are Arbitraged," manuscript, Harvard University.

Papell, D. (1997), "Searching for Stationarity: Purchasing Power Parity Under the Current Float," *Journal of International Economics*, 43, 313-332.

Papell, D. (1998), "The Great Appreciation, the Great Depreciation, and the Purchasing Power Parity Hypothesis," manuscript, University of Houston.

Papell, D. and H. Theodoridis (1998a), "Increasing Evidence of Purchasing Power Parity over the Current Float," *Journal of International Money and Finance*, 17, 41-50.

Papell, D. and H. Theodoridis (1998b), "The Choice of Numeraire Currency in Panel Tests of Purchasing Power Parity," manuscript, University of Houston.

Parsley, D. and S.-J. Wei (1996), "Convergence to the law of One Price without Trade Barriers or Currency Fluctuations," *Quarterly Journal of Economics*, 111, 1211-1236.

Rogoff, K. (1996), "The Purchasing Power Parity Puzzle," *Journal of Economic Literature* 34, 647-668.

Sonora, R. (1997), "Relative Price Level Convergence in the United States," manuscript, The Ohio State University.

Wei, S.J. and D. Parsley (1995), "Purchasing Power *Dis*-Parity During the Floating Rate Period: Exchange Rate Volatility, Trade Barriers and Other Culprits," manuscript, Harvard University.

Table 1**United States City Relative CPI Indexes**

| Numeraire City | α | t_{α} | p-values |
|------------------|----------|--------------|----------|
| Baltimore | -0.015 | -5.397 | .273 |
| Boston | -0.013 | -4.963 | .425 |
| Chicago | -0.018 | -5.125 | .363 |
| Dallas/Ft. Worth | -0.015 | -5.532 | .225 |
| Detroit | -0.019 | -6.176* | .086 |
| Houston | -0.012 | -4.230 | .691 |
| Los Angeles | -0.015 | -4.474 | .609 |
| Miami | -0.013 | -4.727 | .514 |
| New York City | -0.013 | -5.078 | .380 |
| Philadelphia | -0.018 | -5.413 | .267 |
| Pittsburgh | -0.014 | -5.199 | .338 |
| San Francisco | -0.017 | -5.454 | .255 |
| St. Louis | -0.014 | -5.117 | .366 |
| Washington D.C. | -0.014 | -5.248 | .323 |
| Average | -0.015 | -5.152 | .365 |

Note: Kmax = 36 months. Each panel consists of 13 relative CPI series. Number of observations = 228, which span from 1978:05 to 1997:04. We restrict α to be the same across equations. The critical values for t_{α} are -7.19, -6.48, -6.10 for the 1%, 5%, and 10% level of significance, respectively.

Table 2**Canadian City Relative CPI Indexes**

| Numeraire City | α | t_{α} | p-values |
|----------------|----------|--------------|----------|
| Calgary | -0.028 | -4.495 | .245 |
| Edmonton | -0.028 | -4.406 | .273 |
| Montreal | -0.033 | -4.980 | .122 |
| Ottawa | -0.030 | -4.774 | .167 |
| Quebec City | -0.035 | -5.341* | .059 |
| Regina | -0.029 | -4.567 | .226 |
| Toronto | -0.031 | -5.181* | .083 |
| Vancouver | -0.029 | -4.507 | .241 |
| Winnipeg | -0.034 | -5.210* | .077 |
| Average | -0.031 | -4.622 | 0.166 |

Note: Kmax = 36 months. Each panel consists of 8 relative price level series. Number of observations = 226, which span from 1978:09 to 1997:06. We restrict α to be the same across equations. The critical values for t_{α} are -6.08, -5.42, -5.09 for the 1%, 5%, and 10% level of significance, respectively.

Table 3**European Country Real Exchange Rates**

| Numeraire Country | α | t_{α} | p-values |
|-------------------|----------|--------------|----------|
| Austria | -0.028 | -7.378*** | .007 |
| Belgium | -0.029 | -7.225** | .011 |
| Denmark | -0.025 | -6.583* | .052 |
| Finland | -0.023 | -5.721 | .235 |
| France | -0.025 | -6.336* | .087 |
| Germany | -0.025 | -6.838** | .029 |
| Greece | -0.024 | -5.798 | .213 |
| Italy | -0.031 | -6.673** | .042 |
| Netherlands | -0.025 | -6.902** | .025 |
| Norway | -0.029 | -7.025** | .018 |
| Portugal | -0.024 | -5.776 | .221 |
| Spain | -0.025 | -5.963 | .169 |
| Sweden | -0.028 | -6.630** | .046 |
| Switzerland | -0.028 | -6.781** | .033 |
| United Kingdom | -0.024 | -5.605 | .271 |
| Average | -0.026 | -5.773 | .097 |

Note: $K_{max} = 36$ months. Each panel consists of 14 real exchange rates. Number of observations = 230, which span from 1978:01 to 1997:02. We restrict α to be the same across equations. The critical values for t_{α} are -7.27, -6.60, -6.26 for the 1%, 5%, and 10% level of significance, respectively.

Table 4**Power Analysis of the Unit Root Tests**

| Panel | N | T | 1- α | 1% | 5% | 10% |
|----------------------|----|-----|-------------|------|------|------|
| United States Cities | 13 | 228 | .985 | .170 | .425 | .596 |
| Canadian Cities | 8 | 226 | .969 | .225 | .522 | .666 |
| European Countries | 14 | 230 | .974 | .362 | .661 | .793 |

Note: N is the number of elements in each panel and T is the number of observations. $1-\alpha$ is calculated from the average value of α in Tables 1-3. $K_{max} = 36$ months. α is restricted to be the same across equations.