Price Index Convergence Among United States Cities: New Results

Luciano Gutierrez*
University of Sassari,
Department of Agricultural Economics,
Via E. De Nicola 1, 07100 Sassari (Italy)

June, 2006

Abstract

In this paper we analyze the important question raised by Cecchetti *et al.* (2002) on whether price indices for major U.S. cities share a common trend, and, if so, how quickly they revert to that trend following a local shock. Using recent methodologies on panel nonstationary analysis, we are able to decompose the local price dynamics into common and idiosyncratic components. We show, using the same dataset as Cecchetti *et al.* (2002), that the idiosyncratic component is stationary and that this accounts for a small percentage of the total variance of local prices. Summing-up, on average more than 95% of the total variance is accounted for by the common component. After a shock, local price levels among cities mean revert to a higher rate than those estimated by Cecchetti *et al.*(2002). Finally we report that the aggregate consumer price index can be linked to the common component. All these findings may explain why monetary policy makers target consumer price inflation in measurements of aggregate inflation, as the European Central Bank has done with the Harmonized Index of Consumer Prices in recent years, and put less emphasis on the behavior of local price indices.

Key words : Local prices index, Panel nonstationary analysis, Common trends.
JEL classification: C2, C3, C5.

*Corresponding author. E-mail address: lgutierr@uniss.it

1 Introduction

In a recent paper Cecchetti *et al.*(2002) (CMS hereafter) posed the important question of whether local price indices share a common trend, and, if so, how quickly they revert to that trend following a local shock. Given the European Central Bank's (ECB) objective of fixing the year-on-year change in the Harmonized Index of Consumer Prices (HICP) at not more than 2%, the main reasons for the analysis were to arrive at a better understanding of how large EU regional price indices can deviate from the Euro-wide-area average price index, and secondly to study how long these deviations are likely to persist. As a long-term series of regional EU data is lacking, they studied the dynamics of consumer price indices for 19 major U.S. cities over the period from 1918 to 1995. They used panel econometric methods such as Levin *et al.* (2002) and Im *et al.* (2003) panel unit root tests for cross-sectional independent panels. The main reason to adopt these methodologies was that panel unit root tests, by combining the information in time and cross-section dimensions, have proved to be a promising way of increasing the power of single time series unit root tests. Summarizing, they found that price index divergences across U.S. cities are temporary but surprisingly persistent, with a half-life of nearly nine years. Some explanations for incomplete relative price-level adjustment have been proposed in CMS's paper, such as transportations costs or the presence of nontraded-goods in price indices. In synthesis they find only fragmentary evidence that these factors can account for the slow adjustment of the overall consumer price index.

The main contribution of this paper is empirical. We document, using the same dataset as CMS, significant cross-sectional correlation among the U.S. price indices. In this case, demeaning the series, as was done in CMS's paper in order to remove cross-section dependence, could not work when the pair-wise cross-section covariances of the error terms differ across the individual series (as shown in Gutierrez (2006)). This means that the Levin *et al.* (2002) and Im *et al.* (2003) tests, which are not suited to take into account cross-sectional dependence, are biased towards the alternative hypothesis of stationarity and thus may suffer from strong size distortions. In the paper we also show that the cross-section temporal mean used by CMS to compute the relative exchange rate prices can be interpreted as an unobserved single factor, which affects the local prices with a unit coefficient. Thus if local prices are influenced, for example, by two unobservable factors and/or their impact is different from unity, the temporary component observed in CMS's paper could be overvalued, thus giving rise to possible misspecification of the temporary price index dynamics.

To overcome these problems we apply recent cross-section common stochastic trends analysis for nonstationary panel data. Using this methodology we are able to decompose the consumer log-prices indices for U.S. cities into one, or at most two, nonstationary common trend factors and a stationary idiosyncratic component. The decomposition allows us to measure the relative importance of each component in influencing the log-prices. We show that the total variance of local price indices is dominated by the common components, while the idiosyncratic component has much less relative importance. To be precise, the idiosyncratic component accounts on average for only 3% of the total log-prices variance. Furthermore the deviations from the common factors show a half-life of 5 years which is lower than CMS's estimated half-life. They reported it as nearly nine years for the full sample period 1918-1995. Finally we show, using the recent inferential theory proposed in Bai (2004), that the aggregate consumer price index can be associated to the common nonstationary components.

To sum up, we find that local consumer price indices are mainly influenced by the aggregate consumer prices index and local shocks exert a much smaller effect. After a shock local prices revert to the average of the consumer price index in a half-life of five years. These findings may explain why monetary policy makers target consumer price inflation in measurements of aggregate inflation, as the European Central Bank has done with the HICP index in recent years, and put less emphasis on the behavior of regional price indices.

In the following section we briefly present recent nonstationary panel data analysis, showing how it can be used to decompose local prices into common and idiosyncratic components, and how test statistics can be derived in order to analyze if one or both components are nonstationary or stationary. An important question in this case is defining how many factors drive the common component. To discover this we survey recent developments proposed in Bai and Ng (2002) and Bai (2004). In section 3 we provide the empirical results for the panel of price indices of U.S. cities. Finally, section 4 concludes.

2 Common trends and panel unit root tests for cross-sectionally correlated panels.

Assume that the following factor model holds for the log-price level of city i, at time t, p_{it}, with $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$,

$$p_{it} = \alpha_i + \sum_{j=1}^{r} \lambda_{ij} f_{jt} + e_{it} = \alpha_i + \lambda_i^{'} f_t + e_{it} \tag{1}$$

where α_i is a city-specific constant, f_t are r-vectors of unobservable common factors, $\lambda_i = (\lambda_{i1}, ..., \lambda_{ir})^{'}$ is the corresponding non random vector of factor loading coefficients and e_{it} is an idiosyncratic error process for which we allow weak cross-section dependence and that can be serially correlated for each i . From (1) it is simple to see that the log-prices p_{it} will be $I(1)$ processes if the common factor component f_t and/or the idiosyncratic component e_{it} are $I(1)$ processes.

If we set all $\lambda_i = 1$ and we fix $r = 1$, equation (1) is basically the equation analyzed in CMS's paper. To this end CMS subtract the cross-sectional mean $p_{.t} = (1/N)\sum_{i=1}^{N} p_{it}$ from both sides of (1) in order to eliminate the common time effect and take into account for the cross-sectional dependence across prices. If we apply their strategy to equation (1) we obtain

$$q_{it} = p_{it} - p_{.t} = (\alpha_i - \bar{\alpha}) + \left(\lambda_i - \bar{\lambda}\right)' f_t + (e_{it} - e_{.t}) \tag{2}$$

with $\bar{\alpha} = (1/N)\sum_{i=1}^{N} \alpha_i$, $\bar{\lambda} = (1/N)\sum_{i=1}^{N} \lambda_i$ and $e_{.t} = (1/N)\sum_{i=1}^{N} e_{it}$.

From (2) it is simple to see that the CMS strategy will be in effect only when $\lambda_i = \bar{\lambda}$ for all i. Otherwise, we will have that in general for any $i \neq j$

$$E\left[q_{it}q_{jt}\right] = \left(\lambda_i - \bar{\lambda}\right)' E\left(f_t f_t'\right)\left(\lambda_j - \bar{\lambda}\right), \tag{3}$$

i.e. the relative log-prices will be cross-sectional correlated. In addition, we can see from (2) that, assuming $(e_{it} - e_{.t})$ is $I(0)$, q_{it} will be $I(0)$ if either f_t is $I(0)$ or if f_t is $I(1)$ but $\lambda_i = \bar{\lambda}$.

CMS use Levin *et al.*(2002)(LLC) and Im *et al.* (2003) (IPS) panel unit root tests to assess if price indices in major U.S. cities revert to the mean trend $p_{.t}$ following a local shock to price index, i.e. they look for nonstationarity or stationarity of q_{it}. Unfortunately both tests are not suitable when the variables, in this case the price indices, are cross-sectionally correlated. O'Connell (1998) shows that not controlling for cross-sectional dependence dramatically raises the significance level of tests with a nominal size of 5 percent to as much as 50 percent. Pesaran (2006), by using a Monte Carlo simulation that with a high cross-section dependence, calculates that the size of the IPS test when N=20 and T=100 (i.e. approximately the dimension of the panel of price indices used in CMS's paper) is equal to 0.246. [1] Thus by not taking into account cross-section dependence the LLC and IPS tests will tend to over-reject the null hypothesis of nonstationarity, often by a substantial amount.

In order to investigate if the extend of the correlation is determined by the factor loading coefficients, we follow Bai and Ng (2004). We assume the following process for the common component: $(I - L) f_t = C(L) v_t$, with $C(L)$ a polynomial matrix in lag L and v_t an error component with finite fourth moments. When the rank of $C(1)$ is $r_1 < r$, we will have r_1 common trends and $(r - r_1)$ stationary common factors. The idiosyncratic component is modeled as

$$(1 - \rho_i L) e_{it} = D_i(L) \varepsilon_{it} \tag{4}$$

[1] This result is obtained assuming the following uniform distribution for the factor loadings$\lambda_i \sim U[-1, 3]$. Interestingly the IPS test has the correct size when $\lambda_i \sim U[0, 0.2]$ which is reported in Pesaran's paper as a low cross-section dependence case.

where ε_{it} is an error term and $D_i(L) = \sum_{k=0}^{H} D_{ik} L^k$ is a polynomial lag of order H. From (4) we note that each idiosyncratic component e_{it} will be stationary when $|\rho_i| < 1$ and non-stationary, or, equivalently, integrated of order one $I(1)$, for $\rho_i = 1$.

To compute both the common as well as the idiosyncratic components we define with $\mathbf{p}$ the $(T \times N)$ matrix of log-prices p_{it}. Bai and Ng (2004) propose estimating the (differenced) common factor(s), Δf_t, by applying the method of principal components to the first differenced data $\Delta\mathbf{p}$. The factor loading, λ_i, are then obtained as the product of (transposed) matrix $\Delta\mathbf{p}$ and common factor(s) Δf_t and the (differenced) idiosyncratic terms [2] in (4) are computed as

$$\Delta\hat{e}_{it} = \Delta p_{it} - \hat{\lambda}_i' \Delta\hat{f}_t. \tag{5}$$

Finally, the estimates of the level of common factors and idiosyncratic terms are obtained simply by integrating $\hat{f}_t = \sum_{k=2}^{T} \Delta\hat{f}_k$, and $\hat{e}_{it} = \sum_{k=2}^{T} \Delta\hat{e}_{ik}$.

Now $\hat{f}_t$ and $\hat{e}_{it}$ can be used to test if common and/or idiosyncratic, or none of the two components have unit roots or, in other words, we can ascertain whether nonstationarity of log-prices comes from the common or the idiosyncratic component, or both. To test the nonstationarity of the common component, Bai and Ng (2004) suggest using the standard Augmented Dickey-Fuller (ADF) in the case of a single factor. A modified version of the common trend test MQ_c of Stock and Watson (1988), which analyzes the rank of the long-run covariance matrix of $\hat{f}_t$, is suggested in the case of multiple factors. The limiting distribution of the MQ_c statistic is nonstandard and the critical values are presented in Bai and Ng (2004) for up to 6 factors.

For the idiosyncratic components Bai and Ng (2004) propose pooling residuals from individual ADF tests and combining thep_i significance level (p-values) as in Maddala and Wu (1999). To be precise, they use a Fisher-type test to analyze the null hypothesis of $\rho_i = 1 \quad \forall i$, against the alternative of $\rho_i < 1$ for some i. Gutierrez (2006) shows that when combining the p-values from the Dickey-Fuller-GLS proposed in Elliott *et al.* (1996), the power of the test sensibly augmented. In the case of ADF tests we estimate the usual regression

$$\hat{e}_{it} = \rho_i \hat{e}_{it-1} + \sum_{j=1}^{k_i} \gamma_{ij} \Delta\hat{e}_{it-j} + \varepsilon_{it}. \tag{6}$$

For each cross-section $i = 1, ..., N$, we compute the p-values, p_i, associated to the t_i statistic for testing the null hypothesis$\rho_i = 1$. Using the p-values,

[2] As highlighted in Bai and Ng (2004) differencing is needed to preserve a consistent estimation of the common factors when the idiosyncratic components are nonstationary.

the following panel unit root statistic for testing the null hypothesis on nonstationarity of the idiosyncratic component can be used

$$P_{\hat{e}}^{c} = \frac{\left[\left(-2\sum \log p_i\right) - 2N\right]}{\sqrt{4N}}. \tag{7}$$

Choi (2001) shows that for $(N, T \to \infty)$ the statistic converges to a standard normal distribution and under the alternative the test diverges to negative infinity and therefore has to be compared to the left tail of the distribution.

Crucial to the previous method is determining the number of factors. Bai and Ng (2002) employ information criteria essentially based on the difference idiosyncratic residuals (5). The criteria are similar to the well known AIC and BIC criteria for time series. The true number of factors r is found minimizing

$$CR\left(r\right) = V\left(r, \hat{f}_t^r\right) + rG\left(N, T\right) \tag{8}$$

where $V\left(r, f_t^r\right)$ is a measure of the fit, which basically depends on the number r of factors $\hat{f}_t^r$ estimated, and the function $G\left(N, T\right)$ is a penalty function related to the size of the panels. Bai and Ng (2002) propose different criteria. When $N \geqslant 20$ all seem to perform well in estimating the true number of factors, with the exception of $15 \leqslant N < 20$, as in the case of the price indices used in CMS's work, Gutierrez (2006) shows by using the Monte Carlo simulation that the following criterion function has to be preferred

$$BIC_3\left(r\right) = \hat{\sigma}_{\Delta\hat{e}}^2\left(r\right) + r\hat{\sigma}_{\Delta\hat{e}}^2\left(r_{\max}\right)\frac{N+T-r}{NT}\ln\left(NT\right) \tag{9}$$

where $\hat{\sigma}_{\Delta e}^2\left(r\right)$ is the ratio of the sum of squared differenced idiosyncratic residuals computed as in (5) on NT and $\hat{\sigma}_{\Delta e}^2\left(r_{max}\right)$ is the same measure, but estimated for the maximum number of factors permitted in (8).

Recently Bai (2004) has proposed a new method to account for the number of possible nonstationary common factors. Unlike the modified MQ_c test statistic, the method is based on information criteria similar to those previously presented but in this case applied to the levels of the series rather than the first differences. The residuals are computed as $\tilde{e}_{it} = p_{it} - \alpha_i - \lambda_i^{'} f_t$ and now the counterpart of the previous BIC criterion (9) to estimate the total number of nonstationary common trends can be computed as

$$\widetilde{BIC}_3\left(r\right) = \hat{\sigma}_{\tilde{e}}^2\left(r\right) + r\delta_T\hat{\sigma}_{\tilde{e}}^2\left(r_{\max}\right)\frac{N+T-r}{NT}\ln\left(NT\right) \tag{10}$$

with $\delta_T = T/4\ln\ln\left(T\right)$. In synthesis the criteria (9) will be used to compute the total number of factors and the information criteria (10) to estimate the number of nonstationary factors.

After obtaining the estimated factors we can ask if the aggregate consumer price index is one of the underlying factors. Bai (2004) proposes a simple method for testing the previous hypothesis if factors are non stationary with stationary idiosyncratic components. Assuming $\bar{p}_t\,(t = 1, ..., T)$ is the observed log aggregate consumer price index, the method consists of testing if $\bar{p}_t$ can be expressed as a linear combination of the factors. To do this the procedure estimates the fitted value $\hat{\bar{p}}_t = \delta' f_t$ where$\hat{\delta}$ is the least-square estimator. Basically Bai's (2004) pointwise test involves computing a confidential band for the linear combination of factors$\hat{\bar{p}}_t$and for each t. If the observed series $\bar{p}_t$ lies inside the confidence intervals throughout the period of analysis, we can accept the hypothesis that the aggregate consumer price index $\bar{p}_t$ is one of the underlying factors.

3 The Econometric Analysis

Before presenting the test results, it is a useful to look at the cross correlation of prices indices. In table 1 we show the cross-correlation table computed using the price indices of the 19 U.S. cities observed during the period 1918-1995 and analyzed in CMS's paper. [3] In the table we include the short-run cross-sectional correlation of the log demeaned prices, i.e the variable q_{it}.

Table 1 about here

Note how this matrix is far away from the hypothetical diagonal matrix assumed by the standard panel unit root tests which do not allow for cross-sectional dependence. We also consider, but do not present for brevity, the long-run correlation matrix computed by using the Newey-West covariances matrix estimates which allow one to obtain the long-run cross-correlation coefficients. As in the case of the previous correlation matrix, we have cross-section dependence of log-prices. Thus panel nonstationary analysis which allows for cross-sectional dependence must be preferred when inferring if divergences of log-prices from a common trend are temporary or not.

The first step in the analysis is computing the number of factors in equation (1). To do this, in the first half of table 2 we present the number of factors estimated using the criterion (9) and fixing the maximum number of factors as up to 6 factors. In the second half of table we show the estimated number of nonstationary factors using the modified MQ_c test proposed in Bai and Ng (2004) and the $\widetilde{BIC_3}$ criterion suggested by Bai (2004). Note that we report the number of factors both for the log-prices p_{it} and the cross-sectional demeaned log-prices q_{it}.

Table 2 about here

[3]The cities in the sample are Atlanta, Baltimore, Boston, Chicago, Cincinnati, Cleveland, Detroit, Houston, Kansas City, Los Angeles, Minneapolis, New York City, Philadelphia, Pittsburgh, Portland, San Francisco, Seattle, St. Louis, and Washington D.C.

As shown in table 2, the evidence suggests two nonstationary factors for the log-prices p_{it} when the maximum number of factors allowed is greater than 4. When setting the maximum number of factors as less than 4 both statistics select only a single nonstationary factor. Thus the analysis suggests that the log-prices panel of U.S. cities is characterized by 1 or at most 2 nonstationary factors. Interestingly, looking at the statistic $\widetilde{BIC_3}$ the demeaned log-prices q_{it} also seem to be characterized by at least one nonstationary factor. The modified MQ_c test suggests two factors when the maximum number of factors is permitted to be greater than 4. Thus we interpreted the previous results as evidence that demeaning the log-prices using the cross-section time average does not allow the common factor component to be removed from the right side of equation (1). The reasons are presumably the different impacts exerted by the $\lambda_i^{'} f_t$ component on log-prices.

In the first two columns in table 3 we display the panel unit root test statistics (7) calculated on the idiosyncratic components. We report both the test statistics computed by combining the p-values of the ADF test and by combining the p-values of the DF-GLS statistics. The lags k_i for the differenced variable in (6) for the ADF or DF-GLS test have been computed using the procedure suggested in Ng and Perron (2002). We examine both the full sample and two shorter subsamples that allow us to investigate if excluding years such as the Second World War or the 1973 oil-prices shock could alter the results.

If we allow for two nonstationary factors, all test statistics for the full sample as well as for the periods 1956-1995 and 1976-1995 reject the null of nonstationary, i.e. we conclude that the idiosyncratic component is a stationary process. [4] Interestingly, if DF-GLS test statistics are pooled they do not reject the null when there is only one stationary factor and for the shorter periods it is included in (1). In synthesis the test statistics presented in Table 3 show that shocks to the idiosyncratic component are temporary and these results confirm those of CMS.

Table 3 about here

In table 4 we present statistics which highlight the relevance of the idiosyncratic component on the total variance of log-prices $\left[\left(\sigma_{\hat{e}}^2/\sigma_p^2\right) * 100\right]$ and the half-life values associated to a shock to the idiosyncratic component. The approximate half-life, $h = \ln(0.5)/\ln(\rho_i)$, of a shock to $\hat{e}_{it}$ has been computed using the estimated $\hat{\rho}_i$ coefficients in (6).

Table 4 about here

[4] This result is also confirmed by the ADF and DF-GLS tests computed on the single time series residuals.

In the first three columns we display the average, minimum and the maximum values of the percent of variance accounted for by the idiosyncratic component over the total variance of log-prices. When only one nonstationary factor is included in equation (1), the idiosyncratic component on average accounts for only 3-5 percent of the total variance of log prices, with percentages which range from 1.4 to a maximum of 9 percent whatever the period of analysis. Thus only a small fraction of log-prices variance is explained by the idiosyncratic component and the remaining variance is accounted for by the common nonstationary component. If we assume that log-prices are driven by two nonstationary factors, as seems to be the case using some of the previously presented criterion, the percentage accounted for by the idiosyncratic component is naturally smaller. However the reduction is marginal, and this means that the total variance of log-prices is basically determined by the first factor. In the second half of table 4 we present the half-life values associated with a shock to the idiosyncratic component. As before, we report the values when only one stationary factor and two nonstationary factors are included in the process. Unlike the previous results, it seems now that the half-life statistic strongly depends on having considered one or two nonstationary factors. When considering one factor the average half-life for the full period is equal to 15 years, higher than the value estimated in CMS, where the value was set to nearly nine years. When the second factor is included the half-life value is reduced to 5 years. Similar results are obtained for the sub-samples 1956-1995 and 1976-1995. Thus it seems that if we do not take into account the second nonstationary factor the idiosyncratic component could be affected by spurious persistence, thus giving rise to an overvalued half-life statistic.

CMS analyze if one explanation for finding slow intercity relative price adjustment may be the presence of nontraded-goods prices in the price indices. To analyze this hypothesis we use the two panels for traded goods and nontraded goods collected by CMS for the period 1967-1995 and for 14 U.S. cities. Although we have only a small number of cross-sections and observations, we apply the previous tools of analysis. The results, however, must be interpreted with care because the estimated number of factors as well as the test statistics may be biased, given the small dimension of both panels.

We find that both panels are driven by two nonstationary factors and a temporary component. The half-life is equal to 2 years for the traded goods and three years for the nontraded goods. In the same period the value for the consumer price dataset of 14 cities show a half-life of three years. Thus our results show that the traded-goods prices adjust more rapidly than both the nontraded-goods prices and the overall index. Thus the slower adjustment of the overall consumer price index may be caused by slower adjustment in the prices of nontraded goods.

We now analyze whether the aggregate U.S. consumer price index is one of underlying factors. If this hypothesis is true, we can postulate that log-prices for the 19 US cities are integrated I(1) variables, as is the log aggregate consumer price index. Second this means that the log prices for the 19 US cities are cointegrated with the log US consumer price index. This conclusion emerges immediately from looking at (1) and from the previous findings that the idiosyncratic components are stationary processes. As reported in section two, Bai's (2004) methodology requires estimating the regression $\bar{p}_t = \delta' f_t + error$, [5] computing the 95% confidence band relative to the fitted value $\hat{\bar{p}}_t$ and finally observing if all values of the true consumer price indexp_t fall in the confidence band. In Figure 1 we plot the observed log of the U.S. consumer price index and the 95% confidence band computed using Bai (2004) methodology.

Figure 1 about here

From the figure we can see that the U.S. consumer price index lies inside the confidence intervals throughout the whole period 1918-1995. Thus we can conclude that the US consumer price index is strictly connected to the common nonstationary components. [6]

4 Conclusions

In this paper we analyze the important question recently raised by Cecchetti *et al.*(2002) of whether local price indices share a common trend, and, if so, how quickly they revert to that trend following a local shock. We use the same dataset to focus on Chechetti *et al.'s*(2002) result that local prices revert to a common trend and shocks are temporary, i.e. stationary, but surprisingly persistent, with a half-life of nearly nine years.

We show that for the sample of 19 cities analyzed in Cecchetti et al. (2002) local log-prices can be modeled as composed of at most two unobservable nonstationary common factors and a stationary idiosyncratic component. Looking at the average variance of prices indices we find that the idiosyncratic component accounts for only a small percentage of the total variance with a value no higher than 5%. All the remaining variance is explained by the nonstationary common trend factors. Moreover our half-life estimates of the temporary component are, on average, equal to 5 years, and thus lower than those provided in Cecchetti *et al.*(2002), who report an

[5] We can include a constant in the regression.

[6] The confidence band presented in Fig. 1 has been computed including a constant in the regression and using two factors. The results are robust to the exclusion of the constant and when we include in the regression only the first of the two possible nonstationary factors.

half-life statistic of nine years. Finally we show that the nonstationary components are strongly correlated with the US consumer price index. Thus monetary policies which measure national consumer price inflation based on measurements of aggregate inflation, as those of the European Central Bank have done in recent years with the HICP index, and with less emphasis on the behaviour of local price indices may be justified, given the large effect exerted by the national consumer price index and the relative minor importance of the temporary component in influencing local prices.

References

[1] Bai, J. (2004) Estimating cross-section common stochastic trends in non-stationary panel data, *Journal of Econometrics*, 70, 191-221.

[2] Bai, J. and Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models, *Econometrica*, 70, 191-221.

[3] Bai, J. and Ng, S. (2004). A PANIC Attack on Unit Roots and Cointegration, *Econometrica,* 72, 1127-1177.

[4] Cecchetti, S.G., Mark, N.C., Sonora, R.J. (2002). Price index convergence among United Sates cities, *International Economic Review*, 43, 4, 1081-1099.

[5] Choi, I. (2001). Unit Root Tests for Panel Data, *Journal of International Money and Finance*, 20, pp. 249-272.

[6] Elliott, G., Rothemberg, T.J and Stock, J.H. (1996), Efficient Tests for an Autoregressive Unit Root, *Econometrica*, 64, pp. 813-836.

[7] Gutierrez L. (2006) "Panel unit root tests for cross-sectional correlated panels: A Monte Carlo comparison", forthcoming *Oxford Bulletin of Economics and Statistics.*

[8] Im, K.S., Pesaran, M.H. and Shin, Y. (2003). Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics*, 115,53-74.

[9] Levin, A., Lin C. and Chu C.J. (2002). Unit Root Tests in Panel Data: Asymptotic and Finite-samples Properties, *Journal of Econometrics*, 108, 1-24.

[10] Maddala, G.S. and Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics*, Special Issue, pp. 631-652.

[11] Ng, S., Perron, P. (2002). Lag Length Selection and the Construction of Unit Root Tests With Good Size and Power, *Econometrica*, 69, 1519-1554.

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

[13] Pesaran, H. M. (2006) A simple panel unit root test in the presence of cross section dependence, Unpublished, Cambridge University & USC.

[14] Stock, J.H. and Watson, M.W. (1988). Testing for Common Trends, *Journal of American Statistical Association*, 83, pp. 1097-1107.

Fig.1 Confindence intervals for testing
U.S. consumer price index as a factor.
Dashed line-Confidence intervals, Solid line-log consumer price index

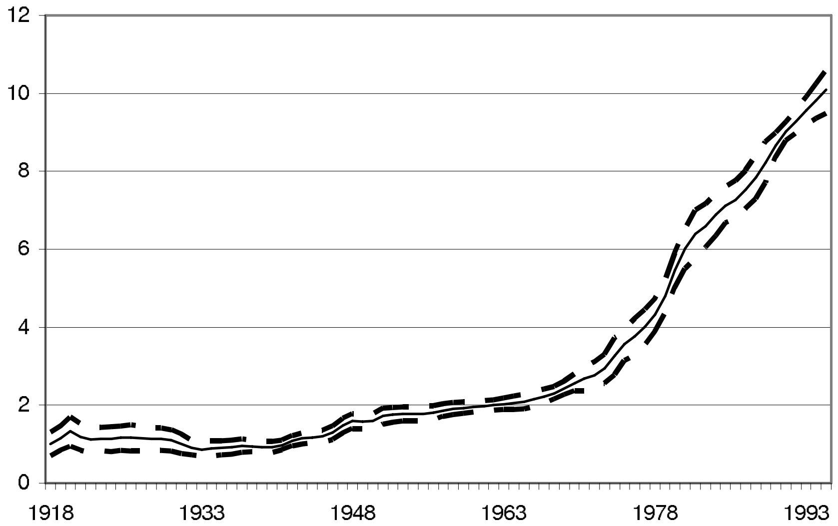

Table 1 Short-run correlation matrix demeaned log price indices of 19 U.S. cities 1918-1995

	NYC	PHI	BOS	CLE	CHI	DET	WDC	BAL	HOU	LAX	SFO	SEA	POR	CIN	ATL	PIT	STL	MIN
PHI	0,75																	
BOS	0,74	0,64																
CLE	-0,27	-0,58	-0,38															
CHI	-0,03	0,41	0,13	-0,20														
DET	-0,23	0,18	-0,18	-0,41	0,39													
WDC	0,52	0,33	0,26	-0,20	-0,39	-0,18												
BAL	0,00	-0,28	-0,29	0,40	-0,55	-0,62	0,02											
HOU	-0,66	-0,55	-0,55	-0,08	-0,24	0,28	-0,31	0,03										
LAX	0,10	0,18	0,29	-0,02	0,55	0,19	-0,36	-0,54	-0,31									
SFO	-0,11	-0,50	-0,15	0,67	-0,46	-0,77	-0,23	0,78	-0,07	-0,15								
SEA	-0,47	-0,72	-0,49	0,54	-0,43	-0,47	-0,31	0,55	0,25	-0,17	0,74							
POR	-0,55	-0,56	-0,62	0,04	-0,43	0,08	-0,23	0,38	0,52	-0,24	0,28	0,68						
CIN	-0,03	0,07	-0,24	0,04	0,00	0,01	0,40	0,00	0,08	-0,47	-0,29	-0,21	-0,23					
ATL	-0,14	0,20	0,17	-0,43	0,40	0,62	-0,07	-0,69	0,08	0,21	-0,62	-0,51	-0,18	-0,19				
PIT	-0,07	0,31	-0,04	-0,31	0,70	0,41	-0,25	-0,43	-0,02	0,13	-0,56	-0,43	-0,32	0,35	0,29			
STL	-0,17	0,24	-0,18	-0,36	0,39	0,39	0,01	-0,29	0,10	-0,11	-0,56	-0,42	-0,24	0,57	0,08	0,63		
MIN	-0,26	-0,60	-0,26	0,60	-0,44	-0,55	-0,02	0,43	0,27	-0,26	0,62	0,47	0,06	-0,01	-0,45	-0,50	-0,28	
KCM	0,01	0,19	0,23	-0,65	0,07	0,49	0,24	-0,66	0,16	0,04	-0,74	-0,52	-0,17	0,05	0,58	0,25	0,39	-0,35

Table 2 Estimated number of factors and nonstationary factors

max=	1		2		3		4		5		6	
	Total number of factors estimated											
	p_{it}	q_{it}	p_{it}	q_{it}	p_{it}	q_{it}	p_{it}	q_{it}	p_{it}	q_{it}	p_{it}	q_{it}
BIC_3	1	1	1	1	1	1	1	2	2	2	3	3
	Total number of nonstationary factors estimated											
	p_{it}	q_{it}	p_{it}	q_{it}	p_{it}	q_{it}	p_{it}	q_{it}	p_{it}	q_{it}	p_{it}	q_{it}
$\widetilde{BIC}_3$	1	1	1	1	1	1	1	1	2	1	2	1
MQ_c	1(*)	1(*)	1	1	1	1	2	2	2	2	2	2

(*) ADF test

Table 3 Panel unit root test statistics idiosyncratic component(*)

	One Factor	Two Factors
	1918 -1995	
P_{ADF}	-16.75(0.00)	-26.03(0.00)
P_{DF-GLS}	-5.93(0.00)	-14.48(0.00)
	1956-1995	
P_{ADF}	-16.10(0.00)	-18.12(0.00)
P_{DF-GLS}	-4.03(0.00)	-5.80(0.00)
	1976-1995	
P_{ADF}	-14.38(0.00)	-20.51(0.00)
P_{DF-GLS}	-1.25(0.11)	-6.66(0.00)

(*) In parentheses the p-values

Table 4 Variance decomposition and half-life

$\left[\left(\sigma^2_{\hat{e}}/\sigma^2_p\right) * 100\right]$			$\ln(0.5)/\ln(\hat{\rho}_i)$		
Avg	Min	Max	Avg	Min	Max
One factor					
1918-1995					
3.4	1.5	7.1	14.9	2.8	53.7
1956-1995					
3.6	1.5	7.1	9.5	1.8	35.2
1976-1995					
6.2	3.7	12.5	4.5	0.8	11.4
Two factors					
1918-1995					
2.4	1.4	3.3	5.2	2.2	10.3
1956-1995					
2.4	1.4	4.7	4.4	1.7	9.5
1976-1995					
4.0	2.3	7.4	2.1	0.7	5.9