Testing for spatial market integration: Evidence for Colombia using a pair-wise approach

Ana María Iregui [1] Jesús Otero [2]

April 2014

[1]Banco de la República, COLOMBIA
[2]Universidad del Rosario, COLOMBIA

Do identical products in spatially separated locations have the same price?

Spatial integration defines the extent of the market, i.e. the area within which producers and consumers determine the price of a product.

Spatial integration has often been assessed through tests of the LOOP.

- Absolute LOOP: Prices of identical products in different locations are the same, once they are converted to a common currency, and after allowing for transportation costs.

- Relative LOOP: Price discrepancies are stationary.

Underlying LOOP is the premise that market participants exploit goods-market arbitrage opportunities by purchasing (selling) a good in one market and selling (purchasing) it in another.

No arbitrage profits are left unexploited.

This paper examines the extent of spatial market integration in Colombia.

The empirical analysis starts off with a time series approach, but in a way that subsequently also employs information from a cross-sectional dimension.

- Use disaggregated CPI data for different cities to assess the validity of (relative) LOOP.
- Adapt the Pesaran (2007) pair-wise approach to analyse market integration. Unit-root tests are conducted on all $(N(N-1)/2)$ price differentials; no need to choose a single reference price.
- Test whether the distance between cities helps explain the speed of adjustment of prices.

Outline

The paper proceeds as follows:

- Provide a brief review of the literature.
- Summarise the econometric modelling strategy.
- Present and discuss the results of the empirical analysis.
- Offer some concluding remarks.

Brief literature review

A large number of authors have empirically tested whether markets are spatially integrated.

For instance, Fackler & Goodwin (2001) survey more than 60 studies with a focus on agricultural products.

However, few have explicitly considered the role of distance as a factor that helps explain the speed of convergence of price discrepancies.

Brief literature review

Engel & Rogers (1996):

- CPI data from 23 North American cities (14 in the US; 9 in Canada) for 14 good categories (1978m6–1994m12).
- Volatility measure of 228 city pairs differentials but no examination of their time series properties.
- Distance between cities and the existence of a border between the US and Canada determine relative price variability.

Parsley & Wei (1996):

- Price data from ACCRA for 51 traded and non-traded goods and services in 48 US cities (1975m1–1992m4).
- Benchmark: New Orleans (New York).
- Price convergence rates are slower for cities farther apart.

Brief literature review

Parsley & Wei (2001):

- 3D-panel with prices of 27 traded goods, across 96 cities in the US and Japan (1976q1–1997q4).
- Benchmark: Louisville (Houston) and Tokushima (Osaka).
- Price dispersion is smaller within countries than between.

See also Baba (2007) for an analysis using Japanese and Korean cities.

- Benchmark: Tokushima and Cheangju.

Brief literature review

Yazgan & Yilmazkuday (2011):

- Apply the Pesaran pair-wise approach to the analysis of price convergence in the US.
- ACCRA price data (1990q1–2007q4).
- Price-level convergence rates depend on the type of good.
- However, the role of distance (if any) is not examine.

Key feature: We extend the Pesaran pair-wise analysis by testing whether the speed of adjustment of price differentials is positively related to the distance between cities.

Econometric modelling strategy

Pesaran's pair-wise approach

Let:

- $p_{i,k,t}$ be the price of good k in city i at time t, and
- $g_{ij,k,t} = p_{i,k,t} - p_{j,k,t}$ be the price differential in city pair (i, j).

The pair-wise approach examines the order of integration of $g_{ij,k,t}$, where $i = 1, ..., N - 1$, $j = i + 1, ..., N$, and N is the total number of cities.

Advantage: This approach does not involve choosing a "base city" for the computation of the price differentials.

Econometric modelling strategy

Pesaran's pair-wise approach

We consider two unit root tests:

- ADF of Dickey and Fuller (1979)
- $\text{ADF}_{\text{máx}}$ of Leybourne (1995)

We apply the ADF ($\text{ADF}_{\text{máx}}$) test to $g_{ij,k,t}$, and define $Z_{ij,k} = 1$ if the unit-root null is rejected, at a significance level α.

Next, let us define $\bar{Z}_k$ as the fraction of rejections:

$$\bar{Z}_k = \frac{2}{N\,(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} Z_{ij,k}, \tag{1}$$

Pesaran shows that under the null of divergence (i.e. non-stationarity), the expected value of $\bar{Z}_k$ must be equal to the nominal size of the underlying unit root test statistic, α.

By contrast, under the null of convergence, the proportion of rejections $\bar{Z}_k$ is expected to be high and approach 100 % as $T \to \infty$.

Econometric modelling strategy

Developing further Pesaran's pair-wise approach

What does explain the speed of adjustment of price differentials?

To address this question:

- Consider the cases for which $g_{ij,k,t}$ is stationary (using the ADF test).
- Estimate the half-life of a shock based on $hl_{ij,k} = -\ln(2/(1+\hat{\delta}))$, where $\hat{\delta}$ is the AR coefficient in the corresponding ADF test regression.

Econometric modelling strategy

Developing further Pesaran's pair-wise approach

Lastly, estimate the model:

$$\ln(hl_{ij,k}) = \beta_1 + \beta_2 \ln(popd_{ij}) + \beta_3 \ln(X_{ij}) + \text{Product Dummies} + \varepsilon_{ij,k}, \quad (2)$$

where:

- X_{ij} is some measure of distance between cities (i.e. $dist_{ij}$ or $durt_{ij}$)
- $\ln popd_{ij} = |\ln pop_i - \ln pop_j|$
- Product dummy variables are also included

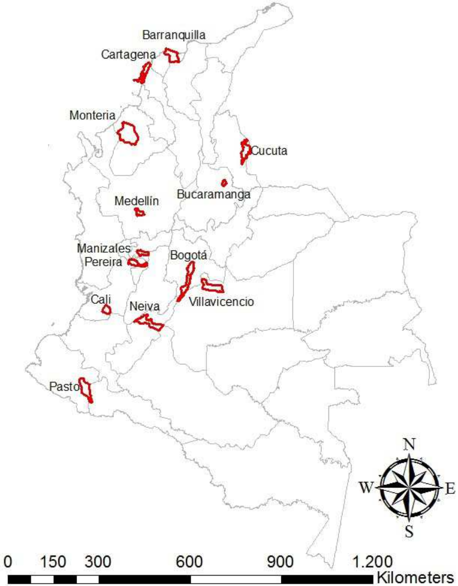
Data

We use CPI monthly data for 153 products in 13 cities:

- ⋆ Bogotá
- ⋆ Medellín
- ⋆ Cali
- ⋆ Barranquilla
- ⋆ Cartagena
- ⋆ Cúcuta
- ⋆ Bucaramanga
- ⋆ Pereira
- ⋆ Villavicencio
- ⋆ Pasto
- ⋆ Monteria
- ⋆ Manizales
- ⋆ Neiva

Sample period is 1999m1 - 2013m2 ($T = 170$). The source is DANE.

Location of the 13 main cities in Colombia



Data

To facilitate the presentation of our results, we adopt the product classification used by the Banco de la República in its quarterly inflation report:

- Food (54)
  - Unprocessed (17)
  - Processed (37)
- Traded (71)
- Non-traded (28)

The products under consideration account for more than 90 % of the overall CPI (excluding products whose prices are regulated, e.g. utilities, petrol and bus fares, among others).

Implementing Pesaran's pair-wise approach

- Unit root regressions include trend if significant at the 5 % level
- Number of lags selected using AIC ($p_{max} = 3$)
- ADF critical values from Cheung & Lai (1995)
- ADF_{max} critical values from Otero & Smith (2012)

Proportion of stationary price differentials

Summary table

Category	ADF	ADF_{max}
Unprocessed food	0.860	0.883
Processed food	0.536	0.481
Tradeable	0.359	0.250
Non-tradeable	0.396	0.240

For each category, the proportion of stationary price differentials corresponds to the simple average of the proportion of stationary price differentials of the corresponding products.

Estimated half-life of a shock

Summary results by product category

Category	Duration in months				
	hl	0 to 3	3 to 6	6 to 12	12 to +
Unprocessed food	3.4		$\star$		
Processed food	5.5		$\star$		
Traded	8.0			$\star$	
Non-traded	8.1			$\star$	

The half-life of each category is the simple average of the half-lives of the individual products.

Determinants of the half-life of inter-city price differentials

	Food							
	Unprocessed		Processed		Traded		Non-traded	
Intercept	0.361	0.392	1.090	1.148	2.020	2.054	2.256	2.251
	(0.169)	(0.157)	(0.145)	(0.135)	(0.197)	(0.192)	(0.327)	(0.311)
ln $durt_{ij}$	0.086		0.100		0.051		-0.027	
	(0.025)		(0.021)		(0.020)		(0.040)	
ln $dist_{ij}$		0.082		0.092		0.046		-0.027
		(0.023)		(0.020)		(0.018)		(0.037)
ln $popd_{ij}$	-0.029	-0.030	0.008	0.007	0.011	0.011	0.007	0.007
	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.019)	(0.019)
Obs.	1141	1141	1546	1546	1990	1990	864	864
R^2	0.554	0.554	0.225	0.225	0.228	0.228	0.273	0.273
F-stat.	94.030	93.918	11.550	11.509	11.459	11.466	21.328	21.299

The numbers in parentheses are White's heteroskedastic consistent standard errors. All regression include product-specific dummy variables.

Conclusions

We use disaggregated CPI data for 153 goods to investigate the extent of market integration in Colombia.

The empirical exercise draws on a pair-wise time series approach, but in a way that also utilises information from a cross-section dimension.

- Examine the time series properties of all possible city-pair price differentials.
- Test whether the distance between cities is a factor that explains the speed of adjustment of prices.

Two alternative measures of distance:

- Kilometers through the road network.
- Time it takes to go from one city to another.

Conclusions

Main findings regarding the (relative version of the) LOOP:

1. Strong support for unprocessed and processed foods.
2. Less support for non-traded and traded goods.

Main findings regarding speed of adjustment:

1. Distance (regardless of how it is measured) is a key factor for unprocessed, processed and traded goods.
2. Speed of adjustment of unprocessed food prices increases when cities are of different sizes.