Measuring Spatial Integration in Tomato and Onion Markets of Pakistan:  
An Application of Error Correction Model in the Presence of Stationarity

Heman D. Lohano  
Department of Economics and Finance  
Institute of Business Administration, Karachi, Pakistan  
loha0002@umn.edu

Fateh M. Mari  
Water Sector Improvement Project  
Planning and Development Department  
Government of Sindh, Karachi, Pakistan  
fatehpk@yahoo.com

Selected Paper prepared for presentation at the Agricultural & Applied Economics 

Copyright 2012 by Heman D. Lohano and Fateh M. Mari. All rights reserved. Readers may 
make verbatim copies of this document for non-commercial purposes by any means, provided 
that this copyright notice appears on all such copies.
Measuring Spatial Integration in Tomato and Onion Markets of Pakistan: An Application of Error Correction Model in the Presence of Stationarity

Heman D. Lohano and Fateh M. Mari

Abstract

This paper empirically evaluates the degree of integration among spatially separated markets using monthly wholesale real price series of tomato and onion from regional markets located in each of the four provinces of Pakistan. Unit root test indicates that the price series are stationary, and are represented as autoregressive model for each region. Results of the error correction model show that the markets of both tomato and onion are spatially integrated, however, the overall degree of integration in onion markets is relatively stronger than in tomato markets. This difference may be attributed to relatively lower degree of perishability in onion as compared to tomato.

Introduction

Markets are said to be perfectly spatially integrated if price changes in one market are fully reflected in alternative markets (Goodwin and Schroeder 1991). Prices are the signals that direct and coordinate not only the production and consumption decisions but also the marketing decisions over time, form, and space (Kohls and Uhl 2001). In spatially separated markets, when the price difference between different markets exceeds transportation and transactions costs, the arbitrage activities involve the purchase of commodities from lower-price regional markets and the subsequent resale in higher-price regional markets. Competition among arbitragers will ensure that a unique equilibrium is achieved where local prices in regional markets differ by no more than transportation and transaction costs (Goodwin and Schroeder 1991). Prices in spatially integrated markets are determined simultaneously in various locations, and information of any change in price in one market is transmitted to other markets (Gonzalez-Rivera and Helfand 2001). Moreover, the improved information between regional markets contributes significantly to spatial price convergence, so explicit trade between each pair of markets may not be necessary in order for regional price adjustments to take place (Serra et al. 2006). The analysis of spatial market integration, thus, provides indication of competitiveness, the effectiveness of arbitrage, and the efficiency of pricing (Sexton, Kling, and Carman 1991).

Markets that are not integrated may convey inaccurate price signal that may result in distortion in producers’ marketing decisions and inefficient product movement (Goodwin and Schroeder 1991), and traders may exploit the market and may benefit at the cost of producers and consumers. In more integrated markets, farmers specialize in production activities in which they are comparatively proficient, consumers pay lower prices for purchased goods, and society is better able to reap increasing returns from technological innovations and economies of scale (Vollrath 2003).

Spatial market integration of agricultural products has been widely used to indicate overall market performance (Faminow and Benson 1990). Market integration of agricultural products has retained importance in developing countries due to its potential application to policy making.
Based on the information of the degree of market integration, government can formulate policies of providing infrastructure and information regulatory services to avoid market exploitation.

Agricultural products especially vegetables are very perishable in nature, and are supplied to market within a short time period after harvesting. Tomato and onion are the most common vegetables in Pakistan and other South Asian countries. The demand for tomato and onion is relatively inelastic in Pakistan, where it is used in cooking with other vegetables and meat in addition to consumed as salad. Due to inelastic demand and perishable nature of tomato and onion, we observe frequent price variations as well as trade between regional markets depending on their supply position in each market. Although tomato and onion are produced in each of the four provinces of Pakistan (Sindh, Punjab, Khyber Pakhtunkhwa, and Balochistan), these products are frequently traded among these provinces mainly due to variation in the supply in each province.

Trading perishable products from one province to another involves risk of product losses in quality as well as in quantity due to long distance transportation without refrigerating or canning them. Higher risk associated with trading more perishable products is likely to induce risk averse traders to involve in trading less risky products. Tomato is more perishable in nature than onion. For policy evaluation, it is important to compare the degree of spatial market integration between tomato and onion.

The objective of this paper is to analyze spatial market integration among four regional markets of Pakistan using monthly wholesale real price series of tomato and onion. First we apply the unit root test to check for the stationarity in the price series, and then evaluate the degree of spatial integration among the regional markets for each tomato and onion using the error correction model. The results for tomato and onion are then compared to investigate how far these markets diverge from perfect spatial integration. The rest of the paper is organized as follows. The next section describes data used in this study, followed by sections on the model and estimation methods. We then present the empirical results of spatial market integration. Finally, the last section draws conclusion.

**Data**

Four regional markets including Hyderabad, Lahore, Peshawar, and Quetta are selected for this study, as these cities are large primary distributing centers of vegetables in the country, and are taken from each of the four provinces of Pakistan including Sindh, Punjab, Khyber Pakhtunkhwa, and Balochistan, respectively. Data used in this study are monthly wholesale price in rupees (Rs.) per 40 kilogram of tomato for the period from January 1988 to June 2010, and onion for the period from July 1981 to June 2010 published in various issues of *Fruits, Vegetables, Condiments Statistics of Pakistan* and *Agricultural Statistics of Pakistan* (Government of Pakistan 1998a, 1998b, 2006a, 2006b). The nominal prices of tomato and onion are transformed into the real prices by deflating them using the monthly Wholesale Price Index published in various issues of *Pakistan Economic Survey* (Government of Pakistan 1993, 1998c, 2011c). Monthly wholesale real prices (in December 2005 rupees) of tomato and onion in the four regional markets including Hyderabad, Lahore, Peshawar, and Quetta are plotted in Figures 1 and 2 in Appendix. Data indicate a large volatility in the price in each market and overall no
trend in the series. These prices are volatile across time due to supply shocks, perishable nature, storage costs, and relatively inelastic demand for tomato and onion.

The Model

If the geographically separated markets are integrated, then there exists an equilibrium relationship among these markets (Goodwin and Schroeder 1991; Gonzalez-Rivera and Helfand 2001; Sexton, Kling, and Carman 1991). The long-run equilibrium relationship for analyzing spatial market integration used in the previous studies (e.g., Goodwin and Schroeder 1991) is specified as:

\[ P_t^1 = \alpha + \lambda P_t^2 \]  

(1)

where \( P_t^1 \) and \( P_t^2 \) represent commodity prices of a homogenous good in two alternative regional markets at time \( t \), and \( \alpha \) and \( \lambda \) are unknown parameters. If two markets are perfectly spatially integrated, then \( \lambda = 1 \). In this case, price changes in one market are fully reflected in the alternative market. When \( \lambda \neq 1 \) (\( \lambda < 1 \) or \( \lambda > 1 \)), then the degree of integration may be evaluated by investigating how far is the deviation of \( \lambda \) from unity.

The above model of long-run relationship between markets, however, may not satisfy at each time period. To integrate short-run dynamics with long-run relationship, we specify the error correction model representation of Equation (1) as:

\[ \Delta P_t^1 = \beta_0 + (\beta_1 - 1)(P_{t-1}^1 - \alpha - \lambda P_{t-1}^2) + \gamma_0 \Delta P_t^2 + \varepsilon_t \]  

(2)

where \( \Delta P_t^i \) represents change in the price \( (P_t^i - P_{t-1}^i) \) at region \( i = 1, 2 \); \( \beta_0 \), \( \beta_1 \) and \( \gamma_0 \) are unknown parameters, and \( \varepsilon_t \) is the error term. In this model, \( (P_{t-1}^1 - \alpha - \lambda P_{t-1}^2) \) measures the extent to which the long-run relationship is not satisfied at time period \( t-1 \). The parameter \( (\beta_1 - 1) \) is interpreted as the proportion of the resulting disequilibrium adjusted in the next period. Therefore, the term \( (\beta_1 - 1)(P_{t-1}^1 - \alpha - \lambda P_{t-1}^2) \) is the error correction term. The adjustment process makes sense if \( 0 \leq \beta_1 < 1 \). When \( \beta_1 \) is close to 0, the speed of adjustment to long-run equilibrium is very fast. When \( \beta_1 \) is close to 1, the speed of adjustment is very slow (Davidson and MacKinnon 2004).

Estimation Methods

Appropriate method for estimating the error correction model of spatial market integration specified in Equation (2) depends on the underlying stochastic process of price series in each region. If the price series are nonstationary with unit root, under certain conditions the equation may be estimated as a cointegration model developed by Granger (1983) and Engle and Granger (1987). However, the price series in this study indicate stationarity, which is checked by the Augmented Dickey-Fuller (ADF) test for unit root. Thus, Equation (2) is the error correction model with stationary time series. This section briefly describes the Augmented Dickey-Fuller
(ADF) test for unit root, and the method for estimating the error correction model in the presence of stationarity (Davidson and MacKinnon 2004).

Unit Root Test

Stationarity of price series is checked by the Augmented Dickey-Fuller (ADF) test for unit root. Data show that the price series of both tomato and onion have no trend as illustrated Figures 1 and 2 in Appendix, and the average value of price series is positive for each region. In this case, the null hypothesis (H₀) in the ADF test is unit root autoregressive model with no drift, and the alternative hypothesis (Hₐ) is autoregressive model with constant term (Hamilton 1994). The ADF test is carried out by estimating the following equation:

\[ P_t = \delta + \phi P_{t-1} + \theta_1 \Delta P_{t-1} + \theta_2 \Delta P_{t-2} + \ldots + \theta_{p-1} \Delta P_{t-p+1} + e_t \]  \hspace{1cm} (3)

where \( P_t \) is commodity price at time \( t \) at a region, \( \Delta P_t \) represents change in the price \( (P_t - P_{t-1}) \), \( \delta, \phi, \theta_1, \theta_2, \ldots, \theta_{p-1} \) are unknown parameters, \( p \) is the order of autoregressive model, and \( e_t \) is the error term.

Using the ADF test, we test the null hypothesis that \( \delta = 0 \) and \( \phi = 1 \). Under the null hypothesis, Equation (3) is a unit root autoregressive model with no drift, and the time series is nonstationary. The alternative hypothesis is stationary autoregressive model with constant term. The OLS F test is performed for testing the joint null hypothesis that \( \delta = 0 \) and \( \phi = 1 \). The OLS F statistic is computed as follows:

\[ F = \frac{(\hat{e}_e \cdot \hat{e}_e - \hat{e}_\hat{e})/(k - k_e)}{\hat{e}_\hat{e}/(n - k)} \]  \hspace{1cm} (4)

where \( \hat{e}_e \) is vector of residuals from \( H_0 \), \( \hat{e}_\hat{e} \) is residual sum of squares (RSS) from \( H_0 \), and \( k_e \) is its number of parameters. Similarly, \( \hat{e}_\hat{e} \) is vector of residuals from \( H_\lambda \), \( \hat{e}_\hat{e} \) is RSS from \( H_\lambda \), \( k \) is its number of parameters, and \( n \) is the number of sample observations. For this test, the OLS F statistic is compared with the critical values provided by Dickey and Fuller (1981). If the F statistic is larger than the critical value, we reject the null hypothesis.

For performing the above unit root test, the order of autoregressive model, \( p \), must be specified in estimating Equation (3). The appropriate order of autoregressive model is such that the error term \( e_t \) is a white noise process. The Ljung-Box test is conducted for checking that the error term is a white noise process (Ljung and Box 1979).

Estimation of Error Correction Model

For describing the method of estimating the error correction model specified in Equation (2), we denote \( y_t \equiv P_t^1 \) and \( x_t \equiv P_t^2 \), and rewrite the equation as:
\[ \Delta y_t = \mu + (\beta_1 - 1)(y_{t-1} - \lambda x_{t-1}) + \gamma_0 \Delta x_t + \epsilon_t \]  \hspace{1cm} (5)

where \( \mu \equiv \beta_0 - \alpha (\beta_1 - 1) \). Although Equation (5) is nonlinear in parameters, its reparameterization yields a linear equation. Denote:

\[ \lambda \equiv \frac{\gamma_0 + \gamma_1}{1 - \beta_1} \]  \hspace{1cm} (6)

Then, Equation (5) can be written as:

\[ y_t = \mu + \beta_1 y_{t-1} + \gamma_0 x_t + \gamma_1 x_{t-1} + \epsilon_t \]  \hspace{1cm} (7)

The model in Equation (7) is referred to as an autoregressive distributed lag (ADL) model, and its point estimates may be computed by the ordinary least squares (OLS) method when both variables are stationary (Davidson and MacKinnon 2004; Enders 2004). As lagged dependent variable is one the regressors in Equation (7), the regressors are only contemporaneously independent of the error term, and are not independent of the error term at each time period. Therefore, the OLS estimator can be justified asymptotically (Hamilton 1994, p. 215).

The point estimates of nonlinear Equation (5) can now be computed using Equation (6) and point estimates of Equation (7). For testing the significance of estimated parameters, the variance of parameters in Equation (5) is estimated using the Gauss-Newton regression (GNR) method as suggested by Davidson and MacKinnon (2004, p. 579). By the GNR approach, we specify the following regression:

\[ \hat{\epsilon}_t = Xd_t \; b + \nu_t \]  \hspace{1cm} (8)

where \( \hat{\epsilon}_t \) is the residual from Equation (7) (which are equal to residuals from Equation (5)), \( Xd_t \) is a vector of partial derivatives of nonlinear regression function in Equation (5) with respect to its parameters, \( b \) is a vector of parameters, and \( \nu_t \) is the error term. Equation (8) is estimated by the OLS method, which will yield \( \hat{b} = 0 \) and can have no explanatory power. The interest of running this regression is to estimate the variance of \( \hat{b} \), \( \hat{\text{Var}}(\hat{b}) \), which is identical to the variance of parameter estimates of Equation (5) as shown in Davidson and MacKinnon (2004, p. 239).

As time series data are used, there may be serial correlation. So, in implementing the GNR method, we estimate heteroscedascity and autocorrelation consistent covariance estimator developed by Newey and West (1987). The Newey-West estimator is a robust estimator for the covariance of the OLS estimator, and constitutes the generalized method of moments (GMM) estimator (Greene 2012, p. 939).
Deviation from Perfect Spatial Integration

If markets are perfectly spatially integrated, then the market integration parameter $\lambda = 1$, which indicates that the price changes in one market are fully reflected in the alternative market, as mentioned in the model. When $\lambda \neq 1$ ($\lambda < 1$ or $\lambda > 1$), then the degree of integration is evaluated by investigating how far is the deviation of $\lambda$ from unity. After estimating the error correction model, the deviation is computed as the absolute value of the difference between 1 and the estimated value of the market integration parameter:

$$\text{Deviation} = |1 - \hat{\lambda}|$$  \hspace{1cm} (9)

The average value of this deviation is computed for tomato market and onion market each to compare the degree of integration between them.

Empirical Results

Results of Unit Root Test

Stationarity of the wholesale real price series of tomato and onion for each region is checked by the Augmented Dickey-Fuller (ADF) test of unit root. In this test, the null hypothesis is a unit root autoregressive model with no drift, and the alternative hypothesis is stationary autoregressive model with constant term. The $F$ test is performed for testing the null hypothesis. Results of the Augmented Dickey-Fuller test are presented in Table 1.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{Market (City)} & \textbf{Order of Autoregression} & \textbf{$F$ statistics} & \textbf{Order of Autoregression} & \textbf{$F$ statistics} \\
\hline
Hyderabad & 3 & 15.29 & 4 & 29.75 \\
Lahore & 2 & 29.58 & 2 & 29.05 \\
Peshawar & 2 & 34.69 & 2 & 31.97 \\
Quetta & 2 & 32.14 & 2 & 33.02 \\
\hline
\end{tabular}
\caption{Augmented Dickey-Fuller Test for Unit Root}
\end{table}

The 5 % critical value is 4.63 and 1 % critical value is 6.52 for each case.

For performing the above unit root test, the order of autoregressive model must be specified in estimating the model. The order of autoregressive model, AR($p$), is determined by checking that the error term is a white noise process using the Ljung-Box test. Results in Table 1 show that tomato price series are described as AR(3) for Hyderabad, and AR(2), for Lahore, Peshawar, and Quetta, whereas onion price series are described as AR(4) for Hyderabad, and AR(2) for Lahore, Peshawar, and Quetta. Table 1 also reports the $F$ statistics for testing the null hypothesis. For each region, the results show that the null hypothesis of unit root is rejected as the $F$ statistic is
much greater than 6.52, which is the critical value at 1% significance level provided in Dickey and Fuller (1981) as reported in Hamilton (1994). Thus, these results indicate that the wholesale real price series of tomato and onion in each region represent stationary autoregressive model with constant term.

**Empirical Results of Market Integration**

Spatial market integration is analyzed by estimating the price relationship between spatially separated markets using the error correction model specified in Equation (5). Given four regional markets, each market has been regressed with the other market from the remaining three markets. In this way, a total number of 12 regressions of the error correction model are run for each tomato and onion, and are presented in Table 2.

**TABLE 2**

<table>
<thead>
<tr>
<th>Results of Error Correction Model for Tomato and Onion Markets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depend. variable (Price in market)</td>
</tr>
<tr>
<td>-----------------------------------</td>
</tr>
<tr>
<td><strong>Results for Tomato</strong></td>
</tr>
<tr>
<td>Hyderabad</td>
</tr>
<tr>
<td>Hyderabad</td>
</tr>
<tr>
<td>Hyderabad</td>
</tr>
<tr>
<td>Lahore</td>
</tr>
<tr>
<td>Lahore</td>
</tr>
<tr>
<td>Lahore</td>
</tr>
<tr>
<td>Peshawar</td>
</tr>
<tr>
<td>Peshawar</td>
</tr>
<tr>
<td>Peshawar</td>
</tr>
<tr>
<td>Quetta</td>
</tr>
<tr>
<td>Quetta</td>
</tr>
<tr>
<td>Quetta</td>
</tr>
<tr>
<td><strong>Results for Onion</strong></td>
</tr>
<tr>
<td>Hyderabad</td>
</tr>
<tr>
<td>Hyderabad</td>
</tr>
<tr>
<td>Hyderabad</td>
</tr>
<tr>
<td>Lahore</td>
</tr>
<tr>
<td>Lahore</td>
</tr>
<tr>
<td>Lahore</td>
</tr>
<tr>
<td>Peshawar</td>
</tr>
<tr>
<td>Peshawar</td>
</tr>
<tr>
<td>Peshawar</td>
</tr>
<tr>
<td>Quetta</td>
</tr>
<tr>
<td>Quetta</td>
</tr>
<tr>
<td>Quetta</td>
</tr>
</tbody>
</table>

Numbers in parentheses are t-statistics of estimated parameters. The 5% critical value is 1.96, and 1% critical value is 2.576 for t test in each case.
Table 2 presents the estimates of parameters of the error correction model including $\mu$, $\beta_1$, $\gamma_0$, and $\lambda$, their t statistics, and adjusted $R^2$ of these regressions. In this model, $\mu$ is the intercept, $\beta_1$ is parameter that measures the speed of adjustment to long-run equilibrium, $\gamma_0$ is the slope on $\Delta x_i$, and $\lambda$ is market integration parameter. If the two markets are perfectly spatially integrated, the value of spatial integration parameters $\lambda$ is 1, as discussed in the model. Results for tomato market in Table 2 show the estimated value of $\lambda$ is 1.06 in the regression of Lahore on Peshawar. This indicates that when tomato price changes by one rupee in Peshawar, it also changes in Lahore by Rs. 1.06. In the regression of Peshawar on Quetta, the estimated value of $\lambda$ is 0.95. Results on these two pairs of tomato markets show very strong spatial market integration. Result for onion market in Table 2 show that the estimated value of $\lambda$ is 0.99 in the regression of Peshawar on Quetta, and in the regression of Quetta on Lahore. These pairs of markets exhibit very strong spatial market integration.

For evaluating and comparing the degree of spatial integration between tomato and onion, the deviation from perfect spatial integration is computed as the absolute value of the difference between 1 and the estimated value of the market integration parameter. The results show that the average deviation is 0.25 (25 percent) in tomato markets, and 0.20 (20 percent) in onion markets. These results indicate that the overall degree of integration in tomato and onion markets is 75 and 80 percent, respectively. Thus, the overall degree of integration in onion markets is relatively stronger than in tomato markets. This difference may be attributed to relatively lower degree of perishability in onion as compared to tomato.

Table 2 also presents the estimates of adjustment parameter $\beta_1$. When $\beta_1$ is close to 1, the speed of adjustment to long-run equilibrium is very slow. When $\beta_1$ is close to 0, the speed of adjustment is very fast. Results show that the estimated parameter ranges from 0.21 to 0.54 for tomato markets, and from 0.46 to 0.70 for onion markets. These results indicate that the speed of adjustment to the long-run equilibrium is moderate in both markets, but it is relatively slower in onion markets despite relatively strong spatial market integration in the long-run equilibrium.

**Conclusion**

Spatial market integration has been examined by estimating the price linkages among geographically separated tomato and onion markets of Pakistan. Data used for the analysis are monthly wholesale real price series in four regional markets, namely Hyderabad, Lahore, Peshawar, and Quetta cities, which are taken from each of the four provinces of Pakistan including Sindh, Punjab, Khyber Pakhtunkhwa, and Balochistan, respectively.

For each location, the units root test indicates that the price series are stationary, and are represented as autoregressive model. Spatial price linkages between regions are evaluated by estimating the error correction model in the presence of stationarity. Results of the model show that the markets of both tomato and onion exhibit the strong spatial integration where most of the trade takes place. The overall degree of integration in onion markets is relatively stronger than in tomato markets. This difference may be attributed to relatively lower degree of perishability in onion as compared to tomato.
References


Figure 1: Monthly Wholesale Real Prices (Rs./40kg) of Tomato from Jan. 1988 to Jun. 2010

Figure 2: Monthly Wholesale Real Prices (Rs./40kg) of Onion from Jul. 1981 to Jun. 2010