The Extent, Pattern, and Degree of Market Integration: A Multivariate Approach for the Brazilian Rice Market

Gloria González-Rivera and Steven M. Helfand

The extent, pattern, and degree of integration are analyzed in a multivariate system with cointegrating restrictions. The extent of the market is found by identifying locations that are linked by trade and where prices share identical long–run information (permanent component). The pattern of integration characterizes interdependence and is analyzed by estimating a vector error correction model. The degree of integration is calculated with persistence profiles of the long run relations. We demonstrate that bivariate models are inadequate for capturing the spatial dynamics of price adjustment. The methodology is applied to the Brazilian rice market and policy implications are discussed.

Key words: Brazil, market integration, permanent component, persistence profiles, rice.

Prices in an integrated spatial market are determined simultaneously in numerous locations. An important empirical question, with relevance for the spatial design of economic policy, is how the information contained in prices is transmitted from one location to another in the short and long run. The multi-location nature of the market suggests that a multivariate approach is necessary for answering this question. Nevertheless, most studies of market integration have employed a bivariate approach. This is true of studies based on linear cointegration and of those that use switching regime methodologies.[1] The objective of this artlce is to illustrate the advantages of a multivariate analysis and to stress the limitations of a bivariate approach to market integration. Within the framework of multivariate cointegration, this article introduces two novel features to the analysis of market integration: (1) the search for the geographic boundaries of the market, and (2) the use of persistence profiles to study the degree of integration of different locations that belong to the market.

While there is general agreement that market integration somehow relates to the flow of goods and information across space, time, and form, the provision of a widely accepted definition with testable components has proven to be an elusive task. We propose a definition that relies on two related dimensions: trade and information. For a market to be called integrated, we require that the set of locations share both the same traded commodity and the same long run information. In a cointegration framework, this second condition is equivalent to requiring the existence of one and only one integrating factor that is common to all series of prices. Thus, given a set of locations, we propose a sequential procedure based on Johansen (1988, 1991) to search for the single common factor. This multivariate search for the extent of the market differentiates our article from previous studies of market integration.

A single common factor implies that there must be $n-1$ cointegrating vectors in an n location market. If we were to normalize the $n-1$ cointegrating vectors with respect to a given location, we would find that all loca-

Gloria González-Rivera is associate professor and Steven M. Helfand assistant professor in the Department of Economics at the University of California, Riverside.

This article was written as part of a larger project on spatial economics called NEMESIS (*www.nemesis.org.br*). The project is based at IPEA in Rio de Janeiro and has been supported by a grant from the Brazilian government under a program called PRONEX. We also thank the Center for Agricultural Studies at the Getúlio Vargas Foundation for providing us with the producer prices, Jesús Gonzalo for useful conversations and computer code that assisted us in estimating the permanent component, Chris Barrett, two anonymous referees, and Kusum Mundra for valuable research assistance.

[1] Among the most important bivariate contributions to this literature are Ravallion; Goodwin and Schroeder; Sexton, King and Carman; Alexander and Wyeth; Dercon; and Baulch. Two notable multivariate exceptions are Goodwin; and Asche, Bremnes, and Wessels.

Amer. J. Agr. Econ. 83(3) (August 2001): 576–592

tions were cointegrated pair-wise. However, this is not sufficient to justify a bivariate analysis of the market for at least two reasons. First, it would be very difficult to determine which locations belong to the same market with a bivariate approach. Of the $n(n-1)/2$ pairwise combinations, only $n-1$ are relevant. The exercise would be unnecessarily complicated and would likely lead to inconclusive results. Second, a cointegrated system can be written as a vector error correction (VEC) model. In a system with n locations, each equation of the VEC is likely to contain error correction terms and lags from numerous other locations in the market. A bivariate model necessarily restricts each equation of the VEC to have at most one error correction term, and lags only from the two states considered. In all but very special market structures, this would grossly misspecify the model.

Most of the literature on market integration has focused on estimating the cointegrating vectors. Since the integrating factor is eliminated when the cointegrating relation is estimated, no attention has been paid to finding the common long run component that gives rise to cointegrated prices. In this paper, we present the estimation of the integrating factor according to the methodology proposed by Gonzalo and Granger. This methodology is particularly attractive because the common factor is associated with observable variables and allows for the identification of the location(s) that contribute to the long run behavior of prices.

Finally, we propose to study market integration as a question of degree. At one end of the continuum are locations that do not belong to the market. Within the market, we seek to offer a ranking of all locations from less to more integrated. We define the degree of integration between locations that belong to the same market as the reaction time to remove disequilibria. A measure of reaction time that is commonly used in the literature are impulse response functions. An important limitation of these functions is that they are not uniquely identifiable when shocks to the system are correlated. In a study of spatial prices it is unreasonable to expect uncorrelated errors because the time series of prices are highly correlated. The usual "solution"—a Cholesky decomposition—imposes a recursive ordering on the variables in the system. The impulse response functions, however, are not invariant to the ordering. For every order we would calculate a different reaction path. We propose to use a different measure that is robust to any ordering of the variables in the system. It was developed by Pesaran and Shin and is called a persistence profile. A persistence profile characterizes the response of a cointegrating relation to a system-wide, rather than to an individual, shock. It measures the reaction time of each long run equilibrium relation to absorb a system-wide shock. Persistence profiles are unique functions that allow us to quantify the degree of integration of all locations that belong to the same economic market.

The article is organized as follows. First, we describe the methodology for estimating the extent, pattern, and degree of integration. In the following section, we apply the methodology to the Brazilian rice market for the period 1970–1997. Finally, we provide conclusions.

Characteristics of an Integrated Market

The Extent of the Market

There is general agreement that market integration somehow relates to the flow of goods and information across space, time, and form. The provision of a widely accepted definition, however, has proven to be an elusive task. To avoid confusion, we begin by defining explicitly what we mean by market integration.

A market with n geographically distinct locations will be considered integrated if the following two conditions are satisfied:

(1) There must be physical flows of goods connecting all n locations either directly or indirectly.

(2) The n locations must have a corresponding vector of prices $\{p_{1t}, p_{2t}, \ldots, p_{nt}\}$ that can be decomposed as $p_{it} = a_i f_t + \tilde{p}_{it},\ i = 1, \ldots, n$, and $a_i \neq 0$, where f_t is the integrating factor that characterizes the permanent (long run) component of the price, and $\tilde{p}_{it}$ is the transitory (short run) component for each location.

The basic elements of this definition are the existence of trade and that f_t is common to all series of prices. The physical flow of goods via trade is important to ensure that arbitrage occurs, but by itself does not guarantee integration because there may exist markets with very thin or intermittent trade for which a common integrating factor for all i and t does not exist. Similarly, the existence of an integrating factor, by itself, does not ensure integration because there may be physically isolated markets that exhibit co-movements of prices that result from seasonal patterns or policies. The definition does not imply that all the participating locations process the relevant information simultaneously. It does require that all locations should be connected either directly or indirectly through trade and long run information.

Our definition provides an operational framework to search for the extent, or geographic boundaries, of an integrated market. The first step is to identify the set of locations that is connected either directly or indirectly through continuous unidirectional trade. Because domestic trade data do not exist for many developing countries, we begin by estimating annual trade flows for each location in the market. This allows us to exclude locations that experience trade reversals (exporters that become importers and vice versa). It also permits us to identify locations that are close to self-sufficiency and are thus candidates for experiencing discontinuous trade.[2] Once we identify the set of locations that is tied together through trade, we begin the search for those states that share a common integrating factor.

A novel feature of this article is our focus on the relevance and implications of searching for a single common integrating factor. Most of the literature on market integration has focused on estimating and testing the cointegrating vectors, thus neglecting the information contained in the integrating factor(s). Cointegrating vectors and integrating factors, however, are intimately related. The existence of *one and only one* integrating factor for all prices implies that (1) prices must be cointegrated, and (2) there must be $n-1$ cointegrating vectors. If transactions costs are non-stationary, then $n-1$ cointegrating vectors must be found when prices are measured net of transactions costs. A formal analysis of the implications of a single integrating factor follows.

Consider an $n \times 1$ non-stationary $I(1)$ vector of log-prices $P_t = \{p_{1t}, p_{2t}, \ldots, p_{nt}\}$ where p_{it} is the log-price of a commodity at time t in market i. Suppose that P_t can be decomposed into two components as follows

$$(1) \quad P_t = A_{n\times s} f_t + \widetilde{P}_t$$

where f_t is an $s \times 1$ vector of $s(s < n)$ common unit root factors and $\widetilde{P}_t$ is an $n \times 1$ vector of stationary components. Every element in the vector P_t can be explained by a linear combination of a smaller number of $I(1)$ common factors f_{jt} (permanent component) plus an $I(0)$ or transitory component (for instance $p_{it} = \sum_{j=1}^{s} a_{ij} f_{jt} + \tilde{p}_{it}$). In the long run, the variables p_{it} move together because they share the same stochastic trends. The representation (1) is known as the common factor representation and its existence is guaranteed if and only if there are $n-s$ cointegrating vectors among the elements of the vector P_t (Granger representation theorem in Engle and Granger). A major result of the Granger representation theorem is that a cointegrated system can be written as a VEC model

$$(2) \quad \Delta P_t = \mu + \Pi P_{t-1} + \Gamma_1 \Delta P_{t-1} + \Gamma_2 \Delta P_{t-2} + \cdots + \Gamma_{p-1} \Delta P_{t-p+1} + \varepsilon_t$$

where Γ and Π are $n \times n$ matrices and Π has reduced rank $n-s$. The matrix Π can be written as $\Pi = \alpha\beta'$, where α is an $n \times (n-s)$ matrix of coefficients, and β is an $n \times (n-s)$ matrix of cointegrating vectors. Using this expression for Π, we have $\Pi P_{t-1} = \alpha\beta' P_{t-1} = \alpha Z_{t-1}$. The error correction term, also known as short run disequilibrium, is $Z_{t-1} = \beta' P_{t-1}$, and α is the matrix of adjustment coefficients. The elements of the matrix β cancel the common unit roots in P_t and, in the long run, link the movements of the elements of P_t.

In this context, our definition of the extent of an integrated market requires that $s = 1$ because we are searching for locations that share the same long run information.[3] The common factor representation (1) becomes $p_{it} = a_{i1} f_{1t} + \tilde{p}_{it}$, $i = 1, \ldots, n$. Searching for just one common factor is equivalent to searching for $n-1$ cointegrating vectors. This is a key point because it differentiates our article from previous studies on market integration. In our approach, the *economic*

[2] When trade reversals or discontinuities are important, then a switching regime model would be required. Barrett, Baulch, Li and Barret, and McNew and Fackler have stressed this point.

[3] If there were more than one common trend, for example two, some prices could be generated by the first common trend, some by the second, and some by a combination of the first and second trends. We would not call these markets integrated because the long run movements in prices would be governed by different components.

market is not given a priori by the set of locations where a good is produced and/or consumed. Nor is the existence of cointegrated prices sufficient to find the market. It needs to be found through a multivariate search for a single common factor. In the case of Brazilian rice, although we have 19 locations we show that only 15 belong to the same economic market.

The search for the largest set of locations that share $n-1$ cointegrating vectors is conducted in a multivariate framework: the reduced rank VAR proposed by Johansen (1988, 1991). Johansen's test for the number of cointegrating vectors focuses on testing the rank of Π. The process of testing for the rank of Π occurs jointly with the estimation of the cointegrating vectors and vector error correction model. Thus, in contrast to the two-stage Engle–Granger methodology, Johansen's approach is a one-stage procedure. When the number of cointegrating relations is identified, we have not only estimated the cointegrating vectors but have also estimated the short run dynamics of the system given by equation (2).

The existence of $n-1$ cointegrating vectors implies that the vectors can be normalized in such a way that all locations will be cointegrated pair-wise. This is not sufficient, however, to justify a bivariate analysis of the market because the true vector error correction model is still a multivariate system. Thus, a bivariate system will in general be misspecified due to the omission of potentially relevant variables. This leads to inconsistent estimates of the parameters of the bivariate VEC as well as of any other estimator based on it.

To determine which locations belong to the same market, we recommend starting with the maximum set of locations, n, and testing for $n-1$ cointegrating vectors. We do this by performing Johansen's likelihood ratio test based on the trace statistic. If the number of cointegrating vectors is less than $n-1$, we need to identify those locations that should be removed from the system. In order to do so we implement a sequential procedure. We start with a core of m locations ($m < n$) and test for the number of cointegrating vectors. If the number is $m-1$ we add an additional location. With $m+1$ locations, either the new one shares a common trend with the previous m locations or it does not. In the first case, we should find m cointegrating vectors, while in the second, we should continue to find $m-1$, thus adding a second common trend to the $m+1$ locations. If we find one common trend, we repeat the procedure by adding locations one at a time. If not, we exclude the location that added a second trend and repeat the procedure until the number of locations is exhausted. This sequential procedure may be subject to some pre-testing problems. Future research should study the econometric problems of sequential exclusion. To ameliorate potential problems, we have considered different orders. In our application, the exclusion of locations is invariant to the order in which they have been analyzed.

Finally, after finding the $n-1$ cointegrating vectors, we proceed to estimate the common factor. We follow the methodology proposed by Gonzalo and Granger to estimate f_{1t}. This methodology is particularly attractive because the common factor is associated with observable variables and it allows for the identification of the location(s) that contribute to the long run behavior of the market price. The estimation of the common factor is easily derived from the specification of the error correction model (2). Two conditions are needed to identify the common factor. The first one imposes that f_{1t} be a linear combination of the elements of the vector of prices $\{p_{1t}, p_{2t}, \ldots, p_{nt}\}$ so that f_{1t} is observable. The second condition imposes that, in equation (1), the transitory component $\widetilde{P}_t$ does not Granger-cause the permanent component Af_{1t} in the long run. Thus, any shock that affects the transitory component is not transmitted to the long run forecast of P_t. This condition implies that in the vector error correction model the only linear combination of $\{p_{1t}, p_{2t}, \ldots, p_{nt}\}$ such that $\widetilde{P}_t$ does not have any long run effect on P_t is

$$(3) \quad f_{1t} = \alpha'_{\perp} P_t$$

where $\alpha'_{\perp}\alpha = 0$. This orthogonality condition means that the vector $\alpha_{\perp}$ eliminates the error correction term $Z_{t-1} = \beta' P_{t-1}$ from the vector error correction model, guaranteeing no effect of the transitory component on the long run forecast of P_t. Equation (3) can be used to reveal the locations that contribute to the transmission of long run information. This is important for the design of economic policy. Price support, or stabilization policies, for example, could be targeted at those locations that form f_{1t}. The transmission of policy to the rest of the market would be guaranteed.

The Pattern of Interdependence

In this article, the pattern of interdependence refers to the set of relationships among the different locations of the market as revealed through an analysis of the vector error correction model. The VEC in equation (2) summarizes the short run dynamics of the vector P_t as a function of a proportion α of past disequilibria Z_{t-1} plus $p-1$ lags of each Δp_i. In this model, the matrix α of adjustment coefficients is of particular interest because it contains the necessary information to uncover the spatial structure of the market. Furthermore, this matrix provides the key to choosing between a bivariate and a multivariate analysis of the system.

There are different patterns that could be observed in a VEC. Several examples follow. Suppose that we were to find that all elements of the matrix α were statistically significant. Then we would have a system in which each location reacts to every single disequilibrium or error correction term of every other location. This would be a case of extreme interdependence where the information contained in prices is generated in every single location. In such a market, it is obvious that a bivariate analysis would be grossly misspecified because it would be omitting numerous relevant variables.

As a second example, suppose that there was an exogenous central location i that dominated the long run behavior of the system. In this case, we should observe that in the equation of the VEC for location i all α_{ij}, $j = 1, \ldots, n-1$ should be statistically zero. This is a test for weak exogeneity with the null hypothesis $H_0 : \alpha_{ij} = 0, j = 1, \ldots, n-1$. A failure to reject the null hypothesis suggests the existence of an exogenous location that by itself would be the integrating factor of the system. Even in this case, however, a bivariate analysis would be inappropriate unless further tests were performed. A bivariate VEC would only be justified if it were also true that each location only adjusted to its own disequilibrium with respect to the exogenous location. Thus, in addition to an exogenous location, all $\alpha_{jk} k \neq i$ would have to be statistically zero.

Between the two extremes described above, many other patterns are possible. In order to reveal the pattern of interdependence in a market, or to determine if a bivariate specification is adequate, it is necessary to begin with a multivariate vector error correction model. The tests for weak exogeneity and for further restrictions can then appropriately reduce the system. At the end of the empirical section of this article, we compare bivariate and multivariate estimations of the VEC in order to expose the biases that could occur due to the misspecification of the model.

The Degree of Integration

Many studies have attempted to answer questions about the degree of market integration based on partial measures derived from a bivariate VEC model. It has been customary to look at the size of the adjustment coefficients (α) or the statistical significance of the lag structure (Γ). Our goal is to jointly evaluate the estimates of equation (2) and summarize them in a single measure that defines the degree of integration. Impulse response functions have been used extensively for this purpose. They trace the impact over time of a shock in location j on the price of location i. The main drawback of impulse response functions is that they are not unique when the shocks are correlated. In a study of spatial prices it is unreasonable to expect to have orthogonal shocks because the time series of all prices are highly correlated. The solution adopted in the literature has been to orthogonalize the shocks with a Cholesky decompositon of the covariance matrix of errors. This decomposition is not invariant to the ordering of the variables of the system and consequently, for every order, we have a different impulse response function. Imposing a recursive ordering on the variables is a very strong identifying assumption, and not justifiable in most studies of market integration. It is because impulse response functions are likely to be misleading and difficult to interpret that we propose an alternative measure that does not require the imposition of an ordering on the system.

The long run equilibrium among prices can be written as:

$$(4) \quad p_{1t} = -(c_i/\beta_{1i}) - (\beta_{2i}/\beta_{1i})p_{2t} - \cdots - (\beta_{ni}/\beta_{1i})p_{nt} + z_{it} \quad i = 1, \ldots, (n-s)$$

where c_i is a constant and all other variables are as defined above. Suppose that there is a shock to the underlying VAR that disturbs the long run equilibrium among the p_{it}, that is $|z_{it}| \neq 0$. Because equation (4) is a cointegrating relation, the vector Z_t is stationary. This implies that the effect of the shock will be transitory and eventually die out, and the long run equilibrium will be restored. We define the degree of integration as the reaction time for each of the long run equilibrium relations to absorb a system-wide shock. This depends on all of the estimated coefficients of α, β and Γ. By analyzing the joint impact of these coefficients, it becomes possible to construct a consistent ranking of markets based on reaction times. We adopt the methodology of Pesaran and Shin and construct persistence profiles.

A persistence profile characterizes the response of the cointegrating relation $Z_t = \beta' P_t$ to a system-wide, rather than to an individual shock, where the response is measured in units of variance. A system-wide shock is understood as a draw from the multivariate distribution of the vector $\varepsilon_t = \{\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{nt}\}$. The advantage of considering a system-wide shock is that the persistence profiles are unique functions and there is no need to orthogonalize the individual shocks. At time t, the variance–covariance matrix of the shock ε_t is Ω. We study the propagation through time $(t+1, t+2, \ldots)$ of the variance of the shock, conditioning on information up to time $t-1$. Thus, with an initial shock to the economy at time t, and considering the information up to time $t-1$, the persistence profile focuses on the *incremental* variance of the disequilibrium error at time $t+k$, as the time horizon increases by one period. In stationary systems, a shock will eventually die out. This implies that its incremental variance becomes smaller as time passes and approaches zero as time goes to infinity. Pesaran and Shin define the (unscaled) persistence profile as

$$(5) \quad H_z(k) = \mathrm{Var}\Big(Z_{t+k} \mid \psi_{t-1}\Big) - \mathrm{Var}\Big(Z_{t+k-1} \mid \psi_{t-1}\Big) \qquad k = 0, 1, 2, \ldots$$

where ψ_{t-1} is the information set containing information up to time $t-1$, Var $(Z_{t+k} \mid \psi_{t-1})$ is the variance of Z_{t+k} conditional on the information set, and k is the time horizon. The definition (5) has an appealing interpretation if we observe that $\mathrm{Var}(Z_{t+k} \mid \psi_{t-1})$ is also the variance of the $k+1$ step ahead forecast error of Z_t. We can write $\mathrm{Var}(Z_{t+k} \mid \psi_{t-1}) = E\{[Z_{t+k} - E(Z_{t+k} \mid \psi_{t-1})] \mid \psi_{t-1}\}^2$ where $Z_{t+k} - E(Z_{t+k} \mid \psi_{t-1})$ is the $k+1$ forecast error of Z_t. According to this interpretation, definition (5) says that a persistence profile is the change in the variance of the forecast of Z_{t+k} with respect to the variance of the forecast of Z_{t+k-1} based on the information set ψ_{t-1}.

From equation (1) and $Z_t = c + \beta' P_t$, we have $Z_t = c + \beta' A f_t + \beta' \widetilde{P}_t = c + \beta' \widetilde{P}_t$, where the last equality follows from $\beta' A = 0$ because Z_t is stationary. Consequently, we have

$$H_z(k) = \beta' \big\{\mathrm{Var}(\widetilde{P}_{t+k} \mid \psi_{t-1}) - \mathrm{Var}(\widetilde{P}_{t+k-1} \mid \psi_{t-1})\big\}\beta$$

where $k = 0, 1, 2, \ldots$ To facilitate the comparison among different profiles, we scale $H_z(k)$. For $k = 0$, $H_z(0) = \beta'\{\mathrm{Var}(\widetilde{P}_t \mid \psi_{t-1})\}\beta = \beta'\Omega\beta$. Define a diagonal matrix G that contains the inverse of the square root of the diagonal elements of $H_z(0)$, $G = \mathrm{diag}\{H_{11}(0)^{-1/2}, \ldots, H_{n-s,\, n-s}(0)^{-1/2}\}$. The scaled persistence profile is defined as

$$(6) \quad h_z(k) = G H_z(k) G = \{h_{ij}(k)\} \qquad k = 0, 1, 2, \ldots$$

Upon impact, at time $k = 0$, the profile $h_{ii}(k) = 1$ for $i = 1, \ldots, n-s$.

The Brazilian Rice Market

Determining the Extent of the Market

The spatial pattern of production, consumption, and trade. Rice production in Brazil is concentrated in a small number of states. In the 1970s, five of Brazil's twenty-five states produced 65% of the country's rice. By the 1990s, these same areas (reconstituted in seven states) had increased their share to 75% of national production. While production data in Brazil are available on an annual basis, data on consumption are virtually nonexistent. In order to estimate inter-state trade flows, we first had to estimate state level consumption on an annual basis with data on population and per capita rice consumption.[4]

[4] The production and population data come from the *Anuário Estatístico do Brasil*, Fundação Instituto Brasileiro de Geografia e Estatística (IBGE), various years. The consumption data come from official IBGE consumption/expenditure surveys conducted in 1974, 1987, and 1996.

Table 1. Estimated Inter-State Trade of Rice for Selected States

State	Trade[a] (Percent of National Production)			Index of Self-Sufficiency (Production/Consumption)		
	1970–79	1980–89	1990–95	1970–79	1980–89	1990–95
North (N)	**−2.1**	**−0.5**	**−0.5**	**0.6**	**0.9**	**0.9**
Acre (AC)	−0.1	0.1	0.2	0.6	1.5	2.0
Pará (PA)	−1.4	−1.2	−0.9	0.5	0.6	0.7
Northeast (NE)	**−2.6**	**−8.7**	**−13.6**	**0.9**	**0.6**	**0.5**
Maranhão (MA)	9.5	5.8	2.2	5.8	2.4	1.4
Ceará (CE)	−1.9	−4.0	−4.2	0.3	0.2	0.3
Rio Grande do Norte (RN)	−0.9	−1.3	−1.5	0.1	0.0	0.0
Paraíba (PB)	−1.2	−1.7	−1.9	0.2	0.1	0.1
Pernambuco (PE)	−3.1	−2.4	−2.5	0.0	0.1	0.1
Bahia (BA)	−4.2	−4.1	−4.6	0.1	0.1	0.2
Sergipe (SE)	−0.3	−0.3	−0.4	0.5	0.5	0.4
Southeast (SE)	**−32.3**	**−34.3**	**−33.9**	**0.4**	**0.3**	**0.3**
Minas Gerais (MG)	−3.3	−5.1	−6.0	0.8	0.6	0.5
Espirito Santo (ES)	−1.1	−1.2	−1.3	0.5	0.4	0.4
Rio de Janeiro (RJ)	−8.8	−8.0	−9.4	0.1	0.1	0.1
São Paulo (SP)	−19.1	−20.0	−17.3	0.3	0.2	0.2
South (S)	**16.3**	**25.5**	**40.3**	**1.9**	**2.8**	**4.3**
Paraná (PR)	1.1	−1.8	−2.6	1.1	0.7	0.5
Santa Catarina (SC)	0.1	2.0	4.0	1.0	1.7	2.6
Rio Grande do Sul (RS)	15.2	25.4	38.9	3.1	5.6	8.9
Center-West (CW)	**20.6**	**18.0**	**7.7**	**4.2**	**3.2**	**1.8**
Mato Grosso do Sul (MS)[b]	–	2.5	0.7	–	2.8	1.4
Mato Grosso (MT)	11.8	7.5	4.8	6.6	6.5	3.6
Goiás (GO)[c]	9.7	8.9	1.2	3.8	3.2	1.3

[a] Positive values indicate exports and negative values indicate imports.
[b] Mato Grosso do Sul was created in 1977. Prior to 1977 it was part of Mato Grosso.
[c] In 1988 Goiás was divided in two and Tocantins was created. In the 1990–95 period, Tocantins exported 2.6% of national production.

Table 1 presents the estimates of inter-state trade for the 19 states for which we have continuous price data. These states accounted for over 90% of production and consumption. The first three columns show the difference between a state's share of national production and its share of national consumption for three sub-periods, thus providing estimates of exports (positive numbers) and imports (negative numbers) as a share of national production. The final three columns show an index of self-sufficiency, defined as the ratio of a state's production share to its consumption share. A ratio close to one implies that a state is close to self-sufficient.

Table 1 shows that although rice is often considered to be a non-tradable good for Brazil as a nation, it was traded extensively within the country. Approximately half of Brazilian rice was traded across state borders throughout the period. Since the demand for rice is constant throughout the year, and rice is stored predominantly in producing regions, trade occurred continuously with a steady flow of trucks transporting rice from surplus to deficit regions.[5]

The Southeast of Brazil is home to more than 40% of the population. This region consistently imported over 30% of national rice production, with most of the deficit coming from São Paulo. The Northeast was the only other region with a significant shortage of rice. With the exception of Maranhão, all of the other Northeastern states were clear importers. Regardless of how small the absolute size of their deficits, the self-sufficiency index reveals that none of these states produced more than half of their consumption, and most produced only 10–20%. Steady inflows of rice from as far away as Rio Grande do Sul have always been necessary.

[5] Continuous trade flows are confirmed by Ereias for the state of Rio Grande do Sul. Interviews conducted by the authors in the Rio Grande do Sul Rice Institute (IRGA) and in the Getúlio Vargas Foundation (FGV) also confirmed that continuous trade is the normal state of affairs.

The distant Northern part of the country is relatively isolated and as a region it is close to self-sufficient. The physical isolation and poor infrastructure of this region led us to expect that states located here were unlikely to belong to the national economic market. Acre, in addition, exhibits a clear trade reversal as it transitions from being an importer in the 1970s to an exporter in the 1990s. For this reason it would be inappropriate to include Acre in the VEC.

Other than Maranhão, the principal surplus states were located in the Center-West and South. The Center-West accounted for a larger share of trade than the South in the 1970s, yet by the 1990s the South—especially Rio Grande do Sul—was exporting five times as much as the Center-West. Along with its neighbor Santa Catarina, Rio Grande do Sul was different from the other states in two important ways. First, it produced irrigated rice that was subject to far less production variability than the rain-fed rice produced in other states. Second, it produced higher quality rice. Both of these facts have important implications and will be discussed below.

We have demonstrated that with the exception of Acre, rice trade occurred with no reversals and was apparently continuous. In the early 1970s this might not have been true for Paraná and Santa Catarina. Our annual trade estimates indicate that these two states hovered around self-sufficiency in the first half of the decade. Price differentials, in contrast, were consistent with the pattern of trade that prevailed throughout the rest of the period. The possibility that discontinuous trade might distort our econometric estimates led us to conduct additional tests for parameter constancy that will be described below. We conclude that the inclusion of the early 1970s does not generate a problem for our model.[6] A plausible explanation is that even if the excess supply and demand in these two states was small, it was still sufficient to keep their prices close to the parity levels. These two states were, in addition, constantly exposed to competitive pressures because large quantities of rice flowed through their territories.

Searching for a single common trend. We conducted tests for unit roots in the log-prices of rice in nineteen states. We work with real monthly producer prices that were obtained from the Getúlio Vargas Foundation (FGV).[7] By having our sample run from 1973:01 to 1997:08 we were able to include a total of nineteen states. At a later stage in the analysis, after determining that several states did not belong to the system, the sample period was extended back to 1970:01. Two versions of the Augmented Dickey–Fuller (ADF) test were performed, one which excluded and one which included a constant in the regression. The time series were not smooth enough to entertain the possibility of a deterministic trend in the regression. We also conducted F tests for the joint null hypothesis of a constant equal to 0 and a unit root. The optimal number of lags in each regression was chosen according to the AIC and SIC criteria.

For the states in the Center-West, South, and Southeast, we could not reject the hypothesis of a unit root at the one percent significance level with any of the tests. In the Northeast and in the North, the statistical evidence was mixed. In particular, for Maranhão, Paraíba, and Sergipe in the Northeast, and for Pará and Acre in the North, the introduction of a constant in the regression made a difference for the results of the tests. With a constant, we rejected the unit root at the one percent significance level, but without a constant we did not. Furthermore, the estimated values of the roots were the smallest among all the states, ranging from 0.89 to 0.94. With the exception of Maranhão, these states were very small in terms of production and consumption of rice. At this point in the analysis, we maintain the unit root hypothesis for all 19 states. Additional evidence is found at a later stage for removing most of these borderline states.

In table 2, we implemented the sequential procedure described above to determine which states shared the same common stochastic trend. Column 1 of table 2 shows the sequence of locations that were analyzed. We started with a core of ten important states in the Center-West, South, and Southeast. Different sequences were analyzed and the results were invariant to the ordering. The value of the likelihood ratio test is shown in parentheses for those cases in which the null hypothesis could not be rejected. At the

[6] The same conclusion is reached for the next most likely candidates to experience discontinuous trade: MG in the 1970s, and MA, MS, and GO in the 1990s.

[7] All prices are in constant *reais*, the Brazilian currency, of 12/1995. The monthly producer prices were deflated by the General Price Index (IGP-DI) of the Getúlio Vargas Foundation.

Table 2. Johansen's Likelihood Ratio Test for the Number of Cointegrating Vectors (Trace Statistic, 1973:01–1997:08) $H_0 : r = h$; $H_1 : r > h$

Series Included	Significance Level (%) 20	10	5
Center-West + South + Southeast = 10	10	9	9 (6.43)
10 + MA = 11	10	10	10 (5.78)
11 + BA = 12	11	11	11 (5.60)
12 + SE = 13	12	12	12 (5.56)
13 + PE = 14	13	13	13 (5.46)
14 + CE = 15	14	14 (4.94)	13 (19.22)
15 + RN + PB = 17	16 (4.24)	13 (50.44)	8 (205.80)
17 + AC + PA = 19	13 (93.12)	9 (242.50)	8 (290.48)

Note: r is the number of cointegrating vectors. The values of the likelihood ratio tests are in parentheses. The critical values are from MacKinnon, Haug and Michelis (1996), and extrapolation. See footnote 8 in the text.

5% significance level, we found one common trend among the original ten states. Regardless of the order chosen, we continued to find one common trend when four Northeastern states were added (MA, BA, SE and PE). Ceará (CE) entered the set of one-common-trend markets at the 10% significance level, and at the 20% significance level Rio Grande do Norte (RN) and Paraíba (PB) could also be included. For Acre (AC) and Pará (PA), the two Northern states, thirteen cointegrating vectors were found at the 20% level, implying six common trends.

The conclusion that we draw regarding the extent of the market is that fifteen states belonged to the same economic market: those in the Center-West, South, and Southeast, plus MA, BA, SE, PE and CE in the Northeast. All fifteen states were shown to engage in a significant amount of unidirectional interstate trade. They also shared a single common trend at a significance level smaller than 10%.[8] Thus, the rice from these fifteen states were substitutes for each other to some degree and arbitrage through trade tied their prices together.

Four states did not appear to belong to this market. Acre was excluded on the grounds that it experienced a trade reversal. Together with Pará, Acre was also found not to share a single common trend with the other states. For Rio Grande do Norte and Paraíba, the one-common-trend hypothesis was only accepted at the 20% significance level, implying a very high probability of committing a Type I error. Furthermore, at the 5–10% significance levels, the inclusion of any of these four states actually reduced the number of cointegrating vectors to a smaller number than that of the original set, implying more than one common trend.

The fact that producer prices in four states did not share a common trend with the other fifteen should be interpreted carefully. First, in the case of Acre this could be due to the trade reversal and the results might differ for sub-periods. A second observation is that, due to a lack of time series data on transactions costs, the failure to find a single common trend could indicate either a lack of integration or non-stationary transactions costs.[9] In fact, the two Northern states are in a remote region of the country in which transportation is more difficult in the rainy months of the year. Similarly, the two excluded Northeastern states had rice prices in the 1990s that were far too high to be consistent with distance and average transactions costs. It is our view, however, that even if we are not able to pinpoint the cause for not finding a single common trend, with the exception of Acre the result is still meaningful. It suggests that there is something qualitatively different about the rice market in these states and it is likely to be related to high and/or unusual

[8] For the fifteenth series (CE), the value of the likelihood ratio test (19.22) is very close to the 5% critical values, 19.96 and 20.26, that are taken from the tables in Osterwald-Lenum and MacKinnon, Haug, and Michelis. Critical values are calculated for a maximum of 11 random walks in Osterwald-Lenum and 12 in MacKinnon, Haug, and Michelis. Our system contains up to 19 variables. In order to calculate critical values we have fitted a quadratic polynomial on the number of random walks to MacKinnon, Haug and Michelis critical values and extrapolated the critical values corresponding to 13 up to 19 random walks. The R^2 of this regression is equal to 1. For cointegration in large systems see Gonzalo and Pitarakis.

[9] Fackler makes a similar point about the difficulty of interpreting the failure to find cointegration. Goodwin provides an empirical example by only finding cointegration in international wheat markets when transportation costs are included.

transactions costs. Even if we were to find that net of transactions costs the prices in these states did share the same trend with the other fifteen, the need to net transactions costs out only for these states would still indicate a significant difference with the rest. Policy implications related to transportation and marketing could still be drawn, and they would likely extend beyond the rice market.

Estimation of the integrating factor. In this section the permanent component is estimated according to equation (3). We estimated the integrating factor as

$$\begin{aligned} f_t = &-0.034 p_{\mathrm{MS},t} - 0.036 p_{\mathrm{MT},t} \\ &+ 0.373 p_{\mathrm{GO},t} - 0.102 p_{\mathrm{MA},t} \\ &+ 0.267 p_{\mathrm{CE},t} + 0.017 p_{\mathrm{PE},t} \\ &- 0.070 p_{\mathrm{BA},t} - 0.037 p_{\mathrm{SE},t} \\ &+ 0.000 p_{\mathrm{PR},t} - 0.316 p_{\mathrm{SC},t} \\ &- 0.081 p_{\mathrm{RS},t} - 0.279 p_{\mathrm{MG},t} \\ &+ 0.310 p_{\mathrm{ES},t} - 0.361 p_{\mathrm{RJ},t} \\ &+ 0.890 p_{\mathrm{SP},t} \end{aligned}$$

We tested the null hypothesis that the coefficients corresponding to MS, MT, PE, BA, SE, PR and RS were statistically 0. The test statistic equals 0.39 and is distributed as a χ^2 with seven degrees of freedom. The p-value associated with the test is 0.99. Consequently we could not reject the hypothesis that these coefficients equal 0. The integrating factor, re-estimated with the imposed restrictions, is

$$\begin{aligned} f_1 = &\ 0.363 p_{\mathrm{GO},t} - 0.110 p_{\mathrm{MA},t} \\ &+ 0.280 p_{\mathrm{CE},t} - 0.250 p_{\mathrm{SC},t} \\ &- 0.274 p_{\mathrm{MG},t} + 0.341 p_{\mathrm{ES},t} \\ &- 0.389 p_{\mathrm{RJ},t} + 0.817 p_{\mathrm{SP},t} \end{aligned}$$

The estimated permanent component shows the role of the different states in shaping the long run behavior of the price of rice. The contribution of São Paulo (SP) to the permanent component of the domestic price dominates the other states. Public policy targeted at São Paulo would have the greatest impact on the long run component of prices in Brazil. Furthermore, it is clear that the long run component of the price is driven by two forces: the production side of the market represented by states from the Center-West (GO), the Northeast (MA), and the South (SC), and the consumption side of the market, which mainly involves the Southeast (SP, RJ, ES, and MG).

The Pattern of Interdependence

Cointegration. In table 3, we present the normalized cointegrating vectors as in Phillips.[10] The normalization is done with respect to the São Paulo (SP) market. The 14 cointegrating vectors are readily interpretable because they consist of 14 pair-wise relationships. We thus explain the long run equilibrium between pairs of markets (MS and SP, MT and SP, etc.). For our system, the long run equilibrium relations shown in equation (4) become

$$p_{it} = \hat{c}_i + \hat{\beta}_i p_{\mathrm{SP},t} + \hat{z}_{it} \qquad i = 1, \ldots, n-1$$

where $p_{\mathrm{SP},t}$ is the price in the São Paulo market. Thus, in the case of Goiás for example, we have: $p_{\mathrm{GO},t} = -0.43 + 1.05 p_{\mathrm{SP},t} + \hat{z}_{it}$. The values of $\hat{\beta}_i$ range from 0.58 to 1.08. In most cases, the hypothesis that $\hat{\beta}_i = 1$ cannot be rejected at the 1% level. There are several cases that clearly diverge from this pattern. Cointegrating vectors can differ from $(1, -1)$ as a result of transactions costs (Dercon), as well as for other reasons.[11] In Brazil, for example, most rice is stored and milled in producing regions, and then shipped to deficit areas to meet demand. It follows that spatial arbitrage at the producer level occurs indirectly through the wholesale market for milled rice. As a result, the differences in state-level producer prices measure elements of arbitrage across form as well as space. Thus, in addition to transactions costs, the cointegrating vectors capture regional differences in policies, technologies, and product quality.

The coefficients in the cointegrating vectors suggest that large transactions costs or other factors differentiated several states from the rest. There is no evidence that discontinuous trade was the source of these results. Although the cointegrating vector for Rio Grande do Sul has the second smallest $\hat{\beta}_i$, we have already demonstrated that it exported rice continuously. The $\hat{\beta}_i$ in the equation for Maranhão is the smallest in table 3. In order to ensure that this was

[10] The lag length of the VAR was chosen by performing a series of F-tests on the lag structure. With three lags, the residuals seemed to behave like white noise.

[11] See *Companhia de Financiamento da Produção* for a snapshot of transfer costs in Brazil around 1980. It provides evidence of both additive costs such as freight and proportional charges such as inter-state taxes, sales commissions, and financial fees.

not caused by the possibility of discontinuous trade in the 1990s when it became only a modest exporter, we re-estimated the system through 12/1989. The estimated $\hat{\beta}_i$ was virtually identical (0.57 rather than 0.58). We reached the same conclusion for Paraná and Santa Catarina in the 1970s, and for all other states that were only marginal traders during sub-periods of the sample.

The coefficients in the equation for Maranhão can be explained by high transactions costs. Maranhão is the farthest state from São Paulo, with a distance of 2970 km between their capital cities. The transactions costs were reflected in the average prices in these two states. Maranhão had the lowest price of all 15 states, while São Paulo had the highest. The cointegrating vector for Rio Grande do Sul, which is no farther from São Paulo than the Center-West, reflects the fact that this state produced a higher quality rice, used a different technology (irrigation), and was subject to a somewhat different policy environment. These factors also explain the estimates for Santa Catarina. Producers in these two states had a different support price than in the other regions and relied on storage credit to a much higher degree.

Cointegrating vectors that diverge from the general pattern reflect structural differences with the other states. These states did, nevertheless, belong to the same economic market. As we demonstrate below, differentiated cointegrating vectors do not necessarily imply a lack of interdependence or a low degree of integration.

The Vector Error Correction Model. Table 4 presents the adjustment coefficients from the restricted vector error correction model that we estimated. This table permits us to highlight the problems of misspecification that would have arisen in a bivariate model. It also allows us to analyze the pattern of interdependence in the Brazilian rice market. Before discussing table 4, we explain the testing that led to the restricted specification.

We began by estimating an unrestricted vector error correction model for the 15 states as a system of seemingly unrelated regressions (SUR). Every equation in the system had the same number of variables on the right-hand side: the 14 error correction terms from table 3, 2 lags for each Δp_i, and a dummy for the outlier January 1990.[12] To ensure that the system was not misspecified, we performed Lagrange multiplier tests for serial correlation, RESET tests for functional form, GARCH and White's tests for heteroskedasticity, and CUSUM and CUSUMSQ tests for model stability.[13] In the overall system, there was no evidence of either serial correlation or seasonal patterns in the residuals of the VEC. This confirmed that the lag structure was appropriate to capture the dynamics of prices. Similarly, a linear specification of the VEC was found to be satisfactory. The CUSUM test, based on the cumulative sum of the recursive residuals, did not indicate any stability problem in the conditional mean. The CUSUMSQ test, based on the cumulative sum of squared recursive residuals, pointed toward a more volatile period between 1990 and 1995 for the states in the South and Southeast. This was in agreement with the mild heteroskedasticity that we found in the same states and is attributable to a combination of high inflation and a reduction of support prices in these years.[14] We re-estimated the model for the period 1970–1989 and found that while the estimation results remained essentially unchanged, most of the heteroskedasticity disappeared. Heteroskedasticity by itself does not affect the consistency of the estimates of the conditional mean, but it does affect the standard errors. As a result, we used heteroskedasticity-consistent standard errors.

We proceeded to explore the spatial pattern of interdependence by conducting a series of F-tests for weak exogeneity and Granger causality on the estimated coefficients from the unrestricted VEC. These tests permit us to determine if there are one or more exogenous states—a necessary finding in order to justify the use of a bivariate model. The tests also permit us to remove unnecessary terms from the VEC and to estimate a more parsimonious restricted specification. Weak exogeneity of location A with respect to the j locations in region B, for example, implies that the price in location A does not respond to disequilibria in region B. Consequently, in equation (2) for location A we should find that the adjustment coefficients α_{Aj} corresponding to the error correction terms from region B all equal 0. If we found a state to be weakly exogenous

[12] Production fell by 33% in 1990 and the average price for Brazil rose substantially.

[13] Results from the unrestricted model and all tests are available from the authors.

[14] Unlike Shively, the heteroskedasticity that we find does not appear to be attributable to storage. It is present in consuming and storing states, and is largely confined to the early 1990s.

Table 3. Normalized Cointegrating Vectors: Johansen's Method (1970:01–1997:08)

	Center-West			Northeast					South			Southeast		
	MS	MT	GO	MA	CE	PE	BA	SE	PR	SC	RS	MG	ES	RJ
$State_i$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
SP	−0.97	−1.06	−1.05	−0.58	−0.89	−0.89	−0.95	−0.73	−1.08	−0.79	−0.65	−0.91	−0.97	−0.85
	(0.04)	(0.08)	(0.06)	(0.12)	(0.10)	(0.08)	(0.09)	(0.07)	(0.04)	(0.08)	(0.08)	(0.03)	(0.08)	(0.07)
Constant	0.04	0.63	0.43	−2.15	−0.64	−0.57	−0.26	−1.41	0.57	−1.02	−1.89	−0.49	−0.07	−0.81
	(0.27)	(0.47)	(0.34)	(0.71)	(0.56)	(0.46)	(0.54)	(0.43)	(0.26)	(0.44)	(0.48)	(0.20)	(0.44)	(0.42)

Note: (1) Standard errors in parentheses. (2) Error $\text{Correction}_1 = P_{MS,t} - 0.97P_{SP,t} + 0.04$; Error $\text{Correction}_2 = P_{MT,t} - 1.06P_{SP,t} + 0.63$, etc.

Table 4. Adjustment Coefficients (α) From The Restricted Vector Error Correction Model (1970:01–1997:08)

Error	Center-West			Northeast					South			Southeast			
Correction	MS	MT	GO	MA	CE	PE	BA	SE	PR	SC	RS	MG	ES	RJ	SP
(MS, SP)	−0.35**			−0.17**										−0.14**	−0.09*
(MT, SP)		−0.20**	0.07**	0.08**										0.07*	
(GO, SP)	0.17**	0.12*	−0.20**									0.11*			0.13*
(MA, SP)		−0.03*		−0.11*									−0.04*		
(CE, SP)		0.04*		−0.11**	−0.13**	0.18**							0.07**		
(PE, SP)						−0.28**		0.13**							
(BA, SP)							−0.22**						−0.05**		
(SE, SP)								−0.30**				−0.04*			
(PR, SP)									−0.20**						
(SC, SP)				0.06*						−0.18**	0.18**			0.10**	
(RS, SP)			−0.07*								−0.32**				
(MG, SP)	−0.21**	−0.15*			−0.09*							−0.49**		−0.12*	−0.19**
(ES, SP)			0.08**	0.14**	0.11**		0.31**					0.13**	−0.12**	0.18**	0.13**
(RJ, SP)			−0.11**	−0.16**	−0.10*	−0.14**	−0.33**							−0.38**	−0.12*
Adj. R^2	0.43	0.28	0.38	0.37	0.27	0.22	0.25	0.22	0.39	0.32	0.30	0.53	0.47	0.50	0.42

Note: Iterative SUR estimation with heteroskedasticity-consistent errors. Coefficients not significant at the 5% level are not shown in the table.
* Statistically significant at 5% level.
** Statistically significant at 1% level.

with respect to other states, we then tested for Granger causality with respect to the same states. The absence of Granger causality implies that the price in location A is not linearly influenced by the lagged variables from region B.

Table 4 presents the coefficients of adjustment (α) from the restricted VEC model. The most important observation relates to the limitations of a bivariate model. A bivariate specification would only be appropriate in the unlikely event that we were to find both a single exogenous state *and* all other locations responding only to error correction terms involving this exogenous state. If this were the case, we should find an empty column in table 4, implying that the state is weakly exogenous, and a maximum of 14 significant error correction terms in the table, with all of them involving the weakly exogenous state. Neither of these conditions was present in the Brazilian rice market. The complexity of adjustment patterns in the market suggests that the estimation of a bivariate VEC constructed from any two of our fifteen states would likely have led to important biases due to the omission of many relevant locations. Even if we were to have limited attention to the most important exporting and importing states in the country—Rio Grande do Sul and São Paulo—both dynamic equations would have been misspecified due to the exclusion (respectively) of one and five statistically significant error correction terms. In the final section of this article we explore the consequences of this type of model misspecification for the estimated path of adjustment by comparing the bivariate and multivariate persistence profiles. For now, the conclusion that we draw is that the pattern of adjustment in a spatially integrated market is likely to be very complex. A bivariate model is only appropriate in a limited number of very special market structures.

Table 4 shows that although there was no state that was weakly exogenous with respect to the entire market, it was also the case that not all states interacted. As the principal consuming region of the country, the states in the Southeast appeared to represent a central location through which price information was processed. These four states were influenced by error correction terms from every other region in the country, and they influenced adjustment in every other region. The lag structure of each equation, which is not shown in the table, also underscored the centrality of the Southeast in the adjustment process. Both São Paulo and Minas Gerais appeared in every equation in the Southeast and South, and one or the other appeared in each equation in the Center-West. It would be incorrect, however, to model the Brazilian rice market as having a central market (as in the Ravallion approach). Not only does the Southeast contain four states, but there were many other important channels through which information was conveyed.

Another important finding is that the three states in the South were the least interdependent in the country. They did not adjust to error correction terms from other regions of the country and there were only three states (GO, MA, and RJ) that responded to them. They did, however, adjust to their own disequilibria with São Paulo and their prices were Granger caused by lagged prices from the Southeast. These results point to a certain degree of market segmentation by quality. This is a somewhat surprising result given the importance of Rio Grande do Sul as a producer for the rest of the country.[15] It suggests that although different qualities of rice were substitutes and had a stationary long run relationship, the degree of substitution was probably low and it only bound their prices together in the long run.

A final observation is that differences in the cointegrating vectors did not appear to be correlated with a state's pattern of interdependence. In spite of similarities in their cointegrating vectors, Maranhão and Rio Grande do Sul had very different forms of insertion in the market.

Persistence Profiles and the Degree of Integration

Figure 1 shows selected persistence profiles that were calculated from the restricted model according to the methodology described above. These graphs show the estimated reaction time for each of the 14 long run equilibrium relations to absorb a system-wide shock. The figure shows the profiles over a twenty month horizon for (MG, SP), (GO, SP), (BA, SP), and (MA, SP). The profiles indicate that disequilibria between Minas Gerais (MG) and São Paulo (SP), for example, are removed rather quickly, while

[15] Barros and Filho find a similar result for the South. They find no causality in either direction between the producer price in Rio Grande do Sul and the retail price in São Paulo.

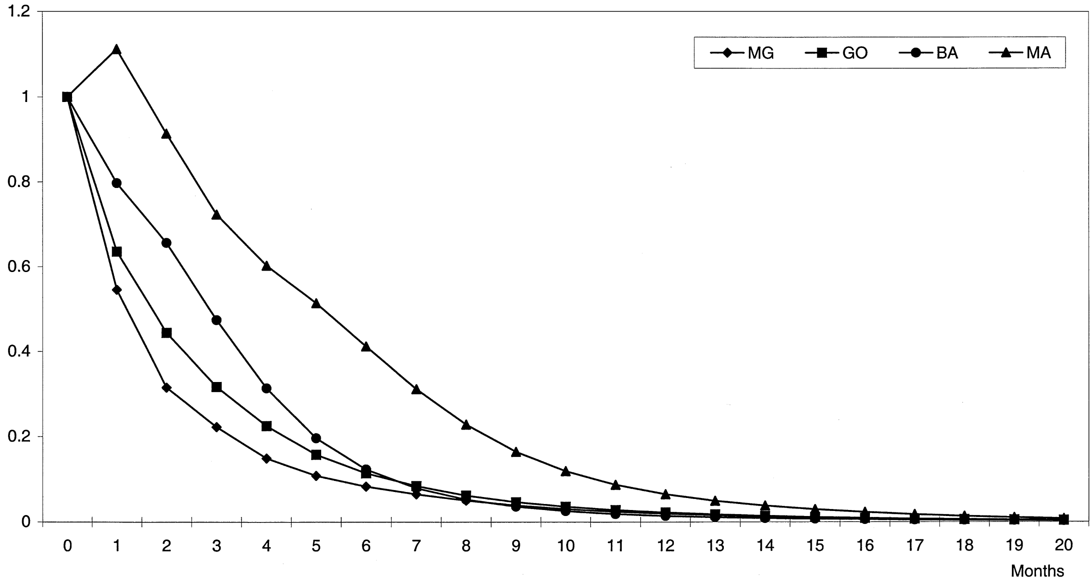

Figure 1. Persistence profiles from the restricted VEC model (1970:01–1997:08)

this is not the case for Maranhão (MA) and SP. When there is a system-wide shock that affects the long run equilibrium between SP and MG, 45% of the adjustments take place in the first month, and nearly 80% within three months. Disequilibria between Goiás (GO) and SP are removed a bit slower, with only 68% of the adjustments occurring within three months. MA actually overshoots at first, and after three months 70% of the effect of the shock remains.

While the profiles capture the entire path of adjustment between a given state and SP, it would be useful to construct a statistic to summarize the information in the graph. For this purpose, we have calculated the median persistence, or half-life, of the effect of the shock for each state with SP, defined as the number of months necessary for 50% of the adjustments to take place. This information is shown in table 5 for the restricted and unrestricted models.

Table 5. Estimated Half-Lives of the Persistence Profiles (months)

	1970:01–1997:08	
State	Restricted	Unrestricted
MG	1.20	1.13
MS	1.35	1.20
GO	1.71	1.83
MT	1.72	1.63
SE	1.77	1.91
PR	1.89	2.15
RJ	2.01	1.83
PE	2.45	2.55
SC	2.63	2.95
ES	2.66	2.72
BA	2.86	3.05
RS	3.16	3.43
CE	3.81	4.16
MA	5.13	4.34

The second column of table 5 shows the half-lives for the restricted model. The states have been divided into three groups. The first group, which has half-lives that are two months or less, includes the three states in the Center-West that export rice to São Paulo (MS, MT, and GO), the three neighbors of São Paulo that are also important consumers (MG, PR, and RJ) and the small state Sergipe (SE). Adjustment between SP and both MG and MS happens the fastest, with half-lives under 1.35 months. Not only does MS supply rice to SP, but they also share a border. MG shares a border with SP as well, and is on the trade route from GO to SP. Thus, all of the rice that comes from GO and the other states in the Center-West must pass through MG.

The second group of states has half-lives between 2.45 and 3.16 months. These states are consumers in the Northeast (PE and BA), producers in the South (SC and RS), and the only state in the Southeast that does not share a border with SP (ES). Since the Northeastern states are only indirectly linked to SP through common suppliers, the reduced speed for Pernambuco (PE) and Bahia (BA) relative to the first group is understandable. Rio Grande do Sul (RS), as we have noted throughout the article, is an exception. Adjustment with SP happens relatively slowly in spite of the strong trade ties that exist. The low degree of integration is most likely due to differences in the quality of the rice that is grown in each state. If middle and high income consumers prefer the higher quality rice that is produced in RS and are hesitant to substitute even when harvests are poor and prices rise, then the cross price elasticity of demand for these products would be low. The channel of transmission from the producer price in one state to the producer price in the other, through consumption decisions in SP, would consequently be rather weak.[16] The third group of states belongs to the distant Northeast. Not only are the links with SP indirect, but the distances are far greater. Finally, the last column of the table shows that the results from the unrestricted model are quite similar.

Multivariate versus Bivariate Models

We estimated 14 bivariate models in order to highlight the problems that could arise with this approach in the context of a spatially integrated market. For comparability with the multivariate model, all 14 bivariate models included São Paulo. The results indicated that the problems with the bivariate approach did not appear to extend to the estimation of the slope coefficient in the cointegrating vectors. In no case did these diverge by more than 6%. Since our system has a single common trend and all states are cointegrated pair-wise, this is not surprising. The adjustment coefficients from the vector error correction models, in contrast, revealed much more substantial discrepancies. The bivariate models appear to estimate the adjustment coefficient between each state and São Paulo with a downward bias. In 13 of the 14 cases the adjustment coefficients were smaller in the bivariate model, and the average difference was −31%. In four cases the coefficient was less than half of what was estimated in the multivariate model, and in one case it was almost double.

We calculated two descriptive statistics from the persistence profiles: the median, or half-life, and the mean persistence. The mean persistence of disequilibria is a weighted average of the information from the entire 20-month horizon, although it weights the most distant months least because the disequilibria at that horizon are negligible. In spite of the smaller adjustment coefficients that would have led us to expect slower adjustment, the comparison reveals that the bivariate models estimated a quicker path of adjustment. On average, the bivariate half-lives were slightly smaller (6.9%), and the bivariate mean profiles were substantially smaller (34%). In three cases the discrepancies between the multivariate and bivariate half-lives were in the 19–33% range, and in eleven of the fourteen cases the bivariate mean profiles were 30–55% smaller. These discrepancies reflect the misspecification of the VEC that arises from excluding relevant error correction terms. It is also related to the loss of explanatory power in the bivariate models. On average, the adjusted R^2 from the bivariate VECs was almost 20% lower. In several cases, such as Pernambuco (PE) and Rio Grande do Sul (RS), the loss of explanatory power was more than 40%.

Conclusions

In this article we posed the question of market integration as one of degree. As a result, we provided a ranking of all locations from less to more integrated. To achieve this objective, we introduced two novel features to the market integration literature. First, we conducted a multivariate search to determine the geography of the market. Second, we used a measure of the degree of integration—a persistence profile—that does not suffer from the drawbacks of impulse response functions.

Our definition of an integrated market requires that the set of locations share the same traded commodity and the same long run information. With information on rice from 19 states in Brazil, we found that only 15 belonged to the same economic market. Two of the four excluded locations were no more physically isolated than other states from the same region of the country. To the extent that these locations were excluded due to poor physical or marketing infrastructure, the consequences are likely to extend beyond the rice market, thus impeding the ability of these two states to improve their

[16] We thank Ignez and Mauro Lopes for this insight. They also observed that the government was more likely to permit imports in response to a poor harvest in RS than in the Center-West due to the importance of this state's rice for urban middle-class consumers. This had the effect of mitigating the impact of events in RS on price transmission to other states.

welfare through specialization and trade. Furthermore, we estimated the common integrating factor for the 15 states in the market as a linear combination of prices in eight locations. These states provide the key to the transmission of long run information. Thus, public policy could be targeted at a relatively small number of locations and still be effective in terms of influencing the entire market.

Once the extent of the market was determined, we then used persistence profiles to measure the degree of integration. We demonstrated that large volumes of trade were not sufficient to generate a high degree of integration. Among other factors, it appears that physical distance and distance in product space (quality) can both lead to a low degree of integration. Future research should focus on explaining the determinants of the degree of integration. We believe that this is an area of inquiry with highly relevant implications for policy.

Finally, we emphasized the shortcomings of a bivariate approach to market integration. The major problem lies in the misspecification of the vector error correction representation of a cointegrated system. When relevant variables are omitted, the estimators become inconsistent. This inconsistency is carried forward to any other statistic that is based on the vector error correction model, including impulse response functions and persistence profiles.

[Received November 1998; accepted September 2000.]

References

Alexander, C., and J. Wyeth. "Cointegration and Market Integration: An Application to the Indonesian Rice Market." *J. Dev. Stud.* 30(January 1994):303–28.

Asche, F., H. Bremnes, and C.R. Wessels. "Product Aggregation, Market Integration, and Relationships Between Prices: An Application to World Salmon Markets." *Amer. J. Agr. Econ.* 81(August 1999):568–81.

Barrett, C.B. "Market Analysis Methods: Are Our Enriched Toolkits Well Suited to Enlivened Markets?" *Amer. J. Agr. Econ.* 78(August 1996):825–29.

Barros, G.S.C., and J.G.M. Filho. "Transmissão de Preços e Margens de Comercialização de Produtos Agrícolas." *Agricultura e Políticas Públicas*. G.C. Delgado, J.G. Gasques, and C.M.V. Verde, eds., pp. 515–65. (IPEA Série 127.) Brasília: IPEA, 1990.

Baulch, B. "Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration." *Amer. J. Agr. Econ.* 79(May 1997):477–87.

Companhia de Financiamento da Produção (CFP). *Análise das Distorções dos Preços Domésticos en Relação aos Preços de Fronteira: Um Estudo Preliminar.* (Coleção Análise e Pesquisa, 30.) Brásilia:CFP, 1983.

Dercon, S. "On Market Integration and Liberalisation: Method and Application to Ethiopia." *J. Dev. Stud.* 32(October 1995): 112–43.

Engle, R.F., and C.W.J. Granger. "Co-Integration and Error Correction:Representation, Estimation, and Testing." *Econometrica* 55(March 1987):251–76.

Ereias, A.C.S. "Análise das Margens de Comercialização do Setor Orizícola Gaúcho." MS Thesis, Federal University of Rio Grande do Sul, 1999.

Fackler, P. "Spatial Price Analysis: A Methodological Review." Unpublished, Department of Agricultural and Resource Economics, NC State, 1997.

Fundação Instituto Brasileiro de Geografia e Estatística (IBGE). *Anuário Estatístico do Brasil*. Rio de Janeiro: IBGE, various years.

——. *Estudo Nacional da Despesa Familiar: Consumo Alimentar, Despesas das Famílias, Dados Preliminares, Tabelas Selecionadas.* Rio de Janeiro: IBGE, 1978.

——. *Pesquisa de Orcamentos Familiares*, 1987 and 1996. From www.IBGE.gov.br.

Gonzalo, J., and C.W.J. Granger "Estimation of Common Long-Memory Components in Cointegrated Systems." *J. Bus. Econ. Stat.* 13(January 1995):27–35.

Gonzalo, J., and J.Y. Pitarakis, "Comovements in Large Systems." Working Paper 95-38, Dept. of Stat. and Econ., Universidad Carlos III de Madrid, 1995.

Goodwin, B.K. "Multivariate Cointegration Tests and the Law of One Price in International Wheat Markets." *Rev. Agri. Econ.* 14(January 1993):117–24.

Goodwin, B.K., and T.C. Schroeder. "Cointegration Tests and Spatial Price Linkages in Regional Cattle Markets." *Amer. J. Agr. Econ* 73(May 1991):452–64.

Johansen, S. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59 (November 1991):1551–80.

——. "Statistical Analysis of Cointegration Vectors." *J. Econ. Dyn. Control* 12(June/September 1988):231–54.

Li, J.R., and C.B. Barrett, "Distinguishing Between Equilibrium and Integration in Market Analysis." Unpublished, June 1999.

MacKinnon, J.G., A.A. Haug, and L. Michelis. "Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration." Working paper, Department of Economics, Queen's University, 1996.

McNew K., and P.L. Fackler. "Testing Market Equilibrium: Is Cointegration Informative?" *J. Agr. Res. Econ.* 22(December 1997): 191–207.

Osterwald-Lenum, M. "A Note with Fractiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics." *Oxford Bull. Econ. Stat.* 54(August 1992):461–72.

Pesaran M.H., and Y. Shin. "Cointegration and Speed of Convergence to Equilibrium." *J. Econometrics,* 71(March 1996):117–43.

Phillips, P.C.B. "Optimal Inference in Cointegrated Systems," *Econometrica,* 59(March 1991):283–306.

Ravallion, M. "Testing Market Integration," *Amer. J. Agr. Econ.* 68(February 1986):102–9.

Sexton, R.J., C.L. King, and H.F. Carman. "Market Integration, Efficiency of Arbitrage, and Imperfect Competition: Methodology and Application to U.S. Celery." *Amer. J. Agr. Econ.* 73(August 1991):568–80.

Shively, G.E. "Food Price Variability and Economic Reform: An ARCH Approach for Ghana." *Amer. J. Agr. Econ.* 78(February 1996):126–36.