Journal of Agricultural and Resource Economics 22(2):191–207

Testing Market Equilibrium: Is Cointegration Informative?

Kevin McNew and Paul L. Fackler

Cointegration methods are increasingly used to test for market efficiency and integration. The economic rationale for these tests, however, is generally unclear. Using a simple spatial equilibrium model to simulate equilibrium price behavior, it is shown that prices in a well-integrated, efficient market need not be cointegrated. Furthermore, the number of cointegrating relationships among prices is not a good indicator of the degree to which a market is integrated.

Key words: market integration, spatial markets, time-series analysis

Introduction

It has become common to apply cointegration techniques to the analysis of spatial price relationships, both to test the law of one price (LOP) and to examine the degree to which different regions are mutually integrated. Cointegration models presuppose that observable variables exhibiting nonstationary behavior will nonetheless maintain long-run relationships. These long-run relationships are conceptually interpreted as stochastic (long-run) economic equilibria. In this view an economy is described as a multivariable dynamic system, equilibria as the attractors of the system (the set of points towards which the system tends to move), and error-correction mechanisms as the forces that move prices toward the attractor.

Among the most common examples of an error-correcting mechanism is arbitrage in a spatial market; this is the example that Engle and Granger use in their introduction to a volume of readings on cointegration. The essence of the argument is that prices of a homogenous good from two different regions should tend to be equal in the long run. More specifically, an extended period with no exogenous shocks would move the two prices towards equality. The error-correction mechanism is the arbitrage process. Profit opportunities arise when the economy is away from the attractor and arbitrage forces prices back towards the attractor defined by the relationship $p_1 = p_2$. This is essentially a statement of the LOP.

A number of studies have used this idea to test the LOP (Ardeni; Goodwin 1992a; Goodwin and Grennes; Michael, Nobay, and Peel).[1] Some studies simply test for cointegration; others contend that not only should cointegration be present but a specific linear price relationship should be stationary. Baffes argues that price movements in one

---

The authors are, respectively, an assistant professor in the Department of Agricultural and Resource Economics at the University of Maryland and associate professor in the Department of Agricultural and Resource Economics at North Carolina State University.

[1] Cointegration has also been used to test forward market efficiency (Corbae, Lim, and Ouliaris) and the rationality of the term structure of interest rates (Hall, Anderson, and Granger). Like the LOP, these involve arbitrage relationships defined by complementarity conditions. The results of this study, therefore, also extend to these concepts.

location should be matched, in the long run, with one-for-one price movements in another location; this is equivalent to the stationarity of price spreads.

A corollary to the above assertion is that the absence of trade between regions (either direct or indirect) should result in prices that are not cointegrated. This has led some researchers to use cointegration as an indicator for the strength of regional connections. For example, Goodwin (1992b) and Goodwin and Grennes suggest that a system of n spatial prices should have at least one cointegrating relationship, and the number of cointegrating relationships among prices can indicate the extent of market integration. Thus, full market integration requires $n - 1$ cointegrating vectors and any number of cointegrating vectors less than $n - 1$ implies markets which are not fully integrated. Similar suggestions are made in Benson et al. (1994a) and Silvapulle and Jayasuriya. Also, some studies perform bivariate tests on price pairs, associating the degree of integration for a given location with the number of other locations exhibiting cointegration (Alexander and White).

The purpose of this study is to suggest caution in the application and interpretation of cointegration models in analyzing spatial price behavior. To demonstrate that such caution is warranted, we develop an equilibrium spatial market model that is used to simulate prices. The simulations allow us to explore the hypothesis that cointegration in prices should occur because the forces of market integration and efficiency tend to result in linearly related prices.

Unlike many studies, we sharply distinguish between the concepts of efficiency and integration. In our treatment, the concepts of efficiency and the LOP are synonymous and taken to mean that arbitrage opportunities are quickly eliminated and therefore negligible in observed variables, including prices. This feature is a necessary condition for market equilibrium and, therefore, is a distinguishing feature of price behavior. Market integration, on the other hand, we define as the extent to which shocks arising in one location are passed on to other locations, a meaning that is consistent with the work of Harriss and Ravallion. The specific definition (stated in a later section) will arise naturally out of the equilibrium model used here.

The analysis indicates that neither efficiency nor market integration necessarily leads to linearly related prices. Our demonstration centers around the arbitrage mechanism, which is shown to be an insufficient force to ensure a simple linear relationship among spatial prices. There are two parts to this demonstration. First we show that if the underlying forces affecting supply and demand in different regions are not cointegrated, arbitrage alone will not guarantee that prices exhibit cointegration, especially as transport rates increase in size and volatility. Second, if the demand and supply forces are themselves cointegrated across regions, an analyst may conclude that prices are cointegrated regardless of whether there are interregional flows of commodities and associated binding arbitrage conditions. Examples of cointegrated economic shocks include weather variability in agricultural production, public policies that have similar impacts across regions, or income and inflation factors. In these instances, the presence of cointegration does not indicate that arbitrage is the source of the error correction mechanism. It therefore follows that the degree of cointegration among prices is not a useful measure of the strength of the interregional market integration.

The suggestion that statistical measures do not always have simple economic interpretations is not new. Harriss and Ravallion argue that high price correlations do not necessarily indicate a high degree of market integration. Transaction costs are also known to influence cointegration tests. For example, Davutyan and Pippenger show that one is less likely to find cointegrated prices or stationary price spreads when transaction costs are large. Goodwin (1992a) suggests that the lack of cointegration for international wheat prices may be due to nonstationarity in ocean freight rates; Hsu and Goodwin provide further evidence on this point. Indeed, Granger also suggests that nonstationary risk premia may explain the lack of cointegrated treasury bond prices in the early 1980s.

In spite of these cautions, the number of studies that pay no attention to the problem greatly outweigh those that even mention it. Furthermore, the issue has not been explored systematically within the context of an economic model of spatial price determination. A novel feature of the current article is that an explicit spatial equilibrium model is used to generate simulated prices with known economic properties. This represents a kind of controlled experiment that allows us to explore whether the interpretations placed on cointegration tests are justified.

To focus attention on the potential problems in applying cointegration to the analysis of spatial prices, we first discuss spatial models of price behavior and their implications for cointegration. Later we discuss a very simple spatial equilibrium model and an associated measure of market integration are developed. The model is used to generate simulated price data to illustrate some of the problems discussed above.

Spatial Equilibrium Models and Price Spreads

A number of spatial equilibrium models currently exist, although many (if not most) can be classified into two groups. The first, originating with Hotelling and Smithies, is a network framework with markets or firms located at network nodes and consumers or commodity producers located continuously along network links. Agents located along the links will choose to transact at the node offering the best price net of transport costs. Such models, which we call agents-on-links models, are often, although not exclusively, used to model spatial oligopoly situations where firms exert some local monopoly or monopsony power but may gain or lose market share to firms located at other nodes. Benson et al. (1994a,b) study the cointegration of spatial prices in the context of such a model, but provide no formal link between the model and cointegration methods. The application of cointegration methods in the analysis of market power has been criticized by Werden and Froeb.

The second group of models is also based on a network structure but the network links serve only for commodity transportation flows. Such point-location models originated with Enke and Samuelson and were popularized by Takayama and Judge. They have mainly been used to model perfectly competitive markets characterized by distinct regions or centers of activity.

Which of these two classes of models is appropriate depends on the nature of the market. For example, agricultural products are often produced in rural areas around isolated processing plants. The plants have local monopsony power, but are constrained by the possibility that producers can ship to competitors' plants. In a study of the prices paid to producers at the plant, the first model would therefore be appropriate. On the other hand, many bulk goods are collected or processed at a small number of points and then transported to a small number of major distribution centers; grain, coal, and gasoline markets are examples. In such cases, as well as in many international trade situations, especially those involving ocean freight or tariffs imposed at the border, point-location models are appropriate.

The issue of whether a perfect competition assumption is appropriate is conceptually different from that of the choice of a spatial structure. An agents-on-links model with a competitive assumption might be appropriate in a situation in which several firms are located at each of the major distribution centers.

In the point-location model each pair of nodes is either linked by trade or it isn't. If the nodes are linked by trade (either directly or indirectly) then prices will differ by an amount that depends only on transport rates. To illustrate, suppose region 1 ships to region 2, so there is a direct trade link. The appropriate arbitrage condition is $p_2 - p_1 = r_{12}$, where p_i is the price in region i and r_{ij} is the cost of shipping from region i to j. The regions may also be indirectly linked. For example, suppose regions 1 and 2 both ship to 3; prices must satisfy $p_2 = p_3 - r_{23}$ and $p_1 = p_3 - r_{13}$, implying that $p_2 - p_1 = r_{13} - r_{23}$. An immediate implication for dynamic price behavior is that price spread stationarity requires transport rate stationarity.[2] This is an empirical question, and indeed, there is some evidence to suggest that ocean freight rates, at least, behave like many other prices in exhibiting nonstationarity (Hsu and Goodwin).[3]

With an agents-on-links model the relationship between the nodal prices is even more complicated. Between each node a boundary exists at which an agent is indifferent between transacting with one node or the other. Thus, the prices at the nodes will generally not differ by the cost of transporting the goods between them, and hence, arbitrage between the two nodal points cannot be the mechanism generating cointegrated prices.

If there is a role for arbitrage as an error correcting mechanism in the agents-on-links model, it must be the arbitrage that occurs in determining the boundary. For example, suppose that regions 1 and 2 are d miles apart and there is a constant per mile transport cost, r, between them. The indifference boundary, B, measured in miles from region 1, satisfies the arbitrage condition $p_1 - rB = p_2 - r(d - B)$; equilibrium prices are therefore related according to

$$p_2 - p_1 = r(d - 2B). \tag{1}$$

In this case the stationarity of the price differentials depends on the stationarity not only of the transport rates but also of the market boundary times the transport rate. The location of the boundary will shift over time as demand, supply, and transportation costs change; thus it is determined in a complex way by all the relevant factors affecting the market. As these are the same factors causing the apparent price nonstationarity, arbitrage alone cannot account for price spread stationarity.[4]

[2] More precisely, for all price spreads to be stationary, all transport rates must be stationary. It is possible, however, for some spreads to be stationary and not others. For example, suppose that r_{13} and r_{23} are nonstationary but that $r_{13} - r_{23}$ is stationary. Consider a market in which locations 1 and 2 consistently ship to location 3, so $p_1 = p_3 - r_{13}$ and $p_2 = p_3 - r_{23}$. This implies that $p_2 - p_1$ is stationary but $p_2 - p_3$ and $p_1 - p_3$ are not.

[3] If transport rate data were available, cointegration among the prices and the transport rates could be analyzed. Such data are generally difficult to obtain and even when transport rates are available, other important transaction costs that affect price relationships may not be.

[4] The theoretical differences between the two spatial models may be less than is commonly supposed. Agents-on-links models differ from point-location models by allowing marginal adjustments at the indifference boundary. In point-location models, shipping patterns represent discrete regimes and thus such marginal adjustments cannot occur. As the number of nodes increases, however, a point-location model can mimic the continuous behavior of an agents-on-links model. Although this is impractical for empirical modeling, it suggests that problems encountered with one model class may be encountered in the other. If a problem arises in the use of a point-location model that does not depend on the number of nodes, it is likely that this problem will also exist in the limiting agents-on-links type model as $n \rightarrow \infty$. Thus, the differences in behavior often attributed to these alternative model formulations may, in fact, have more to do with the competitive structure rather than the spatial structure of the models.

At a minimum, cointegrated prices and stationary price spreads depend on the stationarity of transport rates, as noted in previous studies. Problems with cointegration analysis may still exist, however, even when transport rates are stationary. To show this, we develop a fairly simple model of spatial price determination and use it to demonstrate how such problems can arise. Given that stationary price spreads appear more likely in a point-location than an agents-on-links model, we use the former to make that demonstration.

An Explicit Point-Location Model

In this section an explicit point-location spatial equilibrium model is discussed. The model depicts the simplest of spatial settings with endogenously determined prices and shipment quantities. There are n regions, each with a linear excess demand function that depends on the price in the given region, along with a set of transport rates between each region. The excess demand for a homogeneous commodity in location i we parameterize as:

(2) $$q_i = b_i(a_i - p_i),$$

where q_i is the quantity and p_i is the price. The terms a_i and b_i are strictly positive parameters; a_i is referred to as the autarky price for location i and represents the price at which location i neither imports nor exports. When $q_i > 0$ (<0), location i is a net receiving (shipping) location. Let $r_{ij} \geq 0$ represent the cost of transporting the commodity from location i to location j; transport services have perfectly elastic supply.[5]

The nonnegative amount shipped from location i to j is denoted s_{ij}, with

(3) $$q_i = \sum_j (s_{ji} - s_{ij}).$$

Necessary and sufficient conditions for equilibrium in this model are as follows:

(4) $$\sum_i q_i = \sum_i \sum_j (s_{ji} - s_{ij}) = 0,$$

and

(5) $$p_j - p_i - r_{ij} \leq 0, \qquad s_{ij} \geq 0, \qquad (p_j - p_i - r_{ij})s_{ij} = 0 \qquad \forall i, j.$$

The first of these represents the material balances identity that the sum of imports equals the sum of exports. The second expression is a set of complementarity conditions representing the absence of arbitrage opportunities, namely, the LOP.

As Samuelson first pointed out, this model can be formulated and solved as a quadratic programming problem. With a small number of locations the solution can be directly expressed in terms of the model parameters. For example, with two locations, there will be trade if one of the autarky prices, say a_2, exceeds a_1 by at least r_{12}, in which case

(6) $$p_1 = \frac{1}{b_1 + b_2}(b_1 a_1 + b_2 a_2 - b_2 r_{12}),$$

[5] Although the spatial metaphor is used here and the r_{ij} are referred to as transport rates, other interpretations of the model are possible. For example, it could be taken to be a model of currencies denominated in a numeraire currency, with transaction costs incurred in exchange. In general, this model can represent a market for any good that is transformed into another good at a rate that does not depend on the amount transformed.

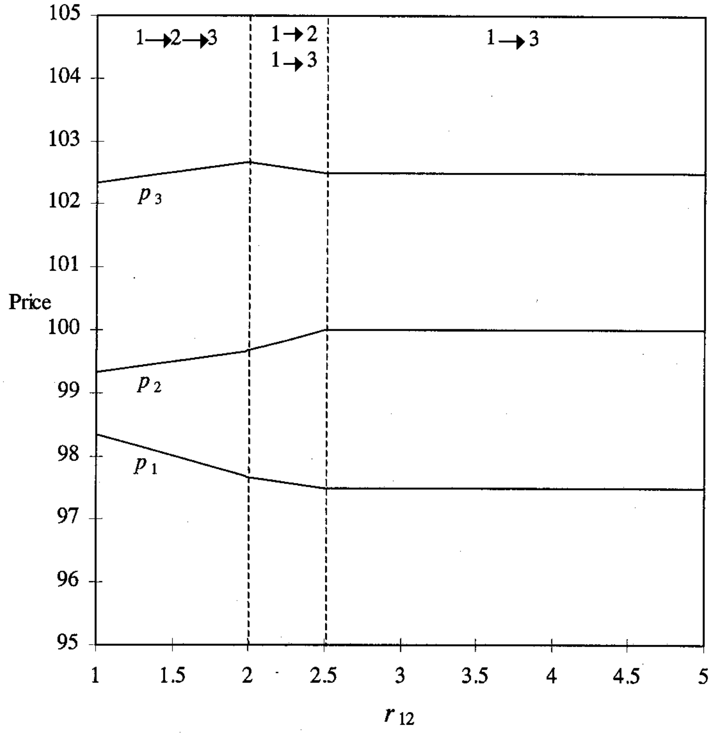

Figure 1. Price responses to changes in the transport rate from location 1 to 2

and p_2 exceeds p_1 by r_{12}. If there is no trade, of course, each price will equal its respective autarky price level. In this simple two-region model there are three different possible shipping patterns: 1 to 2, 2 to 1, and autarky.

Considering only two locations, however, does not provide much richness in terms of interregional price behavior. The simulations performed later in this article make use of a three-location case, which allows regions to be connected by trade indirectly. For example, locations 1 and 3 may both ship to location 2, implying prices at 1 and 3 must differ by exactly the difference in the relevant transport rates ($p_3 - p_1 = r_{12} - r_{32}$). In the three-location case there are 19 different shipment patterns. Details of the equilibrium solutions are given in an appendix.[6]

The nonlinear nature of price response to shocks is illustrated in figures 1 and 2.[7] Both figures demonstrate how the prices in the three locations change in response to a change

[6] Three is the largest number of locations for which it is practical to derive the equilibrium conditions explicitly because the number of possible patterns grows very fast as more locations are added. For example, with four locations the number of possible shipping patterns is 189!

[7] Both figures use the following base case parameters:

$$a = [95\ 100\ 105]^T$$

$$b = [1/3\ 1/3\ 1/3]^T$$

and

$$R = \begin{bmatrix} 0 & 4 & 5 \\ 6 & 0 & 3 \\ 5 & 7 & 0 \end{bmatrix}.$$

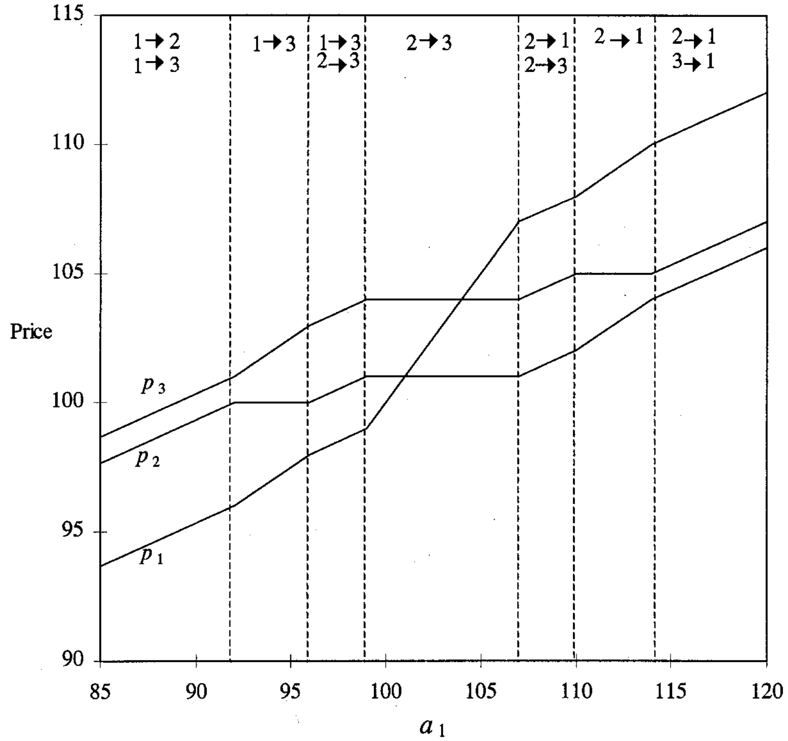

Figure 2. Price responses to excess demand shifts in location 1

in a single parameter. As the parameters change, the market moves through various shipping patterns, which are noted at the top of the figures and are delineated by dashed lines.

Figure 1 demonstrates the response of prices to changes in a single transport rate, r_{12}, the rate for shipping from location 1 to location 2. When this rate is very low, it pays for location 1 to ship directly to location 2, which transships on to location 3. In this case, an increase in r_{12} causes location 1's price to decline and the prices in the other locations to increase (both at the same rate). For $2 < r_{12} < 2.5$, however, it is better to ship directly from 1 to 3, rather than to transship through 2 (the direct cost from 1 to 3 is 5 whereas the transshipment cost is $3 + r_{12}$). Notice that, in this interval, the prices in locations 1 and 3 are linked one-for-one, both dropping as more shipments are diverted away from the increasingly costly 1 to 2 route. For $r_{12} > 2.5$ it is no longer profitable for location 1 to ship to location 2; further increases in r_{12} have no impact on prices because the route is not used.

Figure 2 displays the response of prices to excess demand shifts in location 1. When a_1 is low (less than 92), location 1 is a net surplus region and it supplies both of the other locations. Notice that the slopes of the price response functions are equal, so the prices move in a one-for-one fashion. As the autarky price in location 1 increases (greater local demand or less local supply) it ceases to supply location 2. Thus for $92 < a_1 < 96$, the price in location 2 is not responsive to excess demand changes in location 1. Further increases in a_1, however, stimulate location 2 to become a supplier for location 3, and the three prices become mutually linked again. For $99 < a_1 < 107$, region 1 is no longer a net supplier and its price becomes delinked from those of the other regions.

For a_1 in this interval, the price in location 1 rises quickly and goes higher than the prices in the other locations. When it rises high enough, it becomes profitable for location 2 and, eventually, location 3 to supply location 1.

In addition to demonstrating the highly nonlinear nature of the response of prices to the underlying parameter values, these figures demonstrate several points that have important implications for empirical analysis of price comovements. First, transport rate changes can cause prices to move inversely to one another, even when shipping patterns don't change. The second, and more important, point concerns relative price responses to changes in excess demand. When the markets are all linked by trade, all of the prices rise at the same rate in response to an increase in excess demand. When a region is not linked by trade, however, there is no price response in that region to an increase in excess demand in another region. Although this feature is specific to this simple model, it suggests that trade linkages are related to the transmission of shocks from one region to another.

As we have noted, there is no standard definition of market integration in the literature. Many studies use the term as a measure of market connectedness, but others take integration and market efficiency to be identical concepts. As we already use the term efficiency to mean the lack of arbitrage opportunities, we will use the term integration to refer to the connectedness of the markets. More specifically, we would like to have a measure of how large an impact one region has on another.

We propose, therefore, a definition of market integration between two regions as the degree to which a shock arising in one location is transmitted to the other. Specifically, excess demand shocks originating in one region will have price effects in that location and, to the extent that the region is integrated, will have price effects in other locations. A scaleless measure of the interregional effects of the price effect in the other region relative to the price effect in the region originating the change is the price transmission ratio, defined as

$$PTR_{ij} = (\partial p_j/\partial a_i)/(\partial p_i/\partial a_i).^{8} \qquad (12)$$

In this simple model, if two regions are linked by trade, either directly or indirectly, a small excess demand shock will have the same price impact in both regions, so the price transmission ratio is 1 (this is evident in appendix table A1). Prices must exhibit this relationship to maintain the equality of the price spreads to the relevant transport rates. On the other hand, if the two markets are not connected by trade, there is no mechanism for the transmission of shocks and the price transmission ratio will equal 0.

The price transmission ratio must equal 0 or 1 for infinitesimal changes, but for discrete shifts in excess demand $(\Delta p_j/\Delta a_i)/(\Delta p_i/\Delta a_i)$ can lie anywhere on the [0,1] interval. This can occur when the change in the autarky price, a_i, causes a change in the shipping pattern, as was illustrated in figure 2. To measure the connectedness of two regions the expected price transmission ratio can be used. Market integration, therefore, is defined as

[8] For simplicity, we focus only on shocks causing parallel shifts in excess demand which is equivalent to a change in the autarky price. Equivalent results would be obtained if the price transmission ratio were defined in terms of changes in excess demand slopes.

$$I_{ij} = E[PTR_{ij}] = E[(\partial p_j/\partial a_i)/(\partial p_i/\partial a_i)].^9 \quad (13)$$

Whether regions are linked by trade in any time period depends on the relationship between the autarky prices (a_i) and transport rates (r_{ij}). Statistical relationships among prices and the degree of market integration depend, therefore, on the stochastic nature of these parameters. Furthermore, the nature of this dependence is extremely complicated, even for a three-location market, due to the nonlinearities inherent in the equilibrium conditions. To study the behavior of prices and the relationship between statistical and economic measures, we therefore turn to simulation.

Simulated Prices and Tests for Cointegration

The economic model of the previous section is static. Intertemporal price variation can be introduced through stochastic transport rates and shifts in excess demand. We confine attention to parallel shifts in excess demand, which are equivalent to stochastic variation in the autarky prices. Furthermore, for cointegration tests to be interesting, we require that prices be nonstationary and transport rates be stationary. The latter we assume and the former we achieve by assuming that autarky prices are nonstationary.

Market clearing is assumed to occur in every period and the market is efficient with prices exhibiting no arbitrage opportunities in equilibrium. This is a strong assumption and is more likely to hold for low rather than high frequency data. In actual markets, prices may not adjust instantaneously due to various information and scheduling frictions. Our point here, however, is not to model possible ways in which inefficiencies can arise but rather to demonstrate that, even in well-operating markets, the interpretation of cointegration tests is problematic. It is unlikely that the introduction of short-run frictions will make such tests more interpretable. Furthermore, we are trying to demonstrate that arbitrage is not a sufficiently strong error correcting mechanism to ensure cointegration of spatial prices. As such, we want to explore price behavior when arbitrage is working perfectly; otherwise a lack of cointegration might be attributed to a failure of the arbitrage process.

Having an explicit model of spatial prices permits exploration of the relationship between economic concepts such as efficiency (no arbitrage opportunities) or market integration and statistical tests like cointegration. By simulating prices derived from an economic model, the cointegration properties of prices can be explored in a known economic setting.

The simulations are structured to bring out two points. First, stochastic transport rates can cause market integration to be less than perfect, resulting in cointegration tests that fail to reject unit root price behavior even when the LOP holds instantaneously. This problem becomes worse as transport rates become more volatile and increase in proportion to the value of the delivered good. Second, even in situations where price spreads are stationary, cointegration tests can fail to be informative about the degree of market integration.

[9] In our simple model I_{ij} has another interpretation. The price transmission ratio is either 0 if two regions are not linked by trade or 1 if there is a trade linkage. In expected value, the price transmission ratio is therefore equal to the expected frequency that a trade linkage exists.

Simulation Design

The basic structure for the economic fundamentals is

(7) $a_t = \mu + Ca_{t-1} + \epsilon_t$ (autarky prices),

(8) $\epsilon_t \sim N(0, \Omega)$ (autarky price shocks),

(9) $r_t \sim N(\theta, \Sigma)$ (transport rates),

(10) $E[\epsilon_t r_t] = 0$, and

(11) $b_1 = b_2 = b_3 = 1/3$ (excess demand slopes),

where a_t is a three-vector of autarky prices for period t, r_t is the vector of transport rates for period t, and each slope, b_i, is constant for all periods. In all the simulations it is assumed, for simplicity, that transport rates are symmetric ($r_{ij} = r_{ji}$); r therefore is defined as the three-vector: $r = (r_{12}\ r_{23}\ r_{13})'$. Also, for simplicity, the slopes of demand functions (the b_i) are equal (11), and the autarky price shocks and transport rates are independent (10).

The simulations are not designed to mimic the far richer time-series properties of actual commodity prices or transport rates. Although possible to include higher order lags, seasonal or drift terms, such complications would include extraneous elements not related to the basic issues addressed. Instead we have tried to make the simulations as simple as possible in order to simplify the interpretation of the results.

Two sets of simulations are performed that differ mainly in the stationarity properties of the demand shocks. Simulation parameter values are given in figure 3. In simulation 1 the demand shocks are independent unit-root processes. This represents a situation in which prices wander freely in the absence of trade, but when trade occurs, prices are tied together by the arbitrage boundaries. Simulation 1, therefore, represents a situation in which the only candidate error correction mechanism is the arbitrage process. In this way the strength of the arbitrage process and the effect of transport rates on it can be isolated and examined.

Three different transport rate scenarios are used in this simulation: cases 1A, 1B, and 1C, representing low, medium, and high transport rate situations, respectively. The transport rate levels were chosen to represent approximately 2, 6, and 10% of the initial price of the commodity; such rates are not unusual for bulk commodities. With low transport rates, the regions are more likely to be fully integrated than in the case of high transport rates. Low transport rates result in more periods of binding arbitrage and hence a greater probability that price spreads appear stationary. As transport rates go to zero, price spreads would also be zero and, trivially, the nonlinearities in price behavior would be eliminated. On the other hand, as transport rates increase in mean size, the degree of market integration declines as regions are more often isolated from one another. This leads to more time spent within the arbitrage bands and therefore with prices delinked and free to exhibit nonstationary behavior.

In simulation 2, the autarky prices are nonstationary but are cointegrated in such a way that the autarky price spreads are stationary. This implies that price spreads will also be stationary, although they will not necessarily be representable by a simple linear time-series model. An important implication of price spread stationarity is that the system

General Form:

$$a_t = \mu + Ca_{t-1} + \epsilon_t$$

$$\epsilon_t \sim N(0, \Omega)$$

$$\Omega = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$r_t \sim N(\theta, \Sigma)$$

$$\Sigma = \frac{1}{4}\begin{bmatrix} 1 & 0.75 & 0.75 \\ 0.75 & 1 & 0.75 \\ 0.75 & 0.75 & 1 \end{bmatrix}$$

Simulation 1:

$$a_0 = \begin{bmatrix} 95 \\ 100 \\ 105 \end{bmatrix}$$

$$\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\theta_a = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, \quad \theta_b = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}, \quad \theta_c = \begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix}$$

Simulation 2:

$$a_{0A} = \begin{bmatrix} 102\frac{1}{3} \\ 95 \\ 102\frac{2}{3} \end{bmatrix}, \quad a_{0B} = \begin{bmatrix} 103\frac{1}{6} \\ 92\frac{2}{3} \\ 104\frac{1}{6} \end{bmatrix}, \quad a_{0C} = \begin{bmatrix} 105.5 \\ 85.5 \\ 109 \end{bmatrix}$$

$$\mu_A = \begin{bmatrix} 1.75 \\ -3.75 \\ 2.00 \end{bmatrix}, \quad \mu_B = \begin{bmatrix} 2.5 \\ -5.5 \\ 3.0 \end{bmatrix}, \quad \mu_C = \begin{bmatrix} 4.0 \\ -10.5 \\ 6.5 \end{bmatrix}$$

$$c = \begin{bmatrix} 0.50 & 0.25 & 0.25 \\ 0.25 & 0.50 & 0.25 \\ 0.25 & 0.25 & 0.50 \end{bmatrix}$$

$$\theta = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$

Figure 3. Parameter values for simulations

has a tendency to return to the specific trading pattern associated with the mean autarky price spread. In this simulation three different cases, denoted 2A–2C, are examined. The stationary point for the three simulations was chosen so the degree of trade among the locations varied while maintaining an average price of 100 at the stationary point.[10] Specifically, simulations 2A, 2B, and 2C represent low, medium, and high degrees of market integration, respectively.

Both sets of simulations use the following procedure. First, 200 time-series observations for a_t and r_t are drawn (typical sample sizes in the studies surveyed ranged from 140 to 180). Second, using these 200 realizations, the equilibrium shipping patterns and the associated prices for each period are computed. Third, using the shipping patterns, sample measures of market integration are computed. The sample market integration indices, I_{ij}, are computed as the frequency that locations i and j are linked by trade. Also

[10] In simulation 2, initial values of the a_i were chosen to coincide with the long-run average a_i.

calculated is a joint integration index, I_{123}, defined as the frequency with which all three locations are linked by trade. Fourth, using the sample spatial prices, price level correlations, cointegration tests on price levels, and unit-root tests on price spreads are calculated. These steps are replicated 6,000 times (to coincide with calculations made by Johansen and Juselius for the limiting distribution of cointegration test statistics) and average results are computed and reported.

The cointegration methods follow the work of Johansen and Juselius. Prices are assumed to be autoregressive of order one with no seasonality or drift terms. This specification coincides with autarky price behavior. The maximum eigenvalue test is used for inferences about the cointegrating rank of spatial prices based on a 95% critical value. The critical values differ slightly from those in Johansen and Juselius appendix table A2 because the sample size used here is 200 as opposed to 400 observations as used by Johansen and Juselius.[11]

To test for stationary price spreads, autoregressive models of order one were estimated for spatial price spreads. This analysis also assumes no seasonality or drift terms; the procedures and limiting distribution of the *t*-statistic used here can be found in Fuller (chapter 8). The price spread stationarity tests provide evidence on whether prices not only exhibit cointegration, but more specifically, on whether the cointegrating relationships imply one-for-one price movements in the long run.

Results

The mean and standard deviation of the market integration indices and simple price correlations over the 6,000 replications are reported in table 1. For simulation 1 transport rates increase from simulation 1A to 1C. The higher transport rates lead to more instances of market isolation (less market integration) and, because of independent excess demand shocks, lower price correlations. Simulation 2, on the other hand, demonstrates that low levels of integration, as in case 2A, do not necessarily imply low price correlations.[12]

Test results for the number of cointegrating relationships and for stationary price spreads are given in table 2. A test rejection using λ_i leads to the conclusion that at least $i + 1$ cointegrating relationships are present. A test rejection using τ_{ij} leads to the conclusion that the price spread $p_j - p_i$ is stationary.

In simulation 1, the probability of finding at least a single cointegrating relationship (reject λ_0) declines as the transport rate increases (from 93.3% to 27.9%). Furthermore, the frequency that two cointegrating relationships are found (reject λ_1) occurs less than 50% of the time, even in the low transport rate case. Similarly, the frequency with which price spreads are found to be stationary declines dramatically as mean transport rates increase (from 76%–87% with low rates to 20%–35% at high rates).[13]

Two points are in order relative to these findings. First, the tests fail to find evidence of the LOP, interpreted as prices exhibiting no arbitrage opportunities. The simulated prices are equilibrium prices at which all arbitrage opportunities have been eliminated

[11] Critical values for samples of size 200 were computed by the authors using the methods described in Johansen and Juselius (p. 207).

[12] This is similar to the problem discussed by Harriss of interpreting price correlations when excess demand shocks are correlated.

[13] The asymmetry in these results arises from the initial conditions for these simulations which make it more likely that either locations 1 or 3 would become isolated.

Table 1. Market Integration Indices and Price Correlations

	Simulation 1			Simulation 2		
	A	B	C	A	B	C
Market integration indices						
I_{12}	0.901	0.714	0.561	0.580	0.820	0.934
	(0.063)	(0.155)	(0.212)	(0.041)	(0.032)	(0.019)
I_{23}	0.901	0.713	0.559	0.673	0.967	0.999
	(0.064)	(0.158)	(0.212)	(0.039)	(0.013)	(0.001)
I_{13}	0.916	0.754	0.609	0.367	0.788	0.934
	(0.065)	(0.164)	(0.228)	(0.038)	(0.033)	(0.019)
I_{123}	0.862	0.618	0.440	0.367	0.788	0.934
	(0.076)	(0.167)	(0.202)	(0.038)	(0.033)	(0.019)
Price correlations						
ρ_{12}	0.961	0.787	0.604	0.961	0.982	0.988
	(0.043)	(0.209)	(0.331)	(0.030)	(0.014)	(0.009)
ρ_{23}	0.960	0.785	0.601	0.970	0.990	0.990
	(0.044)	(0.211)	(0.331)	(0.024)	(0.008)	(0.008)
ρ_{13}	0.961	0.785	0.615	0.944	0.985	0.993
	(0.045)	(0.232)	(0.354)	(0.043)	(0.012)	(0.006)

Note: This table displays the average values of 6,000 replications of the sample market integration indices and price correlations (standard deviations are given in parentheses). I_{ij} measures the fraction of the observations in which locations i and j are linked by trade. I_{123} measures the fraction of the observations in which all three locations are linked by trade.

Table 2. Proportion of Test Rejections

Test Statistic[a]	Simulation 1 A	Simulation 1 B	Simulation 1 C	Simulation 2 A	Simulation 2 B	Simulation 2 C
Number of cointegrating relationships						
λ_2	0.1	0.0	0.0	0.6	0.5	0.5
λ_1	42.3	5.1	1.0	97.2	97.9	99.1
λ_0	93.3	47.7	27.9	100.0	100.0	100.0
Stationarity of price spreads						
τ_{12}	86.5	40.3	20.0	99.8	99.8	100.0
τ_{23}	86.7	40.8	20.0	99.8	99.9	100.0
τ_{13}	76.4	42.1	34.6	99.9	99.7	100.0

[a] The cointegration tests, λ_k, are Johansen and Juselius maximum eigenvalue tests. The null hypothesis associated with the test is that the system of three price levels has k cointegrating vectors with corresponding alternative hypothesis of $k + 1$ cointegrating relationships. All test statistics are evaluated at the 95% confidence level. The stationarity tests τ_{ij} are Dickey-Fuller tests that the price spread is $p_j - p_i$ is nonstationary (unit-root) process.

and the LOP holds by construction. Therefore, a perfectly arbitraged market may not exhibit price spread stationarity, and indeed, it need not exhibit any cointegrating relationships, much less $n - 1$ of them. Thus, the finding that price spreads are not stationary or that prices fail to exhibit cointegration is a poor indication of whether the market is operating efficiently.

A second point relates the cointegration and stationarity tests to the degree of market integration. As market integration declines, there is a corresponding decline in both the number of cointegrating relationships detected and in the frequency of price spread stationarity. It would be tempting, therefore, to conclude that these tests provide reasonable indications of the degree of market integration. Simulation 2 will show, however, that such an interpretation is problematic.

In each case for simulation 2, table 2 indicates that the cointegration tests and price spread stationarity tests yield virtually the same results. Based on these tests one would conclude that there are two cointegrating relationships, that all price spreads are stationary, and that the markets are fully integrated. Recall from table 1, however, that these three cases differ dramatically in their degrees of market integration. Thus, the stationarity properties of prices reveal little about the degree to which markets are actually linked by trade. Market integration and price cointegration, therefore, are quite different concepts.

There is an important economic distinction between simulation 1 and 2. In simulation 1, prices are free to wander in the absence of trade. This implies that barriers to trade, such as high transport rates or border controls, lead to prices that exhibit no discernible relationship. Cointegration, in this case, can only be due to the presence of binding equilibrium arbitrage restrictions and, therefore, from some degree of market integration. Simulation 2, by comparison, exhibits price relationships even in the absence of trade, and cointegration tests are incapable of distinguishing the degree to which these relationships are due to market integration rather than to economic shocks with common sources.

The problems this raises for the applied researcher is that situations with distinct economic meanings produce similar time-series test results. Without further information about the true time-series structure of the economic fundamentals or adequate data on trade flows or transport rates, the analyst is unable to distinguish these two diametrically different economic environments. Specifically, if one is trying to assess the degree of market connectedness, cointegration tests are not very informative unless the cointegration properties of the unobserved autarky prices are known. On the other hand, if one is attempting to test whether prices satisfy arbitrage conditions (the LOP), one would need to know that markets were well integrated and that transport rates are stationary to have faith in cointegration tests.

The link between cointegration on the one hand and market efficiency and integration is particularly problematic when arbitrage costs are large and volatile. This tends to change the pattern of price arbitrage relationships, leading to highly nonlinear price behavior. For prices that exhibit such conditions, linear time-series methods appear to have limited usefulness for making inferences about market efficiency and integration.

Conclusions

The familiar admonition that "correlation does not imply causation" can be modified to that of "cointegration does not imply integration." At best, the relationship between the economic concepts of the law of one price and market integration and the statistical concept of cointegration is a complex one. We have argued that arbitrage alone is not necessarily a powerful enough error-correcting mechanism to produce cointegrated spatial prices and have suggested a number of situations in which the relationship between cointegration and arbitrage conditions is not consistent. First, if transport rates and other costs of arbitraging between regions are nonstationary then cointegration is unlikely, even when regions are engaged in trade. Second, if autarky prices (prices that would occur in the absence of trade) are not cointegrated, then the trading patterns among regions are likely to shift over time and there may be periods in which regions are not linked by trade and are therefore not strongly integrated. This can result in prices that are not cointegrated but never violate arbitrage bounds. Third, although it may be tempting to conclude that the lack of cointegation signals a lack of market integration, this connection does not hold if there are forces causing the autarky prices themselves to be cointegrated. This can cause price spreads to appear to be stationary regardless of the strength of the trading ties between the regions.

Although this article has focused on deficiencies in an increasingly common method, it also contains a framework that may offer more promising approaches to spatial market analysis. Equilibrium simulation methods, as those used here, can foster our understanding and development of such methods. By constructing economic equilibrium models, an analyst must lay bare the assumptions underlying the interpretation to be placed on the results. Furthermore, by simulating such equilibria, complex nonlinearities can be explored in a way that would be essentially impossible to analyze analytically. In this way, methods that explicitly model the nonlinearities or that incorporate more information concerning the cost of arbitrage could be explored.

[*Received November 1996; final revision received June 1997.*]

References

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

Ardeni, P. G. "Does the Law of One Price Really Hold for Commodity Prices?" *Amer. J. Agr. Econ.* 71(1989): 661–69.

Baffes, J. "Some Further Evidence on the Law of One Price: The Law of One Price Still Holds." *Amer. J. Agr. Econ.* 73(1991):1264–273.

Benson, B. L., M. D. Faminow, M. H. Maquis, and D. G. Sauer. "The Impact of Provincial Marketing Boards on Price Relations and Pricing Dynamics in the North American Slaughter Hog Market." *Appl. Econ.* 26(1994a):677–88.

———. "Intra-national Effects of a Countervailing Duty on the United States/Canadian Hog Market." *Rev. Agr. Econ.* 16(1994b):187–201.

Corbae, D., K. G. Lim, and S. Ouliaris. "On Cointegration and Tests of Forward Market Unbiasedness." *Rev. Econ. and Statis.* 74(1992):728–32.

Davutyan, N., and J. Pippenger. "Testing Purchasing Power Parity: Some Evidence of the Effects of Transactions Costs." *Econometric Rev.* 9(1990):211–40.

Engle, R. F., and C. W. J. Granger. *Long-Run Economic Relations.* New York: Oxford University Press, 1991.

Enke, S. "Equilibrium among Spatially Separated Markets: Solution by Electric Analogue." *Econometrica* 19(1951):40–47.

Fuller, W. A. *Introduction to Statistical Time Series.* New York: John Wiley and Sons, 1976.

Goodwin, B. K. "Multivariate Cointegration Tests and the Law of One Price in International Wheat Markets." *Rev. Agr. Econ.* 14(1992a):117–24.

———. "Multivariate Cointegration Tests and the Law of One Price: A Clarification and Correction." *Rev. Agr. Econ.* 14(1992b):337–38.

Goodwin, B. K., and T. J. Grennes. "Real Interest Rate Equalization and the Integration of International Financial Markets." *J. Internat. Money and Finan.* 13(1994):107–24.

Granger, C. W. J. "Developments in the Study of Cointegrated Economic Variables." *Oxford Bull. Econ. and Statis.* 48(1986):213–28.

Hall, A. D., H. M. Anderson, and C. W. J. Granger. "A Cointegration Analysis of Treasury Bill Yields." *Rev. Econ. and Statis.* 74(1992):116–26.

Harriss, B. "There is Method in My Madness: Or Is It Vice Versa? Measuring Agricultural Market Performance." *Food Res. Institute Stud.* 16(1979):197–218.

Hotelling, H. "Stability in Competition." *Econ. J.* 39(1929):41–57.

Hsu, J. L., and B. K. Goodwin. "Dynamic Relationships in the Market for Ocean Grain Freighting." Paper presented at AAEA Summer Meetings, Orlando FL, 1993.

Johansen, S., and K. Juselius. "Maximum Likelihood Estimation and Inference on Cointegration—With Applications to the Demand for Money." *Oxford Bull. Econ. and Statis.* 52(1990):169–210.

Michael, P. A., R. Nobay, and D. Peel. "Purchasing Power Parity Yet Again: Evidence from Spatially Separated Commodity Markets." *J. Internat. Money and Finan.* 13(1994):637–57.

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

Samuelson, P. A. "Spatial Price Equilibrium and Linear Programming." *Amer. Econ. Rev.* 48(1952):283–303.

Silvapulle, P., and S. Jayasuriya. "Testing for Philippines Rice Market Integration: A Multiple Cointegration Approach." *J. Agr. Econ.* 45(1994):369–80.

Smithies, A. "Optimum Location in Spatial Competition." *J. Polit. Econ.* 49(1941):423–39.

Takayama, T., and G. G. Judge. *Spatial and Temporal Price and Allocation Models* Amsterdam: North Holland, 1971.

Werden, G. J., and L. M. Froeb. "Correlation, Causality, and All That Jazz: The Inherent Shortcomings of Price Tests for Antitrust Market Delineation." *Rev. Indus. Organization* 8(1993):329–53.

Appendix: Solving the Three-Location Model for Prices

The solution technique to the three-location spatial equilibrium problem begins with imposing a particular shipment pattern among the three locations. Conditional on a given shipment pattern, equilibrium prices and shipment levels can be solved for in terms of the excess demand parameters and transport rates. Using these solutions, one can impose the necessary and sufficient requirements for prices and shipments and, thereby, derive the necessary and sufficient conditions for each equilibrium shipping pattern.

To demonstrate, consider the general shipping pattern of location i shipping to both locations j and k. The equilibrium condition along with the two arbitrage conditions for this shipping pattern are

$$q_i + q_j + q_k = 0,$$

$$p_j - p_i = r_{ij}, \quad \text{and}$$

$$p_k - p_j = r_{jk}.$$

After substituting in the linear excess demand functions, these conditions can be presented more conveniently in matrix notation as

$$\begin{bmatrix} b_i & b_j & b_k \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} p_i \\ p_j \\ p_k \end{bmatrix} = \begin{bmatrix} \sum a_i b_i \\ r_{ij} \\ r_{ik} \end{bmatrix}.$$

Table A1. Parametric Conditions and Prices for Equilibrium Shipping Patterns

Shipment Pattern	Prices[a]	Conditions
1. Autarky	$p_i = a_i,\ p_j = a_j,\ p_k = a_k$	$a_j - a_i - r_{ij} \leq 0\ \forall\ i \neq j$
2. $i \Rightarrow j, k$	$p_i = (\omega_i a_i + \omega_j a_j - \omega_j r_{ij})/(\omega_i + \omega_j)$ $p_j = (\omega_i a_i + \omega_j a_j + \omega_i r_{ij})/(\omega_i + \omega_j)$ $p_k = a_k$	$a_j - a_i - r_{ij} > 0,$ $\omega_i(a_i - a_k) + \omega_j(a_j - a_k) > \max[\omega_j(r_{ij} - r_{ik}) - \omega_i r_{ik}, -\omega_j r_{jk} - \omega_i(r_{ij} + r_{jk})]$ $\omega_i(a_i - a_k) + \omega_j(a_j - a_k) < \min[\omega_i r_{ki} + \omega_j(r_{ij} + r_{ki}), -\omega_i(r_{ij} - r_{kj}) + \omega_j r_{kj}]$
3. $i \Rightarrow j \Rightarrow k$	$p_i = \tilde{a} - (\omega_j + \omega_k)r_{ij} - \omega_k r_{jk}$ $p_j = \tilde{a} + \omega_i r_{ij} - \omega_k r_{jk}$ $p_k = \tilde{a} + \omega_i r_{ij} + (\omega_i + \omega_j)r_{jk}$	$r_{ik} - r_{ij} - r_{jk} > 0$ $\omega_j(a_j - a_i) + \omega_k(a_k - a_i) > (\omega_j + \omega_k)r_{ij} + \omega_k r_{jk}$ $\omega_i(a_k - a_i) + \omega_j(a_k - a_j) > (\omega_i + \omega_j)r_{jk} + \omega_i r_{ij}$
4. $i \Rightarrow j,\ i \Rightarrow k$	$p_i = \tilde{a} - \omega_j r_{ij} - \omega_k r_{ik}$ $p_j = \tilde{a} + (\omega_i + \omega_k)r_{ij} - \omega_k r_{ik}$ $p_k = \tilde{a} - \omega_j r_{ij} + (\omega_i + \omega_j)r_{ik}$	$r_{ik} + r_{kj} - r_{ij} > 0,\ r_{ij} + r_{jk} - r_{ik} > 0$ $\omega_i(a_k - a_i) + \omega_j(a_k - a_j) > (\omega_i + \omega_j)r_{ik} - \omega_j r_{ij}$ $\omega_i(a_j - a_i) + \omega_k(a_j - a_k) > (\omega_i + \omega_k)r_{ij} - \omega_k r_{ik}$
5. $i \Rightarrow k,\ j \Rightarrow k$	$p_i = \tilde{a} - (\omega_j + \omega_k)r_{ij} + \omega_k r_{kj}$ $p_j = \tilde{a} + \omega_i r_{ij} + \omega_k r_{kj}$ $p_k = \tilde{a} + \omega_i r_{ij} - (\omega_i + \omega_j)r_{kj}$	$r_{jk} + r_{ij} - r_{ik} > 0,\ r_{ik} + r_{ji} - r_{jk} > 0$ $\omega_i(a_i - a_j) + \omega_k(a_k - a_j) > (\omega_i + \omega_k)r_{jk} - \omega_i r_{ik}$ $\omega_j(a_j - a_i) + \omega_k(a_k - a_i) > (\omega_j + \omega_k)r_{ik} - \omega_j r_{jk}$

[a] $\tilde{a} = \Sigma\omega_j a_j$ and $\omega_j = b_j/(\Sigma b_i)$.

The solution for this example is

$$p = \begin{bmatrix} 1 & -\omega_j & -\omega_k \\ 1 & \omega_i + \omega_k & -\omega_k \\ 1 & -\omega_j & \omega_i + \omega_j \end{bmatrix} \begin{bmatrix} \omega^T a \\ r_{ij} \\ r_{kj} \end{bmatrix},$$

where $\omega_i = b_i/(\Sigma b_j)$. Using these solutions in the original system of excess demand functions, one can solve for the quantity of shipments in equilibrium. The shipment solutions for this example have the form:

$$\begin{bmatrix} s_{ij} \\ s_{ik} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \mathrm{diag}(B)(p - a),$$

where diag(B) is a diagonal matrix with the slope coefficients along the diagonal. Using the equilibrium price and shipment solutions, the necessary and sufficient conditions are imposed and used to derive the parametric shipping pattern conditions. These conditions, along with the equilibrium price solutions, are given in table A1 for the general shipping pattern cases.