Statistical Analyses of the Geographic Market Delineation with an Application to the U.S. Natural Gas Markets

John T. Cuddington and Zhongmin Wang

gas_survey(12.21.99).wpd

* Georgetown University, Washington, D.C. 20057-1036. Corresponding author: Cuddingj@gunet.georgetown.edu
* The research for this paper was funded by the Federal Energy Regulatory Commission. We thank Bill Meroney at the FERC for his comments, and access to the data series used in the empirical analysis.

Abstract

Statistical Analyses of the Geographic Market Delineation with an Application to the U.S. Natural Gas Markets

This paper has three objectives. The first is to review various statistical approaches for using price data to analyze the geographic extent of a market and to discuss their strengths and limitations from the point of view of modern time series econometrics. Second, the paper surveys the application of these techniques to the U.S. wholesale market for natural gas following deregulation in the 1980s. Third, we carry out a new statistical analysis of the natural gas market using a more comprehensive daily dataset. This dataset has over 50 pricing points and covers the period 1993-1997. We propose a superior approach to those in the gas market literature, namely the estimation of an error correction model (ECM). The ECM provides direct estimates of stationary arbitrage relationships among prices. In addition, it yields an estimate of the speed of adjustment toward the no-arbitrage-profit equilibrium.

Cointegration techniques, including the use of ECMs, are especially appealing when the price series under consideration are I(1). Much to our surprise, however, our unit root tests on daily gas prices found overwhelming evidence that these prices are I(0) processes, not I(1) processes as cointegration requires. These results differ sharply from those in the literature. In spite of this feature of our dataset, we provide an illustration of how to estimate an ECM and then employ impulse response analysis to assess the speed with which the no-arbitrage equilibrium is restored.

JEL Classifications:

L95: Industry Studies: Gas Utilities, Pipelines, Water Utilities
F15: Economic Integration
L11: Production, Pricing, and Market Structure
Q40: Energy: General

Statistical Analyses of the Geographic Market Delineation with Application to the U.S. Natural Gas Markets

1. Introduction

In the past decade, the U.S. natural gas industry has seen dramatic market-oriented regulatory changes. All three segments of the industry -- production, transmission, and distribution -- have undergone significant changes.[1] Following Congressional passage of the Natural Gas Policy Act in 1978, the Federal Energy Regulatory Commission (FERC), by a series of orders, acted to deregulate the pipeline shipment of natural gas. A number of studies have subsequently argued that deregulation has made the U.S. natural gas markets more integrated.

There is now a huge literature, spanning several fields in economics, analyzing the contemporaneous and temporal interactions among prices of otherwise homogeneous goods across various geographic markets. The objective is to assess the extent to which these markets are 'segmented' or conversely the extent to which they are, in effect, an integrated single market. More generally, these analyses attempt to determine the geographic extent of the market for a particular good. That is, which locations encompass a single market? Which markets fall outside this market, but perhaps lie within a second geographic market? This literature spans applied microeconomics

[1] Under current federal regulations, open access policy stops at the city gate and state public utility commissions have the right to allow or deny bypass of local gas distributors. See Walls (1994) for a detailed discussion.

(especially antitrust analysis[2]), international macroeconomics (e.g. tests of the law of one price for internationally tradeable goods, tests of purchasing power parity, studies of the degree of regional integration), the new '*intra*national' economics, and international financial markets (e.g. covered interest parity).

At the most basic level, this research involves examining the extent to which arbitrage activity links various markets for particular goods, services or financial assets. Conversely, it investigates the extent to which transactions costs, transportation costs, and other natural, manmade, or government-imposed barriers to trade have segmented the global, national, or regional markets for particular products.

This paper has three objectives. First, Section 2 reviews and critiques the various statistical approaches for using price data to analyze the geographic extent of a market. Key concepts of market integration, market segmentation, the law of one price (LOOP), the role of arbitrage in bringing about LOOP, and price convergence and common trends

---

[2] Werden and Froeb (1993) write "many economists have taken it for granted that the antitrust law concept of a relevant market must be the same as the economic notion of a market. Since the late 1970s, a substantial economics literature developed in which price data has been relied upon to delineate antitrust relevant antitrust markets by empirically implementing the definitions offered by classical economists." Presumably an an 'economic market' can be defined by using concepts such as the law of one price (LOOP). Antitrust markets, however, are delineated differently, they explain: "The antitrust analysis of a proposed merger is an effort to ascertain whether the merger would create or enhance market power. Under the Merger Guidelines, market delineation is a tool specially designed for use in that effort. Market delineation under the Guidelines entails the determination of the optimal price increase for a hypothetical, profit-maximizing monopolist over various groups of products and areas, each of which is termed a candidate market. ... A candidate market is considered an antitrust market if the optimal monopoly price increase above prevailing levels exceeds some significance threshold, generally taken to be five percent." In a footnote, they elaborate: "Werden (1981, pp. 720-22; 1983, pp. 514-16) first argued that the two concepts were essentially unrelated. Scheffman and Spiller (1987, pp. 127-28) and Spiller and Huang (1986, pp. 132-33) also noted that the two concepts were different, suggesting along the way that the classical definition of a market can and should be empirically implemented but then used for something other than antitrust analysis. They did not indicate the purpose to which one would put a market so delineated."

in price series are discussed. The importance of determining whether prices are stationary in levels (i.e., integrated of order zero, I(0)) or in first-differences (i.e., I(1)) for deciding on appropriate econometric approaches in highlighted. The methods employed in the literature include simple price correlations (in levels, first differences, log-levels, or log-differences), correlations on residuals from regressions of prices on some common factors, Granger causality, and cointegration, among others.

Second, the paper surveys the application of these techniques to the U.S. wholesale market for natural gas following deregulation in the 1980s. This is covered in Section 4, after summarizing recent regulatory and structural changes in the U.S. natural gas industry in Section 3.

Third, the paper carries out a new statistical analysis of the natural gas market using a more comprehensive daily dataset. This dataset has over 50 pricing points and covers the period 1993-1997. It is described and compared to alternative datasets used in the literature in Section 5. This Section also contains unit root analyses on our individual series. Section 6 explores and use and interpretation of cointegration analysis.

In Section 2, we propose a superior approach to those in the literature assessing the geographic extent of markets, namely the estimation of an error correction model (ECM). The ECM provides direct estimates of stationary arbitrage relationships among prices. In addition, it yields an estimate of the speed of adjustment toward the no-arbitrage-profit equilibrium. We stress that cointegration techniques, including the use of ECMs, are especially appealing when the price series under consideration are I(1). Much to our surprise, however, our unit root tests in Section 5 provide overwhelming evidence that daily U.S. natural prices over the 1993-97 period are I(0) processes, not I(1) processes as cointegration requires. These results differ sharply from those in the literature. In spite of this feature of our dataset, Section 6 illustrates the estimation of an ECM and then employs impulse response analysis to assess the speed with which the no-arbitrage equilibrium is restored. Section 7 concludes.

2. Market Segmentation, Arbitrage, and Integration: A Conceptual Framework

Consider geographically separate markets for a good (such as natural gas) that is assumed to be homogeneous in every respect other than location. For the moment, suppose there are only two such markets and that they are ***not*** linked in a way that permits arbitrage activity. The implications of relaxing this assumption are considered below. The nominal prices of the good in locations 1 and 2 at time t are denoted P_{1t} and P_{2t}, respectively.

In general, we hypothesize that P_{1t} and $P_{2t,}$ say, are determined by *k* supply and demand factors, $X_t \equiv [X_{1t}, X_{2t},...,X_{kt}]'$. Some of these factors or determinants, $X_{it,}$ are 'global' (i.e. identical across the two markets) while others are market/location-specific. Elements of the X_t vector presumably include the overall price level, global factor prices for any mobile factors (e.g. unskilled labor), factor prices for location-specific factors, the level and type of technology used in each location, underlying determinants of demand in each location (including income and tastes), among other things. Some of these factors may be easy to measure empirically; others may be thought of as unobservable factors. The long-run dependence of P_{1t} and P_{2t} on these underlying determinants in vector X_t can be captured by the following reduced-form equations:

$$P_{1t} = \beta_1 X_t + u_{1t} \tag{1}$$

$$P_{2t} = \beta_2 X_t + u_{2t} \tag{2}$$

Equations (1)-(2) only characterize the long-run equilibrium. Short-run dynamic interactions between the prices are captured in the error processes, which are assumed to be stationary, but may be correlated, and may exhibit serial correlation and heteroskedasticity.

Careful statistical analysis of the dynamic interactions between P_{1t} and P_{2t} first requires an assessment of whether the series are mean and variance stationary. Standard statistical and econometric techniques are generally valid only when working with stationary variables. If these techniques are applied to nonstationary series, spurious correlations and spurious regressions are likely to result. Nonstationarity in the mean of these price series can be caused by deterministic time trends, stochastic time trends or unit roots, and/or structural breaks in the underlying determinants of prices. Series that are stationary in levels, without the need to first-difference, are said to be I(0) or integrated of order zero. Series that must be first-differenced to render them stationary are said to be I(1), integrated of order one. Equivalently, they contain a (single) unit root.[3]

[3] Series that must be differenced twice to achieve stationarity are said to be I(2), or "integrated of order two."

There are a number of commonly employed unit root tests for pretesting prices series for the presence of one or more unit roots. Two of these, the augmented Dickey-Fulller (ADF) test and the Phillips-Perron (PP) test are used in Section 5 below. Note that for (price) series that exhibit subperiods of high volatility and other subperiods with low volatility, the PP test is more appropriate.

Before proceeding to unit root testing, one might ask: do we have *theoretical* priors about whether nominal spot prices for commodities like natural gas ought to be I(0) or I(1)? *Ex ante*, one would probably expect at least *some* of the determinants of prices in (1)-(2), e.g., real income, the overall price level, the level of technology currently used in production, to exhibit deterministic or stochastic trends. Other determinants of prices are likely to be stationary. These might include weather conditions, other seasonal factors, cyclical fluctuations in demand over the business cycle, etc. If at least some X_t's in (1)-(2) are I(1), prices must be I(1) as well.

On the other hand, one might conjecture that the *long-run* supply curve for natural gas is infinitely elastic. This might be the case if producers pump gas from the ground whenever the wellhead price exceeds a (relatively constant) marginal cost. For marginal cost to be constant, one would have to argue that the impact of increasing resource scarcity over time (as existing reserves are deleted) is just counterbalanced by the emergence of new technologies that aid the finding process.[4] Although some nonstationary X_t's may affect the demand for natural gas, one might argue that only stationary long-run supply considerations ultimately determine prices. If this situation

[4] On this, see Cuddington and Moss (1999).

prevails, prices should be I(0).

Consider the various possibilities regarding the stationary of the X_t's. Suppose, for the moment, that ALL of the X_{it} variables that determine prices are stationary or I(0) variables. As just mentioned, this implies that P_{1t} and P_{2t} must also be I(0). The reduced form equations in (1)-(2) could then be estimated with OLS. Note that the covariance matrix VAR(P_{1t} ,P_{2t}) will depend on the covariance structure among the X_t's and the error processes. As the X_t's and the error processes may both exhibit serial correlation, conditional and unconditional covariances among prices may be very different. This suggests that simple correlation analysis is not very meaningful because it ignores dynamic interactions between P_{1t} and P_{2t} .

Note that if *any* of the price determinants, i.e., the X_t's in (1)-(2), are I(1) variables, then P_{1t} and P_{2t} *must* be I(1) processes. In this case, the series P_{1t} and P_{2t} will not be mean variance stationary. The variances of P_{1t} and P_{2t} in levels or log-levels are not defined in this case. Spurious correlations are *inevitable*. Hence, it does not make statistical sense to examine correlation between the *levels* of the two I(1) series. Note that the spurious regression problem exists as long as *at least one* of the price determinants is I(1), regardless of whether P_{1t} and P_{2t} share any common or 'global' factors.

Suppose both price series are both I(1). Can the spurious correlation and spurious regression problems be avoided by considering the first-differences of prices?[5] No, this

[5] Stigler and Sherman (1985), for example, analyze the correlation of price in both levels and first-differences in determining the extent of the market.

too is inappropriate if there is any long-run equilibrium or ‘cointegrating’ relationship linking the two prices. The process of first-differencing filters out of the data the long-run equilibrium relationships and the associated adjustment dynamics towards this equilibrium that we are trying to detect!

If at least one of the X_{it}’s that enters each of the price equations is I(1), P_{1t} and P_{2t} should, in theory, be I(1). If the individual price series are I(1), something we can test using now-standard unit root tests, we can go on to ask: are P_{1t} and P_{2t} cointegrated? That is, is there any linear combination of the two prices that is stationary? More formally, if there exists a value of λ that makes the linear combination, v_t, stationary:

$$P_{1t} - \lambda P_{2t} = v_t \quad , \tag{3}$$

P_{1t} and P_{2t} are, by definition, cointegrated.

Substituting (1) and (2) into (3) we can see what, if anything, the presence of cointegration implies about the β_1 and β_2 vectors in the long-run price determination equations. It is easy to demonstrate that ***ALL*** X_{it} factors that are I(1) variables must affect P_1 and P_2 with the same proportionality factor (i.e. $\lambda=\beta_{1i}/\beta_{2i}$ ***for all i associated with I(1) variables***), in order for P_{1t} and P_{2t} to be cointegrated. This proportionality is extremely unlikely to arise as a fluke! Thus, if there are I(1) determinants of prices in one market that do not affect the second market, or if all of the I(1) factors affecting one market do not affect the other market in a way that is proportional to their impact on the other market, P_{1t} and P_{2t} can not, in theory, be cointegrated.[10]

Now relax the assumption that the two markets are segmented. Suppose that, in addition to the possibility of common factors determining P_1 and $P_{2,}$ the markets are linked by active arbitragers.[6] Geographic or cross-market arbitrage involves buying a good where it is cheap and reselling it where it is dear in order to make risk-free profit. Provided that: (i) the good in question is homogeneous, (ii) there are no transportation or other transactions or information-collection costs,[7] and (iii) there is free entry and exit into the market so that industry is perfectly competitive (including freedom to distribute and resell), unencumbered arbitrage should ensure that the so-called ***law of one price (LOOP)*** holds. That is, there will be "more or less continuous" price equalization across markets.

If arbitrage is sluggish or is sometimes infeasible in the short-run due to capacity constraints, say, then arbitrage opportunities may persist for a period of time. Nevertheless, if P_{1t} is the price of the commodity in location 1 at time t and P_{2t} is the price of the same good in location j at time, a no-arbitrage-profit equilibrium requires that ultimately:

$$D\,P_{2,t} \equiv P_{2,t} - P_{1,t} = 0. \qquad (4)$$

D is defined here as a difference operator to indicate a ***cross-location*** difference between the price at location i and the price at location j (at time t). We refer to DP_{2t} as a "price gap." Δ will be used to indicate a first-difference over time; i.e., $\Delta P_{1,t} \equiv P_{1,t} - P_{1,t-1}$.

---

[6] The concept of ***arbitrage*** is central in many areas of economics and finance. Geographic or cross-market arbitrage is the most straightforward application of this concept. Financial economics text has many more complicated applications. The famous Modigliani-Miller debt irrelevance propositions and the Black-Scholes options pricing formula, for example, are implications of "no arbitrage" conditions.

[7] Transactions and transportation costs imply that prices can differ by these costs before arbitrage becomes profitable.

With N locations where the good is traded, unrestricted arbitrage across all locations would imply that, for all markets i and j:

$$P_{i,t} - P_{j,t} = 0 \tag{5}$$

In this case, one might justifiably think of the N locations as parts of a "single market" or "integrated market" for the good. In contrast, if price equalization across the N markets

$$P_{1,t} = P_{2,t} \neq P_{3,t} \neq P_{4,t} =\ldots= P_{N,t}$$

was incomplete and persistent, e.g.,: the global market for the good is segmented. In the example immediately above, one might argue that there are, in effect, three separate markets. The first encompassing location 1 and 2, the second is location 3, and the third is locations 4 through N.

In many applications, it will be trivial to verify empirically that (4) or (5) does not hold exactly in each and every time period. Certainly, this is the case for the daily frequency data on U.S. wholesale gas prices examined below. Conceptually, we would not expect these "no arbitrage" conditions to hold perfectly. After all, if they really held exactly and continuously, there would never be any arbitrage opportunities to exploit. Hence, arbitragers would not survive. But then why would the zero-arbitrage condition hold?! This is the Grossman and Stiglitz (1975) argument on the impossibility of complete market efficiency.

In practice, it is necessary to ask: how severe -- in terms of size and/or persistence

-- are the deviations from the no-arbitrage-profit equilibrium condition in (4)? To hypothesize (4) holds, at least as a long-run equilibrium condition, is the same thing as saying the price gap series:

$$DP_{2,t} = P_{2,t} - P_{1,t} \tag{6}$$

is stationary. If P_1 and P_2 are both I(1) variables and there is a stable linear relationship like (4) that links them in the long run, P_1 and P_2 are cointegrated -- here with cointegrating vector [1, -1].[8]

Suppose the no arbitrage condition (4) holds, at least as a long-run equilibrium condition. (Speed of adjustment to equilibrium is discussed below.) What, if anything, does this imply about β_1 and β_2 in the long-run price determination equations (1)-(2)? It is easy to demonstrate that unless ALL of the X_{it} factors affect P_{1t} and P_{2t} with the same β coefficients (i.e. $\beta_{1i}/\beta_{2i} =1$ for all i associated with I(1) X variables), P_{1t} and P_{2t} can not be cointegrated with the cointegrating vector [1,-1]. This suggests that the common practice in the price correlations literature of trying to eliminate the common influences on P_{1t} and P_{2t} before examining their correlation is fundamentally misguided. *If P_{1t} and P_{2t} are, in fact, prices in different parts of a single integrated market, ALL long-run factors that determine them must be the same!* For example, the beta coefficients in (1) must be proportional to those in (2).[9]

---

[8] Alternatively, the natural logarithm of the price ratio P_1/P_2 might be examined in order to get the price gap expressed in percentage terms. The logarithmic transformation may also be helpful in addressing heteroscedasticity in the data or regression residuals.

[9] Note that this is general enough to account for the Werden-Froeb (1993, p.339) example comparing the price of one and two pound boxes of Brand X laundry detergent.

Speed of Adjustment to the No-Arbitrage Equilibrium

With very active arbitrage activity and low transactions costs, one might expect (4) to hold more or less continuously. That is, deviations from the "long-run" equilibrium, e_{2t} in (4), would not just be stationary. They would be white noise, as any disequilibrium reflecting profitable arbitrage opportunities would be eliminated in a matter of hours or days, not months or years. On the other hand, if arbitrage occurs only slowly over time, and is periodically prevented by transportation capacity constraints, $DP_{2,t}$ may exhibit substantial serial correlation.

In practice, it will often be interesting to test whether the zero-arbitrage condition (4) holds in the long run and, if it holds, to consider the speed of adjustment to get a sense of how 'long' the long run is! This is easy to assess using an error correction specification. Engle and Granger showed that, if two prices are cointegrated, it is always possible to express the short-run dynamic transitional dynamics for each of the prices, as they move into long run equilibrium, by using an error correction model (ECM). That is, for P_{1t}:

$$\begin{aligned}\Delta P_{1t} &= \alpha_{10} + \alpha_{11}(P_{1t-1} - P_{2t-1}) \\ &+\gamma_{11}(L)\Delta P_{1t-1} + \gamma_{12}(L)\Delta P_{2t-1} + n_{1t}\end{aligned} \tag{7}$$

Analogously, for P_{2t}:

$$\begin{aligned}\Delta P_{2t} &= \alpha_{20} + \alpha_{21}(P_{1t-1} - P_{2t-1}) \\ &+\gamma_{21}(L)\Delta P_{1t-1} + \gamma_{22}(L)\Delta P_{2t-1} + n_{2t}\end{aligned} \tag{8}$$

The lag polynomials for ΔP_{1t} and ΔP_{2t} capture short-run dynamics. The lagged price gap ($P_{1,t-1}$ - $P_{2,t-1}$) in (7) and (8) are the error correction terms, with their respective coefficients α_{11} and α_{21} indicating the speeds with which P_{1t} and P_{2t} adjust in response to deviations from the long-run equilibrium. α_{11} and α_{11}-α_{21} should be negative to insure that the price gap approaches zero over time. (This does not preclude α_{21} being positive.)

In the special case where P_1 and P_2 are cointegrated with cointegrating vector [1,-1], it is possible to specify the dynamics of the ***price gap*** $DP_{2t} \equiv P_{2t}-P_{1t}$, rather than each price separately, in terms of a single-variable (i.e. DP_{2t}) – rather than two-equation-- error correction model:[15]

$$\begin{aligned} \Delta DP_{2t} &= \alpha_0 + \alpha_1 (DP_{2t-1}) \\ &+ \gamma(L) \Delta DP_{2t-1} + n_t \end{aligned} \tag{9}$$

The speed of adjustment α_1 must be negative and less than one in absolute value to insure convergence to the zero-arbitrage-profit equilibrium where $\Delta DP_{2t} = 0$. Its magnitude measures the speed of adjustment toward the equilibrium. If (4) holds *continuously*, we should find that $\alpha_1 = -1$. In addition, if continuous equilibrium prevails, one would expect $\gamma_{1i} = \gamma_{2i}$ for all i=1,2 in (7)-(8), so that the lag polynomial in (9) collapses to $\gamma(L)[P_{2t-1}-P_{1t-1}]= \gamma(L)DP_{2t-1}$.

In sum, if prices are I(1), testing for the presence of cointegration and then estimating an ECM, which provides an estimate of the speed of adjustment toward the long-run equilibrium, seems to be a very natural way to proceed. Indeed it is difficult to think of a superior approach for assessing the geographic extent of the market using only information on prices.

Consider, now, a situation where P_1 and P_2 are I(0) variables. In this case, the two variables can not be cointegrated. One might, however, argue that it is still appropriate to examine the univariate time series properties of the price gap, assessing its stationary and speed of transitional dynamics using (9). Arbitrage implies that the price gap must be stationary (and smaller in magnitude than the marginal cost of arbitrage). The speed of adjustment indicates how quickly arbitrage restores the no-arbitrage equilibrium. In fact, this argument would suggest the desirability of examining the extent of a geographic market by examining (6)or (9) directly, rather than working with individual prices series whose order of integration is uncertain. Unfortunately, as we will show below, ***any*** and all linear combinations of two stationary time series are stationary. No arbitrage needed.

It is noteworthy that the single-equation error correction specification for the price gap in (9) is very similar to that recommended in Horowitz (1981, p.9), although his approach is not motivated by or derived from modern time series econometric concepts. His specification [his equation (1)] assumes that the deviation of the price gap from its long-run equilibrium level follows an AR(1) process. Thus, it is a simplification of (9) where the coefficients on all of the lagged changes in the price gap terms are assumed to equal zero; i.e., $\gamma(L)\equiv 0$. The AR(1) coefficient equals the speed of adjustment in his framework.

Werden and Froeb (1993, 340-1) are highly critical of Horowitz's methodology, arguing that his price gap analysis is bound to produce spurious estimates of the speed of adjustment towards a no-arbitrage equilibrium:

> Horowitz' measure of the absolute speed of adjustment suffers from two serious problems. First, it is highly sensitive to the frequency of observations. What appears to be instantaneous adjustment in annual or quarterly data may appear to be very slow adjustment in monthly or weekly data. Second, as a practical matter, Horowitz' adjustment parameter, λ, is not a measure of speed of adjustment. The estimate of will tend to be significantly greater than zero whenever there is significant autocorrelation in the price data.24 [Fn. 24: Slade (1985, 294) and Baker (1987, 42-43) made similar points.] Since Horowitz did not propose to pre-whiten the data, that is most of the time. This is easily seen in a very simple example. Assume two price series that are both statistically and economically unrelated. Assume also that both series follow a first-order autoregressive process, and that the two have the same autoregressive parameter:
>
> $$p_{1t} = \beta p_{1,t-1} + u_t$$
> $$p_{2t} = \beta p_{2,t-1} + v_t$$
>
> Because λ is estimated by regressing $(P_1\text{-}P_2)_t$ on $(P_1\text{-}P_2)_{t-1}$, the probability limit of the estimate of λ is β. A high degree of autocorrelation will lead to the erroneous conclusion that adjustment is slow, and a low degree of autocorrelation will lead to the erroneous conclusion that adjustment is rapid. In fact, there is no adjustment at all.

This critique is worthy of careful consideration. If β=1, the P_1 and P_2 are nonstationary. In this case, the price gap equation is a spurious regression unless P_1 and P_2 are cointegrated. The Werden-Froeb critique is not damning in this case. Indeed, the AR(1) model of price adjustment is widely used in the international finance literature testing PPP. In this context, however, the underlying variables -- national price levels, exchange rates -- are unquestionable nonstationary in levels. See, e.g., Obstfeld and Taylor (1997; working paper version):

> Our starting point is the standard model, one used in many empirical analyses of purchasing power parity (PPP) and the law of one price (LOOP), namely the AR1 model. Let P_t^1 and P_t^2 be the log price levels of a good (or

composite good, or basket of goods) in two locations at time t. Adjustment models are concerned with the dynamic behavior of the price gap $z_t = P_t^2 - P_t^1$.

... Accordingly, we define x_t to be the detrended component of the price difference z_t, given by $z_t = \alpha + \beta t + x_t$, where x_t may be estimated as an OLS residual.

... In the standard model x_t is assumed to follow an AR1 process

$$\Delta x_t = \lambda x_{t-1} + e_t \tag{1}$$

where e_t is $N(0, \sigma^2)$ and λ, expected to be between zero and minus one, is called the convergence speed. Note that x_t is already detrended and demeaned, so there is no constant term in (1). Thus, price differentials are diminished by a fraction λ in each period, plus an error term. This type of model has been used countless times in the analysis of PPP and LOOP**.** The convergence speed is usually interpreted as a measure of the integration of markets or the efficiency of arbitrage between spatially separate locations, and is expected to depend on the good or composite goods under consideration, the nature of transaction and transportation costs for these goods, and other aspects of economic distance between locations.

What about situations where prices are stationary variables? Assume that the stochastic processes describing prices have constant terms:

$$p_{1t} = \alpha_1 + \beta_1 p_{1,t-1} + u_{1t}$$
$$p_{2t} = \alpha_2 + \beta_2 p_{2,t-1} + u_{2t}$$

If the prices series are stationary, -1<< β_1, β_2 <<1. The steady state levels for p_1 and p_2, denoted by bars, are, respectively:

$$\bar{p}_1 = \alpha_1 / (1 - \beta_1)$$

$$\bar{p}_2 = \alpha_2 / (1 - \beta_2)$$

β_1 is the speed with which p_1 approaches its long-run equilibrium; β_2 is the speed with which p_2 approaches its long-run equilibrium. The adjustment process for each variable is independent of the deviation of the other variable from its equilibrium. Nevertheless, there is a long-run equilibrium price gap. It equals:

$$\bar{p}_1 - \bar{p}_2 = \alpha_1 / (1 - \beta_1) - \alpha_2 / (1 - \beta_2) \quad .$$

To determine whether this price gap reflects arbitrage across the two markets would presumably require independent information on the size of arbitrage transaction costs relative to the equilibrium price gap. This is much less of an issue in situations where the prices series are I(1). In this case it is extremely unlikely that a linear combination of the two prices will be stationary unless the market are truly linked by arbitrage.

Evolution towards integrated markets following deregulation

Cointegration analysis, along with other standard econometric techniques, assumes that the relationship being estimated is stable over time. This may be problematic when analyzing the ongoing process of increasing market integration following deregulation. To facilitate the study of how initially segmented markets become more integrated, we can adapt the concepts of convergence and common trends that have been used to discuss the cross-country convergence of GDP growth rates in the economic growth literature. Adapting Bernard and Durlauf's (1995) approach, we define *price* convergence and common trends. Price series i and j are said to converge if the long run forecasts of these two series are equal at a fixed time t:

$$\lim_{k\to\infty} E(P_{i,t+k} - P_{j,t+k} \mid I_t) = 0$$

This definition can easily be generalized to the case of N price series. We can say that N price series converge if the long run forecasts of gas price in these N markets are equal at a fixed time t.

Price series i and j are said to contain a "common trend" if the long run forecasts of these two price series are proportional at a fixed time t:

$$\lim_{k\to\infty} E(P_{i,t+k} - \alpha P_{j,t+k} \mid I_t) = 0$$

We can also say that N price series contain a single common trend if the long-run forecasts of these price series are proportional at a fixed time t, let $\overline{P_t} = [\, p_{2,t} \;\; p_{3,t} \cdots p_{n,t} \,]$

$$\lim_{k\to\infty} E(\, p_{1,t+k} - \alpha' \overline{p}_{t+k} \, / \, I_t) = 0$$

Bernard and Durlauf show that their definition of convergence will be satisfied if the cross-location difference in two series, $p_{i,t+k} - p_{j,t+k}$, is a stationary process. If price series do not converge in the sense defined above, they may still respond to the same long run driving forces, i.e., they may face the same permanent shocks with different weights. If so, they will contain a single common trend.[10]

It is interesting to note that the requirement that the price gap is stationary holds

---

[10] These definitions lend themselves well to testable implications from the unit root/cointegration literature. For example, for price series i and j converge, they must be cointegrated with cointegrating vector [1,-1]. Price series i and j have a common trend if they are cointegrated with cointegrating vector [1,-1].

regardless of whether prices themselves are I(0) or I(1). This is an advantage in that it obviates the need to do extensive, and often inconclusive, unit root tests on a huge set of price series.

An alternative approach to coping with structural change is to estimate econometric models with time-varying parameters using Kalman filter methods. See the discussion of the King and Cuc (1996) article in Section 4 below.

3. Regulatory and Structural Changes in the U.S. Natural Gas Industry

Before deregulation, gas pipelines were regulated as merchant carriers who were required to purchase the gas they transported. Pipeline companies were prohibited from offering pure transportation services to their customers. FERC Orders No. 436 and 500 allowed pipelines to unbundle natural gas, the commodity, from its transportation. Pipeline companies became open access contract carriers. FERC Order No. 636 formalized that interstate pipeline companies must separate gas sales from transportation, and mandated open-access to interstate storage facilities. These policy changes led to a restructuring of the gas transportation network and presumably contributed to the integration of the national gas market.[11]

Market Institutions. Gas market institutions such as spot markets, future markets, and market centers,[12] have developed, leading to an increase in gas trading and

---

[11] See U.S. Department of Energy, Energy Information Administration (DOE/EIA) "Natural Gas 1996: Issues and Trends", and "Deliverability on the Interstate Natural Gas Pipeline" for details.

[12] See DOE/EIA "Natural Gas 1996: Issues and Trends" Chapter 3 for a discussion of market centers and market hubs.

related transactions. Spot markets began to emerge in the mid-1980s as open access spread through the pipeline network. They facilitate direct short-term (usually less than 30 days) gas trading among buyers and sellers. *Gas Daily*, an industry newsletter, currently reports daily prices for over 100 spot markets. Many of these spot markets have evolved into electronic trading centers.

Gas futures like NYMEX Henry Hub future contracts can be used for hedging price risks.[13] This presumably encourages gas trading.

As was the case following airline deregulation, market ‘hubs’ emerged and have been developing quite rapidly since deregulation. Market hubs or centers provide many services that facilitate gas trading and transportation, such as gas loans and parking, balancing, rerouting, and information services like electronic bulletin boards. By 1998, some 36 market centers had been developed throughout the North American pipeline grid.

A "release market" for transportation capacity also developed, whereby unused firm capacity can be sublet by other shippers. Companies such as gas distributors and utilities hold long-term firm contracts on pipeline transportation rights. When they have excess rights during a particular bid week, they can sell the unused capacity on the capacity release markets. These markets, therefore, increase the deliverability or utilization of the pipeline network.

Development of Pipeline Infrastructures. Pipeline expansion has also contributed to the integration of gas markets. From 1990 through 1997, total capacity on

[13] See Brinkmann and Rabinovitch (1995).

the interstate pipeline system increased by more than 15 percent. The network has become more interconnected, and the number of bottlenecks has been reduced.[14] In addition, access to underground storage has improved in recent years. Underground storage sites have become closely associated with market centers as they complement the short-term loaning and parking services offered there.

Market Intermediaries. Gas marketers emerged in mid-1980s. They bring gas sellers and buyers together, and facilitate various gas transactions.[15]

With the creation of the above-mentioned market institutions, as well as the emergence of gas brokers to coordinate gas and transportation transactions, gas producers (users) have more choices in selling (or buying) natural gas. More importantly, open access made it possible for market participants to carry out arbitrage if are large price differentials across locations emerge. Consequently, previously separated field markets have undoubtedly become more integrated.

4. Review of the Literature on Integration of the U.S. Natural Gas Market

There are now a number of papers studying the effects of deregulation on the natural gas markets: De Vany and Walls (1993, 1994), Walls (1994a, 1994b), Doane and Spulber (1994), King and Cuc (1996), and Serletis (1997). Most of these articles conclude that gas markets have gradually become more integrated since the deregulation of the mid 1980s. Serletis (1997) is the notable exception. Various empirical methods

---

[14] DOE/EIA "Deliverability on the Interstate Natural Gas Pipeline," p. 31.
[15] Dahl and Matson (1998), p. 398.

have been used to test market integration or the extent of the market. Approaches include price correlation analysis, Granger causality, linear cointegration. These approaches are commonly used in the literature to define market boundaries for trading markets where LOOP applies. See Table 1 for an overview of the papers discussed below.

More generally, there is a huge literature on LOOP and PPP (Purchasing Power Parity) in the international finance literature. See Rogoff (1996) for a comprehensive survey. There is also a large literature on the study of food market integration in the developing countries. Baulch (1997) reviews the commonly used econometric tests for integration of the food markets and assesses the statistical reliability of those tests by a series of Monte Carlo experiments. These tests are also used to delineate antitrust markets. Werden and Froeb (1993) provide a rather damning critique of this literature.

Table 1: Summary of Literature Applying Time Series Methods to the Extent of the U.S. Natural Gas Market

Author and Paper	Methods	Data	Results
De Vany and Walls (1993) Energy Journal	Unit Roots tests; Engle-Granger cointegration tests.	Daily spot prices for 20 field markets; Interval: July 1987-June 1991; Segmented into four one-year samples; Source: Gas Daily	More pairs are becoming cointegrated over the samples; Gas markets became more strongly integrated from 1987 to 1991.
Doane and Spulber (1994) Journal of Law and Economics	Price Correlation; Granger Causality; Unit root tests; Engle-Granger cointegration tests;	Monthly spot prices in five regions; Interval: 1984-91. Source: EIA.	Open access broadened the geographic scope of the spot market.
Walls (1994a) Review of Industrial Organization	Unit root tests; Johansen cointegration tests for price pairs; Cointegrating vector test.	Daily spot prices at 20 markets; Interval: 1989-1990; Source: Gas Daily.	The hypothesis of perfect market integration could not be rejected in most of the market pairs; A geographic pattern exists.
Walls (1994b) Energy Journal	Unit Roots tests; Engle-Granger cointegration tests.	Daily spot prices for 20 field markets and 6 pooling points and city gates; Interval: July 1990 –June 1991; Source: Gas Daily.	City markets are generally not as well integrated into the pipeline network as field markets.
De Vany and Walls (1994) Energy Policy	Institutional analysis; Pearson correlation coefficient.	The regional data set: monthly spot prices in five regions; Interval: Jan 1984 to Dec 1989; Source: EIA[1]. The second data set: bid-week spot prices; Interval: Feb 1988 to 1990; Source: Gas Daily	Spot prices became highly correlated after deregulation. A national market developed.
De Vany and Walls (1996) Journal of Regional Sciences	Network law of one price model; Vector Autoregression and likelihood ratio test	Spot price at city gates and hubs; Interval: 1990-1991; Source: Gas Daily.	Gas market is a national market except for arbitrage path and capacity constrained pairs.
King and Cuc (1996) Energy Journal	Unit root tests; Engle-Granger cointegration tests; Kalman filters.	Bid week prices at monthly frequency; Interval: various-Sep 1995 Source: Inside FERC's Gas Market Report; Canadian Enerdata Inc.	Price convergence occurring everywhere, but not equally; Evidence of east-west split; not single North American market yet.
Serletis (1997) Energy Journal	Unit roots test: Engle-Granger cointegration tests and multivariate Johansen test	Monthly bid-week prices Interval: June 1990 to Jan 1996. Source: Brent Friedenberg Associates in the Canadian Natural Gas Focus.	No cointegration relationship among price pairs at all. No East-West split.

Price Correlation Analysis

In their paper, "The Extent of the Market," Stigler and Sherwin (1985) suggested analyzing the correlations among prices at different locations, in both levels and first-differences, as a way of delineating relevant geographic markets for antitrust analysis. Their presumption is that a greater price correlation is associated with greater likelihood that the products are in the same geographic market. The Stigler-Sherman methodology has been applied to the U.S. natural gas market by De Vany and Walls (1994) and Doane and Spulber (1994).

This method has several shortcomings. Horowitz (1981, pp. 8-9), for example, discusses four important dangers of interpretation when using price correlations as market delineators:

> The danger that is most obvious and easily avoided is that of the analyst's spuriously inferring causality from correlation. In particular, for example, when *all* prices in a geographic area are increasing over time, say as part of an overall inflationary trend, then the prices ... will be positively correlated. The most obvious and appropriate solution is to identify the independent variables to which product prices are linked, and then to remove their effects through partial-correlation analysis.
>
> A second and related danger is that while, *ceteris paribus*, a high partial correlation might imply a single market, the absence of any [contemporaneous] correlation will not necessarily imply that there is not a single market...prices for the product in each area *tend* to equality, although at any point in time they may differ.
>
> A third concern...products or areas are not necessarily one in any meaningful sense. Instead, each market is subjected to the "restraining influence" of the other. Hence there are persistent price differences between the two markets [defined by transportation and other transactions costs], with the price in one market providing a ceiling to the price in the other...
>
> A fourth danger is that the simplistic correlation approach neglects the time dimension...The issue then is whether the markets do in fact constitute a single market over the long run, and the length of the long run. This and other issues can be dealt with through a readily applied and well-known regression approach.

Horowitz's observations here are interesting and his methodology anticipates well the error correction methodology in the modern time series econometric literature, which is appropriate when the price series being analyzed are I(1) series.

If the price series are I(1), whether this is caused by a common I(1) inflationary factor or something else, then simple correlations in the levels of prices are spurious. Correlations in first-differences ignore the dynamic interaction of prices as they move toward long-run equilibrium. If prices are I(0), then correlations in levels are well-defined, but again they ignore temporal interactions among prices.

When P_1 and P_2 exhibit significant serial correlation around their respective (deterministic or stochastic) time trends, the time window used in the correlation analysis is critical. Should one look at contemporaneous correlations or lead-lag relationships as well? If price equalization is not attained continuously, but only gradually over time, contemporaneously correlations may be less meaningful than those at longer lag windows. Consider the change in P_1 over s time periods defined as:

$$\Delta_s P1_t = P1_t - P1_{t-s}$$

Autocorrelation and cross-correlation functions would be useful for capturing the dynamic behavior of individual price and their interaction with other prices.

Are price correlations meaningful in EITHER levels or first differences? If the prices are I(1) and cointegrated, then the answer is no. When prices are cointegrated, there exists an error correction (ECM) specification. The ECM shows that the conditional variance of the two elements of the P vector will not be time invariant.

Rather, it will depend on lagged P and X and lagged differences of these variables. Common factors that impact both prices will certainly have an influence on the conditional variance. This suggests that simple correlation analysis in either levels or first-differences is based on unconditional variances, and hence can not address the question of whether prices are moving because of common determining factors, or in response to arbitrage activity from previous-period disequilibrium.

In addition to considering simple correlations, De Vany and Walls (1994) and Doane and Spulber (1994) follow Stigler and Sherman(1985) by regressing prices on so-called "common factors" that may impart a common trend to the regional price series. They then analyze the correlations of the residuals from the resulting price regressions to see whether their correlation is high or low. De Vany and Walls (1994) adjusted for seasonal demand factors. Doane and Spulber (1994) took into account the oil prices, producer price index, and seasonal dummy. Both papers find that the price correlation between regional spot markets increased substantially over the time of deregulation.

As mentioned above, however, it is inappropriate to attempt to filter out common factors when attempting to assess the extent of the market. If two markets are completely integrated, this implies that ALL factors that impact prices at one location have exactly the same impact on prices at a second location, which is part of the same market. Cointegration analysis allows the researcher to test for the number of common stochastic trends that link a number of I(1) variables.

Regression-based Tests

Many of the early tests of LOOP and PPP in the international finance literature were based on the following regression equation:

$$P_t^1 = \alpha + \beta P_t^2 + e_t \qquad \textbf{(10)}$$

Under the null hypothesis of LOOP holds "on average" we should find that $\alpha = 0$ and $\beta = 1$. To say more, one must carefully examine the properties of the residuals in regression equation. Suppose that e_t is serially uncorrelated, then deviations from LOP are very short-lived. Although fleeting, if deviations from LOOP are large in magnitude, relative to the transactions costs in executing arbitrage transactions, there may be temporary, but large, arbitrage profits to be had.

To evaluate some approaches used in the study of market integration, Prakash (1997) builds up a simple demand and supply model with a profit-maximizing arbitrageur. It is assumed there are two markets in the economy with a representative producer and consumer in each market, and an arbitrageur operating between the two markets. Prakash argues that the coefficient of β is some combination of the parameters and the error terms in the markets, and "it is not clear how this restriction($\beta = 1$) tests for market integration."

If e_t is stationary but exhibits serial correlations, arbitrage opportunities persist for several periods but are ultimately dissipated. Examining the speed of adjustment toward the LOOP may be of interest in this situation. During episodes of market liberalization, one would expect deviations from LOOP less persistent, but not necessarily smaller in magnitude (because freeing prices from regulation may result in considerably higher volatility).

There are a number of econometric pitfalls involved in the estimation of (10). In general, prices at the two locations will be jointly determined, i.e., mutually endogenous. If the price series are I(0), this will lead to simultaneity bias when estimating (10) using OLS. The resulting coefficient β is biased.

Modern times series analysis points to another difficulty if the price series P_{1t} and P_{2t} are nonstationary. If the series are not cointegrated, i.e., they are linearly independent of each other, the residual in (10) will be nonstationary. Spurious regressions result, as Granger-Newbold (1974) stressed. On the other hand, if the prices are cointegrated, the error process in (10) must be stationary. In this case, the OLS estimates are super-consistent, although small-sample bias may still be large. Moreover, t-statistics on coefficient estimates from the OLS regression do not have an asymptotically normal distribution, so hypothesis testing becomes much more difficult.[18]

Some early authors addressed the price nonstationarity issue by first-differencing the series before running the regression in (10):

$$\Delta P_t^1 = \alpha + \beta \Delta P_t^2 + e_t$$

This procedure, however, throws the baby out with the bath water. First-differencing eliminates the possible long-run linkage between the two prices that we are trying to identify.

[18] Using the ECM solves this problem as the t-statistics there are asymtotically normal.

Cointegration Analysis

Cointegration analysis has been used in a number of recent papers studying natural gas markets. The papers can be categorized into two types depending on how many price series they analyze simultaneously. Some papers consider pairwise comparisons among price series at different geographic locations; others use multivariate cointegration analysis. They also differ in terms of the cointegration test used. The Engle-Granger (1987) approach is capable of identifying at most one cointegration relationship. The Johansen test may detect more than one cointegrating relationship in situations where more than two prices are considered simultaneously.

Engle and Granger proposed a two-step approach for testing for cointegration and estimating a single long-run linkage between two or more variables. First, they test for the presence of a unit root in each of the variables. If both variables are I(1), they check to see if the variables are cointegrated. This is done by estimating (10) using OLS, then doing a unit root test on the residuals. If the null hypothesis that the residuals, e_t, have a unit root is rejected, the two price series P_1 and P_2 are said to be cointegrated. When two variables are cointegrated, Engle and Granger show that it is always possible to describe the dynamics of each of the variables in terms of an error correction model (ECM) that takes the following form:

$$\Delta P1_t = \alpha_1 + \gamma_1(\varepsilon_{t-1}) + \beta_{11}(L)\Delta P1_t + \beta_{12}\Delta P2_t + u_{1t}$$

$$\Delta P2_t = \alpha_2 + \gamma_2(\varepsilon_{t-1}) + \beta_{21}(L)\Delta P1_t + \beta_{22}\Delta P2_t + u_{2t}$$

ϵ_{t-1} is the lagged residual from the OLS regression in (10). γ_1 and γ_2 are the speeds of adjustment of P_1 and P_2, respectively, toward the long-run equilibrium estimated in (10). Using the Johansen cointegration test, it is possible to estimate the long-run equilibrium (10) and the two error correction equations in a single step. This involves estimating the two-equation system:

$$\begin{aligned} \Delta P_{1t} &= \alpha_1 + \gamma_1 (P_{1,t-1} - \alpha - \beta P_{2,t-1}) \\ &\quad + \beta_{11}(L)\Delta P_{1,t-1} + \beta_{12}(L)\Delta P_{2,t-1} + u_{1t} \end{aligned} \tag{11}$$

$$\begin{aligned} \Delta P_{2t} &= \alpha_2 + \gamma_2 (P_{1,t-1} - \alpha - \beta P_{2,t-1}) \\ &\quad + \beta_{21}(L)\Delta P_{1,t-1} + \beta_{22}(L)\Delta P_{2,t-1} + u_{2t} \end{aligned} \tag{12}$$

using maximum likelihood techniques, while imposing the across-equation constraint that the α and β are equal in (11) and (12). The term in brackets is the deviation from the long-run equilibrium in the previous period.

A. Pair-wise Cointegration Analysis

De Vany and Walls (1993) apply the Engle-Granger cointegration technique to evaluate market integration in the natural gas industry. They use daily spot prices between 190 market-pairs located in 20 producing fields and pipeline interconnections from July 1987 to June 1991. They divide the dataset into four one-year samples and to repeat the cointegration tests to see if more price pairs are cointegrated in later sub-samples, as one might expect during a process of ongoing market integration following deregulation. They found that the number of cointegrated market-pairs increases over time, from 46% in 1987-88 to 66% in 1990-91. Provided two market prices are found to be cointegrated, DeVany and Walls consider the two markets to be integrated with each other. They do not test whether the coefficient β in the long-run equilibrium relationship is one or not.

Walls (1994a) argues that a cointegrating vector of (1, -1) is a "stringent condition" for market integration. He employs the Johansen maximum likelihood technique to test whether the cointegrating vector is (1, -1) between market-pairs. His data are essentially the same data as those in De Vany and Walls (1993) except that the time period is limited to 1989 to 1990. Walls found that the null hypothesis of no cointegration was rejected for all the 19 market pairs at the 5% significant level, but not all of the market pairs have a cointegrating vector of (1, -1).

Walls (1994b) extended the study of natural gas market integration from field markets to include city markets, and examined how gas prices at city markets are linked with producing field markets prices. He adds three pooling hubs and three city gate markets to the sample in Walls (1994a). Walls tests 120 field/city market-pairs, and found evidence that city market prices are generally not cointegrated with field market prices. City markets that have adopted some policy of open access, however, seem to be more integrated with field markets. One example is the Chicago market, which is cointegrated with almost all of the field markets in the sample.

Serletis (1997) studies the east-west split issue raised by King and Cuc (1996). Using cointegration methods, he concludes that only 5 out of 28 price pairs are cointegrated. This is sharply contrast to those reports in the previous papers. He also found that gas prices within eastern or western areas are driven by more than one stochastic trend (see discussion of this concept below), and that there is no apparent east-west split in the markets. All in all, Serletis interprets his results to imply that U.S. natural gas markets are not well integrated. This paper casts doubts on the other reports that support stronger market integration since the inception of deregulation.

B. Multivariate Cointegration Analysis

Stock and Watson (1988) provide a very useful way to understand cointegrating relationships. If the price series in a vector $P_t = [\, p_t^1 \ p_t^2 \cdots p_t^N \,]$ are cointegrated, it implies that the stochastic trend in one variable can be expressed as a linear combination of the trends in the other variables. Each cointegrating vector represents such a linear combination. There may be as many as N-1 cointegrating vectors in a system of N variables. The common trends in a system of equations might be thought of as the dimensions in which the I(1) variables in the vector of variables might wander off; The cointegrating vectors are constraints imposed on the dimensionality of this wandering.

Recall the definition of price convergence in section 2: for all N price series to converge, there must be N-1 cointegrating vectors. Walls (1995) tests the number of cointegrating vectors among various collections or 'networks' of prices in the natural gas market. His data come from *Gas Daily* and only cover 1991. The networks he selected include: (i) nodes within the same geographic region, (ii) nodes across different regions but on the same company's pipeline system, and (iii) nodes across different regions and on different pipeline systems. The maximum number of nodes in his networks is 8 due to the computational constraints (matrix singularity). He finds that there are typically 5-7 cointegrating vectors in his groups of eight prices. He concludes that this finding generally supports the hypothesis of convergence. Corroborating earlier research, he finds that city markets are not well integrated with production field markets.

It should be noted that the convergence test is particularly vulnerable to the assumption that the individual prices series are all I(1). It is possible to show that for each price series that is I(0) rather than I(1), one should expect to find an additional cointegrating vector. If all eight prices chosen are I(0), there should be eight cointegrating relationships, yet this would tell us nothing about price convergence!

While cointegration methods have been used to assess market integration, 'cointegration' and market 'integration' are, of course very distinct concepts. The distinction is made clear in Haldane and Hall (1991), Hall, Robertson and Wickens (1992), and Caporale and Pittis (1993). Cointegration is a statistical concept involving long-run relationships among I(1) variables. Its application relies on the implicit assumption that there is a fixed structural relationship between the variables being considered in the long run. Market integration, on the other hand, often refers to a gradual, ongoing process, rather than the final state of complete market unification or one-ness. Detection of cointegration will be most likely if there is a long span of data generated after markets have already reached a particular degree of integration. The absence of cointegration may be interpreted as saying that prices or markets are still in the process of converging. If the process of market integration has been going on for a while following deregulation, say, ultimate convergence may be detectable using the Bernard-Durlauf (1995) convergence tests described in Section 3. Their Monte Carlo simulation analysis suggests that cointegration tests should be reasonably successful at identifying a yet-incomplete move to GDP convergence among nation states.

Another problem with unit root and cointegration tests is that they require long data sets if long-run equilibrium relationships are to be identified with any degree of precision. When using short data series, unit root tests will have low power, implying that investigators will often fail to reject the null hypothesis of a unit root, even if it is in fact false. The same logic would imply that cointegration relationships will be difficult to identify with short data sets, unless individual series are in fact I(0) rather than I(1). In this case, the cointegration tests should find one cointegrating vector for each stationary variable in the vector being considered. Note that the several of the De Vany and Walls papers use only one year of data in their cointegration analyses. Because their datasets are daily frequency, there are enough observations to consider the use of cointegration techniques. The data span, however, is also important, particularly if the speed of adjustment toward the long-run equilibrium is very slow.

In light of the above discussions, consider Werden and Froeb's (1993, 344-5) critical assessment about the use of cointegration analysis for assessing the geographic extent of the market:

> Clearly, cointegration is not a sensible test for relevant market delineation if [price] series do not have unit roots, and they often will not. If two series do have unit roots, cointegration has problems similar to correlation. If common influences such as costs or inflation, drive two prices, they may be cointegrated even though the prices are not closely linked through demand substitutability. Absent such spurious cointegration, cointegration still is not a sufficient condition for the relevant market delineated for one product *or* [geographic] *area* [our emphasis] to include another. A linear combination of two prices can be stationary because of [rather, even though] demand linkages are too weak to place either product in the relevant market delineated for the other. For example,

> arbitrage may place a finite limit on the divergence of prices and, thus, cause them to be cointegrated, but that limit may be far in excess of the significance threshold used for market delineation [for antitrust purposes]. In addition, cointegration is strictly a long-run concept, so the arbitrage process can operate very slowly – too slowly to prevent a price increase for several years."

We agree with Werden and Froeb that cointegration approach requires the individual prices series be I(1) variables. This is an *empirical* question, one that we consider for each price in our daily dataset in Section 5 by carrying out unit root tests.

In contrast, their discussion about the possibility of "spurious cointegration" (a term that we have not seen in the econometrics literature) seems to be in error: "If common influences such as costs or inflation, drive two prices, they may be cointegrated even though the prices are not closely linked through demand substitutability." As we stressed above, if two markets are linked by arbitrage (at least in the long run, if not continuously), then ALL of their I(1) determinants must be common and must affect the two (or all) prices proportionately. This will not arise by coincidence, although all statistical tests, of course, have associated Type I and Type II errors.

We agree with WF's statement that cointegration identifies long-run equilibrium relationships and hence the presence of cointegration by itself may not be a "significance threshold" for market delineation for antitrust purposes. When variables are cointegrated, however, the estimation of an error correction model enables us to get estimates of the speed of adjustment toward equilibrium. With high frequency data (e.g., the daily data used below), this enables us to get estimates of just how tightly markets are linked over various time frames. (See the impulse response analysis in section 6.)

Granger Causality

The term "Granger causality" is somewhat of a misnomer. It refers to the incremental explanatory or predictive power that variable x has on variable y, after lagged values of y have already been taken into account. Suppose, as a first case, that two price series P_1 and P_2 are both I(0) or stationary. A bivariate equation explaining the behavior of P_1 by considering ***n*** lagged values of P_1 and ***m*** lagged values of P_2 would be the following:

$$p_{1t} = \alpha_0 + \sum_{k=1}^{n} \beta_k \, p_{1t-k} + \sum_{l=1}^{m} \gamma_l \, p_{2t-l} \tag{13}$$

If all γ_1 coefficients on lagged P_2's are equal to zero, we say that P_2 does not Granger-cause P_1. By reversing the role of the two prices, we can similarly test the null hypothesis that P_1 does not Granger-cause P_2. There are several possible outcomes: It is possible to have no Granger causality, P_1 may Granger-cause P_2, P_2 may Granger-cause P_1, or both ("bi-directional causality").

Uri, Howell, and Rifkin (1985) and Slade (1986) proposed tests of Granger causality for defining geographic markets. Doane and Spulber (1994) also used Granger causality as a method of measuring U.S. gas market integration. Doane and Spulber found that Granger causality between only one of 20 pairs of gas prices in five regions prior to open access. In contrast, all 20 tests produced evidence of instantaneous bi-directional Granger causality in the post-open access period. As Doane and Spulber noted, their Granger causality tests are based on monthly data, a time interval that may exceed that needed for all arbitrage opportunities to be exploited.

If prices are I(1), Granger causality tests are trickier. If the series are not cointegrated, the above test may be re-specified in first-differences rather than levels. If prices are cointegrated, an ECM like that in (11)-(12) must be estimated in order to avoid omitted variable bias. Price series P_{2t} does not Granger cause price series P_{1t} if all the coefficients $\beta_{12}(L)$ on lagged values ΔP_{2t-i} equal to zero AND if P_{1t} does not respond to deviations from the long run equilibrium, i.e. $\gamma_1 = 0$. (Thus, only P_2 adjusts to insure that the long-run equilibrium condition is satisfied.) Rejection of the null hypothesis that all $\beta_{12}(L) = 0$ and $\gamma_1 = 0$ indicates that P_2 Granger causes P_1. Doane and Spulber estimated equations like (13) in first-differences, and did not specify an ECM although they found that most of their price series are cointegrated after open access.

Apart from these technical specification issues, there is a deeper question of what one hopes to glean from Granger causality tests when attempting to assess the degree of market integration. Unlike simple correlation analysis, Granger causality (like vector autoregressions and cointegration) take the temporal or dynamic interactions among variables into account. This is clearly desirable. If some shocks to the underlying determinants of prices (the X variables in Section 3) originate in one location and others in the second location, one would expect to find bi-directional causality. If most shocks originate in one location and are subsequently propagated to the second location, arbitrage could certainly be occurring – quickly or sluggishly. If it occurs very quickly, Granger causality tests may well find no LAGGED relationship among prices in the two locations. If arbitrage occurs only slowly over time in response to location-specific price shocks, slower dynamic, uni-directional causality may be observed. Nevertheless, dynamic adjustment issues can be examined directly by focusing on deviations from the no-arbitrage-profit equilibrium directly, and asking how long does it take to restore equilibrium following typically shocks (regardless of where they originate)?

Kalman Filter Methods

Time-varying parameters (Kalman filter) analysis is well suited for modeling changing structural relationships over time, as might be the case before and after changes in regulatory policies. This approach has been by Haldane and Hall (1991), Hall, Robertson and Wickens (1992), and Caporale and Pittis (1993) in analyses of European economic integration. In particular, these authors use the Kalman filter to study either the convergence of the EU economies or Sterling's relationship with the dollar and the Deutschemark.

King and Cuc (1996) apply Kalman filter methods to study spot price convergence between natural gas producing basins in both Canada and the U.S. using bid-week price data at a monthly frequency. Their data ends in September 1995, but the starting point of their data series varies: as early as April 1986 and as late as January 1990. The procedure is to run the Kalman filter on natural gas price pairs and plot the resulting value for β as in the regression equation (3). Values of that are closer to one are assumed to indicate more market integration or more rapid price convergence. (One possible drawback of this approach is the difficulty of constructing formal statistical tests of the significance of β.) Their Kalman filter results indicate that price convergence has been occurring in all North American natural gas spot markets. They also found that there is an east-west split in natural gas markets: Western fields' markets seem to be more strongly linked with each other than with eastern ones. They argue this split is a result of lack of sufficient transportation capacity from the western fields to the east.

Critique of the Existing Methods

As Baulch (1997) notes, all of the approaches discussed above share one crucial assumption: there exists a linear relationship between market prices. However, the presence of transaction costs in a spatial network of markets implies a nonlinear relationship between the price pairs. An important shortcoming of all approaches discussed above, therefore, is that they fail to recognize the pivotal role played by transaction costs and the implied non-linearity between price pairs.

5. The Data

Both monthly and daily datasets have been studied in the gas literature. The daily data is typically obtained from *Gas Daily*, an industry periodical. Different sample periods have been considered in De Vany and Walls (1993), De Vany and Walls (1996), Walls (1994a), and Walls (1994b). There are two monthly data sets. One is the *Energy Information Administration* (EIA) monthly data set, which has been studied in De Vany and Walls (1994), Doane and Spulber (1994), and Kleit (1997). The other is from the *Inside FERC's Gas Market Report* published by McGraw Hill. This data set was used in the King and Cuc(1996) study. In addition to these daily and monthly series, the EIA publishes an annual average wellhead price series of marketed production, which covers a much longer time span: 1930-97[19]. This price series, to our knowledge, has not been studied using time series econometric techniques.

The EIA's monthly data set includes wellhead spot price averages for five producing regions: Appalachia, Louisiana, Oklahoma, the Rocky Mountains, and Texas. The EIA constructed these data by taking the arithmetic mean of the spot prices reported in several monthly industry periodicals. A wellhead price is the price the producer receives for gas before any additional charges. The advantage of this data set is that these five series are the only ones available from the beginning of 1984, allowing the early period of the deregulation to be studied. The dataset has at least two drawbacks, however. First, the data are aggregated across different sources, and this adds noise to the data. Second, there are only five series in the data set, which cannot represent all the gas markets throughout the U.S.

Gas Daily currently reports daily spot prices at over 100 locations throughout the natural gas transmission network. The data set, to which we have access, has 154 pricing locations. Among these price series, 57 series range from January 4, 1993 to December 31, 1997, and have 1,247 observations. We will focus on these 57 price series because these series have the longest data range without any missing observations. All the other series have less than 1,247 observations, either because of missing observations, or the starting date is later than January 4, 1993, or the end date is earlier than December 31, 1997.

---

[19] This series was obtained from the EIA web site: www.eia.goe.gov under "Historical Natural Gas Annual-1930 through 1997."

As mentioned above, De Vany and Walls studied different periods of *Gas Daily* spot prices, as early as 1987 and as late as 1991, in a number of papers (see table 1). One might ask why we only study the spot price after 1993, and why don't we include their sample into our study? Unfortunately, we have not been able to obtain their full dataset. However, David Walls has been able to provide us with the subsample 1990-91 theused in Walls (1994b).[20]

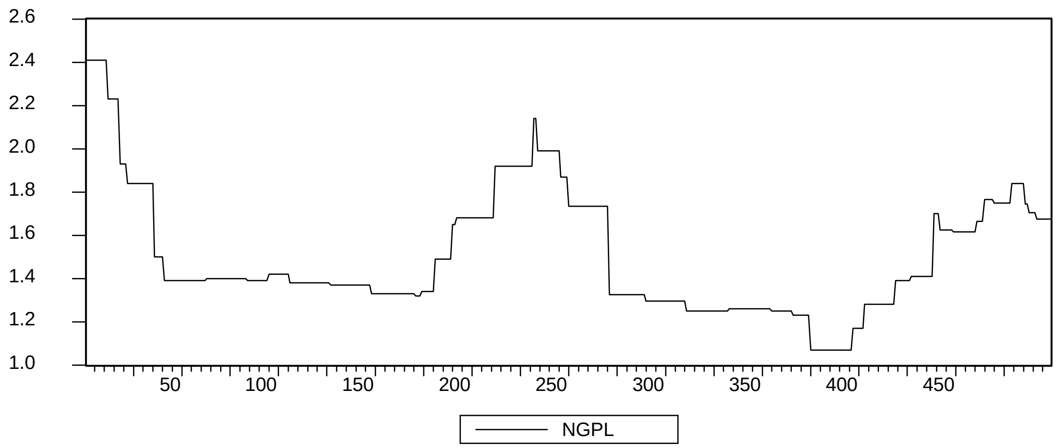
It is also important to note that the data quality in earlier years is questionable.[21] In the early years of the spot markets, gas trading mainly occurred in the bid-week which is the week prior to the pipeline company nomination deadlines. The gas spot markets are not true spot markets since the spot prices are for contracts with a duration of 30 days or less. If we look at a typical price series in Walls (1994b) sample, we see that the price sometimes remains unchanged for 1-3 weeks. Figure 2, for example, shows the price series for NGPL (National Gas Pipeline of America) at East Texas–Houston/Katy area. The lack of fluctuation in the data is probably not the result of the transaction price in an active spot market staying unchanged for that long a period of time. This may be because *Gas Daily* reports the previous trading day's price if there is no trading today. (This is denoted by the absence vs. presence of bold type in the hard copy of the *Gas Daily* publication.) For this reason, many of the daily prices in Walls (1994b) may not be true transaction prices. The statistical properties of such price series may be misleading.

---

[20] We thank David Walls for providing his data.

[21] Bill Meroney at the FERC first cautioned us on this issue. Brinkmann and Rabinovitch (1995) also mention it.

Figure 2: Natural Gas Price for NGPL at West Texas–Houston/Katy from Walls (1994b)1990-1991 daily data

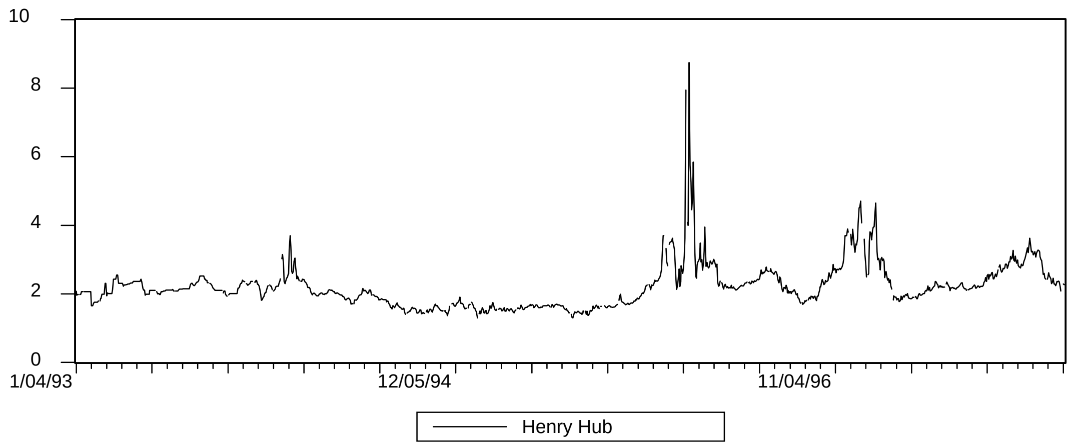


This problem, though much less serious, still plagues us in the observations in 1993. As active trading occurs throughout each month -- a sign of more active spot markets, daily fluctuations in prices become more widespread. This is the norm as of late 1994.

The 57 price series display two different patterns. Forty-nine of these 57 price series increased two or three-fold on February 1, 1996. Most of these 49 price series almost doubled again the next day, February 2, 1996. The cross section mean of these 49 price series jumped from 2.51 ($/mmBtu) on January 31, 1996 to 5.62 on February 1, 1996, then to 9.76 the next day. We will refer to these 49 price series as "pattern 1" series. Most of the 57 price series remain extremely volatile during that whole February.

Eight of the 57 price series do not experience a huge price jump in February 1996. On the contrary, they remain relatively tranquil during that period. We will refer to these 8 series as "pattern 2" series. Figure 3 shows the spot price series, in levels, at Henry Hub and El Paso SJ. It is obvious to see the huge price spikes for the Henry Hub price series in the February of 1996 (For scaling problems, we have eliminated the biggest spike from the graph, which is 14), and no price spikes for the El Paso SJ price series during the same period.

Figure 3: Price Series (in levels) at Henry Hub and El Paso SJ



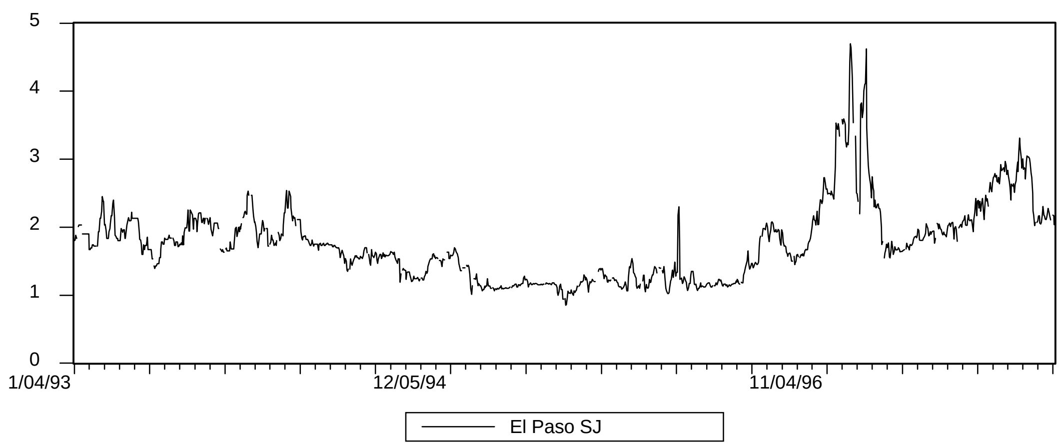
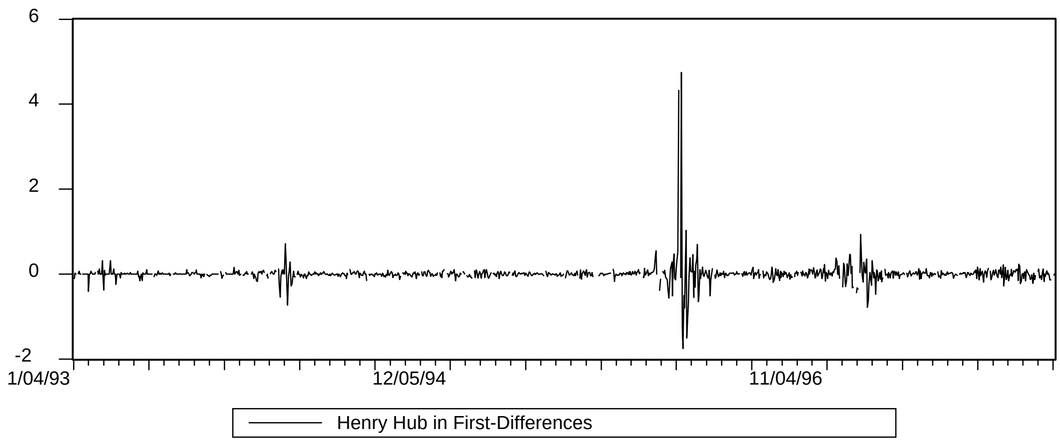
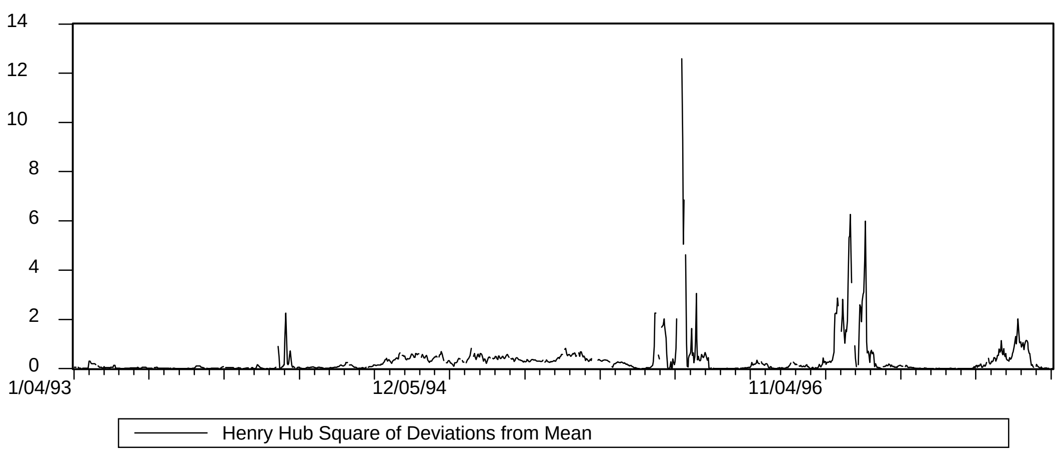


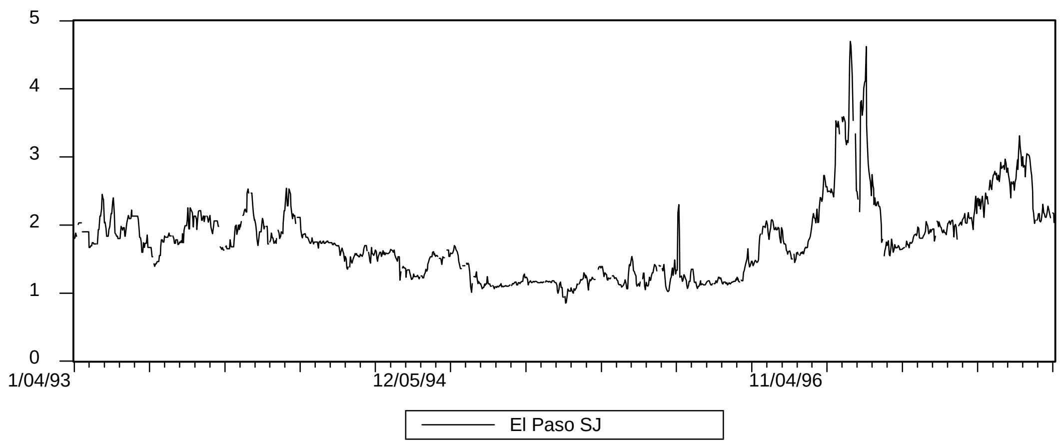
Figure 4: Henry Hub price series in first-differences, and square of deviations from mean





Notice that this classification of price series patterns has great implications for the discussions (section 6) of the gas market integration and segmentation. The eight pattern 2 series are either in the area of "Rockies" (as classified in the Gas Daily groupings), or are intrastate pipeline pricing nodes. This difference of price patterns is stark visual evidence suggesting some form of market segmentation. Details and more formal statistical analyses are to be explored in section 6.

All of our daily price series exhibit huge volatility changes over time. As an illustration, consider Figure 4, which shows the Henry Hub price series in first-differences and square of deviations from mean. Note that, for graphical presentation, we removed a couple of huge outliers to get rid of the scaling problems with the graphs.



Unit Root Tests on Price Series[22]

Section 2 above stressed that it is critical to determine whether price series are mean (and variance) stationary in order to choose appropriate empirical methods for

[22] The two unit tests conducted in this paper both assume the null of a unit root. We also considered the KPSS unit root test (Kwiatkowski, Phillips, Schmidt and Shin (1992)), which differs from the other two unit root tests in that the null hypothesis assumes the series are stationary rather than nonstationary. The KPSS results indicate that the null hypothesis that the prices are I(0) is overwhelmingly rejected, too! This evidence, along with the fact that autocorrelation function decays at a much slower rate than the exponential rate, suggests that spot gas price series may be better modeled as fractionally integrated processes, i.e., integrated with order d, where $0<d<1$. We are currently exploring this possibility.

assessing the degree of gas market integration. Most of the empirical papers summarized in Table 1 test each of the series in their respective datasets for the presence of unit roots. Overwhelmingly, they fail to reject the null hypothesis of a unit root and therefore conclude that gas prices are nonstationary. For example, the null of a unit root cannot be rejected for any of the daily price series tested in De Vany and Walls (1993), and Walls (1994a, 1994b).

Are the spot natural gas price series really nonstationary? We will investigate this issue in great detail below. To anticipate the results, the evidence leads to a sharply contrasting conclusion from those in the literature: The null of a unit root for all 57 of the daily price series considered is overwhelmingly rejected. The conclusions are based on two commonly used unit root test methods, the augmented Dickey-Fuller test (ADF) and the Phillips-Perron test (PP). Both tests have been used in the gas literature. This study differs from the earlier literature in a number of ways. First, we argue that the PP test is the more appropriate test method to use for our sample, though we will do both. Second, we carefully choose whether to include constant term or time trend in the regression. Third, we recognize that where it is appropriate to use the ADF test, one must choose the lag length carefully to eliminate serial correlations in the residuals. In addition, we re-examine the sample in Walls (1994b).

The Dickey-Fuller test assumes that the error terms have a constant variance and are statistically independent. The augmented Dickey-Fuller approach controls for higher-order correlation by adding lagged difference terms of the dependent variable to the right-hand side of the regression equation, but it does not control the heteroskedasticity problem in the error process. Figure 3 and 4 demonstrate that gas price series exhibit huge volatility changes over time, and it is obviously inappropriate to assume the variance is constant over that period. The Phillip-Perron test allows the error term to be serially correlated and heteroskedastic. Therefore, more emphasis should be put on the test results from the PP test.

Consider the results of the PP test first. The null of a unit root is rejected at 1% level for all the 57 price series. For those 49 price series which are extremely volatile during February of 1993, the smallest τ value is -4.8 (the corresponding 1% critical value is -3.43), and the biggest is -14.5, and the mean of the 49 τ values are a resounding -8.7! For the 8 price series which are relatively tranquil during February of 1993, the τ values are much smaller than those for the above 49 series, but are still big enough to reject the null of a unit root. See Appendices II and III for the details of the tests. Note that in the appended tables, we report three cases for the tests with each price series: (1) with a constant term in the regression; (2) with both a constant and time trend in the regression; (3) without any of the two terms in the regression. All of the constant terms in the PP tests are statistically significant, so case (3) is inappropriate for any of the price series. If the time trend is significant, then case (2) is the proper one to look at. If the time trend is insignificant, then case (1) is the appropriate test.

We have argued that the ADF test is not the appropriate test for our sample, nonetheless we report the ADF results for the sake of comparison. The null of a unit root is still rejected for 40 of the 49 pattern 1 series at the 5% significance level. Seven out of the 8 pattern 2 series reject the null of a unit root at 5% level. Note that case (1) is always the proper one since the constant term is always significant for all the price series, but the time trend term is never significant. Also note the number of lags included in the regressions is very long. To illustrate the problem of applying the ADF test for our sample, we present the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)(see table 2) for the residuals and squared residuals from the following regression for the Henry Hub price series:

$$\Delta x_t = C + \alpha x_{t-1} + \beta_i \sum_{i=1}^{18} \Delta x_{t-i} + \varepsilon_i$$

Where the number of lags is chosen by the "general to specific" rule. Such a long number of lags(18) is sure enough to eliminate the serial correlations in the residuals (in absolute levels). However, there are still serial correlations in the squared residuals, which is a sign of ARCH (Autoregressive conditional heteroskedastic) effect.

Table 2: ACF and PACF of the residuals

	AC	PAC	Q-Stat	Prob
1	-0.001	-0.001	0.0018	0.966
2	-0.003	-0.003	0.0128	0.994
3	0.003	0.003	0.0208	0.999
4	0.004	0.004	0.0398	1.000
5	0.007	0.007	0.1010	1.000
6	0.004	0.004	0.1241	1.000
7	0.000	0.000	0.1244	1.000
8	-0.007	-0.007	0.1905	1.000
9	-0.013	-0.013	0.4061	1.000
10	-0.009	-0.010	0.5159	1.000
11	-0.004	-0.004	0.5343	1.000
12	0.005	0.005	0.5605	1.000
13	0.000	0.000	0.5605	1.000
14	0.001	0.002	0.5624	1.000
15	0.003	0.003	0.5743	1.000
16	0.010	0.010	0.7084	1.000
17	0.013	0.013	0.9124	1.000
18	0.033	0.033	2.2622	1.000

ACF and PACF of squared residuals

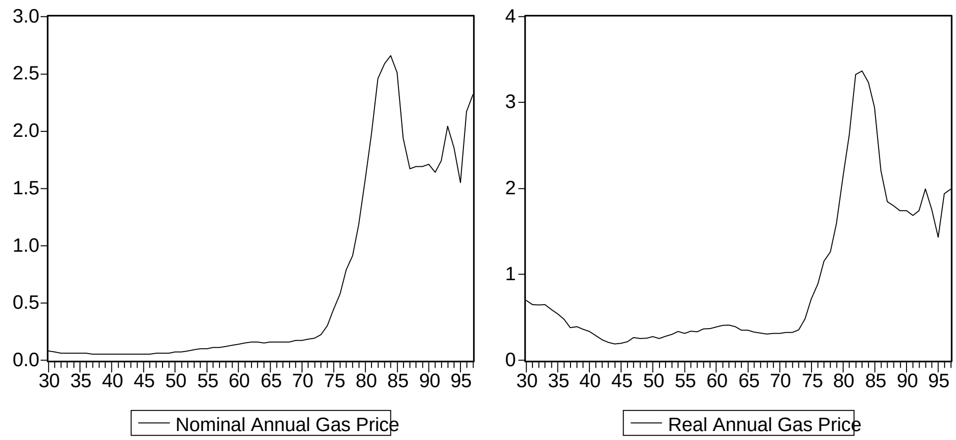
lags	AC	PAC	Q-Stat	Prob
1	0.648	0.648	516.93	0.000
2	0.159	-0.449	548.16	0.000
3	0.008	0.324	548.24	0.000
4	0.000	-0.238	548.24	0.000
5	-0.001	0.175	548.24	0.000
6	0.000	-0.127	548.24	0.000
7	0.000	0.090	548.24	0.000
8	-0.001	-0.064	548.25	0.000
9	0.001	0.048	548.25	0.000
10	0.002	-0.034	548.25	0.000
11	0.001	0.022	548.25	0.000
12	0.003	-0.006	548.26	0.000
13	0.008	0.011	548.35	0.000
14	0.009	-0.006	548.45	0.000
15	0.003	-0.001	548.46	0.000
16	-0.002	0.000	548.47	0.000
17	-0.003	-0.003	548.47	0.000
18	-0.001	0.006	548.48	0.000

The Annual Price Series

We recognize that our sample of daily gas data is a short one (only five-years' data), although we have over 1000 daily observations. Even though we find that the null of a unit root is rejected overwhelmingly for all the price series in our study, we are reluctant to claim that natural gas price, especially in nominal terms, is stationary over time. To be able to answer the question of whether natural gas price series is stationary or not over time more confidently, we turn to the annual average wellhead price of marketed production from 1930 to 1997 as mentioned in section 5. By studying this low-frequency, but long span price series, we can say more about the true underlying gas data-generating process. Figure 5 shows the nominal and real annual gas time series. The real gas series are obtained by dividing the nominal series by the U.S. implicit GDP deflator (1992 = 100).

A striking feature of the annual price series as shown in Figure 5 is that there seems to be a structural break sometime in the 1975-1980 period. If there are structural breaks in a time series, the ADF test and the PP test statistics are biased toward the nonrejection of a unit root (see Perron (1989)). Perron(1990) proposed two different

Figure 5: The nominal and real annual gas prices



procedures to test for a unit root in a time series with a changing mean. Perron writes the null hypothesis and the alternative hypothesis in the form of

$$H_0: x_t = x_{t-1} + A(L)^{-1} B(L)[\delta_1 D_p + \varepsilon_t]$$
$$H_1: x_t = \alpha_0 + A(L)^{-1} B(L)[\delta_2 D_L + \varepsilon_t]$$

where D_p is a pulse dummy variable such that $D_p = 1 if\ t = \tau + 1$ and zero otherwise,

and D_L is a level dummy variable such that D_L =1 if t > τ and zero otherwise. The error process (u_t) is assumed to be an ARMA(p, q) process of the form $A(L)u_t = B(L)\varepsilon_t$ with ε_t ~ iid. Note that the specification in the null and alternative hypothesis imply that the mean change of the series is not instantaneous but depends on the specification of the error process. The first test procedure involves obtaining the residuals $\tilde{x}_t$ from a regression of x_t on a constant and D_L , and then running the following regression:

$$\Delta\tilde{x}_t = \alpha\tilde{x}_{t-1} + \sum_{i=1}^{k} \beta_i \Delta\tilde{x}_{t-i} + \varepsilon_i$$

The second test procedure is to run the following regression in one step:

$$\Delta x_t = \alpha_0 + \alpha x_{t-1} + \delta_1 D_L + \delta_2 D_p + \sum_{i=1}^{k} \beta_i \Delta x_{t-i} + \varepsilon_i$$

where k is the number of lags of x's in first-differences to render the error terms serial uncorrelated. The test is whether α = 0. The test results by procedure 1 and 2 for the real and nominal price series with τ = 1974 are as follows:

	The real series	The nominal series
Procedure 1:		
$\hat{\alpha}$ (t-stat)	-1.19 (-4.25)	-1.98 (-5.66)
# of lags k	9	9
Procedure 2:		
$\hat{\alpha}$ (t-stat)	- 1.21 (-4.16)	-1.98 (-5.66)
# of lags k	9	9

The 1% critical value for the test is - 4.14 (from Table 4 in Perron(1990)), so the null of a unit root is rejected for all the cases. The empirical tests show that the annual gas price series do have a structural break around 1975-1980.[23] Indeed, there was a supply shortage of natural gas in the 1970s (see MacAvoy and Pindyck (1975)). This break may also have something to do with the oil crisis?

Re-examination of Walls(1994b) Sample

We have 13 of the 26 price series studied in Walls(1994b). The data cover the interval from January 1, 1990 to December 31, 1991. We are not sure why Walls chose only the time interval from July 1990 to June 1991 in his study, but if we use either the ADF test or the PP test on this somewhat arbitrary interval, we will not reject the null of a unit root for any of the 13 price series we have. This conforms to the results in Walls(1994b). However, if we apply these two tests to the entire two year samples, then the test results turn out to be quite different from those in Walls (1994b) (see appendix IV for details). Bear in mind the quality problems with the data as emphasized in section 5.

The results for the ADF tests can be summarized as follows. One of the 14 series rejects the null of a unit root at the 1% level. Eight of the 14 series reject the null of a unit root at the 5% significance level. Note that using the "general to specific" lag choosing rules to render the error process serial uncorrelated, we find a lag number

[23] An interesting extension would be to employ the Zivot-Andrews (1992) method for searching the entire sample to find the timing of the break.

around 20 (20, 21, or 22) for all the 13 price series. The constant terms are significant for 12 series, and the time trend is not significant for any of the price series.

The PP test results are quite different from the ADF test results. Two of the price series reject the null of a unit root at 1% level (El Paso and Transwestern). These two series do not reject the null at 5% level for the ADF test. Only three more price series reject the null at 5% level.

6. Cointegration and Price Gaps Analysis

Cointegration Analysis

Cointegration analysis is at the heart of De Vany and Walls' various papers on natural gas markets. This approach presupposes that individual series are all I(1), a hypothesis that the author's unit root tests failed to reject.

Our robust finding that spot gas price series over the period of 1993 through 1997 is stationary casts serious doubt on the appropriateness of cointegration analysis for our longer dataset. Suppose one was to proceed with cointegration analysis even though all variables in the analysis were known to be I(0) rather than I(1) processes. What should we expect to find? In theory, one should find as many cointegrating vectors as the number of I(0) variables being considered. In a bivariate system, there should, in theory, be two cointegrating vectors if both variables are indeed stationary.[24]

---

[24] Price series $P_t = (P_t^1, P_t^2, \ldots, P_t^N)$ is said to exhibit H cointegration relationships if there exist H linear independent vectors $(r_1, r_2, \ldots, r_H)$ such that $r_i' P_t$ is stationary. Any linear combination of $(r_1, r_2, \ldots, r_H)$ would also be cointegrating vectors. If there are N such linear independent vectors,

Although our Dicky-Fuller and Phillips-Perron unit root tests overwhelmingly reject the null of unit root for any price series, we, nonetheless, proceed with the cointegration analysis and expect to see two cointegrating vectors for a bivariate system. In general, the results of Johansen cointegration tests can be quite sensitive to the lag length chosen in the underlying dynamic model being estimated. A procedure for choosing lag length, recommended by Sims and others, is to estimate a VAR model in levels of the variables. The appropriate lag length is then chosen using the Schwartz Bayesian Criterion; i.e. choice the lag length that minimizes that Schwartz criterion.[11] Next, the ECM specification is estimated using a lag length that is one less than the lag length selected for the VAR levels. To see that this makes sense, recall that the ECM is just a restricted form of the unrestricted VAR in levels. Suppose the VAR in levels is estimated with two lags based on the Schwartz criterion:

$$Y_t = \Pi_0 + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \eta_t$$

It is easy to show that this specification can be rewritten as:

$$\Delta Y_t = \Pi_0 + (\Pi_1 + \Pi_2 - I) Y_{t-1} - \Pi_2 \Delta Y_{t-1} + \eta_t$$

Note that the latter expression is a ECM formulation with one less lag than the original specification in levels. The Johansen test involves assessing the rank of the coefficient matrix in front of the Y_{t-1} term..

---

then it is possible to find linear combinations of these N independent vectors that will yield N canonical vectors, and those linear combinations are given by the inverse of matrix $(r_1, r_{2,}, \dots, r_H)$. Since these N canonical vectors must also be cointegrating vectors, each price series must be stationary.

Since there are 57 price series in our dataset, there are 57*56/2=1596 different price pairs. For each of these pairs, we estimated a bivariate VAR system in levels, selected the lag length based on the Schwartz criterion, and estimated the appropriate ECM to carry out the Johansen cointegration rank test for determining the number of cointegrating vectors. Of these 1596 bivariate systems, Johansen tests indicate the presence of two cointegrating vectors for 986 pairs, one cointegrating vector for 573 pairs, and zero cointegrating vector for the remaining 37 pairs. (See appendix V for details.) The finding that there are two cointegrating vectors for the vast majority of the price pairs suggests that the price series are in fact stationary, i.e. I(0) rather than I(1) variables.[25] This, of course, is consistent with our unit roots tests, which overwhelmingly rejected the unit root hypothesis.

As an illustration, Table 3 shows the cointegration test results for the Henry Hub price series and the NGPL (La.) price series, a price pair where one rather than two cointegrating vectors was found. The Table shows the rank (= number of cointegrating vectors) chosen by the Johansen test based on ECMs with different lag lengths. It can be seen that the cointegration test results are quite sensitive to the number of lags included in

---

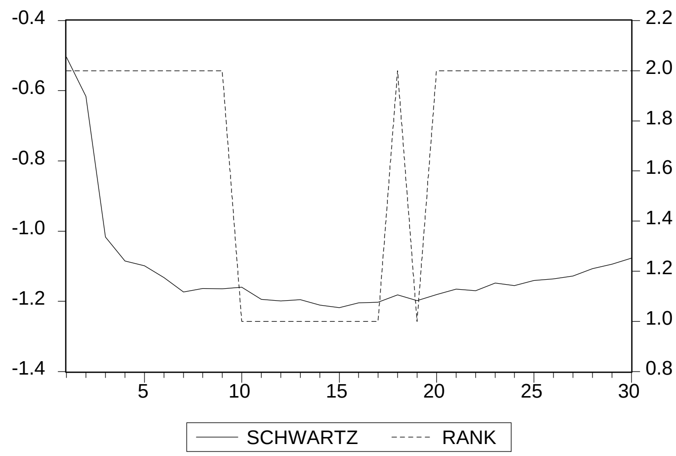
[25]The Johansen results should be interpreted carefully. The Johansen cointegration test assumes Gaussian error terms, but the daily spot gas price series may not be generated by Gaussian stochastic process. First, daily spot gas price series exhibit significant heteroskedasticity as shown in previous sections. Second, it also seems that daily spot gas price series are non-normal. There is substantial positive skewness in these series. For example, in Figure 3, the positive skewness is apparent for the Henry Hub and the El Paso SJ price series, because there is no downward spikes to match the pronounced upward spikes. Daily spot gas price series also display substantial kurtosis, with tails much thicker than those of the normal distribution. The fact that some commodities prices display skewness and kurtosis have been documented by such authors as Deaton and Laroque (1992), Sephton (1992), and Baillie and Myers (1991). Deaton and Laroque (1992) study annual commodity prices while the other two papers study daily commodity prices. Note that daily gas price series exhibit much more skewness and kurtosis than those documented for the other commodity price series in the above papers. Sephton (1992) and Baillie and Myers (1991) also try to model daily commodity price series by various GARCH models.

the ECM. Our procedure chooses a lag length of 15 for the VAR in levels for this price pair, as this lag is associated with the lowest Schwartz value (-1.218006). Since this Schwartz value is from the level VAR estimation, we select 14 as the lag length when estimating an ECM to carry out the cointegration test. The sensitivity of our conclusion that there is a single cointegrating vector to the choice of lag length is shown in Figure 6.

Table 3:
Sensitivity of Cointegration Test Results to Lag Length in the underlying Error Correction Model

Lag #	Schwartz	Rank
1	-0.5033899	2
2	-0.6155192	2
3	-1.0168012	2
4	-1.0851563	2
5	-1.0983091	2
6	-1.1321956	2
7	-1.1733211	2
8	-1.1636095	2
9	-1.1646622	2
10	-1.1596316	1
11	-1.1943512	1
12	-1.1985640	1
13	-1.1952561	1
14	-1.2113701	1
15	-1.2180065	1
16	-1.2042212	1
17	-1.2028633	1
18	-1.1815938	2
19	-1.1984508	1
20	-1.1810913	2
21	-1.1651586	2
22	-1.1702714	2
23	-1.1485849	2
24	-1.1549112	2
25	-1.1408798	2
26	-1.1364335	2
27	-1.1280319	2
28	-1.1068293	2
29	-1.0941630	2
30	-1.0767655	2

Figure 6: Sensitivity of Cointegration Test Results on Rank to Lag Length



Error Correction Specification

As mentioned in previous sections, if prices are I(1), then testing for presence of cointegration is a natural way to proceed. If two prices are cointegrated, then estimating an error correction model (ECM) will provide an estimate of the long run equilibrium and the speed of adjustment toward the long-run equilibrium. Since the unit root test results show that spot natural gas price series are stationary, it makes the application of ECM unattractive here. Nonetheless, the Johansen test shows that there are some pairs of price series have one cointegrating vector. For the sake of illustration, an error correction model is estimated and presented below. The Henry Hub price series and the PEPL price series at Oklahoma are used as the two representative price series.

The estimates of the ECM by Eviews are shown in Table 4. The estimated cointegrating vector is (1, -0.95), and the null of (1, -1) cannot be rejected. The best way to summarize the information regarding adjustment dynamics contained in the ECM is to analyze the impulse response function. Impulse response analysis is used to analyze how a shock to the error term in one of the two equations of the ECM will affect each of the two prices over time. One has to make an assumption about the ordering of the two price series in order to recover the structural shocks. If one chooses the ordering (Henry Hub, PEPL), it means that the price shocks (innovations) originating in Henry Hub are allowed to have an impact on PEPL prices within the same period (i.e. day), while innovations to the PEPL series do not affect Henry Hub until the following day. If, in contrast, one chooses the ordering (PEPL, Henry Hub), price innovations to PEPL will impact Henry Hub prices during the same period, but Henry Hub price innovations do not affect PEPL prices until the following day. Unfortunately, there is no presumption one way or the other in terms of which identifying assumption to choose for the impulse response analysis. If the correlation between innovations in the two prices series is low (say less than 0.2), the ordering will not matter much. When the correlation is high, on the other hand, the impulse response functions can look quite different depending on which ordering is chosen.

Figure 7 shows the impulse response function. Note that the shape of the impulse response function depends on the ordering of the two price series.

Table 4:
Error Correction Model Estimation
for Henry Hub and Pepl Prices

Sample(adjusted): 15 1247
Included observations: 1233 after adjusting endpoints
Standard errors & t-statistics in parentheses

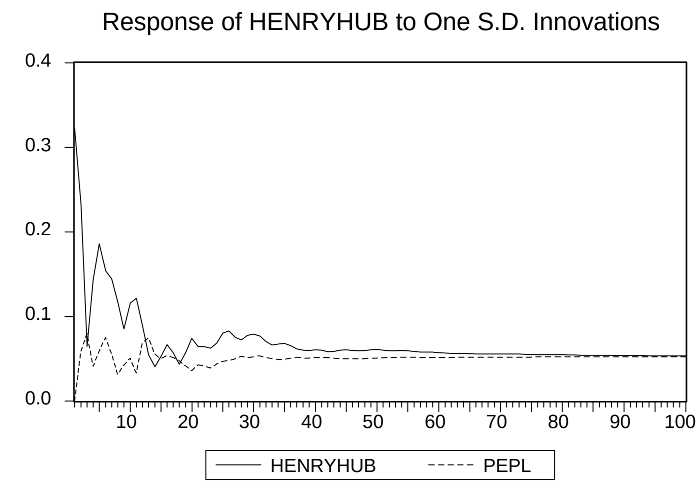
Cointegrating Eq:	CointEq1	
HENRYHUB(-1)	1.000000	
PEPL(-1)	-0.950880	
	(0.10673)	
	(-8.90880)	
C	-0.335636	
	(0.21702)	
	(-1.54659)	
Error Correction:	**D(HENRYHUB)**	**D(PEPL)**
CointEq1	-0.119769	-0.054877
	(0.03203)	(0.02556)
	(-3.73882)	(-2.14716)
D(HENRYHUB(-1))	-0.623919	-0.574156
	(0.08165)	(0.06514)
	(-7.64179)	(-8.81410)
D(HENRYHUB(-2))	-0.734455	-0.340263
	(0.08396)	(0.06699)
	(-8.74773)	(-5.07957)
D(HENRYHUB(-3))	-0.084793	-0.267305
	(0.08683)	(0.06927)
	(-0.97658)	(-3.85866)
D(HENRYHUB(-4))	-0.196810	-0.178731
	(0.08776)	(0.07002)
	(-2.24254)	(-2.55256)
D(HENRYHUB(-5))	0.163688	0.121255
	(0.08301)	(0.06623)
	(1.97189)	(1.83082)
D(HENRYHUB(-6))	0.050791	0.079879
	(0.08250)	(0.06582)
	(0.61565)	(1.21355)
D(HENRYHUB(-7))	0.154460	0.150312
	(0.08231)	(0.06567)
	(1.87650)	(2.28880)
D(HENRYHUB(-8))	-0.161369	-0.051138
	(0.08073)	(0.06441)
	(-1.99881)	(-0.79392)
D(HENRYHUB(-9))	-0.102921	-0.134199
	(0.07816)	(0.06236)
	(-1.31688)	(-2.15215)
D(HENRYHUB(-10))	-0.033942	0.025877
	(0.07763)	(0.06194)
	(-0.43721)	(0.41778)

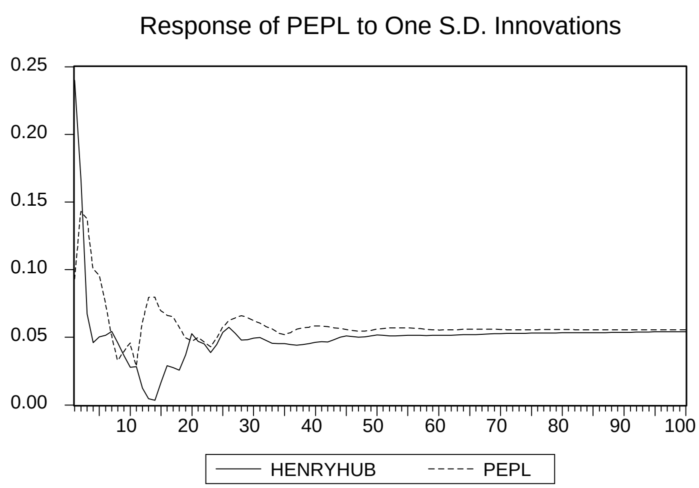
D(HENRYHUB(-11))	-0.355517	-0.404080
	(0.07050)	(0.05625)
	(-5.04281)	(-7.18392)
D(HENRYHUB(-12))	-0.101974	-0.154833
	(0.05850)	(0.04668)
	(-1.74305)	(-3.31717)
D(HENRYHUB(-13))	-0.189150	-0.215878
	(0.05877)	(0.04689)
	(-3.21866)	(-4.60427)
D(PEPL(-1))	0.510303	0.484371
	(0.10071)	(0.08035)
	(5.06716)	(6.02833)
D(PEPL(-2))	0.249434	-0.001494
	(0.10171)	(0.08115)
	(2.45250)	(-0.01841)
D(PEPL(-3))	0.016396	-0.059238
	(0.10105)	(0.08063)
	(0.16225)	(-0.73473)
D(PEPL(-4))	0.302756	0.153945
	(0.10022)	(0.07996)
	(3.02081)	(1.92521)
D(PEPL(-5))	0.047819	-0.153492
	(0.09672)	(0.07717)
	(0.49440)	(-1.98909)
D(PEPL(-6))	0.076910	-0.070072
	(0.09854)	(0.07862)
	(0.78050)	(-0.89128)
D(PEPL(-7))	-0.122207	-0.139596
	(0.09899)	(0.07898)
	(-1.23451)	(-1.76747)
D(PEPL(-8))	0.090422	-0.034604
	(0.09571)	(0.07636)
	(0.94472)	(-0.45315)
D(PEPL(-9))	0.108589	-0.008064
	(0.09113)	(0.07271)
	(1.19152)	(-0.11091)
D(PEPL(-10))	-0.027174	-0.168665
	(0.09033)	(0.07207)
	(-0.30084)	(-2.34040)
D(PEPL(-11))	0.484565	0.369609
	(0.07823)	(0.06241)
	(6.19417)	(5.92182)
D(PEPL(-12))	0.089357	0.167539
	(0.07496)	(0.05980)
	(1.19213)	(2.80151)
D(PEPL(-13))	0.085073	0.100567
	(0.07549)	(0.06023)
	(1.12700)	(1.66983)

R-squared	0.370883	0.271572
Adj. R-squared	0.357320	0.255868
Sum sq. resids	128.1862	81.59773
S.E. equation	0.326022	0.260115
F-statistic	27.34510	17.29306
Log likelihood	-353.9667	-75.50459
Akaike AIC	0.617951	0.166269
Schwarz SC	0.730006	0.278324
Mean dependent	0.000170	0.000195
S.D. dependent	0.406677	0.301537
Determinant Residual Covariance		0.000900
Log Likelihood		824.2588
Akaike Information Criteria		-1.244540
Schwarz Criteria		-1.007978

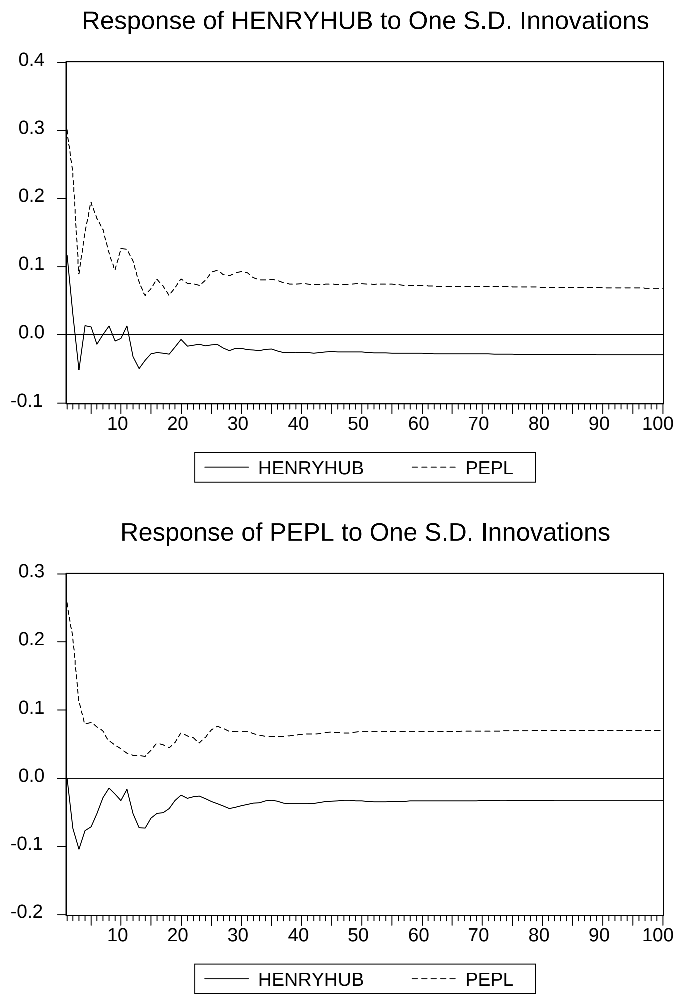
Figure 7:
Impulse Response Function

Ordering: Henry Hub, Pepl





Ordering: Pepl, Henry Hub



The Price Gap Specification

When two price series are cointegrated with a cointegrating vector [1,-1], it is possible to specify the dynamics of the price gap of these two price series by a univariate

error correction specification as in (9). This equation can be estimated directly by OLS, the same procedure used in the augmented Dickey-Fuller unit root test. This method is attractive when the two underlying price series are I(1). When the prices series are stationary, as is the case here, the price gap method is problematic. Recall Werden and Froeb's (1993) critique of Horowitz's (1981) price gap method (discussed above).
The price gap between Henry Hub and PEPL price series is taken as an example. The difference between the natural gas price at Henry Hub and at Pepl is denoted as H_P in the Eviews output in Table 5, which reports the following OLS regression:

$$H\_P = \alpha_1 H\_P(-1) + \ldots + \alpha_p H\_P(-p) + \varepsilon_t$$

where $\varepsilon_t \sim N(0, \sigma^2)$. The number of lags, p, is determined by the general-to-specific method, turned out to be 22 days for this price pair. The estimated impulse response function for this gap is in Figure 8. The half life for the one standard deviation shock, which is 0.37, is about two days. Only one-forth of the shock remains after 7 days. These calculations implicitly assume that transactions costs are zero. To the extent that they are positive, the half life of arbitrage profits will be shorter.

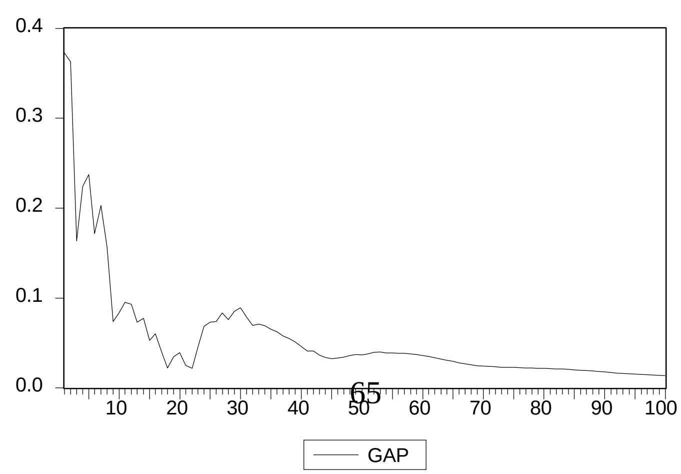
Table 5 , OLS Estimation Results for H_P

Dependent Variable: H_P
Method: Least Squares
Sample(adjusted): 23 1247
Included observations: 1225 after adjusting endpoints

Variable	Coef.	Std Error	t-Statistic	Prob.
C	0.015	0.006	2.271	0.023

H_P(-1)	0.974	0.029	33.852	0.000
H_P(-2)	-0.510	0.040	-12.689	0.000
H_P(-3)	0.672	0.043	15.702	0.000
H_P(-4)	-0.381	0.047	-8.112	0.000
H_P(-5)	0.223	0.048	4.640	0.000
H_P(-6)	-0.035	0.048	-0.717	0.473
H_P(-7)	-0.137	0.048	-2.874	0.004
H_P(-8)	0.012	0.048	0.259	0.796
H_P(-9)	-0.016	0.047	-0.348	0.728
H_P(-10)	0.075	0.047	1.577	0.115
H_P(-11)	0.052	0.047	1.092	0.275
H_P(-12)	-0.085	0.047	-1.794	0.073
H_P(-13)	0.122	0.047	2.578	0.010
H_P(-14)	-0.175	0.047	-3.696	0.000
H_P(-15)	0.175	0.048	3.670	0.000
H_P(-16)	-0.232	0.048	-4.862	0.000
H_P(-17)	0.180	0.048	3.725	0.000
H_P(-18)	-0.084	0.048	-1.746	0.081
H_P(-19)	0.094	0.047	2.002	0.046
H_P(-20)	-0.054	0.043	-1.251	0.211
H_P(-21)	0.004	0.040	0.097	0.922
H_P(-22)	0.067	0.029	2.313	0.021
R-squared	0.799		Mean dependent var	0.24
Adjusted R-squared	0.796		S.D. dependent var	0.37
S.E. of regression	0.169		Akaike info criterion	-0.69
Sum squared resid	34.486		Schwarz criterion	-0.60
Log likelihood	448.515		F-statistic	217.69
Durbin-Watson stat	1.993		Prob(F-statistic)	0.00

Figure 8:
Impulse Response Function for the Price Gap: OLS Estimation



The simple OLS estimation ignores the fact that spot gas price series exhibit significant heteroskedasticity. To account for heteroskedasticity, a GARCH(1,1) model, as developed in Bollerslev (1986), is estimated using maximum likelihood method. In a GARCH(1,1) model, the residual ε_t is modeled as,

$$\varepsilon_t \,/\, I_{t-1} \sim N(0, h_t)$$
$$h_t = \delta + \alpha \varepsilon^2{}_{t-1} + \beta h_{t-1}$$

The variance today depends on yesterday's news about volatility and past forecast variances. Diagnostic checks confirm that there are no further ARCH effects in the residuals after estimating the GARCH(1,1) model. The estimated results are shown in table 6.

Table 6: Maximum Likelihood Estimation for the Price Gap H_P

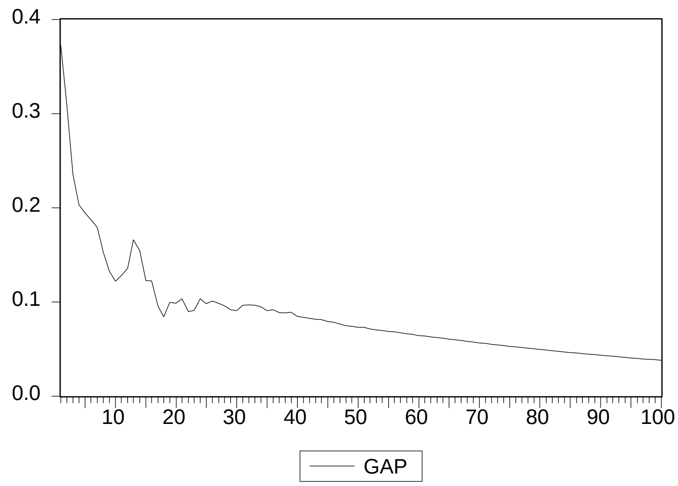
Dependent Variable: H_P
Method: ML - ARCH
Sample(adjusted): 23 1247
Included observations: 1225 after adjusting endpoints
Convergence achieved after 69 iterations

	Coef.	Std. Error	z-Statistic	Prob.
C	0.012	0.002	5.380	0.000
H_P(-1)	0.810	0.034	23.546	0.000
H_P(-2)	-0.053	0.051	-1.040	0.298
H_P(-3)	0.076	0.052	1.468	0.142
H_P(-4)	0.031	0.044	0.712	0.476
H_P(-5)	0.017	0.045	0.377	0.706
H_P(-6)	0.010	0.048	0.205	0.838
H_P(-7)	-0.045	0.041	-1.085	0.278
H_P(-8)	-0.003	0.038	-0.085	0.932
H_P(-9)	0.013	0.041	0.311	0.756
H_P(-10)	0.045	0.037	1.197	0.232

H_P(-11)	0.030	0.032	0.933	0.351
H_P(-12)	0.074	0.032	2.328	0.020
H_P(-13)	-0.069	0.034	-2.058	0.040
H_P(-14)	-0.058	0.036	-1.608	0.108
H_P(-15)	0.045	0.031	1.451	0.147
H_P(-16)	-0.084	0.026	-3.261	0.001
H_P(-17)	0.031	0.035	0.887	0.375
H_P(-18)	0.053	0.034	1.579	0.114
H_P(-19)	-0.004	0.033	-0.127	0.899
H_P(-20)	0.029	0.027	1.052	0.293
H_P(-21)	-0.056	0.024	-2.346	0.019
H_P(-22)	0.037	0.018	2.129	0.033
		Variance Equation		
C	0.000	0.000	7.580	0.000
ARCH(1)	0.365	0.026	14.029	0.000
GARCH(1)	0.700	0.014	51.527	0.000
R-squared	0.749	Mean dependent var		0.242
Adjusted R-squared	0.744	S.D. dependent var		0.375
S.E. of regression	0.190	Akaike info criterion		-2.981
Sum squared resid	43.070	Schwarz criterion		-2.873
Log likelihood	1851.985	F-statistic		143.445
Durbin-Watson stat	1.912	Prob(F-statistic)		0.000

The estimated impulse response function is as follows. The half life of a one standard deviation shock increases from 2 days to 6 days after considering the ARCH effect.

Figure 9:
Impulse Responses Function
based on estimated Model with ARCH Effects



7. Conclusion

This paper reviews the literature on price tests of the geographic extent of the market, and carries out an extensive battery of tests using a daily dataset covering 5 years in the 1990s. Based on the analysis, we conclude that the estimation of error correction models is a very natural way to assess the tightness of linkages across segments of a market provided that the prices under investigation are all I(1) processes. The speed of adjustment towards a zero-profit no-arbitrage equilibrium is easily estimated in this framework. The unresolved task is how to deal with situations where rices are I(0). In this case, all linear combinations of two prices will be stationary, so cointegration tests and ECMs provide little obvious information about the extent of market integration. Spurious estimates of the speed of adjustment are indeed likely, as Werden-Froeb (1993) have emphasized. It is unclear how to resolve this difficulty; hence it remains an important task for future research.

References

Bernard, Andrew B., and Steven Durlauf (1995). 1995. "Convergence in International Output." Journal of Applied Econometrics 10; 97-108.

Brinkman, Emile J., and Ramon Rabinovitch. 1995. "Regional Limitations on the Hedging Effectiveness of Natural Gas Futures." The Energy Journal 16(3):113-124.

Baulch, Bob. 1997. "Testing for Food Market Integration Revisited." Journal of Development Studies 33(4):512-534.

Caporale, Guglielmo M., and Nikitas Pittis. 1993. "Common Stochastic Trends and Inflation Convergence in the EMS." Welwirtschafiliches Archiv 129 (2): 208-215.

Coleman, A.M.G.1995. "Arbitrage, Storage, and the 'Law of One Price.'" Unpublished, Princeton University.

Cuddington, J.T. and D.L. Moss. 1999. "Technological Change, Depletion and the U.S. Petroleum Industry," Georgetown University working paper.

Dahl, Carol A., and Thomas K. Matson. 1998. "Evolution of the U.S. Natural Gas Industry in Response to Changes in Transaction Costs." Land Economics 74 (3): 390-408.

De Vany, Arthur S., and W. David Walls. 1993. "Pipeline Access and Market Integration in the Natural Gas Industry: Evidence from Cointegration Tests." The Energy Journal 14: 1-19.

———, 1994a. "Open Access and the Emergence of a Competitive Natural Gas Markets." Contemporary Economic Policy 12: 77-96.

———, 1994b. "Natural Gas Industry Transformation, Competitive Institutions and the Role of Regulation: Lessons from Open Access in US Natural Gas Markets." Energy Policy 22(9): 755-763.

_____, 1996. "The Law of One Price in a Network: Arbitrage and Price Dynamics in Natural Gas City Gate Markets." Journal of Regional Science 36(4): 555-570.

Doane, Michael J., and Daniel Spulber. 1994. "Open Access and the Evolution of the U.S. Spot Market for Natural Gas." Journal of Law and Economics 37: 477-517.

Dumas, B. 1992. "Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World." Review of Financial Studies 5: 153-180.

Geroski, P.A. 1998. "Thinking Creatively about Markets." International Journal of Industrial Organization 16: 677-695.

Grossman, Sanford. and Joseph Stiglitz. 1980. "On the Impossibility of Informationally Perfect Market." American Economic Review 70(3):393-408.

Haldane, A.G., and S.G. Hall. 1991. "Sterling's Relationship with the Dollar and the Deutchemark: 1976-1989." The Economic Journal 101, 436-443.

Herbert, John H. and Erik Kreil. 1996. "US Natural Gas Markets: How Efficient are They?" Energy Policy 24 (1): 1-5.

Hall, S.G., D. Robertson and M.R. Wickens. 1992. "Measuring Convergence of the EC Economies." The Manchester School of Economic and Social Studies 60: 99-111.

Horowitz, I. 1981. "Market Definition in Antitrust Analysis: A Regression-Based Approach." Southern Economic Journal 48, 1-16.

King, Martin, and Milan Cuc. 1996. "Price Convergence in North American Natural Gas Spot Markets." The Energy Journal 17(2):17-42.

Kleit, Andrew N. 1998. "Did Open Access Integrate Natural Gas Markets? An Arbitrage Cost Approach." Journal of Regulatory Economics 14: 19-33 .

Avoy, P., and R. Pindyck. 1975. The Economics of the Natural Gas Shortage (1960-1980). North-Holland, New York.

Michael, Panos, A.Robert Nobay, and David A. Peel. 1997. "Transaction Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation." Journal of Political Economy 105 (4), 862-879.

Obstfeld, Maurice, and Alan M. Taylor. 1997. "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited," Journal of the Japanese and International Economics 11: 441-479.

Obstfeld, Maurice, and Alan M. Taylor. 1997. "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited." NBER working paper No. 6053.

O'Connell, Paul G. J., and Shang-Jin Wei. 1997. "The Bigger They are, the Harder They Fall": How Price Differences Across U.S. Cities Are Arbitraged." NBER Working Paper 6089.

Perron, Pierre. 1989. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." Econometrica 57, 1361-1401.

Perron, Pierre. 1990. "Testing for a Unit Root in a Time Series With a Changing Mean." Journal of Business & Economic Statistics 8 (2): 153-162.

Prakash, Gauri. 1997. "Pace of Market Integration." Ph.D. dissertation, Northwestern University.

Prakash, Gauri, and Alan M. Taylor. 1998. "Measuring Market Integration: A Model of Arbitrage with An Econometric Application to the Gold Standard, 1879-1913." Photocopy.

Ravallion, M. 1986. "Testing Market Integration", American Journal of Agricultural Economics 68 (1): 102-109.

Rogoff, K. 1996. "The Purchasing Power Parity Puzzle", Journal of Economic Literature 34 (June): 647-668.

Serletis, Apostolos. 1997. "Is There an East-West Split in North American Natural Gas Markets?" The Energy Journal 18(1): 47-63.

Sexton, Richard J., Catherine L. Kling, and Hoy F. Carman. 1991. "Market Integration, Efficiency of Arbitrage, and Imperfect Competition: Methodology and Application to U.S. Celery." American Journal of Agricultural Economics 73 (3): 568-580.

Slade, Margaret E. 1986. "Exogeneity Tests of Market Boundaries Applied to Petroleum Products." Journal of Industrial Economics 34: 291-303.

Stigler, G.J., and R.A.Sherwin. 1985. "The Extent of the Market." Journal of Law of Economics 28:555-585.

Spiller, Pablo T., and Cliff J. Huang. 1986. "On the Extent of the Market: Wholesale Gasoline in the Northeastern United States." Journal of Industrial Economics 35(2):131-145.

Tsay, R.S. 1989. "Testing and Modeling Threshold Autoregressive Processes." Journal of the American Statistical Association 84: 231-240.

Uri, Noel D., and Edward Rifkin. 1985. "Geographic Markets, Causality and Railroad Deregulation." Review of Economics & Statistics 67:422-428.

Walls, David W. 1994a. "A Cointegration Rank Test of Market Linkages with an

Application to the U.S. Natural Gas Industry." Review of Industrial Organization 9:181-191.

———, 1994b. "Price Convergence Across Natural Gas Fields and City Markets." The Energy Journal 15 (4): 37-49.

---------, 1995. "Competition, Prices, and Efficiency in the Deregulated Gas Pipeline Network: A Multivariate Cointegration Analysis." The Journal of Energy and Development 19 (1): 1-14.

Werden, Gregory J., and Luke M. Froeb. 1993. "Correlation, Causality, and All that Jazz: The Inherent Shortcomings of Price Tests for Antitrust Market Delineation." Review of Industrial Organization 8(3): 329-354.

William, J. C., and B. D. Wright. 1991. Storage and Commodity Markets. Cambridge, U. K. Cambridge University Press.

Winston, Clifford. "U.S. Industry Adjustment to Economic Deregulation." Journal of Economic Perspectives 12(3): 89-110.

Zivot, Eric., and Donald W. K. Andrews. 1992. "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis." Journal of Business and Economic Statistics 10 (3): 251-270.