Cointegration Analysis of Commodity Prices: Much Ado about the Wrong Thing?

Mindy L. Mallory
Department of Agricultural and Consumer Economics
University of Illinois at Urbana-Champaign
Urbana, IL 61801
217-333-6547
mallorym@illinois.edu

Sergio H. Lence
368E Heady Hall
Department of Economics
Iowa State University
Ames, IA 50011-1070, USA
Phone: 515-294-6644
shlence@iastate.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2010 AAEA,CAES, & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010

Cointegration Analysis of Commodity Prices: Much Ado about the Wrong Thing?

Mindy L. Mallory and Sergio H. Lence*

Abstract

This study highlights some problems with using the Johansen cointegration statistics on data containing a negative moving average (NMA) in the error term of the data generating process. We use a Monte Carlo experiment to demonstrate that the asymptotic distribution of the Johansen cointegration statistics is sensitive to the NMA parameters and that using the stated 5% critical values results in severe size distortion. In our experiment, using the asymptotic critical values resulted in empirical size of 76% in the worst case. To date a NMA in the error term was known to cause poor small sample performance of the Johansen cointegration statistics; however our study demonstrates that problems associated with a NMA in the error term do not improve as sample size increases. In fact, the problems become more severe. Further, we show that commodity prices in the U.S. tend to exhibit this property. We recommend that researchers pretest data for NMA in the error term before using the standard asymptotic critical values to test for cointegrating rank.

Keywords: cointegration, Johansen cointegration test, moving average

JEL Codes: C32, C15

*Mindy L. Mallory is an Assistant Professor at the University of Illinois at Urbana-Champaign and Sergio H. Lence is Professor and Marlin Cole Chair in International Agricultural Economics at Iowa State University.
The Authors would like to thank Christina Jolejole for technical assistance on this project.

Cointegration Analysis of Commodity Prices: Much Ado about the Wrong Thing?

Mindy L. Mallory and Sergio H. Lence

Economists have formed a cottage industry out of testing for cointegrating relationships among economic variables. Since the development of cointegration tests by Johansen (1988), Johansen and Juselius (1990) and Johansen (1991), the cointegration framework has been used to examine spatial arbitrage and the law of one price for commodities, trade conditions, and purchasing power parity, to name some prominent examples (Corbae and Ouliaris 1988; Enders 1988; Kim 1990; Goodwin and Schroeder 1991; Goodwin 1992; Johansen and Juselius 1992; Chowdhury 1993; Kugler and Lenz 1993; Adamowicz and Luckert 1997; Bahmani-Oskooee 1998; Asche, Bremnes et al. 1999; Boyd, Caporale et al. 2001; Goodwin and Piggott 2001; Caporale and Chui 2002; Sephton 2003; Hoover, Johansen et al. 2008).

Given the widespread use of cointegration analysis in the economics literature, it is noteworthy the lack of studies assessing the robustness of results to slight changes in the assumptions underlying the cointegration tests. This is somewhat puzzling because, for example, there are several studies showing that the Johansen tests possess poor small sample properties when applied to data with a negative moving average (NMA) component in the error term and low power against a persistent alternative (Cheung and Lai (1993), Toda (1995), Johansen (2002), Ahn and Reinsel (1990), and Reimers (1992)).

This study begins to fill a noticeable gap in the existing empirical cointegration literature, by exploring the properties of the Johansen cointegration tests when used to analyze time series like the ones often used by applied economists. The intended contributions of the study are threefold. First, we show that a NMA component in the error term is a common attribute among

U.S. commodity prices. We arrive at this conclusion by performing the corresponding tests on a large number of U.S. price series for important commodities, including major feed grains, other feed products, livestock, and consumer food prices.

Second, we examine the properties of the Johansen cointegration tests as sample size changes for data generating processes (DGPs) containing a NMA in the error term. This study advances the previous small-sample literature by focusing on DGPs that mimic the temporal nature of actual time series for commodity prices. Prior studies did not draw a connection between the simulation experiments and the type of data they intended to mimic. Their simulated DGPs did not correspond to daily, weekly, monthly, quarterly, or annual observations. They merely set the variance within some convenient range, e.g., $\sigma \in [0.25, 1]$. However, this distinction is critically important. While 400 observations of *annual* data is a longer series than social scientists ever enjoy, 400 observations of *daily* data contain scarcely more than a year's worth of information. A year is not likely to be nearly enough time to capture the variation required in the data to test for cointegrating relationships for crops, feedstuffs, and other food-related products with yearly production cycles. Yet, small-sample studies often make recommendations of the nature, "With 100 observations it seems safe to assume that asymptotic results apply."

Because of this, previous studies leave the applied researcher with the question, "How much data is enough to test for cointegrating relationships?" We constructed hypothetical data series designed to mimic "daily data" of varying lengths and varying severity of MA errors. By this we mean that we calibrated the random disturbances of our Monte Carlo experiment's DGP so that the annualized volatility, as measured by standard deviation of prices divided by the mean, is equal to a value typical for commodity prices. By designing the Monte Carlo experiment in this way, we show that NMA errors really are not a small sample issue at all. In fact, the performance of the test statistics becomes worse as sample size is lengthened in the presence of NMA errors (e.g., the empirical sizes corresponding to the stated 0.05 asymptotic critical values are on the order of 0.75). This suggests that the asymptotic distribution of the test statistic is quite sensitive to the MA parameter.

Several researchers have been close to discovering this property, but none really put the NMA error structure and the poor performance of the Johansen statistics together. There is a bit of foreshadowing in the unit root literature, as Phillips and Perron (1988) note that in a univariate unit root testing environment the $T\left(\hat{\alpha}-1\right)$ and related statistics diverge as the MA term approaches -1. Hodgson (1998) develops a residual-based cointegration test that allows for MA errors, but it is unclear from the design of his Monte Carlo experiment whether the asymptotic distribution of the likelihood ratio statistic is affected in the residual based test or not. Lütkepohl and Claessen (1997) estimate cointegrated vector autoregressive moving average processes, citing Saikkonen and Luukkonen (1997) as justification that the cointegrating rank test in this environment follows the distribution of the standard Johansen trace test. However, Saikkonen and Luukkonen showed this result for an infinite order autoregressive process rather than an ARMA process. Therefore, the Johansen critical values are not necessarily appropriate in this context.[1]

The third contribution of this paper is in providing several new examples of how the Johansen test statistics behave for the aforementioned commodity price series. We show that the test statistics are extremely sensitive to the influence of individual data points, making the test statistics very sensitive to the length of the sample. When combined with the Monte Carlo

[1]Since testing for cointegration is analogous to testing for multivariate unit roots, and given the result of Phillips and Perron in the univariate unit root literature, this seems to be a serious misstep of Lutkepohl and Claessen.

analysis, these examples show that the Johansen tests are rendered of little use for the kinds of data agricultural economists are often interested in. This is true because for such data the size distortions are quite large and the test statistics are extremely sensitive to the influence of individual data points. We argue that for many data series it is not a question of whether or not the series are cointegrated, but rather the relevant task is estimating the parameters of a vector error correction model (VECM) with precision. For example, economic theory would predict that if the price of soybeans is high relative to the price of soy oil and meal for a sufficiently long period, soybean processors would exit the industry until processing margins can sustain the remaining firms. Similarly, if the price of soybeans is low relative to the price of soy oil and meal, processing firms will enter the market and eventually drive processing margins down. Therefore, testing for cointegration in the soybean complex is not necessarily an interesting question, whereas estimating the speed at which prices reverse to their long-run equilibrium relationship is a task that arguably has value.

Characterizing the Time Series Properties of Prices for Major U.S. Commodities

We begin our study by characterizing the time series properties of a large number U.S. price series for important commodities, including major feed grains, other feed products, livestock, and consumer food prices. Specifically, we show that a NMA error structure is quite common for such series.

A data dictionary is provided in Table 1 indicating the data source, beginning and end of sample, and frequency of each data set we analyze. Our data series vary in length because we chose to use as much of the data as was available, so that in each case we have the longest data set possible. This is important for us because we argue that the problems associated with a NMA in the error term are not a small-sample concern, but one that persists into the asymptotics. All of the series we examine are quite long, representing 20 to 40 years worth of data. Therefore, to remove the effects of inflation over this long period we deflate each series by the Consumer Price Index (CPI) maintained by the Bureau of Labor Statistics.[2]

To test for MAs in the DGPs of our data, we fit the ARIMA(p,q) model (1) to each series enumerated in the data dictionary.

(1) $$\Delta P_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i \Delta P_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

We use the Bayesian information criterion to select p and q; we specifically note the sign and test the significance of the coefficients corresponding to the lagged error terms. In table 2 we report the parameter estimates and p-values, as well as the R^2 for each data series. Starred values are significant at the 5% level, the parameter estimates of the lagged error terms that are both negative and significant are shaded.

Over 50% of the series investigated (20 out of 36) have a statistically significant NMA component in the error term. The figure increases to over 75% of the series (28 out of 36) when considering all series with negative point estimates on the lagged error terms (regardless of whether they are statistically significant or not). While the series we examined are not an exhaustive list of commodity prices, these findings indicate that MA components in the error term are common and pre-testing for this structure should be part of the routine in estimating cointegrating rank. We underscore the importance of this issue in the sections below.

[2]CPI data can be found here at http://www.bls.gov/data/

Technological Relationships and Issues with the Johansen Tests

Many subsets of the data analyzed above should be expected to contain cointegrating relationships based on long run equilibrium and technological production processes present within industry groups. These include: crude oil, gasoline, and diesel – crack spread; wheat and wheat flour; soybeans, soybean oil, and soybean meal – soybean crush; corn, feeder cattle, and live cattle – cattle crush; lean hogs, pork bellies, and corn; corn and soybeans; corn gluten feed and corn gluten meal, and other feed products.

We perform cointegration tests on these industry groups and report the results in table 3. The Johansen tests clearly indicate in each industrial processing group that the variables are cointegrated except for corn gluten feed and meal. The statistics are generally large so that the null hypothesis of no cointegration is easily rejected. In fact, for wheat flour and the soybean crush the statistics are so large that in sequentially testing for the number of cointegrating relationships the tests conclude that there are as many cointegrating relationships as there are equations. This is equivalent to a joint test that each price series is stationary.

Effect of Sample Size on the Distribution of the Test Statistics

Previous studies conclude that performance problems, when they exist, are small sample issues that can be remedied with more data. We specifically chose variables for which very long data sets exist. Data on the soybean complex covers over 40 years, for example. Any difficulties associated with small samples should not be present in these data.

As a thought experiment, consider what happens to the Johansen test statistics if new data points are added to the time series. New data points will influence the likelihood function and cause the value of the test statistic to be perturbed by a small amount. The larger the dataset the smaller the effect each new data point should have. When the dataset becomes large enough we would expect additional entries to cause the value of the test statistic to vary randomly around the mean of its asymptotic distribution.

To see if this behavior is observed for our data, we calculate the Johansen statistics varying the sample size. We begin with the sample size used to generate the statistics in table 3, then at each iteration we remove the oldest data point, calculate the statistics for $r \leq 0$, and plot them. Therefore with each iteration, looking left to right, the dataset becomes shorter. Figures 1-5 display the results. The horizontal line represents the 95% critical value as a reference point.

These figures clearly do not display behavior we would expect from a statistic sampled from its asymptotic distribution, the values of the test statistic clearly have not converged, and in each case seem to be increasing. Further, this behavior is highly correlated with the presence of a NMA in the error term. That is, if one compares the magnitudes of the NMA parameter estimates, the largest NMA coefficients correspond to the worst behavior in the Johansen cointegration statistics.

Clearly, using the asymptotic critical values to test for cointegration at the 5% and 1% stated levels of significance will result in huge size distortions. In each figure the test statistic seems to be growing with the size of the data set (moving right to left). Comparing this with the 5% critical value represented by the horizontal lines, we see given enough data we will always reject the null hypothesis when a NMA is present.

In the past, researchers in possession of a long dataset have naively assumed they could confidently use the asymptotic critical values. When in fact many of these studies probably have come to erroneous conclusions regarding the cointegrating rank of their data series. In the following sections we develop a Monte Carlo study to further support this finding.

Previous Small Sample Investigations

Since Johansen's trace and maximum eigenvalue tests of cointegrating rank are *asymptotic* likelihood ratio tests it is likely that they have undesirable properties in small samples. It is known that that the tests suffer from size distortion and low power in small samples, especially when the error correction model produces residuals are nearly I(1). Several Monte Carlo studies have been published outlining the severity of these issues.

Cheung and Lai (1993) determine the finite sample sizes of the Johansen tests and quantify the finite sample critical values using response surface analysis. They conclude the Johansen tests are biased toward rejecting a null of no cointegration too often in finite samples compared to the asymptotic distribution of the test statistics. Further, they conclude that the bias worsens as the dimension of the system or length of the lag structure increases. This contrasts with our findings, regarding NMA errors at least, in that increasing the lag length reduces the problem of size distortions.

Toda (1995) performs an independent study of the finite sample performance of the Johansen tests and determines that with 100 observations the simulated distribution of the asymptotic test statistic under the null is fairly good. However, 100 observations are not enough to determine the true cointegrating rank under the alternative if one or more of the stationary roots of the process is nearly 1; that is, the test has low power against a persistent alternative. Unlike Cheung and Lai, Toda asserts that this leads to underestimation of the cointegrating rank because of the nature of sequential testing inherent in the Johansen procedure. Further, he finds the test's performance is affected by initial values of the stationary component of the process. Toda concludes that one needs 300 observations for the test to perform well uniformly over the range of finite sample scenarios he considers. Although, the simulated data is not calibrated to a specific temporal frequency so this recommendation is difficult for the applied researcher to interpret since the evidence is mixed about whether temporal aggregation of the data helps or hinders the power of cointegration tests to detect equilibrium relationships in small samples (see (Shiller and Perron 1985; Hooker 1993; Lahiri and Mamingi 1995; Otero and Smith 2000) for discussion). The level of temporal aggregation will certainly influence the distribution of the test statistics in a small sample, and whether it increases or decreases the tests' power will depend on the nature of the DGP.

Alternative to determining the critical values of the actual finite sample distribution, small sample corrections to the test statistics or critical values have been proposed. Johansen (2002) proposes a correction factor that depends on parameters of the error correction model as well as the sample size. However, the correction is fairly complicated to apply (the components of the correction which depend only on functionals of a random walk are simulated and described in (Johansen, Hansen et al. 2005)); it is not clear that one cannot obtain better estimates of the small sample critical values from simulating the small sample distributions directly. After all, the correction of the statistics developed by Johansen (2002) requires estimating the parameters of the data, just as is required to simulate the small sample distribution correctly.

Ahn and Reinsel (1990) and Reimers (1992) develop a correction that is a simple function of sample size, system dimension and lag order. However, as part of their Monte Carlo analysis Cheung and Lai (1993) conclude that the Ahn-Reinsel method does not yield unbiased estimates of the finite sample critical values.

Therefore, it seems appropriate to investigate how much data is required before the tests statistic appears to approach asymptotic behavior under a range of assumptions about the DGP.

Further, it seems worthwhile to provide small sample critical values that are associated more directly with a particular type of temporally aggregated data set, and for varying lengths of data available.

In the next section we perform a small Monte Carlo study that provides critical values for Johansen's statistics on cointegrating rank. The experiment is tailored explicitly to 'daily data'. We provide both critical values of the distribution of the statistic under the null hypothesis, as well as a small study of the power of the test statistic under the alternative hypothesis.

A Monte Carlo Study

The DGP we use closely resembles that used in prior Monte Carlo studies done by Banerjee et al. (1986) and Haug (1996). In this study, we restrict our attention to a bivariate system. The DGP is

(1) $$y_t - x_t = v_t, \quad v_t = \rho v_{t-1} + w_t,$$

(2) $$y_t + x_t = \psi_t, \quad \psi_t = \psi_{t-1} + r_t, \qquad r_t = \varphi_t + \theta\varphi_{t-1}$$

(3) $$\begin{bmatrix} w_t \\ \varphi_t \end{bmatrix} \overset{iid}{=} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \eta \\ \eta & \sigma_2^2 \end{bmatrix} \right).$$

With this DGP a moving average component in the error terms exists when θ is non-zero. For $\rho = 1$ the data is generated under the null hypothesis of no cointegration, while a value of $|\rho| < 1$ corresponds to the alternative hypothesis that the two series are cointegrated. We choose values of η and σ so that the covariance matrix of $\left[y_t \quad x_t \right]'$ matches levels typical for daily price changes, namely an annualized price volatility of 0.25. We run the simulations for each parameter scenario (ρ, θ), where $\rho \in \{0.85,\ 0.90,\ 1\}$, and $\theta \in \{-0.80,\ 0,\ 0.80\}$. We follow Haug (1996) in the choice of these parameter values, which allows us to illustrate the effect of a moving average component in the error term on the size distortion and power of the Johansen tests.

Deriving the variances and covariance of $\begin{bmatrix} y_t & x_t \end{bmatrix}'$ implied by the DGP in (7) and (8), one finds that the covariance matrix of $\begin{bmatrix} w_t & \varphi_t \end{bmatrix}'$ relates to the covariance of $\begin{bmatrix} y_t & x_t \end{bmatrix}'$ by the equations

(4) $\sigma_1^2 = 2\left(\sigma_x^2 - \sigma_{xy}\right)$

(5) $\sigma_2^2 = 2\left(\sigma_y^2 + \sigma_{xy}\right) / \left(1+\theta^2\right)$

(6) $\eta = \sigma_y^2 - \sigma_x^2$

Where $\Sigma_{yx} = \begin{bmatrix} \sigma_y^2 & \sigma_{yx} \\ \sigma_{yx} & \sigma_x^2 \end{bmatrix}$ is the covariance matrix of the $\begin{bmatrix} \Delta y_t & \Delta x_t \end{bmatrix}'$ series. These equations allow us to tailor the Monte Carlo experiment to mimic data of any temporal aggregation level.

We match the covariance structure of the simulated data series to one that might be found in two price series with cross correlation; we set $\Sigma_{y,x} = \begin{bmatrix} 0.0004 & 0.0002 \\ 0.0002 & 0.0004 \end{bmatrix}$.

We perform the simulation using 100,000 replications, and for sample lengths varying from one month to 100 years. Table 4 contains the finite sample critical values for different assumptions on the error term (values of θ), lag specification of the VECM, and for varying data series lengths. We report the empirical size of the trace and maximal eigenvalue cointegration tests based on the stated 5% asymptotic critical value. We also report the actual 95% critical value of the simulated distribution of the statistics for each sample size.

When the moving average component of the DGP is zero or positive, ($\theta \geq 0$), the small sample size distortion disappears with little more than six months of data, and the statistics have surprisingly small size distortion with as little as three months of data.

On the other hand when data is generated with a NMA in the error term, the size distortion is severe. With a lag length of k = 2 the size distortion persists even after the statistic seems to have converged. Table 4 also shows that increasing the lag length to k = 4 or k = 5 is helpful in reducing the size distortion, but even including 5 lags leaves an asymptotic size distortion of 22 percentage points, and increasing the lag length is costly in terms of precision of the estimate of the mean reverting parameter in the VECM, which is often the primary parameter of interest in these kinds of studies.

Table 5 contains the results of a small power study of the test statistic under the alternative hypothesis that $|\rho| < 1$. The reported values are *size adjusted* powers, so the numbers represent the probability of rejecting the false null hypothesis using the appropriate small sample critical values as calculated in table 4. In other words, the probability of a type II error given that one is testing with the appropriate critical values found in table 4. Given the size distortion reported in table 4 this is an important distinction. Notice that if the asymptotic critical values are used, the probability of committing a type I error grows with sample size, as does the probability of committing a type II error since the mean of the actual distribution under the null hypothesis is increasing as the sample size grows.

By simulating data for $\rho = 0.85$, 0.90, and 0.95 we demonstrate how the test loses ability to discern a cointegrating relationship from increasingly persistent alternative hypotheses. The trace statistic performs relatively well in terms of power also when $\theta \geq 0$. When the DGP has $\rho = 0.85$ the power of the test is relatively good for sample length as small as six months, and has very high power for samples of 2 years or longer. The power becomes weaker when the DGP is closer to a the null hypothesis of $\rho = 1$ with $\rho = 0.90$ and 0.95. In these cases, two years of data are required before the test statistic has reasonable power; however, the power is very good for data series of four years or more.

We get nearly identical power results when the data contain a positive MA in the error term, but the size adjusted power of the test suffers when a NMA is present in the error term of the DGP.

Comparing the Simulated Data to 'Real World' Data Visually

It is useful to construct figures which record the trace and maximal statistics against sample size as we did above for actual data series. These are found in figures 9-11. To construct these figures we simulate one 10 year realization of the DGP for the following scenarios: null hypothesis of no cointegration and no NMA in the error term, null hypothesis of no cointegration and a NMA in the error term, and alternative hypothesis of cointegration and a NMA in the error term. The most instructive graphs are the two scenarios under the null hypothesis. Recall that the figure plots the trace and maximal statistics as sample size is shortened. At each step from left to right we drop the oldest observation, calculate the statistic, and plot it.

Keep in mind that these figures represent a single realization of the DGP and are not necessarily representative of the distribution of the statistic; however, this is still a useful tool since these figures can be directly compared to the figures generated by our actual data series.

Figure 9 is created from simulated data under the null hypothesis of no cointegration and no NMA in the error term. Recall from the Monte Carlo study that under this scenario the Johansen tests perform well; size distortions quickly disappear and power becomes quite good for a modest length sample. The figure is consistent with this as well. Looking at the graph from right to left we see the cointegration statistics varying as new data points are added. Each new data point perturbs the statistic, but the line appears to be quite stationary. Also, recall that the horizontal line represents the 95% asymptotic critical value of the statistic. Since statistics are almost always below the critical value, one would correctly fail to reject the null hypothesis except in a few of the subsamples. This is exactly what we would expect to see when the statistics perform as advertized.

Looking at figure 10 we see a different picture. This figure is generated using 10 years of simulated data under the null hypothesis of no cointegration, but with a NMA in the error term. Notice that the characteristics of this figure are much more like figures 1-8 for our real world data series. The figure is characterized by increasing value of the test statistic as sample size increases (right to left), and for almost all subsamples we would easily (but incorrectly) reject the null hypothesis of no cointegration.

Figure 11 is generated using 10 years of simulated data under the alternative hypothesis and a NMA in the error term. The NMA in the error term causes the value of the cointegration statistics to increase with sample size in this case as well. The magnitude is generally much larger, but both data series generated under the null and under the alternative will easily reject the null hypothesis of no cointegration. This illustrates that when there is a NMA in the error term, it is not wise to test for coinegration by means of the Johansen statistics.

Recommendations for the Applied Researcher

Our results show that it is crucially important when testing for cointegrating rank to determine whether the data being considered contains a NMA error structure or not. When the error term does not exhibit a NMA, small-sample size distortions disappear with little more than six months of daily data, and performs impressively well with as little as three months of simulated daily observations, while if a NMA error structure is present one should not use the published asymptotic critical values for hypothesis testing.

When testing in the presence of a persistent alternative hypothesis (ρ close to 1) the size adjusted power of the Johansen tests are similar regardless of whether a moving average structure was present in the error term or not, but under a more favorable alternative hypothesis (ρ sufficiently less than 1) the size adjusted power of the statistic is worse when a NMA is present than otherwise.

Future Research

There is much work to be done on this issue. The first question is whether or not a NMA in the error term plagues the residual based tests in the same way it does the Johansen tests. Since the residual based cointegration tests are effectively a subset of the unit root testing literature it is also relevant to determine the extent to which a NMA in the error term is responsible for unit root testing problems and if it is an asymptotic or small sample issue here as well.

Due to the severity of this issue and the prevalence of the problem in real world data, new techniques for testing for cointegrating rank would be highly useful. For example, testing for significance of cointegrating terms in a likelihood based estimation of a cointegrated VARMA model may be superior to using the Johansen procedure. Alternatively, it may be possible to add a MA component to the error structure in the auxiliary regressions used to construct the Johansen statistics.

Conclusion

Error structure which contains a NMA is a common characteristic of commodity price series. We provide evidence that a NMA in the error term alters the asymptotic distribution of the Johansen cointegration statistics. This means that the published asymptotic critical values cannot be used in the cointegration testing procedure for data series with this characteristic.

The severity of the size distortion is lessened by increasing the lag length beyond what would be chosen by typical methods like the Akaike information criterion or Bayesian information criterion, but at the cost of precision in parameter estimates. Since the speed of mean reversion is often the question of interest this is an unsatisfactory solution.

A NMA in the error term also has a detrimental effect on the size adjusted power of the statistics, but this seems to be a much less important issue than addressing the size distortion. This is because as the distribution of the statistic grows with sample size, the probability of a type II error goes to 0 while the probability of a type I error goes to 1.

Table 1: Data dictionary

Data Set	Start Date	End Date	Source[a]	Frequency
Sorghum – No. 2 yellow, Kansas City	Sep 1975	Aug 2009	ERS Feed grains database	Monthly
Barley – No. 2 feed Portland, OR	June 1975	May 2004	ERS Feed grains database	Monthly
Oats – white heavy, Minneapolis, MN	June 1975	May 1994	ERS Feed grains database	Monthly
Alfalfa meal dehydrated, 17% protein – Kansas City, MO	Oct 1981	Sep 2009	ERS Feed grains database	Monthly
Corn Gluten Feed 21% - Midwest	Oct 1981	Sep 2009	ERS Feed grains database	Monthly
Corn Gluten Meal 60% - Midwest	Oct 1981	Sep 2009	ERS Feed grains database	Monthly
Cottonseed meal 41% - Memphis, TN	Oct 1981	Sep 2003	ERS Feed grains database	Monthly
Meat and Bone Meal – Central US	Oct 1981	Sep 2009	ERS Feed grains database	Monthly
Urea 42% - Fort Worth, TX	Oct 1981	Sep 2009	ERS Feed grains database	Monthly
Wheat Bran – KC, MO	Oct 1981	Sep 2009	ERS Feed grains database	Monthly
Corn Meal yellow - NY, NY	Sep 1983	Aug 2009	ERS Feed grains database	Monthly
High Fructose Corn Syrup 42% - Midwest	Sep 1983	Aug 2009	ERS Feed grains database	Monthly
Milk, price received by farmers, all milk	Jan 1970	Dec 2004	NASS, Agricultural Prices	Monthly
Milk, wholesale nonfat dry	Jan 1970	Dec 1981	NASS, Agricultural Prices	Monthly
Eggs, grade A large	Jan 1980	Dec 2009	U.S. LS	Monthly
Butter salted, US City average	Jan 1980	Dec 2007	U.S. BLS	Monthly
White flour all purpose, US city average	Jan 1980	Dec 2009	U.S. BLS	Monthly
Ground Beef, US city average	Jan 1984	Dec 2009	U.S. BLS	Monthly
Chicken, fresh whole US city average	Jan 1980	Dec 2009	U.S. BLS	Monthly
Potatoes, white US city average	Jan 1988	Dec 2009	U.S. BLS	Monthly
Sugar, white US city average	Jan 1980	Dec 2009	U.S. BLS	Monthly
Chicken breast, bone in US city average	Jan 1980	Dec 2002	U.S. BLS	Monthly
Chicken legs, bone in US city average	Jan 1980	Dec 1986	U.S. BLS	Monthly
Corn	3/1/1968	8/1/2009	Barchart “cash”	Monthly
Soybeans	3/1/1968	8/1/2009	Barchart “cash”	Monthly
Soy Oil	3/1/1968	8/1/2009	Barchart “cash”	Monthly
Soy Meal	3/1/1968	8/1/2009	Barchart “cash”	Monthly

Wheat	3/1/1968	5/1/2009	ERS Wheat Yearbook	Monthly
Wheat Flour	3/1/1968	5/1/2009	ERS Wheat Yearbook	Monthly
Crude Oil – Cushing, OK	6/15/1986	7/15/2009	U.S. EIA	Monthly
Gasoline – NY Harbor Conventional	6/15/1986	7/15/2009	U.S. EIA	Monthly
Diesel – Los Angeles, CA No2	6/15/1986	7/15/2009	U.S. EIA	Monthly
Live Cattle	1/1/1970	8/1/2009	Barchart "cash"	Monthly
Feeder Cattle	1/1/1970	8/1/2009	Barchart "cash"	Monthly
Lean Hogs	3/1/1968	8/1/2009	Barchart "cash"	Monthly
Pork Bellies	3/1/1968	8/1/2009	Barchart "cash"	Monthly

[a]ERS denotes the Economic Research Service of the U.S. Department of Agriculture (USDA), NASS is the USDA National Agricultural Statistics Service, BLS is the Bureau of Labor Statistics, and EIA is the Energy Information Administration, Barchart denotes spot prices as archived by Barchart.com.

Table 2: ARIMA Results[a]

		φ_1	φ_2	φ_4	φ_4	θ_1	θ_2	θ_3	θ_4
Alfalfa	Coefficient	1.73*	-0.99*			-1.70*	0.96*		
R^2= 0.0866	P-value	0.00	0.00			0.00	0.00		
Barley	Coefficient	0.91*	-0.63*	0.13*		-0.69*	0.37*		
R^2= 0.1450	P-value	0.00	0.00	0.02		0.00	0.00		
Butter	Coefficient	1.32*	-0.95*			-1.30*	0.99*		
R^2= 0.05	P-value	0.00	0.00			0.00	0.00		
Whole Chicken	Coefficient	0.23				-0.23			
R^2= 0.0433	P-value	0.35				0.93			
Chicken Breast	Coefficient	-0.72				0.67			
R^2 = 0.003	P-value	0.14				0.19			
Chicken Legs	Coefficient	0.10				0.04			
R^2 = 0.0174	P-value	0.89				0.96			
Corn Gluten Feed	Coefficient	1.04*	0.06	-0.45*	-0.14*	-1.02*	-0.21	0.68*	
R^2= 0.1061	P-value	0.00	0.85	0.04	0.02	0.00	0.46	0.00	
Corn Gluten Meal	Coefficient	0.87*	-0.07	0.004	0.10	-0.97*			
R^2= 0.0618	P-value	0.00	0.33	0.95	0.10	0.00			
Corn Meal Yellow	Coefficient	-0.35				0.54*			
R^2= 0.0373	P-value	0.13				0.01			
Cotton Seed Meal	Coefficient	-0.23	-0.18	-0.06	-0.13	-0.15			
R^2= 0.1434	P-value	0.61	0.30	0.58	0.06	0.74			
Eggs	Coefficient	1.62	-0.81	-0.11		-1.79	1.11	-0.08	
R^2= 0.1229	P-value	0.00	0.11	0.70		0.00	0.03	0.79	
Ground Beef	Coefficient	-0.02				-0.17			
R^2= 0.0326	P-value	0.94				0.56			
High Fructose Corn Syrup	Coefficient	1.67*	-1.81*	1.50*	-0.87*	-1.55*	1.56*	-1.39*	0.82*
R^2= 0.2659	P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

[a] Starred values significant at 5% level, shaded values are significant and negative

The φ_i and θ_j are the estimated AR and MA coefficients respectively.

Table 2 (continued): ARIMA Results [a]

		φ_1	φ_2	φ_4	φ_4	θ_1	θ_2	θ_3	θ_4
Meat and Bone Meal	Coefficient	-0.99*	-0.34*	-0.30*	-0.13*	0.90*			
R^2= 0.0768	P-value	0.00	0.00	0.00	0.02	0.00			
Milk Farm Gate Price	Coefficient	2.59*	-3.11*	1.94*	-0.62*	-2.12*	2.27*	-1.43*	0.57*
R^2= 0.5196	P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Milk Wholesale Non Fat	Coefficient	-0.07	0.11	-0.21*		0.56			
R^2= 0.2144	P-value	0.88	0.66	0.05		0.29			
Oats	Coefficient	-0.04				0.29			
R^2= 0.0585	P-value	0.88				0.24			
Potatoes	Coefficient	1.38*	-0.49*			-0.96*			
R^2= 0.2555	P-value	0.00	0.00			0.00			
Sorghum	Coefficient	0.37				0.23			
R^2= 0.0627	P-value	0.83				0.22			
White Sugar	Coefficient	0.97*	-0.95*	0.51*	-0.01	-0.43*	0.89*		
R^2= 0.5406	P-value	0.00	0.00	0.00	0.85	0.00	0.00		
Urea	Coefficient	-0.69*	0.83*	0.34*	-0.39*	0.95*	-0.64*	-0.85*	
R^2= 0.2220	P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Wheat bran	Coefficient	1.06*	0.17	-0.67*		-1.28*	-0.03	0.88*	-0.27*
R^2= 0.1600	P-value	0.00	0.67	0.00		0.00	0.93	0.00	0.00
White All Purpose Flour	Coefficient	0.46	-0.45	-0.55		-0.54	0.51	0.48	
R^2= 0.1634	P-value	0.32	0.33	0.23		0.26	0.30	0.32	
Live Cattle	Coefficient	0.11				-0.04	-0.06	-0.20*	-0.19*
R^2= 0.0854	P-value	0.60				0.84	0.18	0.00	0.00
Feeder Cattle	Coefficient	-0.90	-0.98			0.87*	0.98*		
R^2= 0.0403	P-value	0.00	0.00			0.00	0.00		
Corn	Coefficient	-0.13				0.10	0.17*	-0.10	-0.11
R^2= 0.0480	P-value	0.79				0.82	0.00	0.34	0.11

[a] Starred values significant at 5% level, shaded values are significant and negative

The φ_i and θ_j are the estimated AR and MA coefficients respectively, and $\hat{\sigma}^2$ represents the sum of squared residuals in the regression.

Table 2 (continued): ARIMA Results [a]

		φ_1	φ_2	φ_4	φ_4	θ_1	θ_2	θ_3	θ_4
Wheat	Coefficient	0.73				-0.83*	0.10	-0.07	-0.11*
R^2= 0.0523	P-value	0.00				0.00	0.08	0.25	0.04
Wheat Flour	Coefficient	0.89				-0.91*	0.05	-0.25*	0.14*
R^2= 0.0659	P-value	0.00				0.00	0.42	0.00	0.01
Crude Oil	Coefficient	1.27*	-0.23*	-0.18*		-0.90*			
R^2= 0.2436	P-value	0.00	0.03	0.01		0.00			
Diesel	Coefficient	1.10*	-0.10	-0.12*		-0.93*			
R^2= 0.0730	P-value	0.00	0.27	0.06		0.00			
Regular Gasoline	Coefficient	1.90*	-1.53*	0.58*	-0.25*	-0.25*	-1.73*	1.00*	
R^2 = 0.1975	P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Soybeans	Coefficient	0.85*	0.09	-0.12*		-0.94*			
R^2= 0.0426	P-value	0.00	0.11	0.02		0.00			
Soybean Meal	Coefficient	0.13	0.06	-0.20		-0.12			
R^2= 0.0414	P-value	0.56	0.15	0.00		0.60			
Soybean Oil	Coefficient	0.11	0.23	-0.17		-0.36			
R^2= 0.1367	P-value	0.73	0.02	0.00		0.26			
Lean Hogs	Coefficient	0.47*	-0.02	-0.12*		-0.52*			
R^2= 0.0256	P-value	0.03	0.65	0.03		0.02			
Pork Bellies	Coefficient	0.80*	0.15*	-0.09		-0.96*			
R^2= 0.0458	P-value	0.00	0.01	0.06		0.00			

[a] Starred values significant at 5% level, shaded values are significant and negative
The φ_i and θ_j are the estimated AR and MA coefficients respectively.

Table 3: Johansen Cointegration Test Results for Industrial Groups

	H_2:	Trace test	5% c.v.	1% c.v	Max test	5% c.v.	1% c.v
Crack Spread	$r \le 0$	109.79	29.68	35.65	64.93	20.97	25.52
	$r \le 1$	44.87	15.41	20.04	39.09	14.07	18.63
	$r \le 2$	5.77	3.76	6.65	5.77	3.76	6.65
Wheat and Flour	$r \le 0$	42.66	15.41	20.04	34.60	14.07	18.63
	$r \le 1$	8.06	3.76	6.65	8.06	3.76	6.65
Soybean Crush	$r \le 0$	84.35	29.68	35.65	52.04	20.97	25.52
	$r \le 1$	32.31	15.41	20.04	24.19	14.07	18.63
	$r \le 2$	8.11	3.76	6.65	8.11	3.76	6.65
Live Cattle, Feeders, Corn	$r \le 0$	92.15	29.68	35.65	71.37	20.97	25.52
	$r \le 1$	20.78	15.41	20.04	15.74	14.07	18.63
	$r \le 2$	5.05	3.76	6.65	5.05	3.76	6.65
Lean Hogs, Pork Bellies, Corn	$r \le 0$	55.61	29.68	35.65	27.82	20.97	25.52
	$r \le 1$	27.80	15.41	20.04	23.91	14.07	18.63
	$r \le 2$	3.89	3.76	6.65	3.89	3.76	6.65
Corn Soybeans	$r \le 0$	33.72	15.41	20.04	28.64	14.07	18.63
	$r \le 1$	5.08	3.76	6.65	5.08	3.76	6.65
Corn Gluten Feed/Meal	$r \le 0$	19.69	15.41	20.04	16.76	14.07	18.63
	$r \le 1$	2.93	3.76	6.65	2.93	3.76	6.65
Feed Products	$r \le 0$	119.58	69.82	77.82	47.60	33.88	39.37
(Alfalfa meal, Corn gluten feed, Corn gluten meal, Meat and bone meal, Urea, Wheat bran)	$r \le 1$	71.98	47.86	54.68	35.27	27.59	32.72
	$r \le 2$	36.71	29.68	35.65	26.65	20.97	25.52
	$r \le 3$	10.05	15.41	20.04	7.66	14.07	18.63
	$r \le 4$	2.40	3.76	6.65	2.40	3.76	6.65

Critical values from Osterwald-Lenum(1992)
No constant term in the cointegrating vector

Table 4: Finite sample critical values for Johansen's cointegration tests – "Daily data"

	Sample length	Empirical Size based on stated 5% significance level		Actual Trace	Actual Max
	T	Trace	Max	95% c.v.	95% c.v.
$\theta = 0$	1 mo	0.19	0.12	24.17	20.22
k = 2	3 mos	0.07	0.04	19.36	16.17
	6 mos	0.06	0.03	18.71	15.59
	2 yrs	0.05	0.03	18.32	15.28
	4 yrs	0.05	0.03	18.10	15.01
	10 yrs	0.05	0.03	18.11	15.01
	20 yrs	0.05	0.03	18.13	15.08
	100 yrs	0.05	0.03	18.12	15.03
	∞			18.17	16.87
$\theta = 0.8$	1 mo	0.20	0.13	25.33	21.18
k = 2	3 mos	0.07	0.04	19.41	16.18
	6 mos	0.06	0.03	18.54	15.44
	2 yrs	0.05	0.03	17.90	14.92
	4 yrs	0.05	0.03	17.85	14.84
	10 yrs	0.04	0.02	17.74	14.78
	20 yrs	0.04	0.02	17.79	14.83
	100 yrs	0.04	0.02	17.75	14.79
	∞			18.17	16.87
$\theta = -0.8$	1 mo	0.22	0.14	25.23	21.12
k = 2	3 mos	0.40	0.30	28.41	24.86
	6 mos	0.60	0.51	35.19	31.98
	2 yrs	0.73	0.67	52.47	49.67
	4 yrs	0.74	0.70	59.17	56.31
	10 yrs	0.76	0.71	64.23	61.41
	20 yrs	0.76	0.71	66.54	63.74
	100 yrs	0.76	0.71	67.94	64.98
	∞			18.17	16.87
$\theta = -0.8$	1 mo	0.48	0.39	37.32	32.07
k = 4	3 mos	0.18	0.12	23.44	19.85
	6 mos	0.24	0.17	25.44	21.98
	2 yrs	0.34	0.28	30.58	27.43
	4 yrs	0.37	0.30	32.33	29.34
	10 yrs	0.38	0.32	33.67	30.69
	20 yrs	0.39	0.32	34.26	31.25
	100 yrs	0.39	0.33	34.56	31.67
	∞			18.17	16.87
$\theta = -0.8$	1 mo	0.81	0.75	71.82	63.68
k = 5	3 mos	0.17	0.10	23.07	19.35
	6 mos	0.18	0.12	23.45	20.00
	2 yrs	0.24	0.18	26.41	23.18
	4 yrs	0.25	0.19	27.27	24.16
	10 yrs	0.26	0.20	27.93	24.84
	20 yrs	0.27	0.21	28.31	25.22
	100 yrs	0.27	0.21	28.49	25.51

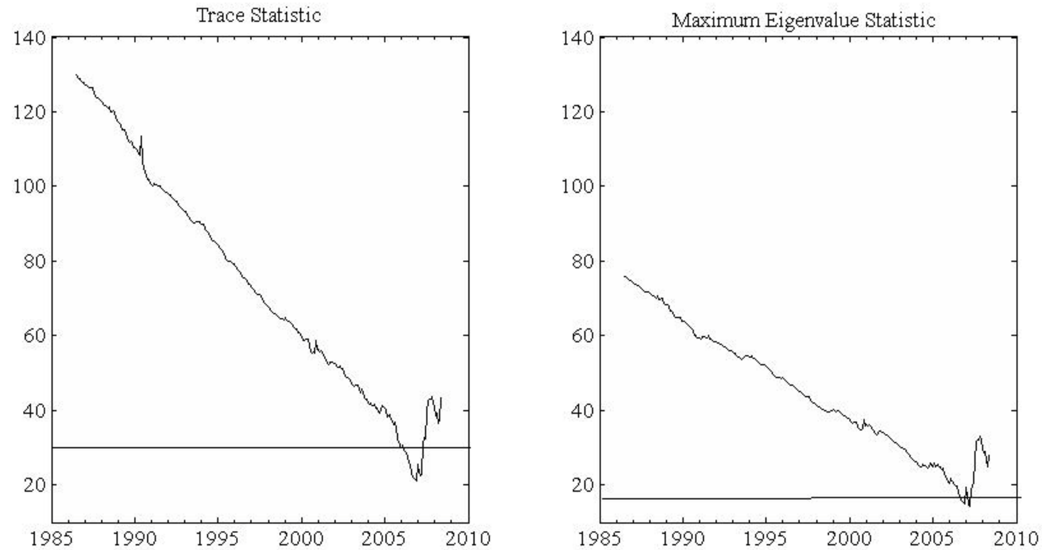
n = 2, k is the number of lags used to perform the tests, 100,000 replications, T assumes 250 trading days in one year. Asymptotic critical values from Osterwald-Lenum (1992)

Table 5: Size-adjusted finite sample power of Johansen's trace cointegration test under the alternative $|\rho| < 1$ – "Daily data"

	Sample length	$\theta = 0$ $k = 2$	$\theta = 0.8$ $k = 2$	$\theta = -0.8$ $k = 2$	$\theta = -0.8$ $k = 4$
$\rho = 0.85$	1 mo	0.05	0.05	0.05	0.05
	3 mos	0.10	0.10	0.08	0.11
	6 mos	0.28	0.27	0.12	0.21
	2 yrs	1	1	0.49	0.99
	4 yrs	1	1	0.99	1
	10 yrs	1	1	1	1
	20 yrs	1	1	1	1
	100 yrs	1	1	1	1
$\rho = 0.90$	1 mo	0.05	0.05	0.05	0.05
	3 mos	0.08	0.07	0.06	0.10
	6 mos	0.15	0.15	0.09	0.16
	2 yrs	0.99	0.99	0.22	0.95
	4 yrs	1	1	0.70	1
	10 yrs	1	1	1	1
	20 yrs	1	1	1	1
	100 yrs	1	1	1	1
$\rho = 0.95$	1 mo	0.05	0.05	0.05	0.05
	3 mos	0.06	0.06	0.06	0.11
	6 mos	0.08	0.08	0.0	0.14
	2 yrs	0.55	0.56	0.09	0.47
	4 yrs	0.99	0.99	0.17	0.98
	10 yrs	1	1	0.90	1
	20 yrs	1	1	1	1
	100 yrs	1	1	1	1

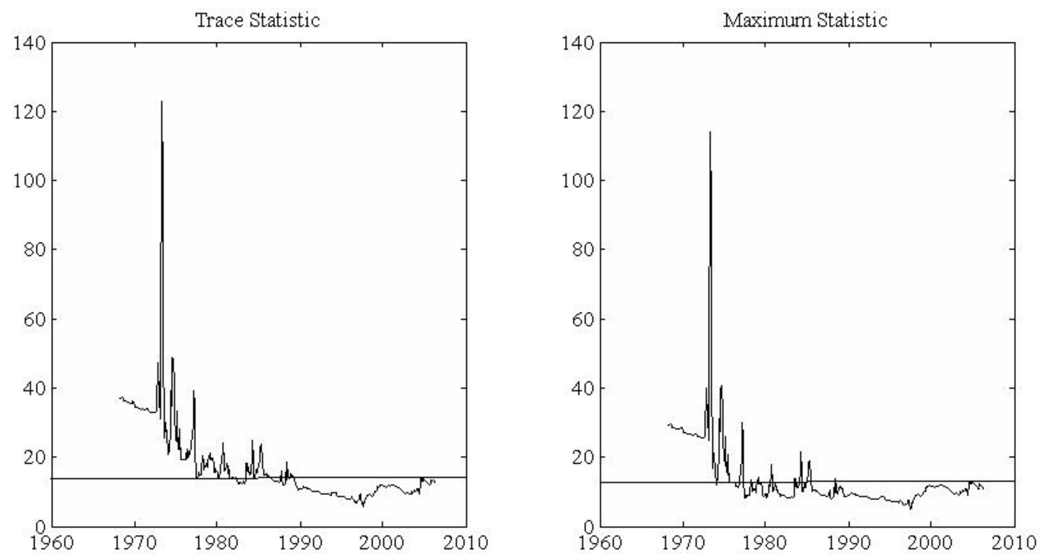
n = 2, k is the number of lags used to perform the tests, 100,000 replications, T assumes 250 trading days in one year.

Figure 1: Crack Spread



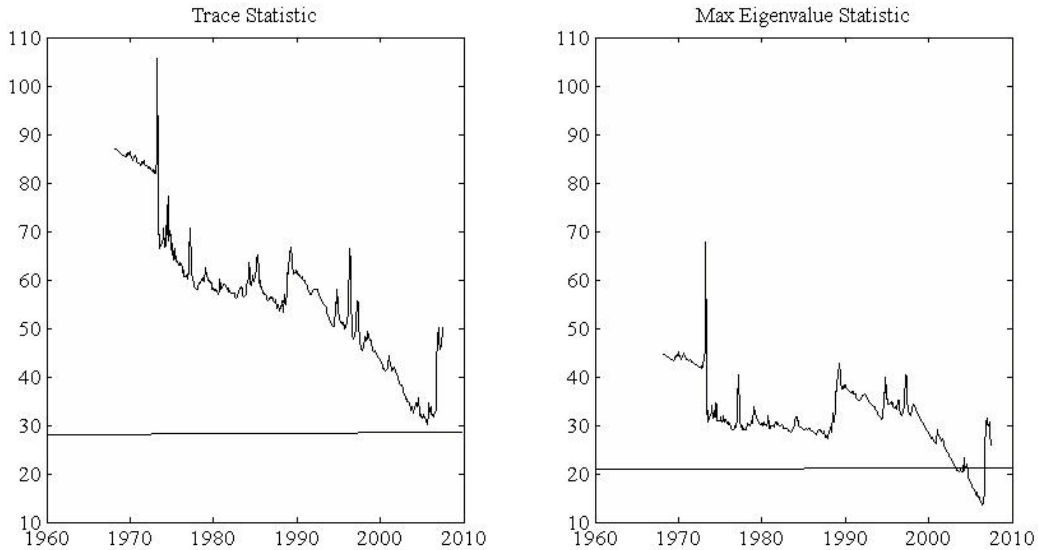
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 2: Wheat, Wheat Flour



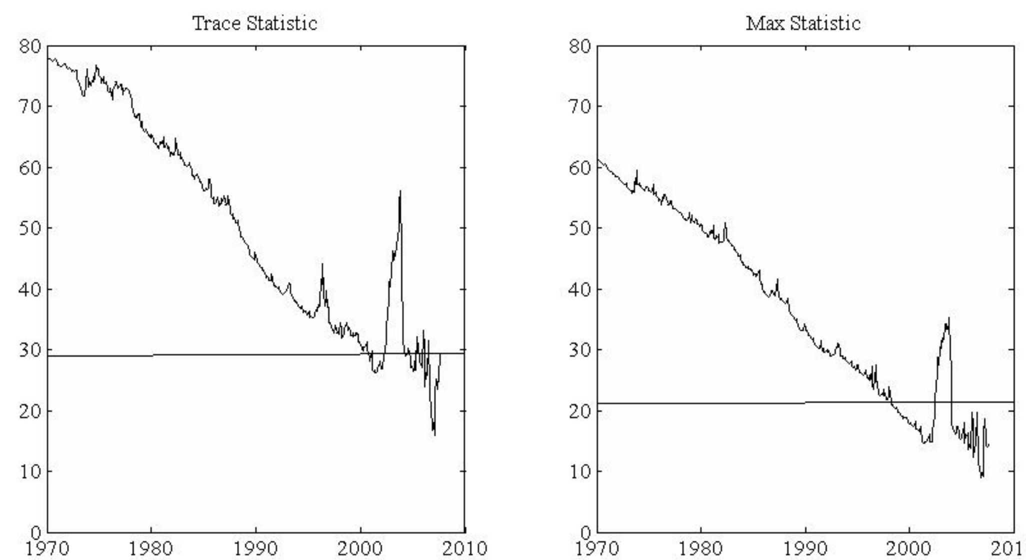
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 3: Soybean Crush



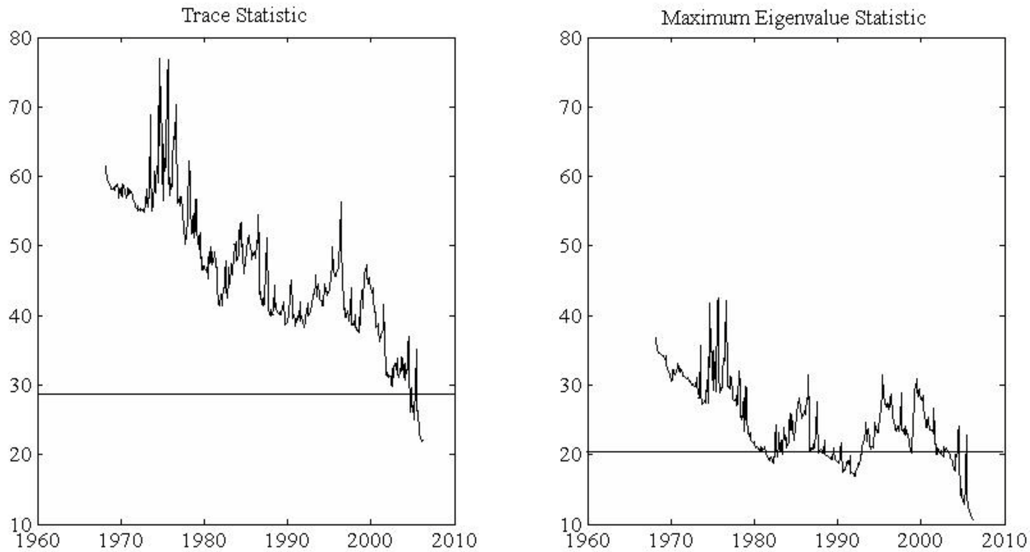
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 4: Cattle Crush



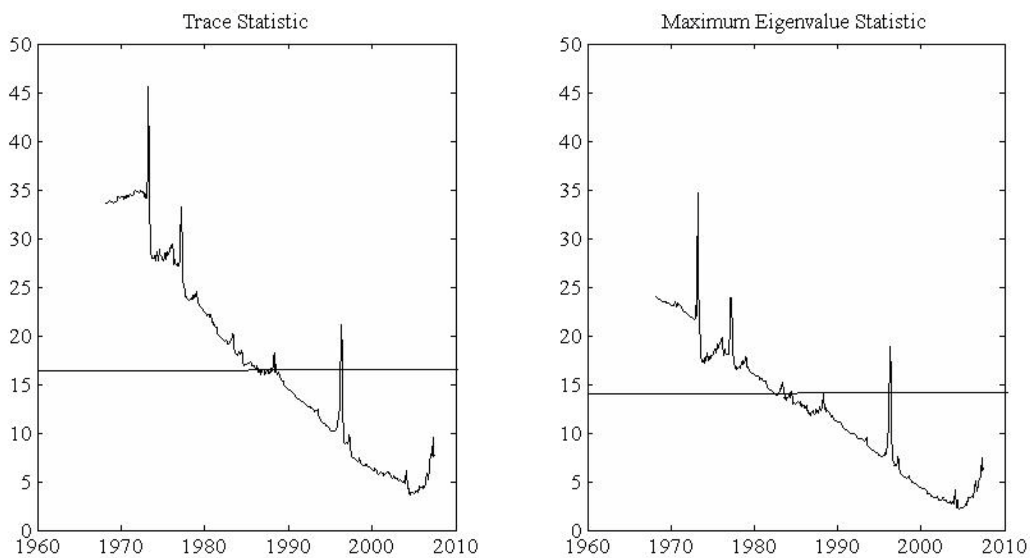
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 5: Lean Hogs, Pork Bellies, Corn



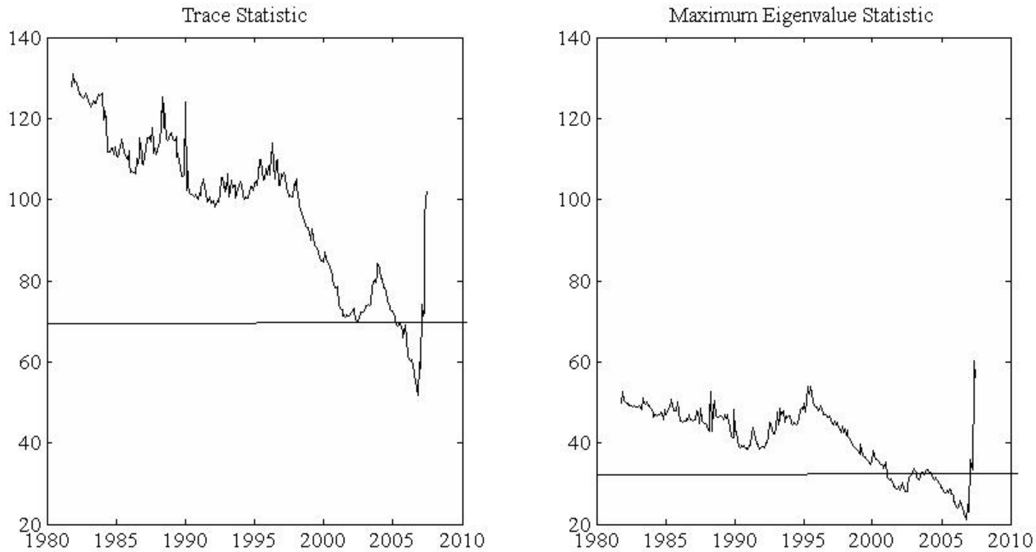
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 6: Corn and Soybeans



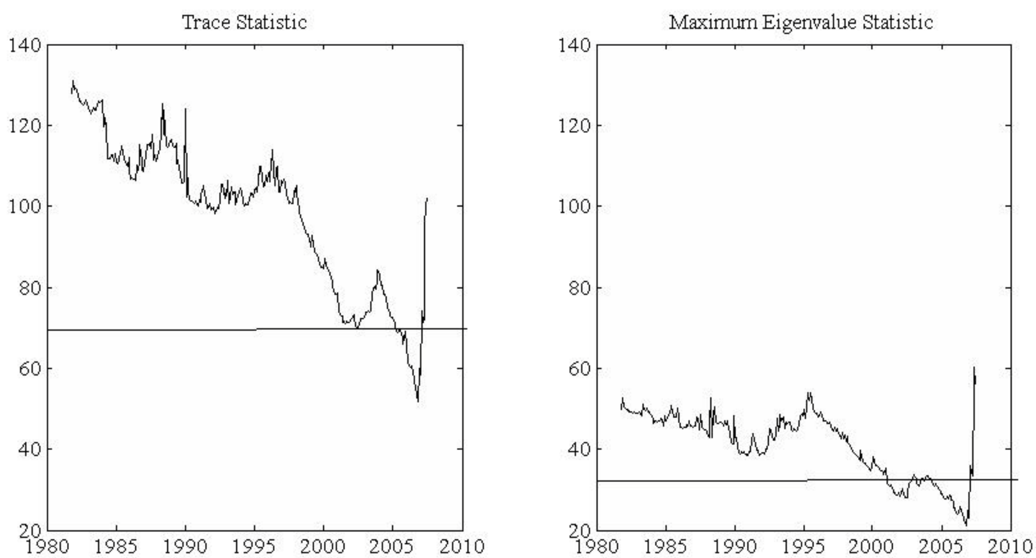
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 7: Corn Gluten Feed and Corn Gluten Meal



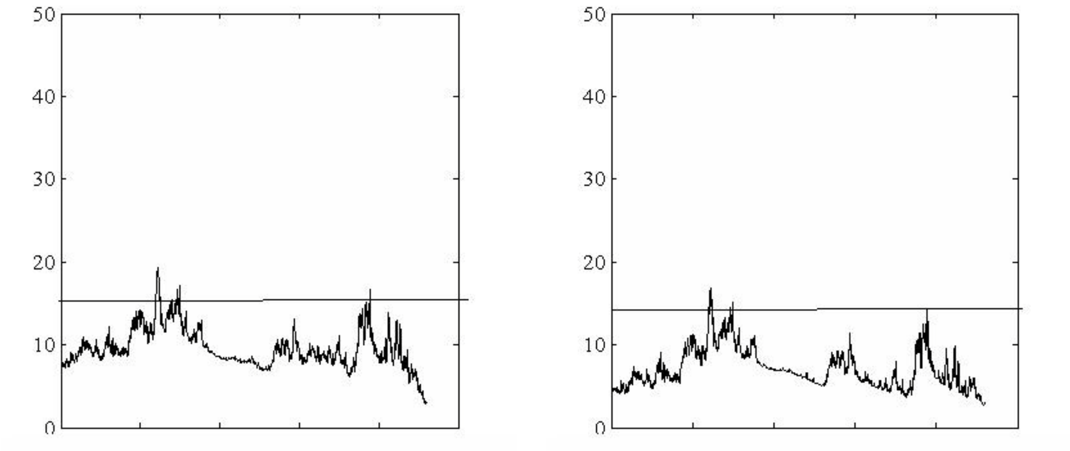
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 8: Other Feed Products



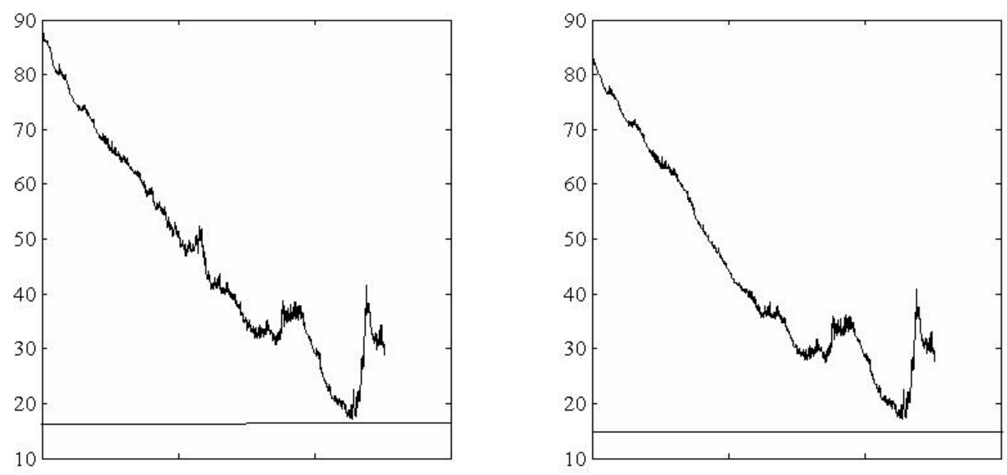
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 9: Simulated Data Under the Null Hypothesis of No Cointegration – No NMA in the Error Term



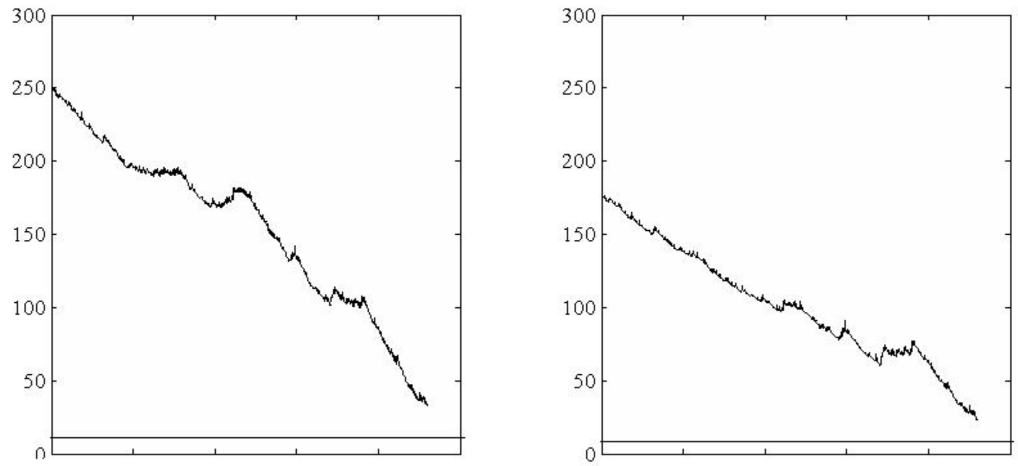
At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 10: Ten Years of Simulated Data Under the Null Hypothesis of No Cointegration – NMA in the Error Term



At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

Figure 11: Simulated Data Under the Alternative Hypothesis of Cointegration – NMA in the Error Term



At each step moving left to right the oldest observation in the dataset is dropped, the trace and maximal statistics are calculated for the remaining subsample, and plotted on the graph.

References

Adamowicz, W. L. and M. K. Luckert (1997). "A Cointegration Analysis of Canadian Wood Pulp Prices." American Journal of Agricultural Economics: 975-986.

Ahn, S. K. and G. C. Reinsel (1990). "Estimation for partially nonstationary multivariate autoregressive models." Journal of the American statistical association: 813-823.

Asche, F., H. Bremnes, et al. (1999). "Product aggregation, market integration, and relationships between prices: an application to world salmon markets." American Journal of Agricultural Economics: 568-581.

Bahmani-Oskooee, M. (1998). "Cointegration approach to estimate the long-run trade elasticities in LDCs." International Economic Journal **12**(3): 89-96.

Banerjee, A., J. J. Dolado, et al. (1986). "Exploring equilibrium relationships in econometrics through static models: some Monte Carlo evidence." Oxford Bulletin of Economics and statistics **48**(3): 253-277.

Boyd, D., G. Caporale, et al. (2001). "Real exchange rate effects on the balance of trade: cointegration and the Marshall-Lerner condition." International Journal of Finance & Economics **6**(3): 187-200.

Caporale, G. and M. Chui (2002). "Estimating income and price elasticities of trade in a cointegration framework." Review of International Economics **7**(2): 254-264.

Cheung, Y. W. and K. S. Lai (1993). "Finite-sample sizes of Johansen's likelihood ratio tests for cointegration." Oxford Bulletin of Economics and statistics **55**(3): 313-328.

Chowdhury, A. R. (1993). "Does Exchange Rate Volatility Depress Trade Flows? Evidence from Error- Correction Models." The Review of Economics and Statistics **75**(4): 700-706.

Corbae, D. and S. Ouliaris (1988). "Cointegration and tests of purchasing power parity." The Review of Economics and Statistics **70**(3): 508-511.

Enders, W. (1988). "ARIMA and cointegration tests of PPP under fixed and flexible exchange rate regimes." The Review of Economics and Statistics **70**(3): 504-508.

Goodwin, B. K. (1992). "Multivariate cointegration tests and the law of one price in international wheat markets." Review of Agricultural Economics: 117-124.

Goodwin, B. K. and N. E. Piggott (2001). "Spatial market integration in the presence of threshold effects." American Journal of Agricultural Economics: 302-317.

Goodwin, B. K. and T. C. Schroeder (1991). "Cointegration tests and spatial price linkages in regional cattle markets." American Journal of Agricultural Economics: 452-464.

Haug, A. A. (1996). "Tests for cointegration: a Monte Carlo comparison." Journal of Econometrics **71**(1): 89-116.

Hodgson, D. (1998). "Adaptive estimation of cointegrating regressions with ARMA errors." Journal of Econometrics **85**(2): 231-267.

Hooker, M. A. (1993). "Testing for cointegration : Power versus frequency of observation." Economics Letters **41**(4): 359-362.

Hoover, K., S. Johansen, et al. (2008). "Allowing the data to speak freely: The macroeconometrics of the cointegrated vector autoregression." American Economic Review **98**(2): 251-255.

Johansen, S. (1988). "Statistical analysis of cointegration vectors." Journal of Economic Dynamics and Control **12**(2-3): 231-254.

Johansen, S. (1991). "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models." Econometrica: Journal of the Econometric Society: 1551-1580.

Johansen, S. (2002). "A small sample correction for the test of cointegrating rank in the vector autoregressive model." Econometrica: 1929-1961.

Johansen, S., H. Hansen, et al. (2005). "A simulation study of some functionals of random walk." University of Copenhagen. Preprint.

Johansen, S. and K. Juselius (1990). "Maximum likelihood estimation and inference on cointegration with applications to the demand for money." Oxford Bulletin of Economics and statistics **52**(2): 169-210.

Johansen, S. and K. Juselius (1992). "Testing structural hypotheses in a multivariate cointegration analysis of the PPP and the UIP for UK* 1." Journal of Econometrics **53**(1-3): 211-244.

Kim, Y. (1990). "Purchasing Power Parity in the Long Run: A Cointegration Approach." Journal of Money, Credit & Banking **22**(4).

Kugler, P. and C. Lenz (1993). "Multivariate cointegration analysis and the long-run validity of PPP." The Review of Economics and Statistics: 180-184.

Lahiri, K. and N. Mamingi (1995). "Testing for cointegration: Power versus frequency of observation—another view." Economics Letters **49**(2): 121-124.

Lütkepohl, H. and H. Claessen (1997). "Analysis of cointegrated VARMA processes." Journal of Econometrics **80**(2): 223-239.

Osterwald-Lenum, M. (1992). "A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics." Oxford Bulletin of Economics and statistics **54**(3): 461-472.

Otero, J. and J. Smith (2000). "Testing for cointegration: power versus frequency of observation—further Monte Carlo results." Economics Letters **67**(1): 5-9.

Phillips, P. C. B. and P. Perron (1988). "Testing for a unit root in time series regression." Biometrika **75**(2): 335-346.

Reimers, H. E. (1992). "Comparisons of tests for multivariate cointegration." Statistical papers **33**(1): 335-359.

Saikkonen, P. and R. Luukkonen (1997). "Testing cointegration in infinite order vector autoregressive processes." Journal of Econometrics **81**(1): 93-126.

Sephton, P. S. (2003). "Spatial market arbitrage and threshold cointegration." American Journal of Agricultural Economics: 1041-1046.

Shiller, R. J. and P. Perron (1985). "Power versus frequency of observation." Economics Letters **18**: 381-386.

Toda, H. Y. (1995). "Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions." Econometric Theory: 1015-1032.