Journal of Agricultural and Applied Economics, 32,1(April 2000):73–87

The Distributional Behavior of Futures Price Spreads

Min-Kyoung Kim and Raymond M. Leuthold

ABSTRACT

The distributional behavior of futures price spreads is examined for four commodities: corn, live cattle, gold and T-bonds. Remarkably different results are found over commodities, time period, and sample size. Actual spread changes for the smaller sample size of gold and T-bonds and for corn produce more normal distributions for weekly than for daily differencing intervals, while all live cattle spreads for actual changes are normally distributed. However, the larger sample size of both gold and T-bonds and the relative spread changes for corn and live cattle do not become more normally distributed under temporal aggregation of the data.

Key Words: *corn, futures price spreads, gold, goodness of fit, live cattle, normality tests, spread distributions, T-bonds.*

The distribution of commodity futures price changes has been widely examined. Several studies (Houthakker; Mann and Heifner; Cornew, Town and Crowson; Blattberg and Gonedes; Hall, Brorsen, and Irwin) suggest that the distribution of price changes is not normal but leptokurtic. However, relatively few studies investigating the nature and the distributional properties of futures price spread (*fps*) changes exist. Identifying the relationships between prices of various futures contracts is crucial in understanding spread trading in futures markets. Spread trading between two futures contracts with different delivery dates provides a mechanism for traders to allocate risk among themselves, and in some cases helps determine carrying charges. Any risk transferred from the spot market to the futures market must be absorbed therein. Billingsley and Chance suggest that spread trading induces risk-averse futures traders to participate in the futures market, and it supplies liquidity to hedgers because spread positions generally carry less price risk than net positions in the market. Spread trading allows traders supplying price insurance to hedgers to reduce and reallocate their risks to other traders, as well as to transmit and interpret new information to the market. Melamed points out that spread trading is the largest contributing source of market liquidity, important for the efficiency and viability of these markets. He notes that spreading is the only mechanism that enables commercial traders to place hedges in distant contract months.

Knowing the distribution of futures price or futures price spread changes is important because most performance norms require that the changes be drawn from a common distribution, usually a normal distribution, with a finite variance. These performance norms typically include measurements of the mean and some measure of variability or risk. Thus, examining the probability distribution is im-

The authors are graduate research assistant and T. A. Hieronymous Professor in the Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

portant in the analysis of futures prices and futures price spreads. Futures market participants, especially speculators, can be successful at using the market to the extent of their knowledge of the probability distribution of the price, or futures price spreads, and can evaluate their risk through the distribution of the changes. The selection of statistical methods is then important in analyzing spread distributions. In this study, the distribution of changes in *fps* is examined and the characteristics of the distribution, skewness and kurtosis, are analyzed. For this reason, the Lagrange Multiplier (LM) and the omnibus tests are selected as methods to investigate the normality of changes in *fps* because they contain skewness and kurtosis properties, good finite sample performance, and the LM test possesses optimum asymptotic power properties (Bera and Jarque). The omnibus test is preferred in the case of small samples, and it is used to verify the LM test.

Although the distribution of changes in futures prices or the distribution of changes in futures price spreads has generally been found to be nonnormal, no study has identified the actual distribution for those variables. Knowing the appropriate distribution may benefit traders in making appropriate trading decisions and in understanding the risk in the futures market.

Therefore, using two normality tests and a goodness-of-fit test, this paper analyzes the distribution of changes in *fps* for gold, Treasury bonds, corn and live cattle, and identifies the 'best-fit' distribution for changes in *fps*. Data characterized by low and high volatility were deliberately selected.

The results of this paper are compared with those of Castelino and Vora and Poitras concerning whether the futures price spread volatility has a positive relationship with the spread length. Increasing spread length increases risks, which means the possible existence of compensating risk premiums (Castelino and Vora).

This paper is structured as follows: a brief discussion of previous research is provided in the next section; the statistical techniques, selection of data and the spread model are presented in Sections III and IV; detailed results follow in Section V; Section VI contains a general summary, and concluding remarks.

Literature Review

Many studies have examined the distributions of changes in price levels and returns of stocks, but the distributions of futures price spreads have rarely been examined. Poitras, likely the first researcher to study spread distributions, showed the 1982, 1983, and 1985 Dec.–June gold *fps* converged to a normal distribution for both actual and relative changes in futures price spreads when the spread differencing interval was changed from daily to weekly. But the 1981 and 1984 gold *fps* did not become more normally distributed with temporal aggregation. The distribution of daily gold *fps* was never normal.

Regarding the distributional effect of spread length, the shorter the spread length, the more likely that daily gold *fps* is peaked and fat tailed[1]. Also, futures price spreads' volatility was found to increase directly with the increment of spread length. This result was consistent with the results of Castelino and Vora who analyzed the effect of spread length on spread volatility for agricultural commodities and found strong evidence that the volatility of spreads increases with its length.

Monroe and Cohn tested market efficiency by investigating whether implied interest rates in gold spreads deviated substantially enough from Treasury bill interest rates to allow traders to earn profits from speculating on changes in the difference between two rates. They examined the frequency distribution of all differences between the implied gold interest rate and the T-bill rate, and found that this distribution exhibited a wide dispersion, and the differences between the gold and T-bill rates were frequently negative, providing significant evidence of market efficiency.

A sizable body of empirical research observes that futures price changes for short time

[1] Spread interval refers to the differencing interval, daily or weekly. Spread length refers to the number of months between the two contracts in the spread.

intervals are not normally distributed but exhibit a high degree of leptokurtosis, and suggests the stable Paretian or a mixture of normal distributions as reasons for the observed leptokurticity. Hall, Brorsen, and Irwin, and Cornew, Town, and Crowson found that the distribution of futures prices of agricultural, financial, and metal commodities was leptokurtic and hence not normally distributed. Also, Officer and Hsu, Miller, and Wichern used stock returns to describe the distribution of rates of returns on common stock. They found that the returns had some properties of a stable process, but the distributions have fat tails compared to the normal distribution. While the above studies suggested nonnormality in futures prices and stock returns, Hudson, Leuthold and Sarassoro found normality for commodity futures price changes. According to their results, futures price changes were found to be random, indicating that futures prices adjust efficiently to information, i.e., when the distributional aspect is considered, their results indicated a move toward normality.

Methodology

This study uses the LM and omnibus (K^2) tests to examine whether futures price spread changes for two agricultural and two nonagricultural commodities are and/or become normally distributed when increasing the differencing interval from daily to weekly (temporal aggregation)[2]. In addition, a skewness measure test is used to assess symmetry of the distributions and a kurtosis measure test assesses peakedness and fatness of tails. Finally, the optimal distribution of changes in *fps* for these four commodities will be found using "Bestfit" software (Palisade Corp.), which assesses how well the observations fit a certain distribution using a chi-square test.

LM Test

In this procedure, the log-likelihood function is maximized subject to a constraint, and a test statistic is constructed from the Lagrange multiplier (LM) for the constrained maximization (Ramanathan). If the constraint is true, then the slope of the log-likelihood function is zero. The LM test examines whether the slope of the log-likelihood function, evaluated at the restricted estimate, is significantly different from zero.

Bera and Jarque tested the power of several prevailing tests for normality and found that the LM may have good relative power even when the distribution is not a member of the Pearson family (normal, gamma, beta, or Student's t distributions). They also found that the power of the LM is not unsatisfactory even when the number of observation is 20 and 50, and suggest two aspects of the LM test as being useful. First, this test has asymptotic power characteristics (asymptotically efficient) including maximum local asymptotic power on the basis of small sample properties. Second, computation of this test is easy.

Assume that there are N independent observations on a random variable x, and that testing the normality of x is of interest. The LM test statistic is given by:

$$(1)\quad LM = N\left[\frac{(\sqrt{b_1})^2}{6} + \frac{(b_2 - 3)^2}{24}\right]$$

where

$$\sqrt{b_1} = \frac{\hat{\mu}_3}{\hat{\mu}_2^{3/2}}, \qquad b_2 = \frac{\hat{\mu}_4}{\hat{\mu}_2^2},$$

$$\hat{\mu}_j = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^j}{N},$$

x_i = ith random variable,

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} \qquad \hat{\mu}_j = j\text{th moment.}$$

N is a number of independent observation, and $\sqrt{b_1}$ and b_2 are, respectively, the skewness and kurtosis sample coefficients. The LM test

---

[2] Increasing the differencing intervals from daily to weekly eliminates an abundance of zeros and small changes. To examine monthly intervals is not possible because of the small number of observations at these wide intervals.

Table 1. Distributional Test for Changes in Gold *fps* (Large Sample Size)

Difference of Futures Price Spread			Skewness	Kurtosis	LM	K^2
1988	Dec.–Dec.	Daily (N = 210)	−0.838*	6.1*	108.12*	41.12*
		Weekly (N = 42)	−1.102*	5.40*	18.12*	14.73*
	Dec.–June	Daily (N = 334)	−2.456*	27.72*	8813.6*	240.39*
		Weekly (N = 68)	−0.121	4.40*	6.24*	4.67*
	Dec.–Feb.	Daily (N = 350)	−3.178*	31.61*	8910.4*	297.39*
		Weekly (N = 70)	−1.093	5.927*	30.58*	21.77*
1992	Dec.–Dec.	Daily (N = 209)	0.384	8.19*	238.18*	36.84*
		Weekly (N = 42)	0.470	3.116	1.53	2.25
	Dec.–June	Daily (N = 358)	−0.664*	11.95*	1218.7*	93.23*
		Weekly (N = 74)	0.124	4.49*	6.90*	5.21*
	Dec.–Feb.	Daily (N = 360)	−1.199*	13.25*	1343.6*	134.98*
		Weekly (N = 74)	−0.096	4.75*	7.61*	6.08*
Rate of Change in Futures Price Spread			**Skewness**	**Kurtosis**	**LM**	**K^2**
1988	Dec.–Dec.	Daily (N = 210)	−0.709*	5.851*	88.27*	34.58*
		Weekly (N = 42)	−0.974*	5.055*	3.69*	12.32*
	Dec.–June	Daily (N = 334)	−0.791*	12.541*	1297.7*	97.75*
		Weekly (N = 68)	0.433*	4.234*	6.92*	6.24*
	Dec.–Feb.	Daily (N = 350)	−1.571*	16.491*	1990.7*	168.48*
		Weekly (N = 70)	−1.626*	9.342*	116.41*	39.65*
1992	Dec.–Dec.	Daily (N = 209)	1.17*	13.31*	968.09*	84.01*
		Weekly (N = 42)	0.636*	4.128*	4.94*	6.19*
	Dec.–June	Daily (N = 358)	−0.262*	6.314*	198.32*	37.46*
		Weekly (N = 74)	0.287	3.25	1.19	1.80
	Dec.–Feb.	Daily (N = 360)	−0.792*	8.067*	341.71*	78.98*
		Weekly (N = 74)	−0.331	4.638*	7.68*	7.06*

* Indicates the null hypothesis of normal distribution is rejected at the 10% level of significance. The numbers in parentheses are the numbers of the observation.

is asymptotically distributed as χ^2 with two degrees of freedom.

Omnibus Test

Among several tests for normality, the omnibus tests might be appropriate for small sample sizes, which is based on the joint use of $\sqrt{b_1}$ and b_2. D'Agostino and Pearson suggested the statistic:

$$(2) \qquad K^2 = X^2(\sqrt{b_1}) + X^2(b_2)$$

as an omnibus test where $X(\sqrt{b_1})$ and $X(b_2)$ are standardized normal equivalent deviates and detect the direction of departure from normality. K^2 is needed to test the departure from normality due to the interaction between skewness and kurtosis. This statistic is based on an assumption of independence between $\sqrt{b_1}$ and b_2, but Bowman and Shenton pointed out that $\sqrt{b_1}$ and b_2 are uncorrelated and nearly independent. Thus, K^2 is approximately a chi-square variable with two degrees of freedom as is the LM test. A Johnson S_U approximation and Anscombe and Glynn approximation are used to estimate $X(\sqrt{b_1})$ and $X(b_2)$. These transformations are applicable for any sample size $N \geq 8$ for skewness and $N \geq 20$ for kurtosis. More details of $X(\sqrt{b_1})$ and $X(b_2)$ can be found in D'Agostino and Pearson, Bowman and Shenton, Pearson, D'Agostino, and Bowman, and D'Agostino, Belanger, and D'Agostino Jr.

Skewness and Kurtosis

The LM and the omnibus test statistics contain two properties of the normal distribution:

skewness and kurtosis. If a vector x follows a normal distribution, which has mean (μ) and variance (σ^2), then:

(3) $E[(x - \mu)^3] = 0$

(4) $E[(x - \mu)^4] = 3\sigma^4$.

Equations (3) and (4) measure skewness and kurtosis, respectively, and define that zero as the third central moment and three times the square of the variance of the fourth central moment as the normal distribution. These moments are used to measure skewness and kurtosis. The ratio of the skewness (kurtosis) to its standard deviation is used to construct tests of significance based on the Student's t distribution. If $\sqrt{b_1}$ test statistic as defined in (1) is not equal to zero, skewness is present and normality is rejected[3]. A random variable with a fourth moment larger than three times the square of the second moment has thicker tails than a normally distributed random variable, which is referred to as *excess kurtosis,* or as leptokurtic (Davidson and Mackinnon).

By observing individual $\sqrt{b_1}$ and b_2, then *fps* can be examined for the direction of departure from normality.

Bestfit Distribution

In the program Bestfit, the χ^2 test of goodness-of-fit is used as a measure of how well the sample data fit the hypothesized probability density function (Palisade Corp.). For a continuous distribution on a certain interval, the hypothesis is tested against the alternative that the distribution is not uniform over all the data (DeGroot). The distribution that has the lowest valued chi-square statistic will have the best fit among 25 different functions or distributions in the program.

Data And Model

Daily and weekly (Friday) closing futures prices were used for the contracts of corn, live cattle, gold, and T-bonds[4]. Rather than analyzing a series of years, coefficients of variation were calculated for every year from 1986 to 1995 to determine extremes in stability and instability. The highest price volatility for corn, live cattle and gold was observed for 1988 and the lowest volatility for 1992. However, for T-bonds, high volatility was observed for 1992 and low variation for 1988, the opposite of the other three commodities. Hence, these two years were selected for analysis.

For consistency in sample size with Poitras, the large samples for gold and T-bonds use data from July 1, 1987 to November 30, 1988 and from July 1, 1991 to November 30, 1992[5]. Meanwhile, a smaller sample size was necessary for corn and live cattle due to the shorter duration of their futures contracts. The data used in this case are from May 26, 1988 to November 30, 1988 and April 7, 1992 to November 30, 1992. In order to directly compare the two nonagricultural commodities with the agricultural commodities, gold and T-bonds are also examined over this same (smaller) sample size as for corn and live cattle. Each sample begins with the starting date of the deferred contract of the spread and ends two weeks before the first delivery date on the spread's nearby contract.

In defining futures price spread (*fps*), three spread lengths are examined, depending on the delivery months available: in the case of gold, one year (Dec.–Dec.), six months (Dec.–June), and two months (Dec.–Feb.); for T-bonds, one year (Dec.–Dec.), six months (Dec.–June), and three months (Dec.–Mar.); for corn, one year (Dec.–Dec.), seven months (Dec.–July), and three months (Dec.–Mar.). For live cattle, data beyond one year forward do not generally exist, so spread lengths examined are six months (Dec.–June), four months (Dec.–April), and two months (Dec.–Feb.). Hence, a total of six daily and six weekly sample futures price spreads for each commodity are examined.

---

[3] $\sqrt{b_1}$ tells how unsymmetric a distribution is around the mean. If $\sqrt{b_1}$ is positive (negative), the distribution is skewed to the right (left) with a long tail in that direction.

[4] Futures price data for corn, live cattle, gold, and T-bonds come from the Chicago Board of Trade, Chicago Mercantile Exchange, New York Mercantile Exchange (COMEX Division) and Chicago Board of Trade, respectively.

[5] Sample sizes are indicated in the tables of results.

For this study, a *futures price spread* is defined as the difference between two futures prices with different delivery dates. Specifically:

F(t, T); the futures price at time t for deferred delivery at time T

F(t, N); the futures price at time t for nearby delivery at time N

$$fps(\mathrm{t}) = \mathrm{F(t, T)} - \mathrm{F(t, N)}, \quad \text{and} \quad \mathrm{T} > \mathrm{N}.$$

Normality of futures price spread will be checked by estimating the distributional parameters, skewness, kurtosis, variance, and by examining the estimated parameters when the differencing interval is increased from daily to weekly. The following transformations for examining distribution of futures price spreads (*fps*) will be used.
Difference of futures price spread:

$$\mathrm{DFPS} = fps(\mathrm{t} + 1) - fps(\mathrm{t}),$$

and rate of change in futures price spread[6]:

$$\mathrm{RFPS} = \frac{fps(\mathrm{t} + 1) - fps(\mathrm{t})}{fps(\mathrm{t})}.$$

Results Of Statistical Tests

Normality Tests

Tables 1–6 show specific results for each statistical test, while Table 7 presents a general summary[7].

Gold and T-Bonds

The distributional behavior of gold and T-bonds is found to be very sensitive to the sample size. For the large sample size, gold and T-bonds did not converge to a normal distribution with temporal aggregation. Meanwhile, gold and T-bonds with small samples do become normally distributed as differencing intervals are increased from daily to weekly. All LM test results are confirmed by the omnibus (K^2) tests, especially for small sample sizes[8].

For the case of large sample gold *fps* (Table 1), only one spread length of each transformation does not reject the null hypothesis of a normal distribution for weekly intervals, but rejects it for daily intervals. Hence, five spreads for each DFPS and RFPS do reject the null hypothesis of normality. This nonnormality characteristic is usually due to significantly negative skewness combined with fat tails. The distributions of ten daily and three weekly spread changes are significantly skewed (negatively) to the left. The kurtosis tests show the distributions of all cases being significantly leptokurtic except two weekly intervals in 1992. The degree of leptokurtosis decreases with temporal aggregation in all cases[9].

Similar to gold, no general trend of convergence to normality is found for T-bonds' *fps* changes with temporal aggregation (Table 2). All T-bonds' *fps* changes for the stable period (1988) reject the null hypothesis of normality, producing significant coefficients of LM and K^2. Meanwhile, two DFPS and one RFPS for the unstable period (1992) do converge to the normal distribution with temporal aggregation.

Nonnormality of T-bonds is due to the combination of skewness and fat tails. However, the direction of skewness is different between DFPS and RFPS. The distributions of most DFPSs are skewed to the left while those of most RFPSs to the right. T-bonds' *fps* changes also exhibit significant leptokurtic distributions except three 1992 weekly intervals. Hence, daily and weekly changes in T-bonds' *fps* are neither generally normally nor lognormally distributed.

---

[6] RFPS is the same as the log difference between two *fps*.

[7] All tests are conducted at the 5% level of significance unless otherwise indicated.

[8] The omnibus (K^2) test confirms the LM test in most cases. Since the results of these two tests are so similar, and have the same implications with respect to this study, discussion of the LM test results is the focus in the text, although the K^2 test results are included in the tables.

[9] Skewness and kurtosis tests were also conducted within the framework of the omnibus test, and their results are very similar to those shown in the tables.

Table 2. Distributional Test for Changes in T-Bonds *fps* (Large Sample Size)

Difference of Futures Price Spread			Skewness	Kurtosis	LM	K^2
1988	Dec.–Dec.	Daily (N = 295)	0.157	9.044*	450.24*	47.58*
		Weekly (N = 60)	−0.445*	5.069*	12.69*	8.68*
	Dec.–June	Daily (N = 334)	−0.576*	8.080*	377.57*	61.91*
		Weekly (N = 68)	−0.499*	6.668*	40.94*	14.88*
	Dec.–Mar.	Daily (N = 350)	−0.770*	9.08*	467.06*	82.43*
		Weekly (N = 70)	−0.470*	6.581*	33.71*	14.55*
1992	Dec.–Dec.	Daily (N = 299)	−0.776*	7.453*	194.54*	62.35*
		Weekly (N = 61)	−0.149	2.459	0.68	0.97
	Dec.–June	Daily (N = 358)	−0.174*	4.690*	44.43*	18.43*
		Weekly (N = 74)	−0.432	4.159*	6.44*	6.27*
	Dec.–Mar.	Daily (N = 360)	−0.171	5.237*	66.17*	24.53*
		Weekly (N = 74)	−0.284	3.705	2.19	3.21
Rate of Change in Futures Price Spread			**Skewness**	**Kurtosis**	**LM**	**K^2**
1988	Dec.–Dec.	Daily (N = 295)	0.671*	11.766*	966.63*	78.70*
		Weekly (N = 60)	1.278*	8.413*	89.57*	28.74*
	Dec.–June	Daily (N = 334)	1.371*	14.977*	2100.8*	143.90*
		Weekly (N = 68)	1.435*	10.924*	201.25*	38.67*
	Dec.–Mar.	Daily (N = 350)	1.593*	16.321*	2227.6*	169.40*
		Weekly (N = 70)	1.338*	10.082*	140.89*	36.15*
1992	Dec.–Dec.	Daily (N = 299)	2.182*	19.954*	2681.6*	189.61*
		Weekly (N = 61)	0.095	2.755	0.17	0.12
	Dec.–June	Daily (N = 358)	0.398*	8.955*	538.39*	62.88*
		Weekly (N = 74)	0.744*	6.308*	40.56*	17.97*
	Dec.–Mar.	Daily (N = 360)	0.380*	9.704*	587.91*	67.07*
		Weekly (N = 74)	0.309	4.444*	6.58*	6.15*

* Indicates the null hypothesis of normal distribution is rejected at the 10% level of significance. The numbers in parentheses are the numbers of the observation.

To be consistent with the subsequent analysis on corn and live cattle, LM and K^2 distributional tests are performed on gold and T-bonds with a smaller sample size. Strikingly different results are found from the larger sample size. Most cases of gold and T-bonds either converge to a normal distribution with temporal aggregation, or are normally distributed for both intervals. In the case of gold in Table 3, all but one DFPS and four out of six of RFPS converge to a normal distribution based on both the LM and omnibus tests. The components of the omnibus test, $X(\sqrt{b_1})$ for skewness and $X(b_2)$ for kurtosis, also confirm the individual statistics of skewness and kurtosis shown in the tables at 10% level of significance. A similar situation is found in T-bonds (Table 4) where all cases of 1992 DFPS and four cases of RFPS (two each for 1988 and 1992) converge to a normal distribution with temporal aggregation[10]. In addition, all cases of 1988 DFPS and two cases of RFPS are normally distributed for both differencing intervals. As in the larger sample size of gold and T-bonds, nonnormality is usually caused by the combination of skewness and fat tails. One feature different in the smaller sample size is that leptokurtosis is found as the only reason causing nonnormality in four gold DFPS out of seven cases of nonnormal distribution, one gold RFPS out of eight cases, and three T-bond RFPS out of four cases. Unlike the larger sample size of gold and T-bonds, negative

[10] In the case of 1992 weekly Dec.–Dec. DFPS, the K^2 test departs slightly from normality at 10% level of significance, which is normal at 5% level of significance.

Table 3. Distributional Test for Changes in Gold *fps* (Small Sample Size)

Difference of Futures Price Spread			Skewness	Kurtosis	LM	K^2
1988	Dec.–Dec.	Daily (N = 129)	−0.513*	4.150*	12.77*	10.65*
		Weekly (N = 26)	0.237	2.976	0.24	0.60
	Dec.–June	Daily (N = 129)	−0.225	4.163*	8.35*	6.24*
		Weekly (N = 26)	0.172	2.921	0.14	0.37
	Dec.–Feb.	Daily (N = 129)	0.805*	10.058*	281.73*	40.61*
		Weekly (N = 26)	−0.375	1.989	1.72	2.88
1992	Dec.–Dec.	Daily (N = 164)	0.314*	9.174*	263.14*	33.02*
		Weekly (N = 33)	0.412	3.906	2.06	3.41
	Dec.–June	Daily (N = 164)	−0.007	6.366*	77.41*	18.53*
		Weekly (N = 33)	−0.213	3.135	0.27	0.79
	Dec.–Feb.	Daily (N = 164)	−0.017	7.896*	163.79*	25.43*
		Weekly (N = 33)	0.50	5.069*	7.26*	6.96*
Rate of Change in Futures Price Spread			**Skewness**	**Kurtosis**	**LM**	**K^2**
1988	Dec.–Dec.	Daily (N = 129)	−0.382*	4.030*	8.84*	7.63*
		Weekly (N = 26)	0.325	3.097	0.47	1.08
	Dec.–June	Daily (N = 129)	−0.082	4.237*	8.37*	5.61*
		Weekly (N = 26)	0.259	3.115	0.30	0.88
	Dec.–Feb.	Daily (N = 129)	1.314*	13.535*	633.66*	62.34*
		Weekly (N = 26)	−0.401	2.102	1.57	2.18
1992	Dec.–Dec.	Daily (N = 164)	1.240*	12.924*	715.14*	70.81*
		Weekly (N = 33)	0.805*	4.835*	8.20*	8.74*
	Dec.–June	Daily (N = 164)	0.832*	9.809*	335.74*	48.57*
		Weekly (N = 33)	0.152	3.759	0.92	2.01
	Dec.–Feb.	Daily (N = 164)	1.072*	12.004*	585.39*	62.81*
		Weekly (N = 33)	1.162*	7.740*	38.32*	18.64*

* Indicates the null hypothesis of normal distribution is rejected at the 10% level of significance. The numbers in parentheses are the numbers of the observation.

skewness does not prevail among the smaller sample size of gold and T-bonds. In sum, these distributional results seem very sensitive to sample size and data period.

In addition to the distributional behavior, the effect of spread length on distribution as well as volatility is examined. Negative correlation is expected between the spread length and nonnormality in daily results because daily *fps* would be dominated by zeros for shorter spread lengths, and hence the distribution would appear peaked and fat tailed. Again, discrepancy is detected between the two sample sizes. The results of gold and T-bonds with the larger sample size in Tables 1 and 2 partially support this expectation in gold DFPS and 1988 RFPS, and T-bonds 1988 RFPS. For instance, the 1988 daily T-bond RFPS kurtosis results show larger coefficients as the spread length moves from Dec.–Dec., Dec.–June to Dec.–Mar. Hence, the shorter the spread lengths, the more likely the daily results appear peaked and fat-tailed. This type of result occurred in three out of four situations for gold. Meanwhile, gold and T-bonds with the smaller sample size in Tables 3 and 4 do not show any consistent pattern of negative relationship between spread length and nonnormality, except for 1988 gold.

Corn and Live Cattle

Changes in corn and live cattle *fps* show substantially different results from the larger sample sizes of gold and T-bonds in case of DFPS, but more similar results in case of RFPS. In Table 5, four out of six cases of corn DFPS converge to the normal distribution as differencing intervals change from daily to weekly, but they show a significant discrepancy be-

Table 4. Distributional Test for Changes in T-Bonds *fps* (Small Sample Size)

Difference of Futures Price Spread			Skewness	Kurtosis	LM	K^2
1988	Dec.–Dec.	Daily (N = 130)	0.194	2.880	0.90	0.89
		Weekly (N = 26)	−0.080	2.324	0.52	0.37
	Dec.–June	Daily (N = 130)	0.121	2.858	0.43	0.36
		Weekly (N = 26)	−0.367	2.318	1.81	1.15
	Dec.–Mar.	Daily (N = 130)	0.025	2.520	1.26	1.51
		Weekly (N = 26)	−0.297	2.514	0.64	0.57
1992	Dec.–Dec.	Daily (N = 163)	0.620*	3.980*	16.95*	14.54*
		Weekly (N = 33)	0.638	4.309	4.59	5.98*
	Dec.–June	Daily (N = 163)	0.352*	3.937*	9.33*	7.94*
		Weekly (N = 33)	0.426	3.034	1.00	1.57
	Dec.–Mar.	Daily (N = 163)	0.476*	3.993*	12.85*	10.92*
		Weekly (N = 33)	0.624	3.073	2.15	2.96
Rate of Change in Futures Price Spread			**Skewness**	**Kurtosis**	**LM**	**K^2**
1988	Dec.–Dec.	Daily (N = 130)	−0.375	4.064*	9.17*	7.74*
		Weekly (N = 26)	0.119	1.905	1.36	2.97
	Dec.–June	Daily (N = 130)	−0.290	4.133*	8.77*	6.87*
		Weekly (N = 26)	0.185	1.951	1.34	2.60
	Dec.–Mar.	Daily (N = 130)	−0.116	3.653	2.61	2.77
		Weekly (N = 26)	0.308	2.580	0.60	0.58
1992	Dec.–Dec.	Daily (N = 163)	−0.416*	3.772*	8.75*	8.23*
		Weekly (N = 33)	−0.185	3.403	0.41	1.25
	Dec.–June	Daily (N = 163)	−0.045	3.845*	4.90*	3.97
		Weekly (N = 33)	−0.212	2.670	0.37	0.33
	Dec.–Mar.	Daily (N = 163)	−0.238	3.423	2.76	3.18
		Weekly (N = 33)	−0.405	2.670	1.05	1.16

* Indicates the null hypothesis of normal distribution is rejected at the 10% level of significance. The numbers in parentheses are the numbers of the observation.

tween stable and unstable periods. All three cases of corn DFPS converge to the normal distribution in the stable period (1992) as determined by the LM test, while only one case (Dec.–Dec.) converges in the unstable period (1988). Meanwhile, five corn RFPS out of six transformations reject the null hypothesis of a normal distribution for both daily and weekly intervals. The omnibus test confirms the LM test results except the case of 1992 Dec.–Dec. DFPS.

The normality tests on live cattle *fps* changes demonstrate a unique pattern of results between the DFPS and RFPS (Table 6). For DFPS, test statistics do not reject the null hypothesis of a normal distribution for any of the 24 tests across both intervals. However, live cattle *fps* do not necessarily converge to a normal distribution because the coefficients of LM and K^2 tests tend to increase with temporal aggregation. This result may stem from the fact that live cattle, as opposed to the other three commodities analyzed, is nonstorable, meaning there is less linkage between different futures contracts and fewer observations of zero, creating a more normal distribution of spread changes. Figure 1 clearly indicates these contrasting features, showing daily DFPS for T-bonds (small sample) and live cattle, the former rejecting normality and the latter failing to reject normality.

Meanwhile, all the distributions of RFPS for live cattle, except 1988 weekly Dec.–April, are positively skewed and fat tailed, resulting in nonnormality for both differencing intervals. Nevertheless, most of the coefficients of all the tests are reduced with temporal aggregation. Thus, both daily and weekly live cattle

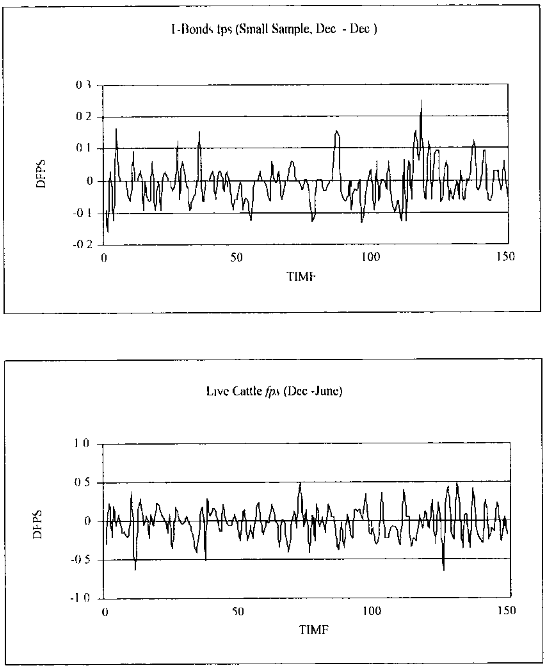

Figure 1. Changes in futures price spreads for T-bonds and live cattle, 1992

futures price spreads are normally distributed by DFPS results, but not lognormally by RFPS results.

The expected negative relationship between nonnormality and spread lengths is found in only two cases of corn and one case of live cattle *fps* changes in Tables 5 and 6. As the spread lengths are shortened, 1988 daily DFPS and weekly RFPS for corn and 1992 daily DFPS for live cattle appear more peaked and fat tailed as shown by the kurtosis tests. Interestingly, there is positive correlation with spread length for 1992 daily and weekly RFPS for corn and 1988 daily DFPS for live cattle.

Spread Volatility

Castelino and Vora studied spread volatility and found, as did Poitras, that *fps* volatility increased with spread length. This phenomenon is supported in this study. Table 8 shows the standard deviations for DFPS for corn, live cattle and the small sample size gold and T-bonds[11]. In all sixteen cases, volatility is smaller for shorter spread lengths. Note also that volatility decreases from the unstable period (1988 for gold, corn and live cattle: opposite of T-bonds) to the stable period (1992). This is as expected given the data selection procedure, and shows evidence of variance nonstationarity. Interestingly, corn spreads are far more volatile than the other three commodities in both periods. T-bond spreads are the most stable.

Goodness-of-Fit Tests

The χ^2 test of goodness-of-fit is performed to determine the best fit distribution of the changes in the four commodities' *fps* over time. The lowest valued chi-square test designates the most appropriate distribution among 25 functions built in the program "Bestfit" (Palisade Corp.).

Table 9 summarizes the best fit distributions of changes in daily *fps* and their χ^2 values with $n-1$ degree of freedom for each sample period and each *fps*[12]. The logistic distributions prevails for corn, gold and T-bonds, which are known to be more peaked and fat tailed than the normal distribution. This finding is consistent with the leptokurtosis, which is one reason that daily futures price spread changes are generally not normally distributed as reported above. For these three commodities, all one-year spread lengths are logistically distributed except the smaller sample size of T-bonds for 1988. In addition, seven cases of six- and seven-month spread lengths over three commodities are distributed logistically too. Meanwhile, for the shortest spread lengths, logistic distributions are accepted as best fit distributions only four times out of ten for these three commodities. In some cases, the correct distribution could not be determined at the 5% level, but could be at

[11] This test cannot be performed on the large sample size gold and T-bonds because the number of observations is not constant across spread lengths (see Tables 1 and 2). This test is not valid for RFPS due to its relative nature.

[12] Table 9 summarized only daily DFPS of each commodity.

Table 5. Distributional Test for Changes in Corn *fps*

Difference of Futures Price Spread			Skewness	Kurtosis	LM	K^2
1988	Dec.–Dec.	Daily (N = 130)	−0.339*	6.082*	53.95*	17.24*
		Weekly (N = 26)	−0.515	3.329	1.27	2.40
	Dec.–July	Daily (N = 130)	−1.142*	10.911*	367.22*	52.13*
		Weekly (N = 26)	−1.383*	6.244*	19.69*	16.09*
	Dec.–Mar.	Daily (N = 130)	−0.741*	11.387*	392.95*	42.57*
		Weekly (N = 26)	−0.915*	5.560*	10.72*	10.46*
1992	Dec.–Dec.	Daily (N = 162)	0.723*	6.144*	80.86*	29.97*
		Weekly (N = 33)	0.679	4.216	4.57	6.06*
	Dec.–July	Daily (N = 162)	−0.058	6.615*	17.69*	19.68*
		Weekly (N = 33)	0.492	3.499	1.67	2.89
	Dec.–Mar.	Daily (N = 162)	−0.216	3.452	2.64	3.03
		Weekly (N = 33)	−0.096	4.345	2.54	3.42
Rate of Change in Futures Price Spread			**Skewness**	**Kurtosis**	**LM**	**K^2**
1988	Dec.–Dec.	Daily (N = 130)	10.867*	121.86*	79084.0*	276.50*
		Weekly (N = 26)	1.333*	6.019*	17.58*	15.15*
	Dec.–July	Daily (N = 130)	−1.655*	13.590*	666.75*	72.34*
		Weekly (N = 26)	−0.602*	7.189*	20.58*	11.56*
	Dec.–Mar.	Daily (N = 130)	6.779*	72.567*	27187.0*	214.87*
		Weekly (N = 26)	−4.377*	21.359*	448.15*	59.39*
1992	Dec.–Dec.	Daily (N = 162)	−3.004*	23.772*	3156.1*	143.23*
		Weekly (N = 33)	−1.126*	9.347*	62.37*	21.04*
	Dec.–July	Daily (N = 162)	0.590*	8.159*	189.03*	35.18*
		Weekly (N = 33)	1.367*	6.643*	28.53*	18.57*
	Dec.–Mar.	Daily (N = 162)	−0.153	3.575	2.86	3.02
		Weekly (N = 33)	0.302	4.020	1.93	3.17

* Indicates the null hypothesis of normal distribution is rejected at the 10% level of significance. The numbers in parentheses are the numbers of the observation.

10%. For example, the distributions of gold with the larger sample size for two months' spread (Dec.–Feb.) are found as triangular and logistic at the latter level. Thus, none of 25 distributions built in the program is appropriate for some spreads at the 5% confidence levels. For corn's three-month spread (Dec.–Mar.), logistic and normal distributions are detected as the best, consistent with Table 5.

For live cattle DFPS, the normal distribution, which was not rejected in the previous LM distributional test, is found as the best four times in the goodness-of-fit test. The logistic distribution is detected as the best for the remaining two spreads. However, the normal distribution cannot be ignored because it is found as the second best with significant test values.

In general, the logistic distribution prevails as the best fit for commodity spreads from these goodness-of-fit tests. It may not be the absolutely correct distribution of these spreads; however, it appears to describe the distributions examined here better than other distributions.

Summary

Two types of tests have been performed to detect normality and confirm the most appropriate distributions on corn and live cattle futures price spread changes in one sample size, and gold and T-bonds futures price spread changes with two different sample sizes. These are examined for selected stable and unstable periods.

The distributional behavior has been examined by conducting skewness, kurtosis, LM, K^2 and standard deviation tests, using the normal distribution as the null hypothesis. In-

Table 6. Distributional Test for Changes in Live Cattle *fps*

Difference of Futures Price Spread			Skewness	Kurtosis	LM	K^2
1988	Dec.–June	Daily (N = 130)	0.206	3.455	2.04	2.53
		Weekly (N = 26)	−0.05	2.384	0.42	0.22
	Dec.–April	Daily (N = 130)	−0.032	2.967	0.03	0.07
		Weekly (N = 26)	−0.180	2.239	0.77	0.79
	Dec.–Feb.	Daily (N = 130)	−0.048	2.849	0.17	0.07
		Weekly (N = 26)	−0.208	2.564	0.39	0.28
1992	Dec.–June	Daily (N = 151)	−0.042	2.988	0.05	0.11
		Weekly (N = 31)	−0.267	2.717	0.47	0.50
	Dec.–April	Daily (N = 164)	0.104	3.133	0.42	0.69
		Weekly (N = 33)	−0.332	2.268	1.34	1.58
	Dec.–Feb.	Daily (N = 164)	0.112	3.410	1.49	1.86
		Weekly (N = 33)	0.020	3.317	0.14	0.81
Rate of Change in Futures Price Spread			**Skewness**	**Kurtosis**	**LM**	**K^2**
1988	Dec.–June	Daily (N = 130)	−0.866*	20.436*	1663.1*	60.89
		Weekly (N = 26)	−2.295*	8.218*	52.33*	28.44
	Dec.–April	Daily (N = 130)	0.563*	3.539	8.43*	8.65
		Weekly (N = 26)	0.538	2.846	1.28	1.78
	Dec.–Feb.	Daily (N = 130)	−0.278	19.243*	1430.8*	47.11*
		Weekly (N = 26)	−1.912*	6.906*	32.38*	22.44*
1992	Dec.–June	Daily (N = 151)	0.977*	6.327*	93.66*	36.77*
		Weekly (N = 31)	1.186*	6.101*	19.68*	15.05*
	Dec.–April	Daily (N = 164)	7.439*	86.229*	48847.0*	272.25*
		Weekly (N = 33)	4.175*	21.150*	548.86*	62.79*
	Dec.–Feb.	Daily (N = 164)	1.281*	10.158*	394.93*	64.70*
		Weekly (N = 33)	1.285*	4.986*	14.51*	13.84*

* Indicates the null hypothesis of normal distribution is rejected at the 10% level of significance. The numbers in parentheses are the numbers of the observation.

Table 7. Summary: The Number of Times Normal Distribution Not Rejected

	1988				1992			
	Difference of Futures Price Spread		Rate of Change in Futures Price Spread		Difference of Futures Price Spread		Rate of Change in Futures Price Spread	
Commodity	Daily	Weekly	Daily	Weekly	Daily	Weekly	Daily	Weekly
Gold (large)	0	0	0	0	0	1	0	1
T-Bonds (large)	0	0	0	0	0	2	0	1
Gold (small)	0	3	0	3	0	2	0	1
T-Bonds (small)	3	3	1	3	0	3(2)*	1(2)*	3
Corn	0	1	0	0	1	3(2)*	1	1
Live Cattle	3	3	0	1	3	3	0	0

* K^2 result in parenthesis differs from LM result.

Table 8. Standard Deviation for DFPS*

		1988		1992	
		Daily	Weekly	Daily	Weekly
Gold	Dec.–Dec.	0.45	0.90	0.24	0.53
	Dec.–June	0.23	0.48	0.15	0.30
	Dec.–Feb.	0.11	0.20	0.07	0.10
T-Bonds	Dec.–Dec	0.06	0.14	0.07	0.16
	Dec.–June	0.04	0.07	0.03	0.08
	Dec.–Mar.	0.02	0.04	0.02	0.04
Corn	Dec.–Dec.	5.19	15.36	1.81	4.23
	Dec.–July	1.99	5.93	0.69	1.76
	Dec.–Mar.	0.85	1.98	0.36	0.64
Live Cattle	Dec.–June	0.41	0.81	0.21	0.37
	Dec.–April	0.35	0.69	0.19	0.31
	Dec.–Feb.	0.24	0.50	0.15	0.26

* See Tables 3–6 for sample sizes. Gold and T-bonds are from the small sample sizes.

creasing the difference interval from daily to weekly is expected to remove zeros and small changes in the data, which then produces a more normal distribution. Quite disparate results are found. Gold and T-bonds in the larger sample size did not produce more normal distributions with temporal aggregation. By contrast, many of the smaller-sized samples of gold and T-bonds became normally distributed for weekly differencing intervals. For agricultural commodities, a discrepancy was found between DFPS and RFPS. In the case of DFPS, there was a trend toward a normal distribution for corn at larger differencing intervals and there exists a normal distribution for both daily and weekly intervals for live cattle. On the other hand, nonnormal distributions dominate for both commodities' daily and weekly RFPS. Clearly, however, for all the data examined, more weekly intervals are normally distributed than daily intervals, as expected.

For the larger sample size of gold and T-bonds, the combination of negative skewness and fat tails was the main reason for nonnormality. Meanwhile, leptokurtosis alone as well as the combination of skewness and fat tails created nonnormality for the smaller sample size of gold and T-bonds and for corn and live cattle. The nonnormality of distributions leads to the question of the correct distribution. This is of interest to spread traders so they can manage risks more efficiently and make informed trading decisions. Most often, the logistic distribution was the best-fit distribution for changes in daily DFPS. Leptokurtosis, which was a main reason causing nonnormality, was confirmed through detection of a logistic distribution, generally known as more peaked and fat tailed than the normal distribution.

Two spread length effects were also examined. The negative correlation expected between spread length and nonnormality was only partially supported for the larger sample size of gold and T-bonds and for corn and live cattle, but not for the smaller sample size of gold. However, reduced volatility as the spread length was shortened was found in all cases, consistent with previous research.

Futures spreads exhibit variance nonstationarity, and optimal spread positions depend on traders' risk attitudes and their subjective estimates of statistical parameters. Daily spreads changes, especially for storable commodities, exhibit a high probability of being zero, but also skewed with fat tails. Meanwhile, spreads held for a week are more likely to be normally distributed. Hence, traders holding spread positions for a week will typically find it comparatively easier to assess and evaluate the relative risk and return of their positions. It is not unexpected that spreads for storable products would exhibit a large num-

Table 9. Best Fit Distributions (Daily Spreads)

		Distribution	χ^2	Distribution	χ^2	Distribution	χ^2
		Dec.–Dec.		Dec.–July		Dec.–March	
Corn	1988	Logistic	18.643	Student's t	23.624	Logistic	48.930
	1992	Logistic	19.763	Logistic	10.491	Normal	53.365
		Dec.–June		Dec.–April		Dec.–Feb.	
Live Cattle	1988	Logistic**	9.390	Normal	2.473	Normal	12.369
	1992	Normal	12.626	Normal	20.431	Logistic**	6.281
		Dec.–Dec.		Dec.–June		Dec.–Feb.	
Gold (Large Size)	1988	Logistic	51.803	Student's t	326.87	Triangular	555.9*
	1992	Logistic	78.339	Logistic	209.31	Logistic	368.8*
		Dec.–Dec.		Dec.–June		Dec.–Feb.	
Gold (Small Size)	1988	Logistic	35.787	Logistic	39.786	Loglogistic	22.932
	1992	Logistic	80.864	Logistic	58.558	Logistic	48.764
		Dec.–Dec.		Dec.–June		Dec.–March	
T-Bonds (Large Size)	1988	Logistic	51.452	Logistic	119.20	Logistic	66.402
	1992	Logistic	119.49	Logistic	62.962	Logistic	47.260
		Dec.–Dec.		Dec.–June		Dec.–March	
T-Bonds (Small Size)	1988	Triangular*	4.247	Normal	28.146	Normal	260.2*
	1992	Logistic	28.449	Logistic	101.33	Logistic	322.6*

* Indicates the null hypothesis of the best fit distribution is rejected at 10% level of significance.
** Indicates that normal distribution is the second best fit with the coefficient of 13.219 for 1988 live cattle, 7.487 for 1992 live cattle, and 9.005 for 1988 T-bonds respectively, which can not be rejected at the 10% level of significance.

ber of zeros for daily changes since expected storage costs would not change in the short run. However, as time increases, the market absorbs and interprets more information, demonstrating more normally distributed spread changes. Futures contracts for nonstorable commodities do not have the same theoretical linkage as storables, hence there is no reason to expect an unusually large number of zero spread changes on a daily basis, as shown here.

A prime function of spread trading, besides contributing to market liquidity, is to allow commercial traders to more effectively hedge in distant contracts, and to permit those traders absorbing hedgers' risk to share it effectively among other speculators. These results do not dispell the notion that the market effectively and efficiently accommodates this function, only that one statistical distribution with predictable parameters does not fit all cases. There is a real mixture of results, meaning traders need to subjectively assess each spread situation individually for trading and pricing opportunities. One clear result is that the greater the length between spreads, the larger the volatility and trading risk, as expected.

Overall, normal and logistic distributions dominate the changes in futures spreads examined here, but results are clearly sensitive to commodity, sample period, sample size, spread length, differencing interval, and spread definition. Spread traders can expect to find normal distributions more often with weekly intervals than daily, and with nonstorable commodities than storable. Whether the data are relatively stable or unstable does not consistently impact the results. Hence, as traders search for the probability distributions of futures price spreads, each spread is likely to have its own unique characteristics, making it difficult for traders to generalize or find common patterns.

References

Bera, A. and C. M. Jarque. "A Test for Normality of Observations and Regression Residuals." *In-*

ternational Statistical Review 55 (1987):163–172.

Billingsley, R. S. and D. M. Chance. "The Pricing and Performance of Stock Index Futures Spreads." *Journal of Futures Markets* 8 (1988): 303–318.

Blattberg, R. C. and N. Gonedes. "A Comparison of the Stable Paretian Distribution and Student Distribution as Statistical Model for Prices." *Journal of Business* 47 (1984):244–280.

Bowman, K. O. and L. R. Shenton. "Omnibus Test Contours for Departures from Normality Based on $\sqrt{b_1}$ and b_2." *Biometrika* 62 (1975):243–250.

Castelino, M. G. and A. Vora. "Spread Volatility in Commodity Futures: The Length Effect." *Journal of Futures Markets* 4 (1984):39–46.

Cornew, R. W., D. E. Town, and L. D. Crowson. "Stable Distributions, Futures Price, and the Measurement of Trading Performance." *Journal of Futures Markets* 4 (1984):531–557.

D'Agostino, R. B., A. Belanger, and R. B. D'Agostino Jr. "A Suggestion for Using Powerful and Informative Tests of Normality." *American Statistician* 44 (1990):316–321.

D'Agostino, R. B. and E. S. Pearson. "Tests for Departure from Normality. Empirical Results for the Distributions of b_2 and $\sqrt{b_1}$." *Biometrika* 60 (1973):613–622.

Davidson, R. and J. G. Mackinnon. *Estimation and Inference in Econometrics.* New York: Oxford University Press, 1993.

DeGroot, M. H. *Probability and Statistics,* 2nd Edition. Reading: Addison-Wesley, 1989.

Hall, J. A., B. W. Brorsen, and S. H. Irwin. "The Distribution of Futures Prices: A Test of the Stable Paretian and Mixture of Normals Hypothesis." *Journal of Financial and Quantitative Analysis* 24 (1989):105–116.

Houthakker, H. S. "Systematic and Random Element in Short-term Price Movement." *American Economic Review* 51 (1961):164–172.

Hsu, D. A., R. B. Miller, and D. W. Wichern. "On the Stable Paretian Behavior of Stock Market Prices." *Journal of American Statistical Association* 69 (1974):108–113.

Hudson, M. A., R. M. Leuthold, and G. F. Sarassoro. "Commodity Futures Price Changes: Recent Evidence for Wheat, Soybeans and Live Cattle." *Journal of Futures Markets* 7 (1987): 287–301.

Mann, J. S. and R. G. Heifner. *The Distribution of Shortrun Commodity Price Movements.* Economic Research Service, USDA, Technical Bulletin No. 1536, 1976.

Monroe, M. and R. A. Cohn. "The Relative Efficiency of the Gold and Treasury Bill Futures Markets." *Journal of Futures Markets* 6 (1986): 477–493.

Officer, R. R. "The Distribution of Stock Returns." *Journal of the American Statistical Association* 67 (1972):807–812.

Palisade Corporation Editors. *Bestfit, Release 1.02,* Palisade Corporation, 1994.

Pearson, E. S., R. B. D'Agostino, and K. O. Bowman. "Tests for Departure from Normality: Comparison of Powers." *Biometrika* 64 (1977): 231–246.

Poitras, G. N. "The Distribution of Gold Futures Spreads." *Journal of Futures Markets* 10 (1990):643–659.

Ramanathan, R. (1993). *Statistical Methods in Econometrics.* San Diego: Academic Press, Inc, 1993.