MEAN REVERSION AND THE COMOVEMENT OF EQUILIBRIUM SPOT AND FUTURES PRICES: IMPLICATIONS FROM ALTERNATIVE DATA-GENERATING PROCESSES

TIAN ZENG

The comovements of spot and futures prices are characterized by six binary variables, including the term structure curvature of futures prices. These variables are used to uniquely identify 48 possible comovement patterns. Among them, 24 cases are associated with *mean reversion*, which is defined as a state when spreads between futures and spot prices are shrinking. These pattern frequencies are then calculated on a daily basis with the futures prices of 10 commodities, including precious metal, agricultural, and financial commodities. The results are further compared to simulation output from three data-generating processes: a bivariate pure random walk, a mixed random walk with first-order autoregression (AR(1)), and an error-correction representation. The mean-reverting frequencies for all 10 commodities are about 50%. Around half of the time,

The author is grateful to Norman Swanson, Ahmet Kocagil, and two anonymous referees for their helpful suggestions. All errors remain the responsibility of the author.
For correspondence, Tian Zeng, 10 State House Square, 9th Floor, Hartford, Connecticut 06103-1205; e-mail: zeng_tian@yahoo.com

Received January 2000; Accepted January 2001

■ *Tian Zeng is an Analyst at Aeltus Investment Management Inc. in Hartford, Connecticut.*

The Journal of Futures Markets, Vol. 21, No. 8, 769–796 (2001)

spot and futures prices are moving toward each other, and the rest of the time they move in the same direction. The symmetry of these results implies that the existence of substantial shocks originated from futures markets; thus, this is consistent with the risk premium view of futures trading. Also, although all simulation models produce similar mean-reversion frequencies, the patterns of comovements of spot and futures prices are different, and the price dynamics depend heavily on whether the market is dominant contango or backwardation. Furthermore, the error-correction model outperforms the random-walk model for agricultural commodities, and the mixed random walk with AR(1) is hardly distinguishable from the pure random walk. Jrl Fut Mark 21:769–796, 2001

INTRODUCTION

Many studies have shown that financial asset prices can be characterized as mean-reverting. For example, Bessembinder, Coughenour, Seguin, and Smoller (1995) found negative correlations in agricultural, financial, and precious metal prices. Fama and French (1998b) identified a predictable component of stock prices and a U-shaped autocorrelation pattern in stock returns. More evidence can also be found in Lo and MacKinlay (1988); Fama and French (1988a); Poterba and Summers (1988); Cecchetti, Lam, and Mark (1990); and Bonomo and Garcia (1994).

Mean reversion is a property of a stochastic process where the variable value tends to revert back to some normal level when it is away from that level. This normal level is often called the *mean* or trend of the variable underlying the stochastic process (for further discussion, see Dixit & Pindyck, 1994). A simple example of a mean-reversion process is

$$p_t - p_{t-1} = \phi(p - p_{t-1}) + \sigma e_t \tag{1}$$

where p_t is the asset price and e_t is white noise with zero mean and unit variance. The parameter ϕ is positive and measures the speed of reversion, and σ measures volatility. Note that p_t is increasing (decreasing) whenever p_{t-1} is smaller (larger) than p and the innovation term e_t is relatively small. Therefore, p_t tends to move toward the trend p over time. It also implies a negative autocorrelation of p_t.

Using such a process to model asset price is appealing because the trade-off between risk (σ) and return ($p - p_0$) is explicitly incorporated. In contrast to a random-walk specification (when $\phi = 0$, Equation 1 becomes a random walk without drift), the mean-reversion process is less volatile because it keeps pulling back toward the mean level, whereas prices in random walks can drift away explosively. However, the price from a mean-reverting process also has a smaller chance of reaching higher levels than the price from a random walk, implying smaller returns.

The difficulty of using a process such as Equation 1 is that the long-run mean p is hard to determine. The literature often focuses on autocorrelation only, at the expense of much richer implications on price movement from a mean-reversion process. In this article, I use the term structure of futures prices to define the mean reversion, following Bessembinder et al. (1995) and Kocagil, Swanson, and Zeng (in press), under the assumption that the futures price shares the same long-run equilibrium price with the spot price. In particular, mean reversion is found only if spot and futures price are moving toward each other or chasing each other, which leads to basis shrinkage.

This article also contributes to the comovement dynamics of spot and futures prices. In particular, the price movement patterns are uniquely defined by six variables, which are used to capture the market type (contango or backwardation), market status (bull or bear), price elasticity, and curvature of price path. As a result, the frequency of mean reversion can be calculated each day. Three data-generating processes (DGPs) are simulated to further compare the actual price movement for 10 commodities. The DGPs include a bivariate pure random walk, a mixed random walk with first-order autoregression (AR(1)), and an error-correction representation. The mean-reverting frequencies for the 10 commodities are about 50%. Around half of the time, the spot and futures prices are moving toward each other, and the rest of the time they move in the same direction. The symmetry of these results implies that the existence of substantial shocks originated from futures markets; thus, this is consistent with the risk premium view of futures trading. The simulation results also indicate that price dynamics depend on whether the commodity markets investigated are dominant contango or backwardation. Furthermore, the error-correction model is shown to outperform the random-walk model for agricultural commodities in terms of matching the price movement patterns.

The article is organized as follows. In the next section, the mean reversion is defined as a state where the spread between futures and spot prices is shrinking. In the third section, a binary vector is developed to uniquely identify the all 48 patterns of price comovement. In the fourth section, three bivariate simulation models and real data are introduced. The fifth section is a summary of the empirical results, and the sixth section provides a conclusion.

MEAN REVERSION AS BASIS SHRINKAGE

The theory of storage developed by Kaldor (1939), Working (1949), Brennan (1958), and Telser (1958), as well as many others, stipulates an equilibrium contemporaneous relationship between spot price and futures price. Let F_t be the futures price at date t and P_t be the spot price at the same time. The cost-of-carry relationship can be stated as

$$F_t = P_t e^{s_t} \tag{2}$$

where s_t is the continuously compounded cost-of-carry rate per period for the remaining maturity of the futures contract multiplied by the time to maturity. More specifically, s_t is a function of the storage cost (+), interest rate (+), time to maturity (+), and the convenience yield (−). If there is no transaction cost and the daily settlement and deliver option features are abstracted from the futures contract, Equation 2 holds with a no-arbitrage condition. Whenever Equation 2 is violated, for example, if F_t is overpriced, an investor can take a short position in the futures contract while buying the underlying commodity at the spot price and holding the commodity until maturity, paying the net storage cost. When the futures contract expires, he or she can deliver the commodity and lock in the risk-free profit. The validity of cost-of-carry equilibrium is a hypothesis maintained in this article.

Equation 2 can be rewritten in the following log forms:

$$f_t = p_t + s_t \tag{3}$$

where $p_t = \log(P_t)$, $f_t = \log(F_t)$, and $s_t = f_t - p_t$. As a measure of the spread between the futures and spot prices, s_t is defined as the basis.

Shrinking basis is indicative of the mean reversion of the anticipated spot price. The cost-of-carry relationship states that the futures price can be taken as the expected future spot price. It is, however, a noisy one; see, for example, French (1986) and Fama and French (1987). Because f_t and p_t are both nonstationary series and s_t is stationary (augmented Dickey–Fuller tests were conducted, and results are available on request; see Dickey & Fuller, 1981) and on the basis of Equation 3, f_t and p_t share the same trend. This is also supported by the fact that the futures price, f_t, converges to the spot price, p_t, near the expiration date of the futures contract. In practice, the dealers in a pure dealer market often simply move up their quotes in the spot price on observing the price jump in the futures. The specialists in exchanges do the same thing, except at probably a slower pace due to the restriction that successive movements have to be within the regulatory limits (see Miller, Muthuswamy, & Whaley, 1994). If spot and futures prices do have the same mean (or trend), shrinking basis implies that spot prices move closer to the mean:

Definition: The spot price is mean-reverting if and only if

$$ds_t/s_t \leq 0 \tag{4}$$

where d is the difference operator.

Several comments are in order.

First, the definition equates mean reversion with the case when the absolute value of basis becomes smaller. When $s_t > 0$ (i.e., contango market, $F_t > P_t$), mean reversion is equivalent to $ds_t < 0$, implying the market anticipates that the futures premium is going to be smaller. However, when $s_t < 0$ (i.e., backwardation market, $F_t < P_t$), mean reversion implies $d(-s_t) < 0$, or that the futures discount is becoming smaller. In summary, the mean reversion is a state when spot and futures prices are moving closer to each other.

Second, the mean reversion is based on the characterization of comovement between futures and spot prices. For example, $ds_t < 0$ is the same as $df_t - dp_t < 0$, or $d(F_t)/F_t - d(P_t)/P_t < 0$, implying the futures price return is less than the spot price returns, or the futures log prices increment is smaller. Generally, the comovement when there is a mean reversion can be classified into the following three categories, as illustrated in Figure 1:

1. F_t and P_t are moving toward each other.
2. F_t and P_t are moving in the same direction, but the futures price moves at a slower pace, and the spot price is chasing the futures price.
3. F_t and P_t are moving in the same direction, but the futures price moves at a faster pace, and the futures price is following the spot price.

Category 2 is an interesting case, where the rate of intertemporal price appreciation (depreciation) decreases as the spot price rises (declines). This is the case used by Bessembinder et al. (1995) to test a sufficient condition of mean reversion. In Kocagil et al. (in press), the mean reversion is extended further to include both Categories 1 and 2. Also, the relative frequency of Category 3 versus Categories 1 and 2 could provide a test to see if shocks from the futures market would be as frequent as shocks from the spot market. Such an assumption is often implicitly assumed in some earlier studies (see Fama & French, 1988a).

Third, the definition is a measure of a certain phase of a mean-reversion process. If a time series has a zero mean, the mean reversion is identified only if the series moves down from above zero or moves up

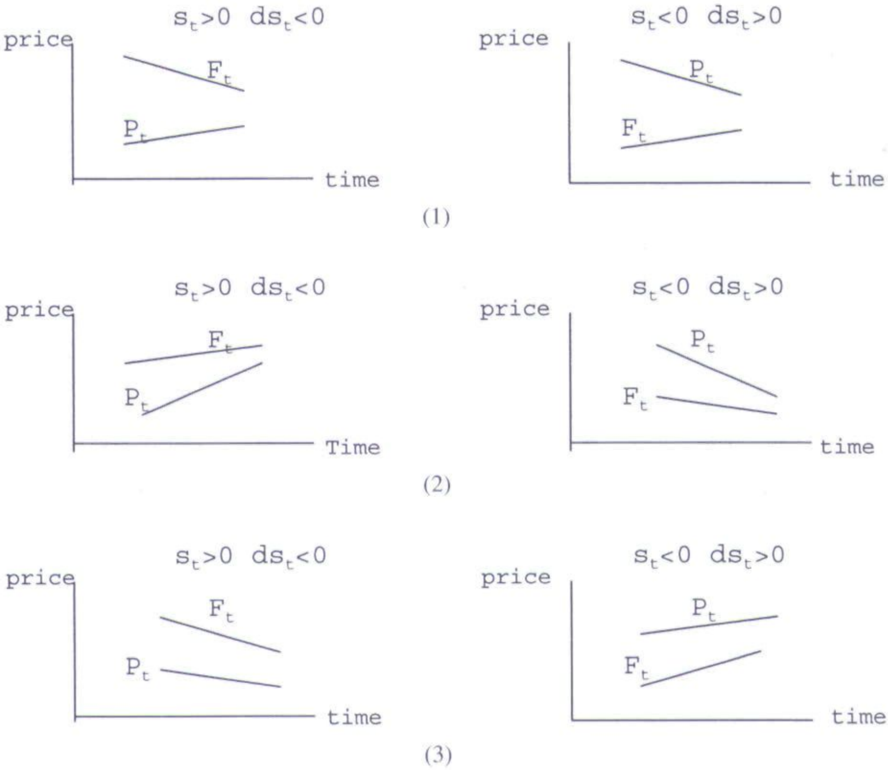

FIGURE 1

Categorical definition of the mean reversion. In Category 1, futures and spot prices move toward each other; in Category 2, spot prices chase futures prices; and in Category 3, futures prices follow spot prices.

from below zero. In this sense, the definition is also a definition of mean reversion of basis.

Finally, a perhaps more important feature of the definition is that it permits the dynamics of mean reversion, where mean reversion can be found in the current sample period but is absent in the next. A stochastic process that fits best in a certain sample period of the data may well be conditioned on some economic conditions that are not present in other periods. Its underlying marginal distribution and the statistical property, including the mean reversion, could be periodically changed as a result of changes in the economic environment. Although the cause of economic sources for mean reversion is not directly addressed in this study, more adaptive concepts such as the aforementioned definition could potentially better facilitate our learning of the price dynamics.

Defining the mean reversion of the asset price as basis shrinkage would not have been possible if there were no futures market for that asset. To fully take advantage of this approach, I develop a six-dimensional vector with binary components in the next section to capture patterns of price comovement, which uniquely links to all the market possibilities. As a result, the frequencies of defined mean reversion can be analyzed in much more detail. Actually, each of the three categories is further classified into 8 different cases, and a total of 24 mean-reversion cases are shown to be among 48 comovement possibilities.

PATTERNS OF SPOT AND FUTURES PRICE COMOVEMENT

A vector with six binary components, symbolized as $\mathbf{ID}_t$ at date t, is defined as

$$\mathbf{ID}_t = [\mathrm{sign}(s_t), \mathrm{sign}(dp_t), \mathrm{sign}(df_t/dp_t), \mathrm{sign}(ds_t), \mathrm{sign}(\delta_{pt}), \mathrm{sign}(\delta_{ft})]$$

where sign (*) is the sign function, which equals one when the argument is positive and has a value of zero otherwise. Also, $\delta_{pt} = d^2p_t/dp_t$, and $\delta_{ft} = d^2f_t/df_t$.

The first component [sign(s_t)] of $\mathbf{ID}_t$ determines whether a commodity market is a contango market ($F_t > P_t$), or a backwardation market ($F_t < P_t$). The second component [sign(dp_t)] ascertains a bull (bear) spot market. When sign (df_t/dp_t) = 1 (Component 3), the futures price and spot price are moving in the same direction. A zero value of Component 3 implies the two prices are moving in opposite ways. Notice that combining Components 2 and 3 can determine if the futures market is a bull or bear market. Component 4 [sign(ds_t)] compares the directional difference between spot and futures log prices. For example, a value of one means either $ds_t = df_t - dp_t > 0$ or $df_t > dp_t$. The signs of df_t and dp_t are important. If both prices go up, a value of one means the futures price increases faster. If both prices go down, a value of one means the futures price decreases at a slower pace. However, if both prices are moving in opposite directions, a value of one for Component 3 implies that the future price is up and the spot price is down.

Component 4 in the $\mathbf{ID}_t$ vector also measures the price elasticity of the futures price with respect to the spot price. The cost-of-carry relationship (Equation 3) can be transformed into the following: $df_t/dp_t = 1 + ds_t/dp_t$. In other words, $ds_t/dp_t < 0$ if and only if $df_t/dp_t < 1$. Therefore, a value of zero is equivalent to the case where the elasticity of the futures price with respect to the spot price is less than one. More intuitively, this corresponds to one scenario of mean reversion, where the rate of intertemporal price appreciation (depreciation) decreases as the spot price rises (declines).

Components 5 [sign(δ_{pt})] and 6 [sign (δ_{ft})] measure the growth rate of spot and futures price returns, respectively. In combination with Component 2, signs of d^2p_t and d^2f_t can be established, indicating the curvature of the price movement. In particular, a positive sign of d^2p_t (d^2f_t) means $p_t(f_t)$ moves in an accelerating convex path over time, whereas a negative sign implies a concave, decelerating path.

The first three components of $\mathbf{ID}_t$ are more directional measures, whereas the last three are magnitude measures. Also, Components 2, 5, and 6 are measures of individual markets (spot or futures), whereas Components 1, 3, and 4 link spot and futures prices. Notice that Component 3 is the elasticity of the futures price with respect to the spot price for $df_t/dp_t = d(\log F_t)/d(\log P_t) = [d(F_t)/F_t]/[d(P_t)/P_t]$, but the measurement differs from Component 4 in that the concern is the direction of the price movement. It can also be taken as the price elasticity with a cutoff threshold of zero, whereas Component 4 is the price elasticity with a threshold of one. To put it differently, a combination of (0,0) of Components 3 and 4 implies that the price elasticity is less than zero; a combination of (1,1) implies that the price elasticity is larger than one; and a combination of (1,0) means the price elasticity is within [0,1]. The price elasticity, when less than one, is indicative of mean reversion. When both futures and spot prices are rising, for example, it implies that intertemporal price appreciation is decreasing (for further discussion, see Bessembinder et al., 1995; Kocagil et al., in press). The elasticity measure is important for other reasons as well. For example, Fama and French (1988a) argued that the futures price adjusts less than the spot price movement when inventory is low (convenience yield is high), whereas the futures price adjusts in approximately the same magnitude following spot price changes when inventory is high (convenience yield is low). The proposition is also related to Samuelson's argument that the futures price varies less than the spot price and the variation of the futures price is a decreasing function of maturity.

The vector $\mathbf{ID}_t$ can take on a total number of $2^6 = 64$ values. However, the appropriate explanation of Component 4 requires information from the previous three components. Actually, a zero value of Component 3, which means the price elasticity is less than zero, could not coexist with a value of one for Component 4 of $\mathbf{ID}_{t,}$ which implies the same elasticity is larger than one. Because there are four other components with binary possibilities, the overall number of incompatible $\mathbf{ID}_t$ cases is $2^4 = 16$. If we subtract this from 64, the number of compatible cases is 48.

In Table I, all 48 cases along with the mean-reversion categories are listed and numbered according to the number and order of value

TABLE I

Classification of the Mean Reversion and Price Comovement

Case 1	$\mathbf{ID}_t = (0, 0, 0, \; 0, 0, 0)$	Mean reversion (category 1)
Case 2	$\mathbf{ID}_t = (0, 0, 0, \; 0, 1, 0)$	Mean reversion (category 1)
Case 3	$\mathbf{ID}_t = (0, 0, 0, \; 0, 0, 1)$	Mean reversion (category 1)
Case 4	$\mathbf{ID}_t = (0, 0, 0, \; 0, 1, 1)$	Mean reversion (category 1)
Case 5	$\mathbf{ID}_t = (1, 0, 0, \; 0, 0, 0)$	No mean reversion
Case 6	$\mathbf{ID}_t = (1, 0, 0, \; 0, 1, 0)$	No mean reversion
Case 7	$\mathbf{ID}_t = (1, 0, 0, \; 0, 0, 1)$	No mean reversion
Case 8	$\mathbf{ID}_t = (1, 0, 0, \; 0, 1, 1)$	No mean reversion
Case 9	$\mathbf{ID}_t = (0, 1, 0, \; 0, 0, 0)$	No mean reversion
Case 10	$\mathbf{ID}_t = (0, 1, 0, \; 0, 1, 0)$	No mean reversion
Case 11	$\mathbf{ID}_t = (0, 1, 0, \; 0, 0, 1)$	No mean reversion
Case 12	$\mathbf{ID}_t = (0, 1, 0, \; 0, 1, 1)$	No mean reversion
Case 13	$\mathbf{ID}_t = (0, 0, 1, \; 0, 0, 0)$	No mean reversion
Case 14	$\mathbf{ID}_t = (0, 0, 1, \; 0, 1, 0)$	No mean reversion
Case 15	$\mathbf{ID}_t = (0, 0, 1, \; 0, 0, 1)$	No mean reversion
Case 16	$\mathbf{ID}_t = (0, 0, 1, \; 0, 1, 1)$	No mean reversion
Case 17	$\mathbf{ID}_t = (0, 0, 1, \; 1, 0, 0)$	Mean reversion (category 2)
Case 18	$\mathbf{ID}_t = (0, 0, 1, \; 1, 1, 0)$	Mean reversion (category 2)
Case 19	$\mathbf{ID}_t = (0, 0, 1, \; 1, 0, 1)$	Mean reversion (category 2)
Case 20	$\mathbf{ID}_t = (0, 0, 1, \; 1, 1, 1)$	Mean reversion (category 2)
Case 21	$\mathbf{ID}_t = (1, 1, 0, \; 0, 0, 0)$	Mean reversion (category 1)
Case 22	$\mathbf{ID}_t = (1, 1, 0, \; 0, 1, 0)$	Mean reversion (category 1)
Case 23	$\mathbf{ID}_t = (1, 1, 0, \; 0, 0, 1)$	Mean reversion (category 1)
Case 24	$\mathbf{ID}_t = (1, 1, 0, \; 0, 1, 1)$	Mean reversion (category 1)
Case 25	$\mathbf{ID}_t = (0, 1, 1, \; 0, 0, 0)$	No mean reversion
Case 26	$\mathbf{ID}_t = (0, 1, 1, \; 0, 1, 0)$	No mean reversion
Case 27	$\mathbf{ID}_t = (0, 1, 1, \; 0, 0, 1)$	No mean reversion
Case 28	$\mathbf{ID}_t = (0, 1, 1, \; 0, 1, 1)$	No mean reversion
Case 29	$\mathbf{ID}_t = (0, 1, 1, \; 1, 0, 0)$	Mean reversion (category 3)
Case 30	$\mathbf{ID}_t = (0, 1, 1, \; 1, 1, 0)$	Mean reversion (category 3)
Case 31	$\mathbf{ID}_t = (0, 1, 1, \; 1, 0, 1)$	Mean reversion (category 3)
Case 32	$\mathbf{ID}_t = (0, 1, 1, \; 1, 1, 1)$	Mean reversion (category 3)
Case 33	$\mathbf{ID}_t = (1, 0, 1, \; 0, 0, 0)$	Mean reversion (category 3)
Case 34	$\mathbf{ID}_t = (1, 0, 1, \; 0, 1, 0)$	Mean reversion (category 3)
Case 35	$\mathbf{ID}_t = (1, 0, 1, \; 0, 0, 1)$	Mean reversion (category 3)
Case 36	$\mathbf{ID}_t = (1, 0, 1, \; 0, 1, 1)$	Mean reversion (category 3)
Case 37	$\mathbf{ID}_t = (1, 0, 1, \; 1, 0, 0)$	No mean reversion
Case 38	$\mathbf{ID}_t = (1, 0, 1, \; 1, 1, 0)$	No mean reversion
Case 39	$\mathbf{ID}_t = (1, 0, 1, \; 1, 0, 1)$	No mean reversion
Case 40	$\mathbf{ID}_t = (1, 0, 1, \; 1, 1, 1)$	No mean reversion
Case 41	$\mathbf{ID}_t = (1, 1, 1, \; 0, 0, 0)$	Mean reversion (category 2)
Case 42	$\mathbf{ID}_t = (1, 1, 1, \; 0, 1, 0)$	Mean reversion (category 2)
Case 43	$\mathbf{ID}_t = (1, 1, 1, \; 0, 0, 1)$	Mean reversion (category 2)
Case 44	$\mathbf{ID}_t = (1, 1, 1, \; 0, 1, 1)$	Mean reversion (category 2)
Case 45	$\mathbf{ID}_t = (1, 1, 1, \; 1, 0, 0)$	No mean reversion
Case 46	$\mathbf{ID}_t = (1, 1, 1, \; 1, 1, 0)$	No mean reversion
Case 47	$\mathbf{ID}_t = (1, 1, 1, \; 1, 0, 1)$	No mean reversion
Case 48	$\mathbf{ID}_t = (1, 1, 1, \; 1, 1, 1)$	No mean reversion

Note. The 48 cases are uniquely defined by $\mathbf{ID}_t$ vectors whose components take the value of 1 if the correspondent component of the vector (s_t, dp_t, df_t/dp_t, ds_t/dp_t, δ_{pt}, and δ_{ft}) is positive or the value of 0 otherwise. The last column reports if the correspondent cases are mean-reverting by the definition of $ds_t/s_t < 0$, where the category in the parentheses is numbered as 1>. Spot and futures prices are moving toward each other and 2>. Spot prices are catching up with futures prices and 3>. Futures prices are catching up with spot prices.

"1" appearing in the $\mathbf{ID}_t$ vectors. To illustrate how to relate to the pattern of price comovement from $\mathbf{ID}_t$, let us look closely at Case 32, where $\mathbf{ID}_t = (0,1,1,1,1,1)$. First, because $s_t < 0$ (Component 1), we know the market is a backwardation market, or $f_t < p_t$. Second, the spot price is going up because $dp_t > 0$ (Component 2). Adding $df_t/dp_t > 0$ (Component 3), we can conclude $df_t > 0$, or the futures price is going up as well. Furthermore, by $ds_t > 0$ (Component 4), it is clear that the futures price is chasing and gaining against the spot price. Note this is the case where the price elasticity is larger than one, yet the basis is shrinking. $\delta_{pt} > 0$ and $dp_t > 0$ imply that $d^2p_t > 0$, or p_t is following a convex path. Similarly, $\delta_{ft} > 0$ and $df_t > 0$ mean $d^2f_t > 0$, or f_t also has a convex path. The comovement pattern of spot and futures prices for Case 32 is illustrated in Figure 2 in the time interval of [2,3]. Note that the futures price is catching up with the spot price, thus being one of the 24 mean-reversion situations defined (Category 3).

We can now proceed to discuss the simulation models and data that are used in this article to calculate the frequencies of each of the 48 market situations and of the three categories of mean reversion defined.

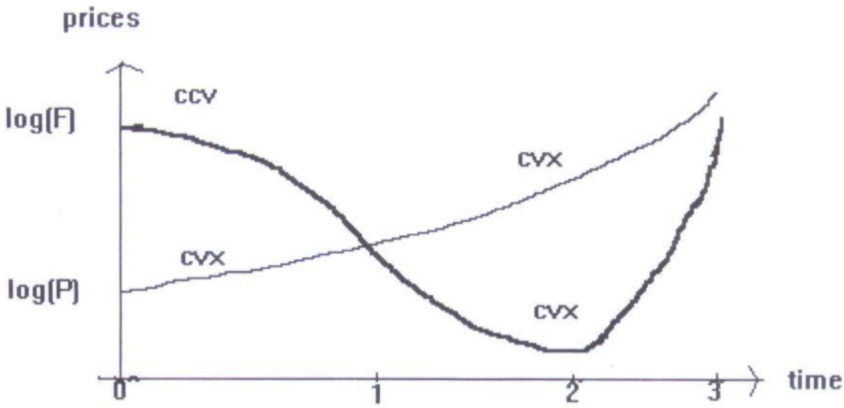

FIGURE 2

Graphical illustration of the comovement of spot and futures prices (cvx = convex and ccv = concave). The thicker curve represents the futures price path; the thinner curve is the path of spot prices. In the time interval (0,1), the pattern corresponds to Case 1, whereas in the time interval (1,2), the pattern corresponds to Case 11, and in the time interval (2,3), the pattern belongs to Case 32. Note that Cases 1 and 32 are all mean-reverting situations, but they fall into different categories, specifically Categories 1 and 3 (see Table I for details).

SIMULATION MODELS AND THE DATA

Alternative DGPs

Three models are considered. One is the random walk with drift (RWWD) DGP. Another is the mix of a random walk with a first-order autoregression (RWAR), similar to that used by Poterba and Summers (1988) and Fama and French (1988b). The last is an error-correction model (ECM) proposed by Engle and Granger (1987) for cointegrated variables.

The purpose of the Monte Carlo exercise is to generate a dual price series (futures and spot prices) and compare the comovement pattern with the real-time data of 10 commodities. The key strategy in constructing the price pairs is to use the cost-of-carry relationship, or Equation 2, as a binding condition. For DGP 1 (RWWD),

$$f_t = p_t + s_t$$

$$p_t = \mu + p_{t-1} + \sigma_1 e_{1t}$$

$$s_t = s_{t-1} + \sigma_2 e_{2t}$$

where e_{1t} and e_{2t} are white noise and we assume $\text{cov}(e_{1t}, e_t) = 0$ for simplicity. The DGP1 produces a RWWD for the log spot price and a RWWD plus a random walk without drift for the log futures prices. The parameters are chosen to mimic the mean and standard error of the commodity gold (see Tables II and III). In particular, $\mu = 0.01$, $\sigma_1 = 0.1$, and $\sigma_2 = 0.015$, and starting values are $p_0 = 6$ and $f_0 = 6$ $(s_0 = 0)$. Notice that σ_1 is chosen such that it dominates σ_2. As a result, volatility in p_t tends to dominate the volatility from s_t. For DGP 2 (RWAR),

$$f_t = p_t + s_t$$

$$s_t = \beta\, s_{t-1} + \sigma_1 e_{1t}$$

$$p_t = q_t + a_t$$

$$a_t = \alpha\, a_{t-1} + \sigma_3 e_{2t}$$

$$q_t = \mu + q_{t-1} + \sigma_2 e_{3t}$$

where e_{1t}, e_{2t}, and e_{3t} are assumed to be identically independently distributed. By the design, both spot and futures prices are following a mixture of RWAR series. Again, the standard error from the random walk is chosen to be much greater than that of the autocorrelated part. More specifically, $\mu = 0.01$, $\alpha = 0.98$, and $\beta = 0.99$, whereas $\sigma_1 = 0.0001$,

TABLE II
Descriptive Statistics of Bases

	SPD	*M*	*SD*	*AUC1*	*AUC2*	*AUC3*	*AUC4*	*AUC5*
Gold								
	12	0.010	0.004	0.965	0.946	0.920	0.899	0.879
	23	0.011	0.005	0.998	0.997	0.995	0.994	0.993
	34	0.011	0.005	0.999	0.998	0.996	0.995	0.994
	45	0.011	0.005	0.999	0.998	0.996	0.995	0.994
	56	0.012	0.005	0.999	0.998	0.996	0.996	0.995
	67	0.012	0.005	0.997	0.996	0.995	0.994	0.993
	78	0.012	0.005	0.996	0.995	0.994	0.993	0.993
Silver								
	12	0.012	0.006	0.733	0.710	0.697	0.669	0.640
	23	0.014	0.004	0.935	0.930	0.926	0.923	0.920
	34	0.014	0.004	0.933	0.927	0.924	0.923	0.921
	45	0.014	0.005	0.900	0.892	0.882	0.860	0.850
	56	0.014	0.004	0.888	0.873	0.863	0.848	0.834
	67	0.013	0.007	0.879	0.844	0.799	0.757	0.712
Platinum								
	12	0.010	0.009	0.888	0.841	0.820	0.806	0.795
	23	0.010	0.009	0.929	0.901	0.911	0.907	0.901
	34	0.011	0.008	0.994	0.991	0.987	0.983	0.980
	45	0.010	0.079	0.972	0.945	0.919	0.895	0.873
Crude oil								
	12	−0.007	0.019	0.887	0.833	0.800	0.782	0.756
	23	−0.005	0.013	0.964	0.938	0.909	0.891	0.879
	34	−0.004	0.010	0.979	0.956	0.940	0.926	0.912
	45	−0.003	0.009	0.982	0.963	0.949	0.936	0.923
	56	−0.003	0.008	0.979	0.962	0.948	0.934	0.920
	67	−0.002	0.007	0.956	0.924	0.902	0.882	0.864
	78	−0.002	0.006	0.956	0.925	0.903	0.881	0.862
Live cattle								
	12	−0.014	0.035	0.976	0.953	0.931	0.909	0.889
	23	−0.009	0.028	0.973	0.945	0.920	0.895	0.870
	34	−0.004	0.023	0.976	0.952	0.928	0.904	0.880
	45	−0.004	0.020	0.974	0.949	0.926	0.903	0.880
	56	−0.003	0.019	0.972	0.946	0.922	0.898	0.874
Wheat								
	12	−0.009	0.037	0.975	0.949	0.925	0.900	0.875
	23	−0.007	0.032	0.982	0.966	0.949	0.933	0.918
	34	−0.004	0.032	0.982	0.966	0.952	0.938	0.925
	45	−0.002	0.028	0.968	0.944	0.923	0.903	0.833
Orange juice								
	12	−0.003	0.030	0.952	0.921	0.894	0.874	0.855
	23	−0.002	0.023	0.979	0.964	0.951	0.939	0.927
	34	−0.001	0.019	0.976	0.960	0.945	0.931	0.917
	45	−0.000	0.016	0.959	0.933	0.911	0.893	0.878
	56	−0.000	0.017	0.950	0.898	0.844	0.784	0.729
	67	−0.008	0.133	0.991	0.982	0.973	0.964	0.954

TABLE II
(Continued)

SPD	M	SD	AUC1	AUC2	AUC3	AUC4	AUC5
World sugar							
12	0.022	0.061	0.972	0.955	0.941	0.927	0.915
23	0.015	0.046	0.982	0.968	0.957	0.946	0.935
34	0.014	0.042	0.988	0.980	0.971	0.963	0.955
45	0.021	0.044	0.986	0.978	0.971	0.964	0.958
Treasury bond							
12	−0.008	0.005	0.986	0.982	0.979	0.975	0.972
23	−0.007	0.004	0.997	0.995	0.993	0.991	0.989
34	−0.006	0.005	0.584	0.584	0.573	0.552	0.593
45	−0.006	0.005	0.548	0.548	0.537	0.514	0.557
56	−0.005	0.003	0.995	0.992	0.989	0.986	0.984
67	−0.005	0.003	0.987	0.978	0.971	0.964	0.961
78	−0.005	0.004	0.966	0.942	0.920	0.902	0.885
S&P 500 index							
12	0.008	0.005	0.982	0.973	0.969	0.962	0.958
23	0.008	0.004	0.989	0.986	0.983	0.980	0.978
34	0.008	0.009	0.893	0.821	0.755	0.695	0.659

Note. The reported entries include means (M) and standard errors (SD) (Columns 2 and 3) of the bases from all the adjacent futures price pairs (Column 1) for all commodities. Also included are the autocorrelation coefficients of these bases with up to five lags (Columns 4–8). SPD column indicates which pair of nearby contract is used to compute log price spread.

$\sigma_2 = 0.1$, and $\sigma_3 = 0.0001$, with starting values of $p_0 = 6$ and $f_0 = 6$ ($s_0 = 0$). For DGP 3 (ECM),

$$p_t = p_{t-1} + \alpha(f_{t-1} - p_{t-1}) + \sigma_1 e_{1t}$$

$$f_t = f_{t-1} - \beta(f_{t-1} - p_{t-1}) + \sigma_2 e_{2t}$$

where the error terms are again assumed to be white noise with zero mean and unit variance. $\alpha > 0$ and $\beta > 0$ are assumed. The error-correction representation was developed by Engle and Granger (1987) to characterize the behaviors of cointegrated variables. The two integrated variables are cointegrated if a linear combination of the two becomes stationary. The cointegration relationship is interpreted as a long-run equilibrium attractor, where the error terms (i.e., the residual of the linear combination) measure or correct the deviations of the relationship. In the ECM specifications, p_t and f_t are assumed to be cointegrated with a cointegration vector of $(1, -1)$, or s_t is stationary (the Engle–Granger procedure is applied, and cointegration is found in all the cases; the results are available on request). This means the basis s_t can be used as an error-correction term. Notice that whenever s_t is greater than 0, p_t (f_t) goes up (down); whenever s_t is less than 0, p_t (f_t) goes down (up). The parameters are specified as $\alpha = 0.008$, $\beta = 0.007$, $\sigma_1 = 0.15$, and $\sigma_2 = 0.10$. The starting values are $p_0 = f_0 = 6$.

TABLE III
Descriptive Statistics of Log Prices

Nby	*M*	*SD*	*AUT1*	*AUT2*	*AUT3*	*AUT4*	*AUT5*
Gold							
1	5.938	0.107	0.995	0.990	0.985	0.980	0.975
2	5.948	0.107	0.995	0.990	0.985	0.980	0.975
3	5.959	0.108	0.995	0.990	0.985	0.980	0.975
4	5.970	0.109	0.995	0.990	0.985	0.980	0.976
5	5.981	0.110	0.995	0.990	0.986	0.981	0.976
6	5.993	0.111	0.995	0.991	0.986	0.981	0.977
7	6.005	0.112	0.995	0.991	0.986	0.982	0.977
8	6.017	0.114	0.995	0.991	0.987	0.982	0.978
Silver							
1	6.370	0.305	0.998	0.996	0.994	0.992	0.991
2	6.383	0.307	0.998	0.996	0.995	0.993	0.991
3	6.396	0.310	0.998	0.997	0.995	0.993	0.991
4	6.410	0.312	0.998	0.997	0.995	0.993	0.991
5	6.424	0.315	0.998	0.997	0.995	0.993	0.992
6	6.438	0.317	0.998	0.997	0.995	0.994	0.992
7	6.452	0.321	0.998	0.997	0.995	0.994	0.992
Platinum							
1	6.022	0.200	0.996	0.993	0.990	0.986	0.983
2	6.032	0.198	0.997	0.993	0.990	0.986	0.983
3	6.042	0.196	0.997	0.993	0.990	0.986	0.983
4	6.054	0.194	0.997	0.993	0.990	0.986	0.983
5	6.063	0.191	0.997	0.994	0.991	0.988	0.985
Crude oil							
1	3.027	0.244	0.995	0.991	0.987	0.984	0.980
2	3.020	0.241	0.997	0.993	0.990	0.987	0.984
3	3.015	0.237	0.997	0.994	0.991	0.988	0.985
4	3.010	0.233	0.997	0.994	0.991	0.989	0.986
5	3.007	0.230	0.997	0.994	0.992	0.990	0.987
6	3.004	0.227	0.997	0.995	0.992	0.990	0.988
7	3.002	0.225	0.997	0.995	0.992	0.990	0.988
8	2.999	0.223	0.997	0.995	0.993	0.991	0.989
Live cattle							
1	4.229	0.100	0.994	0.988	0.982	0.975	0.969
2	4.215	0.097	0.995	0.989	0.985	0.979	0.974
3	4.206	0.094	0.996	0.992	0.989	0.985	0.982
4	4.200	0.091	0.997	0.994	0.991	0.989	0.986
5	4.198	0.090	0.997	0.994	0.992	0.989	0.986
6	4.194	0.087	0.997	0.994	0.992	0.989	0.987
Wheat							
1	5.843	0.160	0.997	0.993	0.989	0.986	0.982
2	5.834	0.158	0.997	0.995	0.992	0.990	0.988
3	5.827	0.154	0.997	0.995	0.993	0.990	0.988
4	5.823	0.152	0.998	0.995	0.993	0.990	0.988
5	5.825	0.150	0.998	0.995	0.993	0.991	0.989

(Continued)

TABLE III
(Continued)

Nby	*M*	*SD*	*AUT1*	*AUT2*	*AUT3*	*AUT4*	*AUT5*
Orange juice							
1	4.862	0.227	0.997	0.994	0.991	0.988	0.985
2	4.858	0.212	0.997	0.994	0.991	0.988	0.985
3	4.856	0.201	0.997	0.994	0.991	0.988	0.984
4	4.855	0.192	0.997	0.994	0.991	0.988	0.985
5	4.855	0.186	0.998	0.995	0.991	0.988	0.984
6	4.855	0.179	0.998	0.995	0.992	0.989	0.985
7	4.847	0.173	0.997	0.993	0.989	0.985	0.980
World sugar							
1	2.167	0.351	0.997	0.993	0.991	0.988	0.985
2	2.189	0.316	0.997	0.994	0.992	0.990	0.987
3	2.204	0.285	0.997	0.995	0.992	0.990	0.987
4	2.218	0.258	0.997	0.994	0.992	0.989	0.986
5	2.239	0.232	0.997	0.994	0.991	0.988	0.985
Treasury bond							
1	4.493	0.176	0.999	0.998	0.997	0.996	0.996
2	4.485	0.175	0.999	0.998	0.997	0.996	0.996
3	4.478	0.174	0.999	0.998	0.997	0.996	0.995
4	4.472	0.173	0.999	0.998	0.997	0.996	0.995
5	4.466	0.173	0.999	0.998	0.997	0.996	0.995
6	4.461	0.172	0.999	0.998	0.997	0.996	0.995
7	4.456	0.171	0.999	0.998	0.997	0.996	0.995
8	4.451	0.170	0.999	0.998	0.997	0.996	0.995
S&P 500 index							
1	5.653	0.425	1.000	0.999	0.999	0.999	0.998
2	5.662	0.423	1.000	0.999	0.999	0.999	0.998
3	5.670	0.421	1.000	0.999	0.999	0.999	0.998
4	5.677	0.420	1.000	0.999	0.999	0.999	0.998

Note. The reported entries include means and standard errors (Columns 2 and 3) of the all nearby log prices of 10 commodities and up to fifth-lag autocorrelation coefficients (Columns 4–8). Nby refers to nearby prices (see the text).

The found mean reversion is determined by the generated sequence of basis s_t for all the DGPs. In DGP3, by subtracting the first equation from the second, we obtain

$$s_t = [1 - (\alpha + \beta)]s_{t-1} + (\sigma_2 e_{2t} - \sigma_1 e_{1t})$$

Thus, s_t is an AR(1) series with close to one coefficient, except for different error terms. In DGP2, s_t is an AR(1) process with a coefficient of $\beta = 0.99$. In DGP1, s_t follows a random-walk process.

In each of the 2000 replications for all the DGPs, dual price series are generated with 120 observations, where only the last 60 (simulating 3-month contracts) are used to calculate the frequencies. The means and standard deviations of these dual price series are listed in Table IV. It can be seen that by design, futures (log) prices in RWWD have a

TABLE IV
Averages of Mean and Standard Errors of the Simulated Dual Price Series

	Prices	*M*	*SE*
RWWD	Spot	8.468439	1.013116
	Futures	8.467245	1.013145
RWAR	Spot	8.458433	1.013115
	Futures	8.458443	1.013114
ECM	Spot	5.964028	0.793370
	Futures	5.972742	0.619250

slightly higher standard error, whereas futures prices in RWAR have slightly smaller standard errors. In ECM, means are lower, and standard errors are also smaller. I also tried a variety of other parameter values with a smaller number of replications and did not find the results significantly different. But a more thorough analysis, such as an imposing covariance structure, is needed. I leave this interesting area for future investigation.

Data

The frequencies of the mean reversion and each of the 48 cases of spot and futures price comovement are calculated with the daily settlement prices for 10 commodity futures markets. All prices are obtained from the CRB InfoTech Commodity Data database of Knight–Ridder Financial. The sample includes four mineral contracts (crude oil, gold, silver, and platinum), four agricultural contracts (wheat, world sugar, orange juice, and live cattle), and two financial contracts [Treasury bonds and Standard & Poor's (S&P) 500 index]. Crude oil and platinum data are from the New York Mercantile Exchange. Gold data are from the New York Commodity Exchange, whereas silver and Treasury bond data are from the Chicago Board of Trade. Live cattle and S&P 500 index data are from the Chicago Mercantile Exchange. World sugar and orange juice data are from the New York Cotton Exchange. Finally, wheat data are from the Kansas City Board of Trade.

The sample period of all the commodities starts on April 1, 1982 and ends on October 10, 1995, except for the S&P 500 index and crude oil, which began trading on April 21, 1982 and March 30, 1983, respectively. The futures prices are collected for the active contract month (or delivery month) of the commodities. The delivery month cycles are different for different commodities. Crude oil delivers in all 12 months. Gold, silver, and live cattle deliver on even months. Platinum and orange juice deliver on Months 1, 4, 7, and 10. Wheat delivers on Months 3, 5, 7, 9, and 12, and world sugar delivers on Months 3, 5, 7, and 10. Financials, including Treasury bonds and the S&P 500 index, deliver quarterly. Futures prices are ordered according to their maturity and are called *nearby prices*, with the nearby one corresponding to the nearest maturity, and so forth. Several nearby price series are created for each commodity. In particular, for crude oil, gold, and Treasury bonds, eight nearby price series are created, whereas seven series are created for silver and orange juice; six are created for live cattle; five are created for platinum, wheat, and world sugar; and three are created for the S&P 500 index.

Following Bessembinder et al. (1995) and Fama and French (1987), we use the first nearby futures contract as a proxy for spot price. For each commodity, multiple bases are calculated corresponding to each pair of adjacent nearbys, or $s_t(m) = f_t(m) - f_t\,(m - 1)$, $m = 2, 3. \ldots, N$, where N is the number of the nearbys for the commodity, with $m = 1$ representing the first nearby and $m = N$ representing the most distant futures contracts in terms of the maturity. The mean-reversion frequencies are calculated accordingly.

HOW FREQUENT COMMODITY PRICES ARE MEAN-REVERTING?

Table V reports the empirical findings. The mean-reversion frequencies, along with the associated categories, are reported for all 10 commodities and simulation models. These frequencies are computed daily according to the $\mathbf{ID}_t$ vectors and case classifications of Table I. For example, the daily price change $dp_t = p_t - p_{t-1}$ is computed as the difference of today's price over the previous day's price. And the difference of price change $d^2p_t = dp_t - dp_{t-1}$ is calculated as the price difference of the price change. Also included in Table V are the frequencies of ones of each component of $\mathbf{ID}_{t,}$ except for the last two columns, where the frequencies of ones of numerators of the correspondent components are shown instead (i.e., reporting the convexity of the price path rather than growth rate of the price return). The histograms of more detailed price comovement patterns are gathered in Figures 3 and 4 for both simulated and real data. In the next section, I discuss the real data results and then proceed to those from simulation models.

TABLE V
Mean-Reversion Frequencies and Patterns of Price Comovement

Commodities	Price Pairs	M Reversion All	M Reversion 1	M Reversion 2	M Reversion 3	Patterns s	Patterns dp	Patterns df/dp	Frequencies ds/dp	Frequencies d^2p	Frequencies d^2f
Simulations											
RWWD		48	24	12	12	48	52	48	24	48	48
RWAR		49	25	12	12	50	53	48	24	49	49
ECM		49	24	12	13	49	48	48	19	47	49
Gold											
	12	46	27	9	10	97	48	42	23	50	49
	23	48	27	10	11	100	48	44	23	50	50
	34	47	27	10	10	99	48	43	23	50	50
	45	47	27	10	10	99	48	43	23	50	50
	56	47	27	10	10	99	48	43	23	50	50
	67	47	27	10	10	99	48	43	23	50	50
	78	47	27	10	10	99	48	43	23	50	50
Silver											
	12	47	27	9	11	97	47	42	22	51	51
	23	47	27	9	11	100	47	42	22	52	52
	34	46	26	9	11	99	46	42	22	52	51
	45	46	27	9	10	98	47	41	22	51	51
	56	47	27	10	10	99	48	42	22	51	52
	67	48	27	10	11	98	48	42	21	52	52
Platinum											
	12	48	26	11	11	85	50	45	23	48	49
	23	48	26	11	11	87	49	45	23	49	49
	34	49	26	12	11	92	50	47	24	50	50
	45	49	26	11	12	57	48	44	19	46	49
Crude oil											
	12	48	25	10	13	30	48	48	25	49	50
	23	48	25	10	13	27	46	44	24	46	46
	34	49	26	11	12	26	48	46	25	48	48
	45	50	26	11	13	26	49	47	26	49	50
	56	51	26	11	14	27	48	47	26	50	50
	67	49	25	11	13	26	46	46	25	49	49
	78	49	26	11	12	25	47	46	25	49	49
Live cattle											
	12	50	26	12	12	34	51	45	23	49	48
	23	48	27	10	11	36	49	43	19	47	45
	34	49	26	9	14	47	48	44	22	48	48
	45	48	26	9	13	49	46	44	21	48	49
	56	49	25	12	12	49	48	45	21	49	51
Wheat											
	12	50	23	12	15	46	49	51	26	50	48
	23	50	25	11	14	50	48	50	26	49	51
	34	47	24	10	13	58	46	47	26	47	48
	45	50	25	11	14	68	48	49	25	51	51

TABLE V
(Continued)

Commodities	Price Pairs	M Reversion All	1	2	3	Patterns s	dp	df/dp	Frequencies ds/dp	d^2p	d^2f
Orange juice											
	12	51	24	12	15	48	49	50	26	49	51
	23	47	25	12	10	51	49	46	22	49	49
	34	46	24	10	12	45	47	45	22	47	48
	45	49	24	12	13	44	48	49	24	48	50
	56	48	25	11	12	45	48	49	25	49	52
	67	52	26	12	14	45	51	48	24	48	49
World Sugar											
	12	49	27	10	12	63	47	43	22	48	50
	23	50	28	10	12	57	48	41	22	48	50
	34	49	28	9	12	54	45	41	22	48	49
	45	49	27	10	12	59	47	41	19	47	50
Treasury bond											
	12	54	28	12	14	6	51	48	25	52	52
	23	51	27	11	13	5	50	46	23	51	51
	34	59	18	18	23	11	52	66	44	51	50
	45	50	26	11	13	10	51	49	24	49	51
	56	50	27	11	12	6	50	45	23	50	50
	67	52	27	12	13	7	50	47	24	51	51
	78	50	27	11	12	7	50	47	22	50	51
S&P 500 index											
	12	47	25	13	9	96	50	45	23	47	47
	23	48	26	13	9	99	52	46	24	49	49
	34	48	25	14	9	96	52	45	22	48	47

Note. Column 3 reports the frequencies of the mean reversion by the definition of basis shrinkage. Columns 4–6 further report the frequency decomposition of mean reversions as Categories 1, 2, and 3 (see Table I). Columns 7–12 are the computed frequencies of contango prices, rising spot prices, spot and futures prices moving in the same directions, futures (log) price changes larger than spot changes magnitudes when they move in the same directions, and the convexity frequencies of spot and futures. The entries include the frequency results for both the simulations and each price pair of the 10 commodities, which have approximately 3,400 daily observations.

Pattern of Mean Reversion and Price Comovement

Approximately 50% of the time, spot prices of all 10 commodities are mean-reverting. Among them, in half the cases the futures and spot prices are moving toward each other (Category 1), and in half the cases they are chasing each other (Categories 2 and 3). In particular, when spot and futures prices are moving in the same direction, they split almost equally into Category 2 and 3 cases.

The symmetry of these frequencies shows that the shocks from the futures market are not rare and are probably as frequent as the shocks from the spot market. It is conceivable that if shocks from the spot market are dominant, the spot price will be moving more violently. As tested

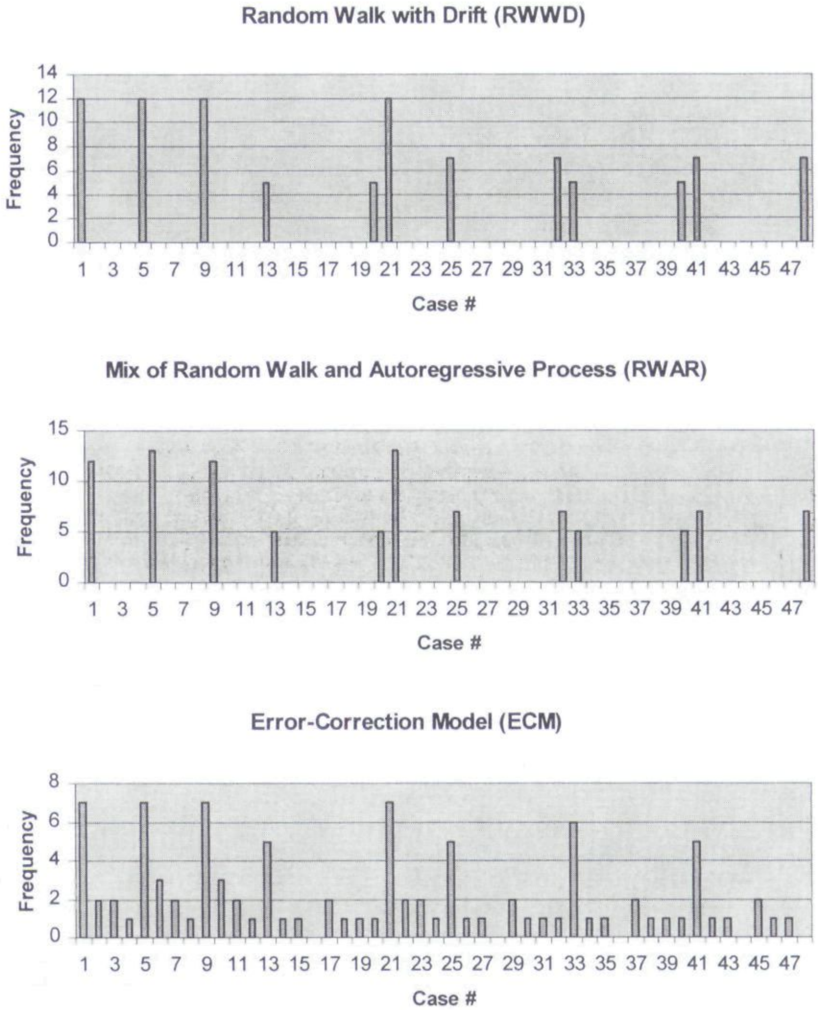

FIGURE 3
Histogram of the price comovement pattern for simulated data. The horizontal axis refers to the case number defined in Table I.

in Fama and French (1988a), the futures price responds in less than a one-to-one fashion to spot price change when inventory is low but adjusts fully when inventory is high. If this is true, mean-reversion cases where the futures price change is bigger will happen more. In other words, Category 2 of mean reversion is dominant over Category 3. In Table IV, however, Category 3 appears a little more frequently (1–5%) than Category 2, except for the S&P 500.

Another noteworthy insight relates to Category 1 of mean reversion when spot and futures prices move toward each other. This condition is the fundamental driving force of mean reversion in Bessembinder et al.

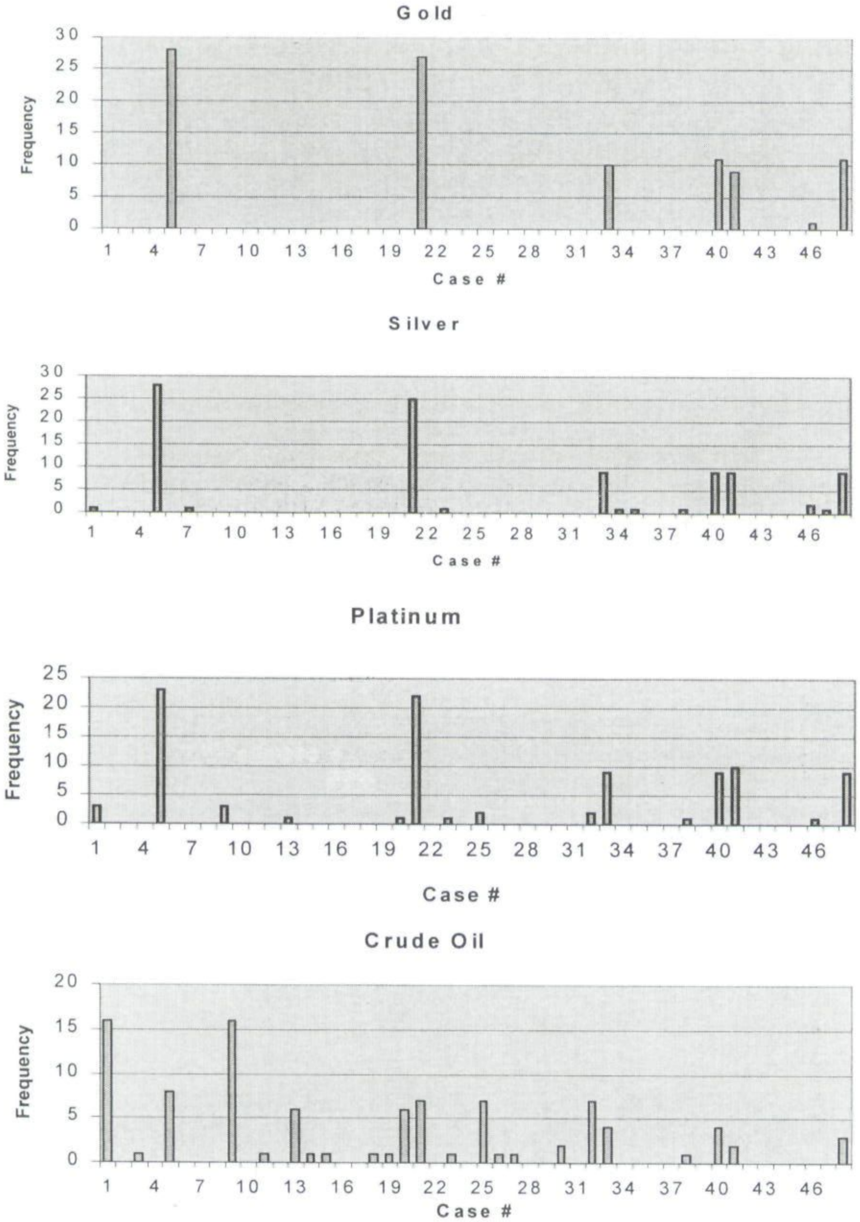

FIGURE 4
Histogram of the price comovement pattern for real data. All frequencies are computed on the basis of nearby 1 and nearby 2 futures prices. The horizontal axis refers to the case number defined in Table I.

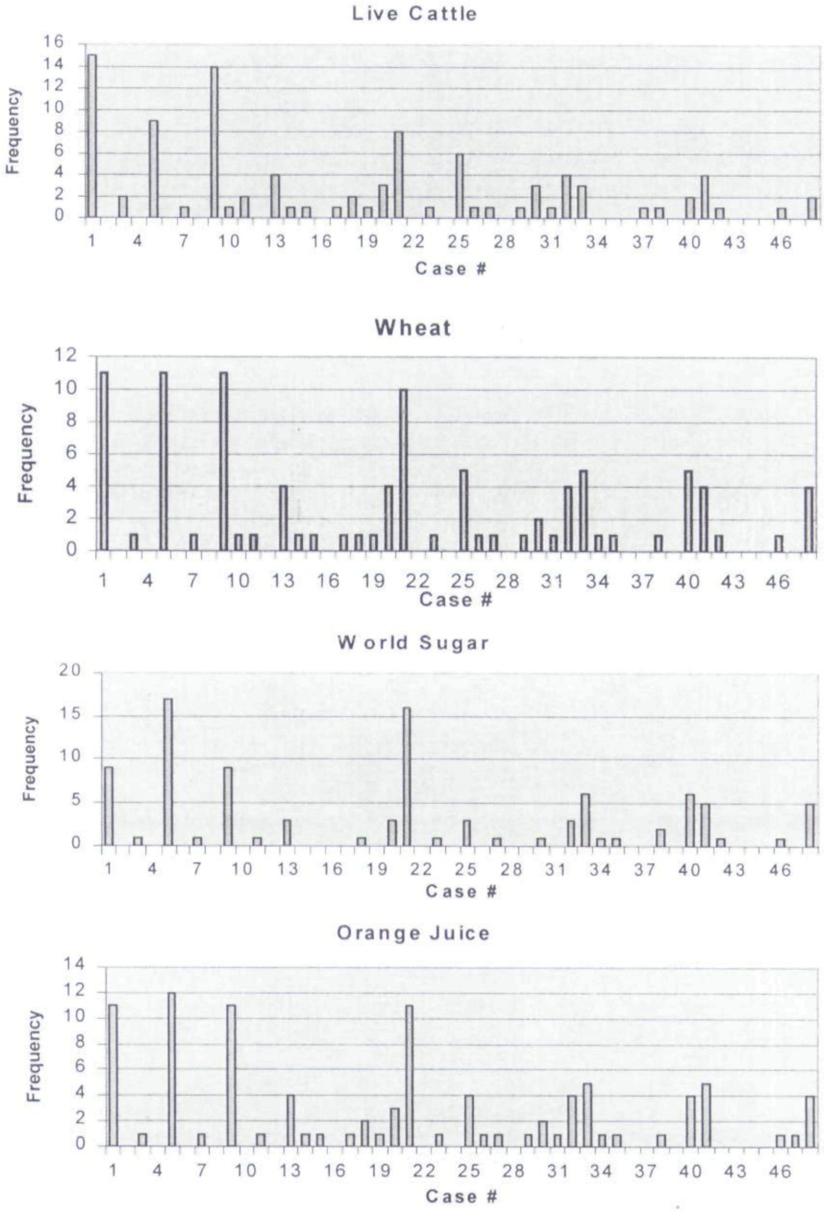

FIGURE 4
(Continued)

(1995), where the condition is translated as a negative correlation between the spot price and interest rate and a positive correlation with a net convenience yield. However, such a characterization is neither necessary nor sufficient for mean reversion under the framework presented here. To observe the difference, first note that Category 1 can also be expressed as cases when the basis and spot price move in opposite directions. Next, the following counterexamples are helpful. More

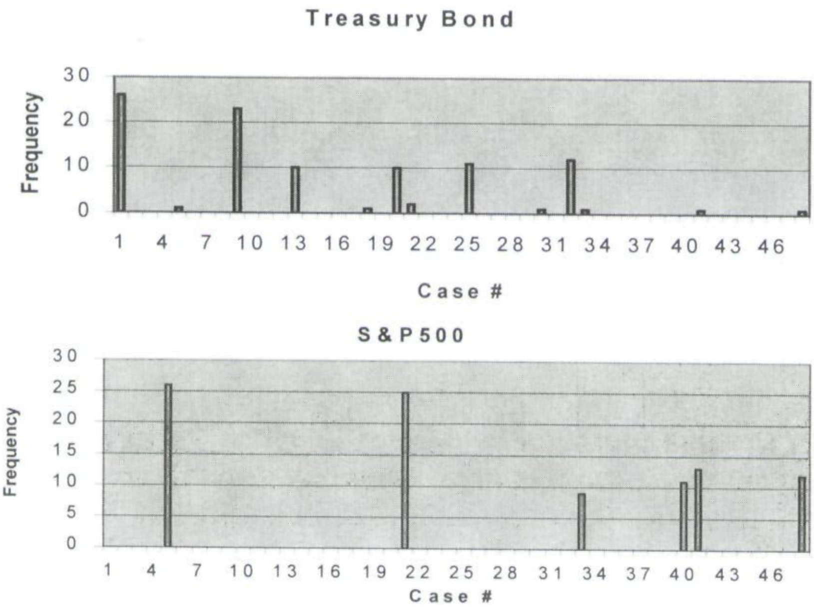

FIGURE 4
(Continued)

specifically, in a backwardation market ($f_t < p_t$), it is a Category 3 mean reversion when both the spot price and basis rise. Under such a scenario, the basis and spot price are positively correlated. However, because the basis is negative, the increase in the basis means the absolute value of the basis is shrinking. On the other hand, also in a backwardation market, it will not be mean-reverting if the spot price increases and the basis decreases, where the correlation between the spot price and basis becomes negative. However, the decreasing basis implies the absolute value of the basis is expanding, or the spot and futures prices are moving apart, although in the same direction. Overall, using a negative correlation between the spot price and basis as the signal for mean reversion fits well with the approach taken here only in a contango market, conditional on shocks from futures demand and supply being relatively rare. Under such circumstances, the intuition that futures prices respond to shocks from a spot market less violently holds precisely.

Table V also reports fundamental characteristics for 10 commodity markets. Column 7 (s) reports the market types (i.e., contango, $f_t > p_t$; backwardation, $f_t < p_t$; or a mix). Precious metals and the S&P 500 index are basically in contango markets, Crude oil and Treasury bonds are mostly in backwardation markets, whereas agricultural commodities interchange between the two. Note that the classification is based on the sign frequencies of the basis. The results are dramatically different if the signs of the average value of s_t across whole sample periods are used;

these are shown in Column 3 (M) of Table II. Using frequencies as classification standards seems to be obvious. Otherwise, the pattern is backwardation about 50% of the time and contango the rest of the time. Wheat, for example, would not have been revealed if the average value were used that identified the wheat as a backwardation market.

Columns 8 (dp) and 9 (df/dp) indicate that spot prices rise half of the time, and half the time futures prices move in the same directions as spot prices for all the commodities and all the maturities. Column 10 (ds/dp) implies that a little less than 80% of the time the futures price elasticity, with respect to the spot price elasticity, is less than one. The last two columns (d^2p and d^2f) show that about half of the time the log spot (futures) price is following convex or concave paths over time or continuously changes in the curvatures.

Figures 3 and 4 demonstrate the histograms of the 10 commodities and simulations for the 48 cases of price comovement. The most striking observation from these charts is that the frequencies are clustered around 4 to 6 cases for each commodity. Table VI lists those cases that appear greater than or equal to 7%. The similarities and dissimilarities among different commodities are illuminating. In particular, the cluster pattern can be examined by the 10 commodities being divided into the following three groups:

1. The dominant contango market, including precious metals and the S&P 500 index.
2. The dominant backwardation market, including crude oil and Treasury bonds.
3. The mixed market, including all agricultural commodities.

TABLE VI

Cluster Patterns of Commodity Price Comovements

Commodity	*Cases with a Frequency Greater Than or Equal to 7%*
Gold	5, 21, 33, 40, 41, 48
Silver	5, 21, 33, 40, 41, 48
Platinum	5, 21, 33, 40, 41, 48
Crude oil	1, 5, 9, 13, 20, 21, 25, 32
Live cattle	1, 5, 9, 21
Wheat	1, 5, 9, 21
Orange juice	1, 5, 9, 21
World sugar	1, 5, 9, 21
Treasury bond	1, 9, 13, 20, 25, 32
S&P 500 index	5, 21, 33, 40, 41, 48

Note. See Table I for the definition of each case.

For Group 1 commodities, the six cases (5, 21, 33, 40, 41, and 48) are frequent. Among them, Cases 5 and 21 imply that futures and spot prices are moving in opposite directions and that the price paths have different curvatures (convexity or concavity). The other cases relate to the scenarios where prices move in the same directions and with the same curvature. Furthermore, Cases 21, 33, and 41 are mean-reverting cases, corresponding to Categories 1, 2, and 3 respectively.

For Group 3, the four cases of 1, 5, 9, and 21 are such that spot and futures prices move in opposite ways with different curvatures. Although Cases 5 and 21 relate to contango markets, Cases 1 and 9 arise in a backwardation market, with log spot prices declining in an accelerating way and log futures prices rising in a decelerating way. Both Cases 1 and 21 are mean-reverting within Category 1. With respect to dominant contango markets (Group 1), only 5 and 21 are overlapping cases, and the market seems to be more concentrated because only four instead of six cases frequently appear. However, relative to a backwardation market, for example, in the case of Treasury bonds in Table VI, the overlapping cases are 1 and 9. For the other backwardation market, crude oil overlaps with the mixed market in all four cases. This is largely because crude oil is a borderline commodity between backwardation and mixed groups. As seen in Table V in Column 7 for the frequency of $s < 0$, crude oil is present around 70% of the time in a backwardation market (or 30% in a contango market), whereas Treasury bonds are backwardation for 90 to 95% of the time. The backwardation market also adds Cases 13, 20, 25, and 32 as cluster members. All of these cases have the type of spot and futures moving in the same direction with the same curvature. For Group 2, Cases 20 and 32 imply mean reversion corresponding to Categories 2 and 3, respectively.

In summary, there are only 16 out of 48 cases that have significant frequencies in the 10 commodities investigated. For agricultural commodities, which have mingled states of both contango and backwardation prices, Cases 1 and 9 are the major sources of a backwardation pattern, whereas Cases 5 and 21 are more likely to give rise to the contango pricing. In all these cases, rising prices are always decelerating (concave path), and declining prices are always accelerating (convex path). These 16 cases can generate all types of price comovements; this further supports that shocks from the demand and supply of futures markets are not scarce. It also implies that price deteriorates faster than appreciation for the agricultural commodities relative to the other markets. However, the dominant contango market and backwardation market each add an extra four different cases where the spot and futures prices are always moving parallel with the same curvature. This is indicative of symmetry impacts for the shocks coming from either futures or spot markets.

Implications From Alternative DGPs

The first three rows of Table V report the frequencies of mean reversion and patterns of price comovement for simulation models. The results show that all three DGPs approximate the data reasonably well except for the dramatic differences in the judgment of contango or backwardation markets. However, all the DGPs are reasonable candidates for agricultural commodities for that matter. It may not be surprising that the mean-reversion frequencies are well approached because by design the basis is following either a random-walk or AR(1) process close to a random walk. However, Table VII shows that the cluster pattern of price comovements are potentially quite different.

First, RWWD and RWAR are indistinguishable in terms of price comovement patterns. Compared to the data in Table VI, they pick up all the cases for a mixed (i.e., the coexistence of contango and backwardation) market and half of the cases from each of the other types of markets. In this sense, both RWWD and RWAR are incomplete representations for any type of market. However, ECM seems to be solely an approximation of the mixed markets. ECM is, therefore, expected to perform better than RWWD or RWAR in agricultural markets. However, for the dominant markets (i.e., metal and financial commodities), RWWD and RWAR might perform better than ECM, depending on how frequently the alternative comovement patterns show up within the sample period.

CONCLUSION

Economic shocks can affect spot and futures prices in complex ways. The cost-of-carry relationship implies that the futures price shares a common trend with the expected spot price, and the basis change

TABLE VII

Cluster Patterns of Data-Generating-Process (DGP) Price Comovements

DGP	*Cases with a Frequency Greater Than or Equal to 7%*
RWWD	1, 5, 9, 21, 25, 32, 41, 48
RWAR	1, 5, 9, 21, 25, 32, 41, 48
ECM	1, 5, 9, 21

Note. See Table I for the definition of each case.

contains important messages about how futures and spot prices respond to underlying shocks. The basis shrinkage is a sign of mean reversion because it implies that the spot price is moving closer to its mean trend.

Empirical results indicate that about 50% of the time spot prices are mean-reverting for all 10 commodities investigated. Among these mean-reverting cases, 50% of the time futures prices and spot prices are moving toward each other, and the rest of the time they move in parallel fashion. I interpret this as evidence that shocks from demand or supply in futures markets are not rare.

The patterns of mean reversion are shown to be possibly generated by the following three processes: a bivariate RWWD process, a bivariate mixed process of RWWD plus an AR(1), and an ECM with the basis as the correction term. All these models are shown to be reasonable in approximating the frequencies of mean reversion, bull or bear markets, and the curvature of price paths. However, these processes can also potentially misrepresent the underlying price dynamics by overemphasizing the importance of certain price comovement patterns.

The spot and futures price dynamics depend heavily on the types of commodity market. In particular, precious metals and the S&P 500 index, which are dominant contango markets, display frequent patterns of parallel movement of spot and futures prices. In the dominant backwardation markets such as Treasury bonds and crude oil, mean reversion can occur often when both futures and spot prices decline along a concave path, which is rare for a contango market. However, in a market where contango and backwardation prices can interchange frequently from one to the other, like agricultural commodities, the spot and futures prices are often moving in a path with different curvatures. In fact, it is more likely that prices go up slowly (concave) but go down quickly (convex).

The frequencies of price comovement patterns show that the ECM performs better for agricultural commodities than the random-walk model but is indifferent to random walk for other commodities. Moreover, a mixture of random walk and AR(1) does not improve on a pure random-walk model.

However, one must be cautious in applying these results because other features may confound the results of this experiment. For example, different results are possible when the covariance structure on the error terms of the simulation models is further specified. It would also be worthwhile to investigate the impacts on the price comovement when the cost-of-carry relationship breaks down. This study is just meant to be a beginning in developing a new approach to characterize price movement and capture the rich dynamics of asset spot and futures price comovement. It is potentially useful for designing arbitrage and hedging strategies and trade implementations. The framework also provides a theoretical challenge to further validate such an approach. All of these issues are left for further research.

BIBLIOGRAPHY

Bessembinder, H., Coughenour, J. F., Seguin, P. J., & Smoller, M. M. (1995). Mean reversion in equilibrium asset prices: Evidence from the futures term structure. Journal of Finance, 50, 361–375.

Bonomo, M., & Garcia, R. (1994). Can a well-fitted equilibrium asset-pricing model produce mean reversion? Journal of Applied Econometrics, 9, 19–29.

Brennan, M. J. (1958). The theory of storage. American Economic Review, 48, 50–72.

Cecchetti, S., Lam, P., & Mark, N. (1990). Mean reversion in equilibrium asset prices. American Economic Review, 80, 399–418.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio tests for autoregressive time series with a unit root. Econometrica, 49, 1057–1072.

Dixit, A. K., & Pindyck, R. S. (1994). Investment under uncertainty. Princeton, NJ: Princeton University Press.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. Econometrica, 55, 251–276.

Fama, E. F., & French, K. R. (1987). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. Journal of Business, 60, 55–73.

Fama, E. F., & French, K. R. (1988a). Business cycles and the behavior of metals prices. Journal of Finance, 43, 1075–1093.

Fama, E. F., & French, K. R. (1988b). Permanent and temporary components of stock prices. Journal of Political Economy, 96, 246–273.

French, K. R. (1986). Detecting spot price forecasts in futures prices. Journal of Business, 59, 39–54.

Kaldor, N. (1939). Speculation and economic stability. Review of Economic Studies, 7, 1–27.

Kocagil, A. E., Swanson, N. R., & Zeng, T. (in press). A new definition for time-dependent mean reversion in commodity markets. Economic Letters.

Lo, A. W., & MacKinlay, C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies, 1, 41–66.

Miller, M. H., Muthuswamy, J., & Whaley, R. E. (1994). Mean reversion of S&P 500 index basis changes: Arbitrage-induced or statistical illusion? Journal of Finance, 49, 479–513.

Poterba, J., & Summers, L. (1988). Mean reversion in stock prices: Evidence and implications. Journal of Financial Economics, 22, 27–59.

Telser, L. G. (1958). Futures trading and the storage of cotton and wheat. Journal of Political Economy, 66, 233–255.

Working, H. (1949). The theory of price of storage. American Economic Review, 39, 1254–1262.