Department of Agricultural and Resource Economics
University of California, Davis

Partially Overlapping Time Series:
A New Model for Volatility
Dynamics in Commodity Futures

by

Aaron Smith

Working Paper No. 04-013

September, 2004

Giannini Foundation of Agricultural Economics

Partially Overlapping Time Series: A New Model for Volatility Dynamics in Commodity Futures

Aaron Smith

Department of Agricultural and Resource Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
Email: adsmith@ucdavis.edu
Ph: 530-752-2138
Fax: 530-752-5614

SUMMARY

In commodity futures markets, contracts with various delivery dates trade simultaneously. Applied researchers typically discard the majority of the data and form a single time series by choosing only one price observation per day. This strategy precludes a full understanding of these markets and can induce complicated nonlinear dynamics in the data. In this paper, I introduce the partially overlapping time series (POTS) model to model jointly all traded contracts. The POTS model incorporates time-to-delivery, storability, seasonality, and GARCH effects. I apply the POTS model to corn futures at the Chicago Board of Trade and the results uncover substantial inefficiency associated with delivery on corn futures. The results also support two theories of commodity pricing: the theory of storage and the Samuelson effect.

1. INTRODUCTION

Futures markets play an integral role in the pricing and distribution of commodities. Many observers perceive these markets to be very volatile, but a full portrait of volatility patterns has been elusive. Creating such a portrait requires incorporating the effects of the time to delivery on the contract, inventory levels of the commodity, and seasonality. Understanding these effects in a unified framework is important in helping firms manage risk and minimize transaction costs. A well-specified model of volatility dynamics is also imperative for pricing options on futures contracts, which comprise an active and growing market.

For most commodities with futures markets, multiple contracts trade simultaneously. These contracts differ by the time to delivery. As time proceeds, some contracts reach delivery and cease to exist, while others are born and begin trading. From an econometric perspective, a set of futures prices presents a potentially large number of partially overlapping time series. Most applied researchers ignore the cross-sectional dimension and reduce the data to a single time series. A common method for such a reduction entails splicing together the nearby contracts, i.e., when a contract matures, take the next observation in the series from the contract that is the next closest to delivery. In many markets, ten or more contracts can be trading at a given point in time, so this strategy excludes most of the information about the commodity.

In this paper, I introduce the partially overlapping times series (POTS) model to model jointly all contracts trading on a given day. The POTS model is a factor model for partially overlapping time series that incorporates time varying conditional heteroskedasticity and time and cross-sectional variation in the factor loadings and

1

innovation variances. For commodity futures, this model captures the effects of the time to delivery, storability, and seasonality.

I apply the POTS model to corn futures at the Chicago Board of Trade (CBOT). The results reveal the dynamic structure of corn futures prices and uncover substantial inefficiency associated with delivery on the contract. The results also corroborate two well-known theories of commodity pricing: the theory of storage and the Samuelson effect. These theories have been studied extensively in isolation, but not within a fully specified dynamic model.

Finally, my results illustrate the nonstandard dynamics in a time series of spliced nearby prices. Erratic price behavior near the delivery month induces complicated nonlinear dynamics around the points where the spliced series moves from contract to contract. I suggest a strategy for avoiding these nonstandard dynamics and creating a well-behaved single time series for the fundamental commodity price. This strategy requires that the contracts be rolled over two to three months before delivery and that one of the contracts be avoided entirely.

The paper proceeds as follows. Section 2 reviews the theory of commodity futures pricing to set the stage for the POTS model, which I introduce in Section 3. Section 4 presents results from the application to CBOT corn, and Section 5 concludes the paper.

2. PRICING AND VOLATILITY IN COMMODITY FUTURES MARKETS

Futures markets allow economic agents to trade future obligations on commodities. These trades enable agents to manage risk, reduce transaction costs, diversify portfolios, and speculate. Two strands of research dominate the theory of commodity futures pricing (Williams 2001). The first is the risk management perspective, which maintains that risk-

averse agents use futures markets to hedge price risk and speculators earn a risk premium for accepting this risk (Keynes, 1930, Stein, 1986). Risk management models imply that futures prices provide a biased forecast of future spot prices. This bias constitutes the risk premium. However, there is minimal empirical evidence of a risk premium in commodity futures markets (Telser, 1958, Gray, 1961, Kolb, 1992, Bessembinder, 1993), implying that futures prices follow a martingale process.

The second strand of research focuses on the theory of storage, where arbitrage relationships primarily determine futures prices (Working, 1948, Working, 1949). A dynamic rational expectations model with risk-neutral agents underscores modern versions of the theory of storage (Williams, 1987; Routledge *et al.*, 2000). Equilibrium occurs in such models when the marginal expected profit from storing an extra unit equals the marginal value of consuming that unit. Therefore, the difference between a futures price and the spot price equals the cost storing the commodity. Such a cost is often referred to as the *cost of carry*, and it includes warehouse fees and foregone interest. The potential to arbitrage the physical commodity against the futures contract enforces a tight link between spot and futures prices.

Under the theory of storage, equilibrium implies that futures prices are in *contango*, i.e., more distant futures prices exceed nearby prices. However, when inventory of the commodity is low, nearby prices can exceed distant futures prices. This phenomenon is known as *backwardation*; it arises when the marginal value of current consumption exceeds the marginal value of storage, but it is impossible to reach equilibrium by consuming more because inventory must be nonnegative. Until inventory is restored, demand shocks will affect only nearby prices. Thus, the theory of storage implies that a

break in the link between nearby and distant futures prices is associated with low inventory and negative cost of carry. Fama and French (1988) and Ng and Pirrong (1994) provide empirical evidence supporting this theory.

A break in the link between nearby and distant prices is resolved when new inventory arrives through a new harvest. For many agricultural commodities, the harvest is seasonal and subject to weather shocks. Because weather shocks typically have a greater effect on intra-year prices than do demand shocks, much of the annual price discovery occurs around harvest time. Consequently, harvest time produces the highest volatility for all contracts (Anderson, 1985). This seasonality in futures price volatility confounds the relationship between volatility and time to maturity. The Samuelson effect (Samuelson, 1965) asserts that volatility should increase as the delivery date approaches. However, if shocks are heteroskedastic, much of the price discovery can occur months before delivery (Anderson and Danthine, 1983). Nonetheless, the Samuelson effect may exist conditional on season if, on a given date, nearby contracts are more volatile than distant ones.

The Samuelson effect should be most pronounced when inventory is low and the link between pre-harvest (i.e., old-crop) and post-harvest (i.e., new-crop) futures prices is broken. In this case, the broken link means that current demand shocks only affect nearby prices, implying that nearby prices are more volatile (Streeter and Tomek, 1992). When stocks are plentiful, spot and futures prices move together and the effect is less evident.

Previous studies of futures market volatility have typically concentrated only on a specific contract and rolled it over upon expiration. Many use the nearby contract, which results in frequent rollovers. However, Goodwin and Schnepf (2000) study the December contract for corn and the September contract for wheat, rolling each one over to the next

year in the delivery month. Streeter and Tomek (1992) go one step further and model jointly the November and March contracts for Soybeans, rolling over to the next year's contract upon maturity. In the following section, I propose an econometric model for all contracts that trade on a given commodity.

3.   PARTIALLY OVERLAPPING TIME SERIES (POTS) MODEL

I begin this section by discussing some features of CBOT corn futures. This discussion serves to motivate the POTS model, which follows in Section 3.1. Contracts on corn at the CBOT mature five times per year: March, May, July, September, and December. The contracts typically start trading a year and a half to two years before delivery, but the start date varies across contracts. Consequently, seven to nine contracts are typically trading on a given day, each with different delivery dates up to two years into the future. The exact number of contracts trading is time varying, but deterministic; it can be treated as exogenous. Thus, CBOT corn futures present a sequence of partially overlapping time series of length 18 months to two years.

State-level data on corn inventory exist only at the quarterly frequency, so they are uninformative about daily volatility. Nonetheless, because inventory only matters to the extent that it ties old-crop and new-crop prices together, the relationship between old-crop and new-crop prices contains all of the relevant information on inventory. Thus, the POTS model includes two common factors for futures prices: one for old-crop contracts and one for new-crop contracts. When inventory is high, the two factors move together, but when inventory is low, the link between them breaks. To elucidate this two-factor structure, Figure 1 shows the structure of the partially overlapping time series for CBOT corn over a three-year period from 1991-1993.   Each horizontal line in the figure

5

indicates a different contract and the span over which it trades. For example, contract 11 begins trading in September 1991 and reaches delivery in March 1993.

Consider the nine contracts trading on November 8, 1991 (indicated by a vertical line in Figure 1). The next four contracts to reach delivery after November 8 are the December, March, May, and July contracts (numbers 5-8 in Figure 1). Each of these four contracts must deliver from the current set of inventory, because the next U.S. corn harvest does not begin until September 1992. Consequently, these four contracts are all priced by the old-crop common factor. In contrast, the December, March, May, and July contracts for the 1992-93 crop year (numbers 10-13 in Figure 1) all deliver after the 1992 harvest and are therefore priced by the new-crop common factor.

The appropriate classification of the September contract on November 8 (number 9 in Figure 1) is less clear because it comes to delivery after the harvest has started. According to U.S. Department of Agriculture data, approximately five percent of the crop is harvested by mid September, so some new-crop corn could potentially be delivered on the September contract. However, sufficient corn is unlikely to have been harvested by the delivery deadline to sufficiently replenish inventory, so a mix of the old-crop and new-crop factors could determine the September price. In summary, on November 8, 1991, four contracts are priced by the old-crop factor (solid line in Figure 1), four contracts are priced by the new-crop factor (dotted line in Figure 1) and one contract is affected by a mix of the two factors. The switch from new-crop to old-crop status occurs on October 1, the first month after the last contract on the previous crop ceases trading.

The price change on a particular contract is a linear combination of the common factors and an idiosyncratic term. As discussed in Section 2, serial correlation in futures

6

price changes is nonexistent because there is no risk premium. Nevertheless, the response of prices to news shocks may vary stochastically and deterministically. I model the stochastic component using a GARCH model (Engle, 1982, Bollerslev, 1986).[1] The deterministic component of volatility arises because news is likely to cause a greater price change if it arrives close to harvest than if it arrives at another time. In addition, the factors may affect distant contracts less than the nearby contracts because of the Samuelson effect. Thus, the POTS model is a factor model where the factor loading depends on the season and the time to delivery. Finally, the proportion of the variance explained by the common factors may not be constant. For example, institutional frictions regarding delivery may cause contracts to have a higher idiosyncratic component close to delivery. Consequently, I allow the variance of the idiosyncratic term to also depend on the season and the time to delivery.

3.1   The POTS model

Given the above discussion, the POTS model possesses the following four features: (i) two common factors, (ii) time varying conditional heteroskedasticity, (iii) both time-to-delivery and cross-sectional variation in the factor loadings, and (iv) both time-to-delivery and cross-sectional variation in the innovation variances.

I index the price of a futures contract, F, with two subscripts; d represents the number of trading days until the first day of the delivery month and t represents the date on which a particular price observation occurs. This (d, t) pair is sufficient to identify any price observation in the sample. The model is

---

[1] See, for example, Goodwin and Schnepf (2000) and Baillie and Myers (1991) for applications of GARCH models to futures markets.

$$\Delta F_{d,t} = \theta_{d,t} c'_{d,t} \varepsilon_t + \lambda_{d,t} u_{d,t},$$ (1)

where $\Delta F_{d,t} = F_{d,t} - F_{d+1,t-1}$ and $\theta_{d,t}$ and $\lambda_{d,t}$ represent the factor loading and innovation

standard deviation, which are deterministic functions of d and t. The 2×1 vector ε_t

denotes the factors and the vector $c_{d,t}$ selects which factor, or linear combination of

factors, applies to the contract defined by (d, t).

For identification, assume

$$\Omega \equiv E(\varepsilon_t \varepsilon'_t) = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

$c'_{d,t} \Omega c_{d,t} = 1$, $E(\varepsilon_s u_{d,t}) = 0$, and $E(u_{d,t}^2) = 1$ for all s, d, t. In addition, assume that

$E(u_{d,t} u_{d',t}) = 0$ for $d \neq d'$. These assumptions imply that the factor loadings and the

innovation variances determine the scale of $\Delta F_{d,t}$, i.e., $E(\Delta F_{d,t}^2) = \theta_{d,t}^2 + \lambda_{d,t}^2$. Based on

the documented lack of a risk premium in commodity futures prices, I assume

$E(\Delta F_{d,t} \mid \mathfrak{I}^{t-1}) = 0$, where $\mathfrak{I}^{t-1}$ denotes the information set, which contains past prices.[2]

This assumption implies that the variance of long-horizon changes in futures prices can

be calculated directly from (1) as

$$E(F_{d,t} - F_{d+k,t-k})^2 = \sum_{i=0}^{k-1} (\theta_{d+i,t-i}^2 + \lambda_{d+i,t-i}^2).$$

Furthermore, the martingale difference sequence assumption on $\Delta F_{d,t}$ implies that

$F_{d+k,t-k}$ is the mean-square optimal forecast of $F_{d,t}$ given information at time t–k; the

forecast error variance from any other forecast of $F_{d,t}$ exceeds $\sum_{i=0}^{k-1} (\theta_{d+i,t-i}^2 + \lambda_{d+i,t-i}^2)$.

This implication underlies traditional tests of futures market efficiency.

---

[2] Note that, because $\Delta F_{d,t}$ is a martingale difference sequence, the change in the log futures price has a negative (and possibly time varying) conditional mean. Thus, it is cleaner to directly model the level price change than the log price change.

For most contracts, it is obvious which factor is relevant on a given day. In such cases $c_{d,t}$ has one element equal to one and the other element equal to zero. However, when the appropriate factor is unknown, such as for the September contract on CBOT corn, $c_{d,t}$ must be estimated under the constraint $c'_{d,t}\Omega c_{d,t}=1$. To clarify this constraint, define $c_{d,t} \equiv (\delta_1 \quad \delta_2)'$ where $\delta_1 \geq 0$ and $\delta_2 \geq 0$, so that $c'_{d,t}\Omega c_{d,t}=1$ implies $\delta_1^2 + 2\rho\delta_1\delta_2 + \delta_2^2 = 1$. Thus, one free parameter δ_1 is to be estimated and the constraint residually determines δ_2. The term δ_1^2 represents the proportion of the variance due to the first factor, δ_2^2 denotes the proportion of the variance due to the second factor, and the remainder of the variance is common to both factors. Future extensions of the POTS model could allow δ_1 and δ_2 to be flexible functions of d and t, rather than holding them constant.

By stacking all observations on a given day, the model in (1) can be written as

$$\Delta F_t = \theta_t c_t \varepsilon_t + \lambda_t u_t,$$

where θ_t and λ_t denote diagonal matrices containing the factor loading and innovation standard deviation terms and c_t is a $n_t \times 2$ vector, where n_t denotes the number of contracts trading on date t. Note that, unlike typical factor models, these data comprise an unbalanced panel.

The common factors exhibit time varying conditional volatility, denoted by $E(\varepsilon_t\varepsilon'_t | \mathfrak{I}^{t-1}) = H_t$, where I model H_t using a BEKK-type GARCH model (Engle and Kroner, 1985). The model is

$$H_t = \omega + \beta H_{t-1}\beta' + \alpha E(\varepsilon_{t-1}\varepsilon'_{t-1} | \mathfrak{I}^{t-1})\alpha', \tag{2}$$

where ω, α, and β denote 2×2 parameter matrices. Because the unconditional variance of each factor equals one, the matrix ω contains only one free parameter. This parameter ρ

measures the correlation between the factors. Assuming that H_t is stationary, it follows that ω depends on ρ, α, and β according to the formula $\omega = \Omega - \beta\Omega\beta' - \alpha\Omega\alpha'$.

Like the latent factor ARCH model of Diebold and Nerlove (1989), equation (2) specifies the conditional variance as measurable-$\mathfrak{I}^{t-1}$ to preserve a recursive relation for H_t. If the latent variable ε_t were included in (2), then new data would bring information about past values of H_t and the dynamic structure would be substantially more complicated. Specifying the conditional volatility as a function of observed data simplifies forecasting and estimation. The measurable-$\mathfrak{I}^{t-1}$ innovation term in (2) is

$$E(\varepsilon_{t-1}\varepsilon'_{t-1} \mid \mathfrak{I}^{t-1}) = \varepsilon_{t-1|t-1}\varepsilon'_{t-1|t-1} + P_{t-1|t-1},$$

where $\varepsilon_{t-1|t-1} \equiv E(\varepsilon_{t-1} \mid \mathfrak{I}^{t-1})$ and $P_{t-1|t-1} \equiv E\big((\varepsilon_{t-1} - \varepsilon_{t-1|t-1})(\varepsilon_{t-1} - \varepsilon_{t-1|t-1})' \mid \mathfrak{I}^{t-1}\big)$. Expressions for $\varepsilon_{t-1|t-1}$ and $P_{t-1|t-1}$ can be obtained from the Kalman filter (Hamilton 1994, Diebold and Nerlove 1989). Defining $\Sigma_t \equiv E(\Delta F_t \Delta F'_t \mid \mathfrak{I}^{t-1}) = \theta_t c_t H_t c'_t \theta_t + \lambda_t^2$ and assuming normality, the Kalman filter yields

$$\varepsilon_{t|t} = H_t c'_t \theta_t \Sigma_t^{-1} \Delta F_t$$

$$P_{t|t} = H_t - H_t c'_t \theta_t \Sigma_t^{-1} \theta_t c_t H_t.$$

Considerable scope exists for extensions to the model in (2). For example, one could allow volatility to react differently to positive shocks than to negative shocks; news about reduced supply or increased demand may affect volatility more or less than news about increased supply or reduced demand. Extending the dynamic specification to allow for long memory or a second lag could also be fruitful. Such extensions will be easiest if the recursive nature of the model is preserved by specifying innovations in H_t as functions of the conditional moments $E(\varepsilon_{t-i}\varepsilon'_{t-i} \mid \mathfrak{I}^{t-i})$ and $E(\varepsilon_{t-i} \mid \mathfrak{I}^{t-i})$.

3.2 Modeling the factor loadings

To model the factor loadings and innovation standard deviations, I use cubic spline functions with a small set of nodes as in Engle and Russell (1998). This type of spline is a flexible parametric model that consists of a sequence of connected cubic polynomial functions. These cubics connect at the nodes, which are chosen *a priori* and are typically spaced evenly across the domain of the function. At each node, the adjoining cubic functions are constrained to have equal value and slope. In the POTS model, the spline functions capture deterministic effects of the season and the time to delivery and are linear in their parameters, which enables standard inference conditional on the nodes.

Commodity futures typically reach delivery only a few times each year and typically only a few contracts trade on any particular day. Therefore, given a date t, the corresponding value of d will be one of a small number of values determined by the contracts trading on date t. It follows that all of the data lie on a small number of lines on the (d, t) plane and not throughout the (d, t) plane. Thus, I do not attempt to model the factor loadings and innovation variances throughout the entire (d, t) plane. Rather, I fit the spline functions only to the observed combinations of (d, t). To this end, I estimate a separate spline for each delivery month. For example, for CBOT corn this strategy requires estimating five different splines; one each for the March, May, July, September, and December contracts. These splines also capture seasonality because the delivery month and time to delivery uniquely determine the date of a particular price observation.

For a given delivery month, the factor loading and innovation standard deviation spline functions take the form

$$\theta_{d,t} = \sum_{j=1}^{K} \left(\phi_{0j} + \phi_{1j}(d_t - k_{j-1}) + \phi_{2j}(d_t - k_{j-1})^2 + \phi_{3j}(d_t - k_{j-1})^3 \right) I_{jt},$$

$$\lambda_{d,t} = \sum_{j=1}^{K} \left(\gamma_{0j} + \gamma_{1j}(d_t - k_{j-1}) + \gamma_{2j}(d_t - k_{j-1})^2 + \gamma_{3j}(d_t - k_{j-1})^3 \right) I_{jt},$$

where $I_{jt} = 1(k_{j-1} \le d_t \le k_j)$ denotes an indicator function and ϕ_{ij} and γ_{ij} denote parameters. The variable d_t denotes the time to delivery on date t for the contract of interest, and the nodes k_0, k_1, …, k_K are chosen *a priori*.[3] The spline constrains the value and slope of adjoining cubics to be equal at the nodes. For example, for the θ spline

$$\phi_{0\,j+1} = \phi_{0j} + \phi_{1j}(k_j - k_{j-1}) + \phi_{2j}(k_j - k_{j-1})^2 + \phi_{3j}(k_j - k_{j-1})^3$$

$$\phi_{1\,j+1} = \phi_{1j} + 2\phi_{2j}(k_j - k_{j-1}) + 3\phi_{3j}(k_j - k_{j-1})^2,$$

for all $j = 1, 2, \ldots, K-1$. Two final constraints are that the slope equals zero at the end points, i.e., $\phi_{11} = 0$ and $\phi_{1K} + 2\phi_{2K}(k_K - k_{K-1}) + 3\phi_{3K}(k_K - k_{K-1})^2 = 0$. Each spline is a linear combination of $2K$ free parameters. Thus, given asymptotic normality of the parameter estimates, asymptotic confidence intervals for the splines follow directly.

3.3   Estimation

To estimate the POTS model, I maximize the Gaussian likelihood function. Conditional on past prices, the first two moments of ΔF_t are $E(\Delta F_t \mid \mathfrak{I}^{t-1}) = 0$ and $E(\Delta F_t \Delta F_t' \mid \mathfrak{I}^{t-1}) \equiv \Sigma_t = \theta_t c_t H_t c_t' \theta_t + \lambda_t^2$, so the likelihood function is

$$L = \sum_{t=1}^{T} \log f(\Delta F_t \mid \mathfrak{I}^{t-1})$$

$$= -\frac{\bar{n}T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log|\Sigma_t| - \frac{1}{2}\sum_{t=1}^{T}\Delta F_t'\Sigma_t^{-1}\Delta F_t, \tag{4}$$

where $\bar{n} = T^{-1}\sum_{t=1}^{T} n_t$. This function can be maximized numerically.

---

[3] Choosing the number of node points, K, is analogous to choosing the bandwidth in nonparametric analysis; it exhibits the same bias-efficiency tradeoff and the same difficulty in consistent estimation. Conditional on the number of nodes, however, the exact location of the node points has little influence on the properties of the estimator in the same way that, conditional on bandwidth, the choice of kernel is not usually important in nonparametric analysis.

The likelihood function in (4) is highly nonlinear because the spline parameters appear in the GARCH equation through $\varepsilon_{t|t}$ and $P_{t|t}$. For this reason, analytic expressions for the gradient and hessian are infeasible. This complication coupled with the potentially large number of spline parameters can make numerical optimization very slow. However, conditional on ε_t and H_t, the estimation problem is much simpler and analytic expressions for the gradient and hessian with respect to the spline parameters exist. I use this feature to form an approximate EM algorithm (Dempster *et al.*, 1977).

In general, the EM algorithm maximizes a likelihood function by alternately computing the expectation of the complete data likelihood with respect to the latent variable (the *E*-step) and maximizing this expected likelihood with respect to the parameters (the *M*-step). The complete data likelihood function for the POTS model is

$$L_c = \sum_{t=1}^{T} \log f(\Delta F_t \mid \varepsilon_t, \mathfrak{I}^{t-1}) + \sum_{t=1}^{T} \log f(\varepsilon_t \mid \mathfrak{I}^{t-1})$$

$$= -\frac{\overline{n}T}{2}\log(2\pi) - \sum_{t=1}^{T}\log|\lambda_t| - \frac{1}{2}\sum_{t=1}^{T}(\Delta F_t - \theta_t c_t \varepsilon_t)'\lambda_t^{-2}(\Delta F_t - \theta_t c_t \varepsilon_t)$$

$$-\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log H_t - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t' H_t^{-1}\varepsilon_t .$$

Conditional on the observed data, the expected complete data likelihood is

$$E(L_c \mid \mathfrak{I}^T) = -\frac{\overline{n}T}{2}\log(2\pi) - \sum_{t=1}^{T}\log|\lambda_t| - \frac{1}{2}\sum_{t=1}^{T}(\Delta F_t - \theta_t c_t \varepsilon_{t|t})'\lambda_t^{-2}(\Delta F_t - \theta_t c_t \varepsilon_{t|t})$$

$$-\frac{1}{2}\sum_{t=1}^{T}\mathrm{tr}\left(c_t'\theta_t\lambda_t^{-2}\theta_t c_t P_{t|t}\right) - \frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log H_t - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_{t|t}' H_t^{-1}\varepsilon_{t|t} - \frac{1}{2}\sum_{t=1}^{T}\mathrm{tr}\left(H_t^{-1}P_{t|t}\right), \quad (5)$$

where tr() denotes the trace operator. To simplify computation, I keep H_t fixed when maximizing (5) in the *M*-step, implying that the resulting parameter estimates merely approximate the maximum likelihood estimates. Thus, I take the final steps to the maximum using the estimates from this approximate EM algorithm as starting values for a numerical algorithm such as BHHH applied to (4).

In summary, I use the following algorithm to maximize the likelihood in (4):

1. Use the Kalman filter to obtain $\varepsilon_{t|t}$ and $P_{t|t}$ (*E*-step).

2. Keeping H_t fixed, maximize (5) with respect to spline parameters using Newton-Raphson (*M*-step).

3. Still keeping H_t fixed, iterate on the *E*- and *M*-steps until convergence.

4. Keeping the spline parameters fixed, maximize the likelihood in (4) numerically with respect to other parameters in the model.

5. Repeat 1-4 until convergence

6. Using estimates from step 5 as starting values, use a numerical algorithm such as BHHH to take the final steps to the maximum of the likelihood in (4).

Steps 1-5 in this algorithm markedly improve computation time because they generate starting values for the numerical algorithm in step 6 that are close to the maximum likelihood estimates. In contrast, a numerical algorithm such as BHHH that begins with arbitrary starting values can be very slow to converge.

4. EMPIRICAL RESULTS FOR CBOT CORN

In this section, I test the theory of storage and estimate the Samuelson effect by applying the POTS model to corn futures on the CBOT. CBOT corn is the most actively traded agricultural futures contract. Active trading leads to liquid markets and observed prices that correspond to the market's valuation of the contract. Thus, CBOT corn provides an ideal medium for illustrating the POTS model.

I estimate the POTS model using daily settlement prices for trading days from January 1, 1991 to December 31, 2000. The sample includes data on 53 different contracts. Specifically, the sample includes contracts that delivered in March, May, July,

September, and December for each year from 1991-2000 and on the March, May, and July contracts for 2001. Because I use trading days from 1991-2000, the sample includes a subset of the trading days for some contracts. For example, the sample includes only the last three months of trading on the March 1991 contract and does not include the last three months of trading on the March 2001 contract.[4] There are 19,745 total observations.

During the sample period, the average settlement price across all trading days and all contracts was 263 cents per bushel, with a minimum of 175 and a maximum of 548. The mean daily price change was close to zero at –0.02 cents, and the standard deviation was 2.96. The average daily trading volume was 57,154 contracts and average number of contracts open at the end of each day was 323,064, where each contract requires delivery of 5000 bushels.

The last trading day on a given CBOT corn contract is the business day prior to the 15th calendar day of the contract month. Holders of a short position can deliver any time between the first day of the delivery month and the second business day following the last trading day of the delivery month. The March, May, and September contracts begin trading approximately a year and a half before delivery. Specifically, they begin in September, December, and May, respectively. The July and December contracts begin trading at various horizons throughout the sample, but typically at least two years before delivery. The mean number of contracts traded per day equals 7.8, and on 85 percent of days in the sample the number of contracts traded is either seven, eight, or nine.

---

[4] In the year 2000, the data include a subset of the contracts actually being traded because I do not have data on the September 2001, December 2001, March 2002, May 2002, and July 2002 contracts, all of which were trading by the end of December 2000. Two atypical contracts also traded during the sample period. These contracts reached delivery in November 2000 and January 2001, but I exclude them from the sample because they are the only contracts for November and January delivery during the sample period.

For the spline functions, I place nodes at zero, 126, and 252 days before delivery. These nodes approximately correspond to the first day of the delivery month, the first day of the month six months before delivery, and the first day of the month one year before delivery. For the July and December contracts, I also include nodes at 378 days from the delivery date (approximately one and a half years before delivery) because these contracts trade for a longer period of time.[5] I use the same set of nodes for λ and θ. I estimate a total of ten splines; one for the factor loading λ and one for the innovation standard deviation θ for each of the five delivery months.

Table 1 presents results from the estimation of a one-factor model and a two-factor model. The two-factor model specifies α and β in (2) as diagonal matrices with nonzero elements α_{11}, α_{22}, β_{11}, and β_{22}. To guarantee a positive and stationary conditional variance, I directly estimate α_{11}^2, α_{22}^2, $\alpha_{11}^2 + \beta_{11}^2$, and $\alpha_{22}^2 + \beta_{22}^2$ constraining each to lie in the (0,1) interval. Table 1 shows that the estimates of $\alpha_{11}^2 + \beta_{11}^2$ and $\alpha_{22}^2 + \beta_{22}^2$ equal 0.987 for the two-factor model and 0.983 for the one-factor model. These estimates indicate a high degree of persistence in the volatility of the common factor(s).

As discussed in Section 3, the September contract reaches delivery after the harvest begins, and thus it may be affected by both the old-crop and new-crop factors. Table 1 shows that the estimated value of δ_1 equals 0.338. Under the scale constraint $c'_{d,t} \Omega c_{d,t} = \delta_1^2 + 2\rho\delta_1\delta_2 + \delta_2^2 = 1$, this estimate implies that the proportion of the variance due to the old crop is $\delta_1^2 = 0.114$ and the proportion of the variance due to the new crop is $\delta_2^2 = 0.425$. These two proportions do not sum to one, because the high correlation

---

[5] I also estimated models with a larger set of nodes, but I do not report the results because these models were inferior by the Bayesian Information Criterion (BIC) and generated the same qualitative conclusions as the reported models.

between the factors means that much of the variation is common to both factors. This estimate indicates that the September contract relates more closely to the new-crop prices than the old-crop prices. In other words, enough corn is harvested by the end of the September delivery period to mitigate most of any inventory shortfall.

Adding the second factor generates a substantial improvement in likelihood over the one-factor model. This improvement indicates that breaks in the link between the old and new crops are empirically relevant. The importance of the second factor is also evidenced by the correlation parameter ρ being significantly less than one. However, although the second factor is significant, the correlation between the factors is high at 0.928. These high correlations go hand in hand with the relatively small increase in the proportion of the variance explained by the factors[6] when the second factor is added; adding the second factor increases the overall proportion from 0.848 to 0.932.

Figure 2 illustrates the estimated spline functions from the two-factor model by showing the unconditional variance as a function of the time to delivery. I compute the unconditional variance as $(\theta_{d,t}^2 + \lambda_{d,t}^2)$. Figure 2 contains separate plots for each of the five contracts with associated 95 percent confidence intervals. The confidence intervals are computed using the delta method and the fact that λ and θ are linear in the estimated parameters. The relatively tight confidence intervals around the spline estimates permit some inference about the relationship between volatility and the time to delivery. Most notably, volatility spikes during the delivery month for all but the December contract, and it is low at long horizons. I discuss these features further in Sections 4.1 and 4.2.

---

[6] This proportion is measured as $\sum_{d,t}(\theta_{d,t}\varepsilon_{t|t})^2 / \sum_{d,t}\Delta F_{d,t}^2$, where the summation is taken over all observations on a given contract when calculating the proportion by contract and over all observations in the sample when calculating the overall proportion.

The March contract is most closely related to the factors, with 97.6 percent of the variation explained by the factors in the two-factor model (see Table 1). The factors explain the least in the September and December contracts, where the proportions of the variation explained are 0.910 and 0.886, respectively. To see how much variation the factors explain through the life of a contract, consider Figure 3. This figure plots the proportion of the model variance in each of the five contracts that is explained by the common factors against the date of trading. I compute this proportion from the spline functions as $\theta_{d,t}^2/(\theta_{d,t}^2 + \lambda_{d,t}^2) = \theta_{d,t}^2/E(\Delta F_{d,t}^2)$. The figure covers three crop years: the delivery crop year (year zero) and the two years preceding it (year one and year two).

According to the theory of storage, the common factors should explain most, if not all, of the variance in prices. Figure 3 shows that, for the December, March, May, and July contracts, the common factors explain over 95 percent of the variation for most of the duration of the contracts. The curves fall substantially below 100 percent in two cases. First, the proportion of the variance explained by the factors drops sharply in the delivery month for all contracts. Second, the proportion of the variance explained by the factors is smaller at points far from delivery for the July, September, and December contracts. I discuss the first of these cases in Section 4.1 below. I then discuss the second case in Section 4.2, before addressing the Samuelson effect and seasonality in Section 4.3, and the theory of storage in Section 4.4.

4.1   Delivery month volatility

Figure 3 shows that the proportion of the model variance explained by the factors drops sharply in the delivery month for all five contracts. For the May contract, the drop begins two months before delivery in early March. For the September contract, the

18

proportion drops similarly in March, six months before delivery. These drops indicate that the influence of the March, May, and September contracts on the old-crop common factor reduces after the beginning of March in year zero. Consequently, the July contract drives this factor from early March until early July, as indicated in Figure 3 by the high proportion of the variance explained by the factor for the July contract during this period. From July until the end of the crop year, the September contract is the only contract with an old-crop component still trading. These volatility patterns imply that the March, May, July, and September contracts are less related to each other in the last half of the crop year than earlier in their lives.

The relatively high idiosyncratic volatility around the delivery period indicates that delivery costs are high. In general, delivery costs will be lowest when a liquid spot market for the commodity exists at the delivery location. In such cases, holders of a short position can easily buy the commodity and deliver on the contract. If the spot market at the delivery location is illiquid, then its price will be volatile for reasons unrelated to aggregate fundamentals, and therefore the futures price for imminent delivery will also exhibit high idiosyncratic volatility. Even if liquid markets exist at points far from the delivery location, the cost of transporting the commodity from this market to the delivery location in time to meet an impending delivery obligation will be prohibitive. Thus, high idiosyncratic volatility around the delivery period indicates inefficiency in the delivery process. For CBOT corn, this inefficiency is largest in the delivery month for all contracts, but it is also significant in the months leading up to delivery for the May, July, and September contracts.

For all but the last year of this sample, delivery on CBOT corn futures could be made in either Chicago Illinois or in Toledo Ohio. A delivery point in St. Louis Missouri was added beginning with the December 1993 contract. Concern that these delivery points were peripheral to the main corn supply channels led to a change in delivery specifications. Beginning with the March 2000 contract, delivery could be made anywhere along the Illinois River between Chicago and Pekin Illinois. My results indicate that such a change in the delivery institution was necessary. As more data becomes available, it will become apparent whether this change was successful in improving the efficiency of the futures contract.

4.2   Contracts far from delivery

Figures 2 and 3 reveal that, for the December, July and September contracts, both the total variance and the proportion of the variance explained by the factors are small at long horizons. This pattern is reinforced by Figure 4, which plots the unconditional variance curves from Figure 2 as a function of the trading date. The figure covers three crop years: the delivery crop year (year zero) and the two years preceding it (year one and year two).

At the times the July and December contracts begin trading, two harvests will occur before delivery. Figure 4 shows that, from the start of trading until the first harvest begins in September of year one, both contracts exhibit relatively low volatility. Volatility then increases during the first harvest, i.e., between September and December. The proportion of the variance explained by the factors mirrors this pattern, as shown in Figure 3. Therefore, it appears that a significant amount of news about the following crop year is irrelevant for contracts on the crop two years ahead. No trade occurs in March or May contracts for corn two crop-years ahead.

20

The September contract also exhibits low volatility and a low proportion of variance explained by the common factors in early trading. In its first four months of trading (May to September), the volatility patterns in the September contract match those of the July and December contracts in the pre-September period of their first year (see the first September curves in Figures 3 and 4). This pattern reflects the high new-crop component in September volatility that is indicated by the low estimated value of δ_1 in Table 1. The final point of note on the September contract is that its volatility drops significantly in the two months before delivery (see Figure 4). This drop may reflect the resolution of uncertainty about how much new-crop corn will be available for delivery. Volatility increases again during the delivery month indicating delivery frictions.

4.3   Samuelson effect and seasonality

Figure 4 demonstrates a nonlinear and nonmonotonic version of the Samuelson effect. Specifically, at a given date, contracts that are closer to delivery typically have higher volatility than more distant contracts. The nonlinearity of the Samuelson effect becomes apparent if volatility is tracked through time. In year two in Figure 4, only the first September contract and the July and December contracts are trading. The volatility of these contracts is relatively low in this period. During year one, volatility steadily increases as the market obtains more information about the year zero harvest. Volatility peaks between September and November of year zero, during the middle of the harvest. The harvest resolves much of the uncertainty about year zero prices, so volatility decreases between November and March. As noted above, the September contract exhibits higher volatility than the other year zero contracts between January and March due to it being more closely related to the new-crop contracts than the old-crop contracts.

21

Beginning in March, volatility increases dramatically for the two remaining year zero contracts, May and July. As indicated in Figure 3, much of this volatility is idiosyncratic and indicates the presence of frictions in delivery arrangements.

Figure 4 reveals substantial seasonality in year zero but little seasonality in year one. The volatility curves exhibit only a slight increase going into the year zero harvest, when a substantial amount of news about the new crop arrives. After the year zero harvest, the variance drops by about 30 percent. One reason for the small seasonal differences in year one is that in many years there were few surprises in the period leading up to harvest and in the harvest season itself. Moreover, when surprises occurred, they often arose at different times of the year. Figure 5 illustrates this point by showing the conditional variance of the two factors as estimated by the GARCH model. The most prominent surprises are indicated by high conditional volatility periods, which occurred in July-September 1991, May-July 1994, January-February 1996, April-October 1996, February-May 1997, July-September 1997, June-July 1998, and July-September 1999.

4.4   Theory of storage

The high proportion of price variation explained by the common factors lends support to the theory of storage. For most of the life of all contracts, the proportion of the variance explained by the factors exceeds 0.95. The theory of storage also has dynamic predictions. Specifically, periods when the old-crop variance exceeds the new-crop variance should coincide with periods of low correlation between the factors, backwardated prices, and low inventory.

Figure 6 illustrates the dynamic predictions of the theory of storage by showing the conditional correlation between the factors, the relative conditional variance, and the

basis slope.[7] The summer of 1996 stands out in Figure 6, as well as Figure 5. Due to a poor harvest in 1995, total stocks of corn were lower during the summer of 1996 than at any point since 1976. In addition, futures prices were in steep backwardation and spot prices reached $5 per bushel, after averaging about $2.50 a bushel in the first half of the decade. Figure 6 reveals that the backwardation in 1996 was indeed associated with a lower conditional correlation and a higher old-crop variance as predicted by the theory.

The fall of 1993 also produced a backwardation, and it also corresponded to lower conditional correlation and a higher old-crop variance. In addition to 1993 and 1996, the conditional correlation decreased and the relative variance increased in the fall of 1991. This period exhibited no prolonged backwardation, but the 1991 crop had a low yield which caused December stocks of corn to be lower than they had been since the shortage of 1984. These low stocks lead to a break in the link between 1991 crop prices and 1992 crop prices. The only other times in the 1990s where December stocks were below their 1991 levels were in 1993 and 1995, corresponding to the other observed periods of backwardation and low conditional correlation. Thus, the dynamic features of the data support the theory of storage.

5. CONCLUSION

In this paper, I develop the POTS model for volatility dynamics in commodity futures markets. The POTS model incorporates common factors across contracts with differing delivery dates as well as time-to-delivery and seasonal effects. Applying the POTS model to CBOT corn futures reveals substantial inefficiency related to delivery. Specifically, all contracts exhibit a large proportion of idiosyncratic variation during the delivery month.

---

[7] I compute the basis slope as the slope coefficient in a regression of the futures prices on a given trading day on the times to delivery on those contracts. I compute a separate basis slope for each day in the sample.

The May, July, and September contracts also show significant idiosyncratic variation in the months leading up to delivery. In addition, the empirical results substantiate the theory of storage and yield estimates of the Samuelson effect.

The POTS model directly applies to futures contracts on any storable commodity with seasonal production. The model could easily be adapted to non-seasonal commodities by specifying the selection variable $c_{d,t}$ to reflect the production and transportation lag, rather than seasonal production patterns. Applying the POTS model in other commodity futures markets will enable additional tests for delivery frictions, the theory of storage, and the Samuelson effect. Also, because it models jointly the whole spectrum of futures prices, the POTS model will be very useful in pricing options on futures contracts and in determining optimal hedging strategies.

I close this paper by using the POTS model to assess the common practice of splicing nearby prices together into one time series to proxy a fundamental spot price. Applied researchers typically treat the resulting series as a continuous, homogeneous time series. Figure 7 shows a plot of the unconditional variance of such a nearby series as estimated by the POTS model. I assume that rollover occurs on the first day of each delivery month. The figure shows that volatility jumps sharply around the delivery period for the May and July contracts and declines steeply for the September contract. The nonlinear heteroskedasticity in this series indicates that it is far from a homogeneous time series.

For applied researchers who desire to represent a fundamental corn price by a single series of futures prices, I make two suggestions. First, the futures contracts should be rolled over two to three months before delivery to avoid delivery month inefficiency. Second, the September contract should be avoided entirely because it has a high new-

crop component. Specifically, I recommend using May contract prices during January and February, prices on the July contract from March though May, prices on the December contract from June through October, and prices on the March contract during November and December. Figure 7 shows that this strategy yields a modified nearby time series with a relatively smooth variance throughout the year, although the variance still increases by 40 percent from the low point in March to the high point in September. Thus, if seasonal heteroskedasticity is accounted for, this modified nearby price series can be treated as a continuous time series. Moreover, the proportion of the variance explained by the factors in the POTS model is close to one throughout the range of the modified nearby series (see Figure 3), so it accurately reflects a fundamental corn price.

REFERENCES

Anderson, R.W. (1985), "Some Determinants of the Volatility of Futures Prices," *Journal of Futures Markets*, 5:331-348.

Anderson, R.W., and J.P. Danthine (1983), "The Time Pattern of Hedging and the Volatility of Futures Markets," *Review of Economic Studies*, 50:249-266.

Baillie, R.T. and R.J. Myers (1991), "Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge," *Journal of Applied Econometrics*, 6:109-124.

Bessembinder, H. (1993), "An Empirical Analysis of Risk Premia in Futures Markets," *Journal of Futures Markets*, 13:611-630.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31:307-327.

Dempster, A.P., N.M. Laird, and D.B. Rubin (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society (Series B)*, 39:1-38.

Diebold, F.X. and M. Nerlove (1989), "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model," *Journal of Applied Econometrics*, 4:1-21.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation," *Econometrica*, 50:987-1008.

Engle, R.F. and K.F. Kroner (1985), "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11:122-150.

Engle, R.F. and J.R. Russell (1998), "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66:1127-1162.

Engle, R.F., V.M. Ng, and M. Rothschild (1990), Asset Pricing with a Factor-ARCH Structure: Empirical Estimates for Treasury Bills," *Journal of Econometrics*, 45:213-237.

Fama, E. and K. French (1988), "Business Cycles and the Behavior of Metals Prices," *Journal of Finance*, 43:1075-1093.

Goodwin, B.K., and R. Schnepf (2000), "Determinants of Endogenous Price Risk in Corn and Wheat Futures Markets," *Journal of Futures Markets*, 20:753-774.

Gray, R.W. (1961), "The Search for a Risk Premium," *Journal of Political Economy*, 69:250-260.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Pres: Princeton.

Keynes, J.M. (1930), *A Treatise on Money (Volume II): The Applied Theory of Money*, MacMillan: London.

Kolb, R.W. (1992), "Is Normal Backwardation Normal?" *Journal of Futures Markets,* 12:75-91.

Ng, V.K. and S.C. Pirrong (1994), "Fundamentals and Volatility: Storage, Spreads and the Dynamics of Metals Prices," *Journal of Business*, 67:203-230.

Routledge, B.R., D.J. Seppi, and C.S. Spatt (2000), "Equilibrium Forward Curves for Commodities," *Journal of Finance*, 55:1297-1338.

Samuelson, P.A. (1965), "Proof That Properly Anticipated Futures Prices Fluctuate Randomly," *Industrial Management Review*, 6:41-49.

Sentana, E. (1998), "The Relation Between Conditionally Heteroskedastic Factor Models and Factor GARCH Models," *Econometrics Journal*, 1:1-9.

Stein, J.L. (1986), *The Economics of Futures Markets*, Basil Blackwell: Oxford.

Streeter, D.H. and W.G. Tomek (1992), "Variability in Soybean Futures Prices: An Integrated Framework," *Journal of Futures Markets*, 12:705-728.

Telser, L.G. (1958), "Futures Trading and the Storage of Cotton and Wheat," *Journal of Political Economy*, 66:233-255.

Williams, J.C. (1987), "Futures Markets: A Consequence of Risk Aversion or Transactions Costs?" *Journal of Political Economy*, 95:1000-1023.

Williams, J.C. (2001), "Commodity Futures and Options," in *Handbook of Agricultural Economics, Vol 1*, B. Gardner and G. Rausser (eds).

Working, H. (1948), "Theory of the Inverse Carrying Charge in Futures Markets," *Journal of Farm Economics*, 30:1-28.

Working, H. (1949), "The Theory of Price of Storage," *American Economic Review*, 39:1254-1262.

Table 1: Estimates of POTS Model

|	1 Factor	2 Factors
ρ		0.928 (0.003)
α_{11}^2	0.097 (0.031)	0.087 (0.004)
α_{22}^2		0.100 (0.006)
$\alpha_{11}^2 + \beta_{11}^2$	0.983 (0.019)	0.987 (0.002)
$\alpha_{22}^2 + \beta_{22}^2$		0.988 (0.001)
$\delta_{1,sept}$		0.338 (0.002)
Diagnostics		
LLF	−45,787,649	−45,783,828
BIC	−45,788,539	−45,784,757
Skewness	0.01	0.06
Kurtosis	8.35	7.26
Q_5 for ε_{1t} (p-value)	6.85 (0.23)	7.04 (0.22)
Q_5 for ε_{2t} (p-value)		6.93 (0.23)
Proportion of Variance Explained by Factors		
Dec contract	0.866	0.886
March contract	0.935	0.976
May contract	0.843	0.946
July contract	0.780	0.946
Sept contract	0.855	0.910
Overall	0.848	0.932

Note: The spline nodes are at six-month intervals from the delivery data, i.e., for 0, 126, 252, 378 days until delivery (the 378 node is only included for July and December). The data contain daily settlement prices on corn futures contracts traded on the CBOT between January 1 1991 and December 31, 2000. Heteroskedasticity consistent standard errors are given in parentheses. BIC is computed as LLF–K log(T), where K equals the number of parameters and T =19,745. The skewness and kurtosis coefficients measure the third and fourth moments of standardized futures price changes $\Delta F_{d,t}\big/\sqrt{E(\Delta F_{d,t}^2 \mid \mathfrak{I}^{t-1})}$. The Q_5 statistics test for serial correlation of up to order five in the standardized factors $\varepsilon_{i,t|t}\big/\sqrt{H_{ii,t}}$ for $i = 1,2$.

# Figure 1: Partially Overlapping Time Series
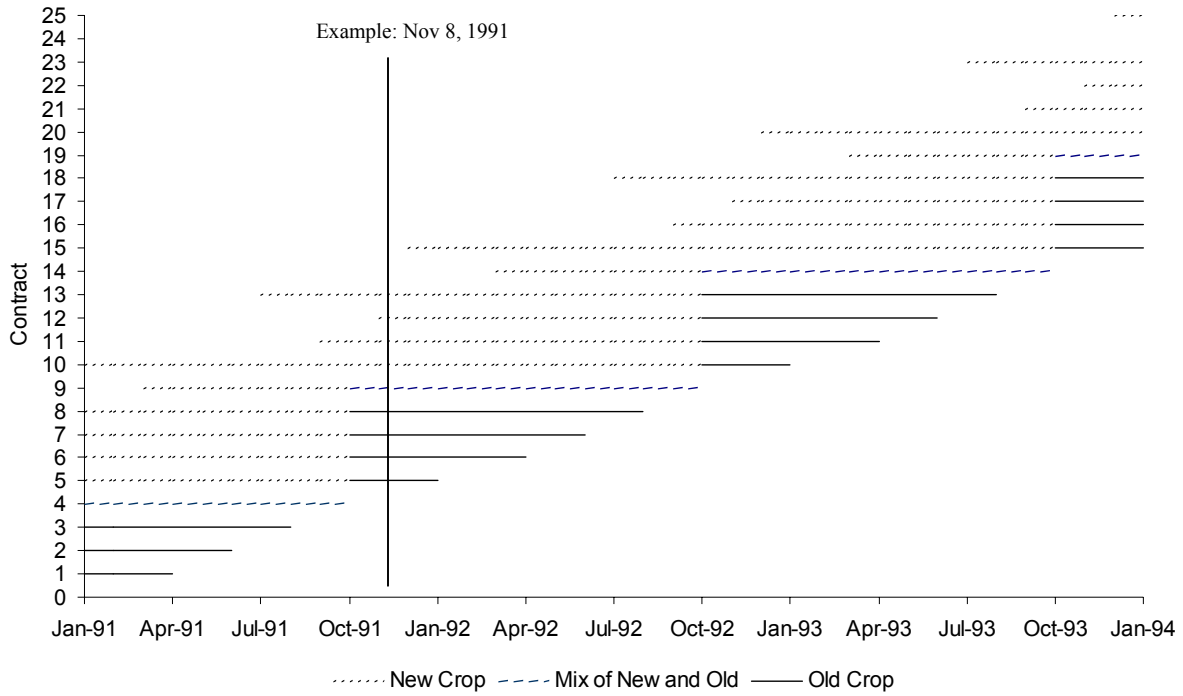
Note: Each horizontal line represents a particular contract on CBOT corn and the span over which it trades. The style of the lines indicates whether a particular contract is for old-crop or new-crop corn, or a mixture of the two. The vertical line at Nov 8, 1991 represents an arbitrary date for illustration purposes.

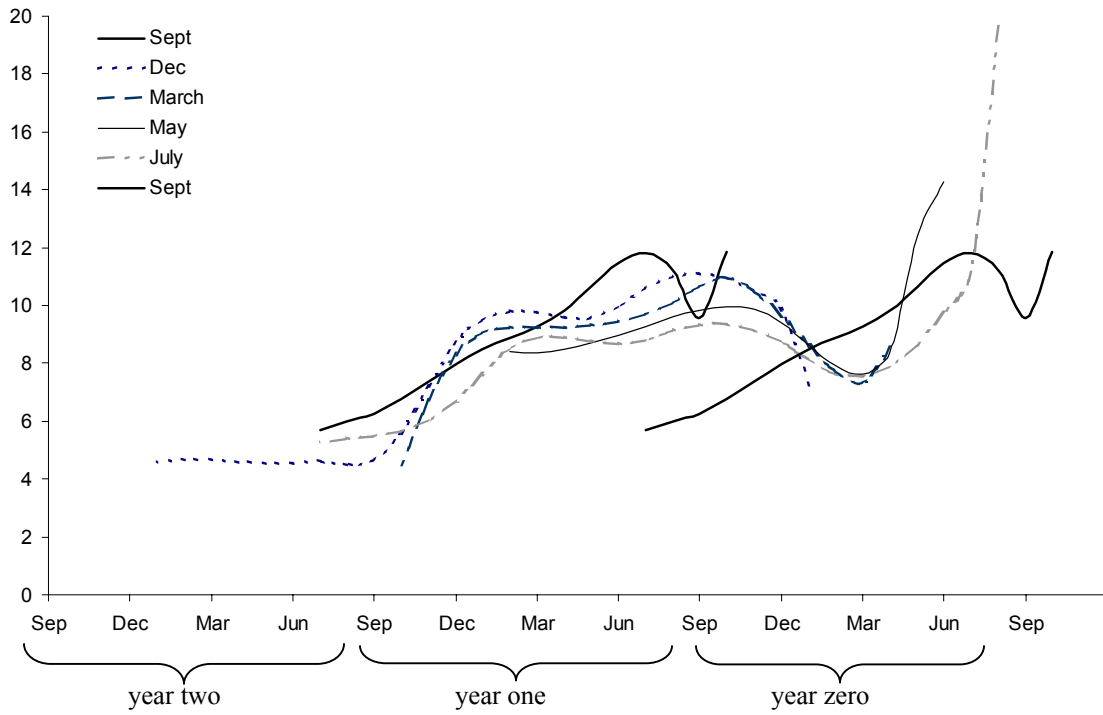# Figure 2: Unconditional Variance as a Function of Time to Delivery

December Contract

March Contract

May Contract

July Contract

September Contract
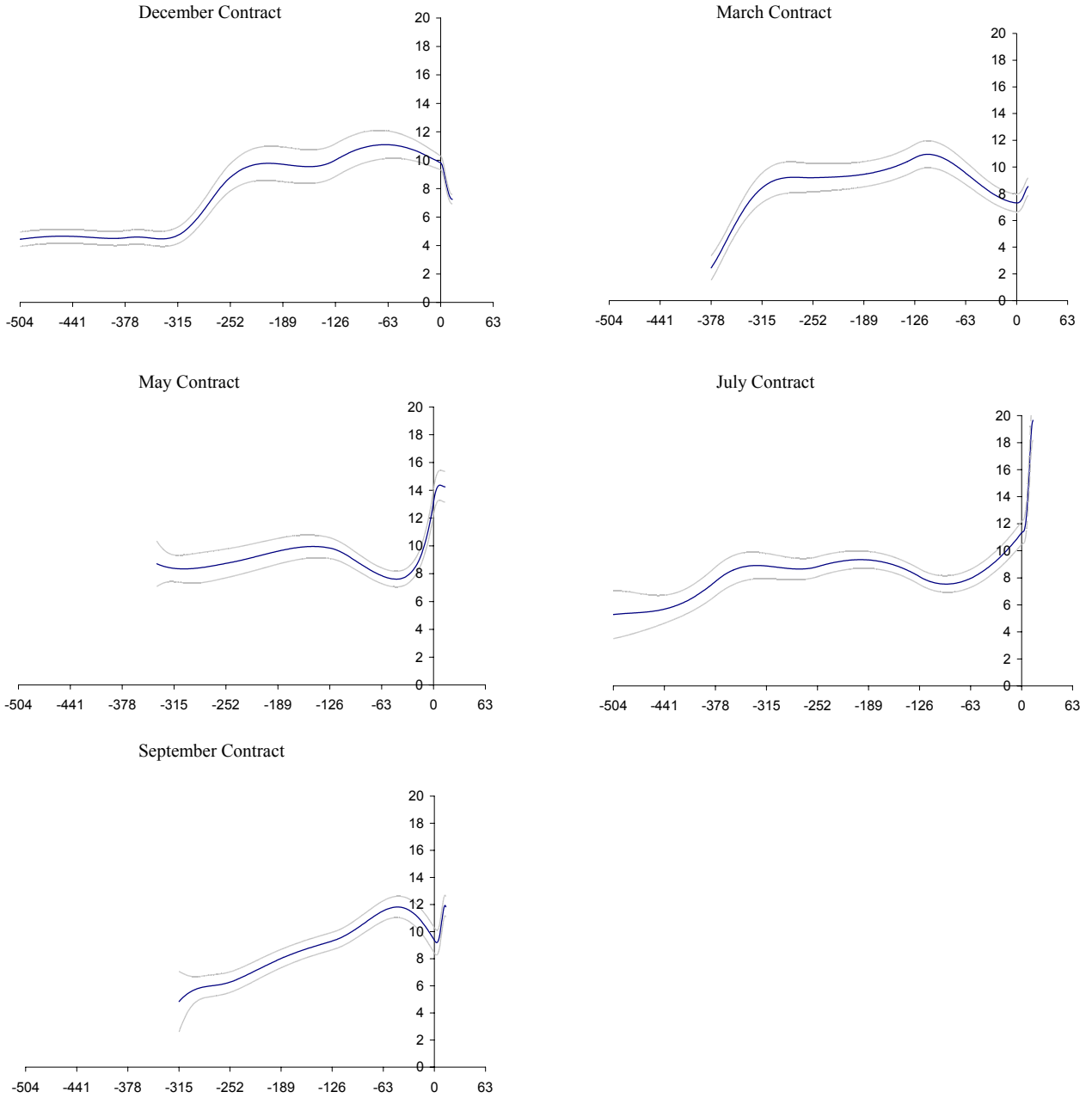
Note: The curves are computed from the estimated spline functions ($\theta^2_{d,t} + \lambda^2_{d,t}$) from the 2-factor model in Table 1. Units on horizontal axis represent time to delivery (d) in number of trading days. One year equals 252 days. Dashed lines indicate 95% confidence intervals.

# Figure 3: Proportion of Model Variance Explained by Common Factors
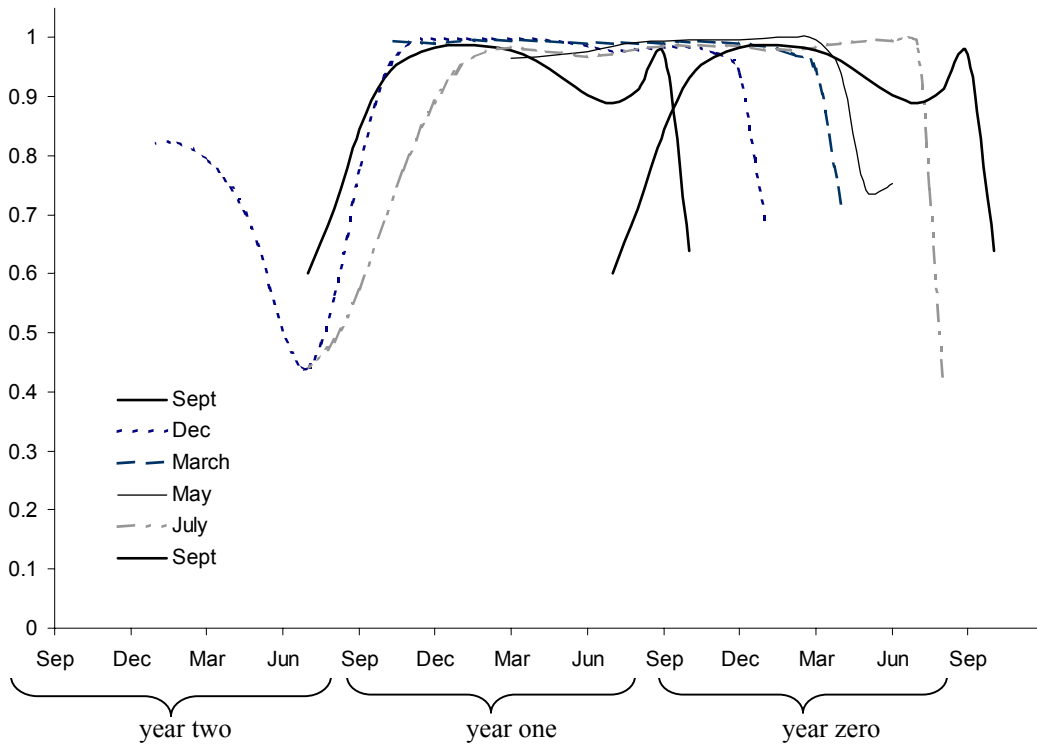
Note: The curves are computed as $\theta_{d,t}^2 / (\theta_{d,t}^2 + \lambda_{d,t}^2)$ from the estimated spline functions in the 2-factor model in Table 1. September is included twice because it is a function of both the old-crop and new-crop factors.

**Figure 4: Unconditional Variance as a Function of Trading Date**

Note: The curves are computed as $\theta_{d,t}^2 + \lambda_{d,t}^2$ from the estimated spline functions in the 2-factor model in Table 1. September is included twice because it is a function of both the old-crop and new-crop factors.

## Figure 5: Conditional Variance of Common Factors
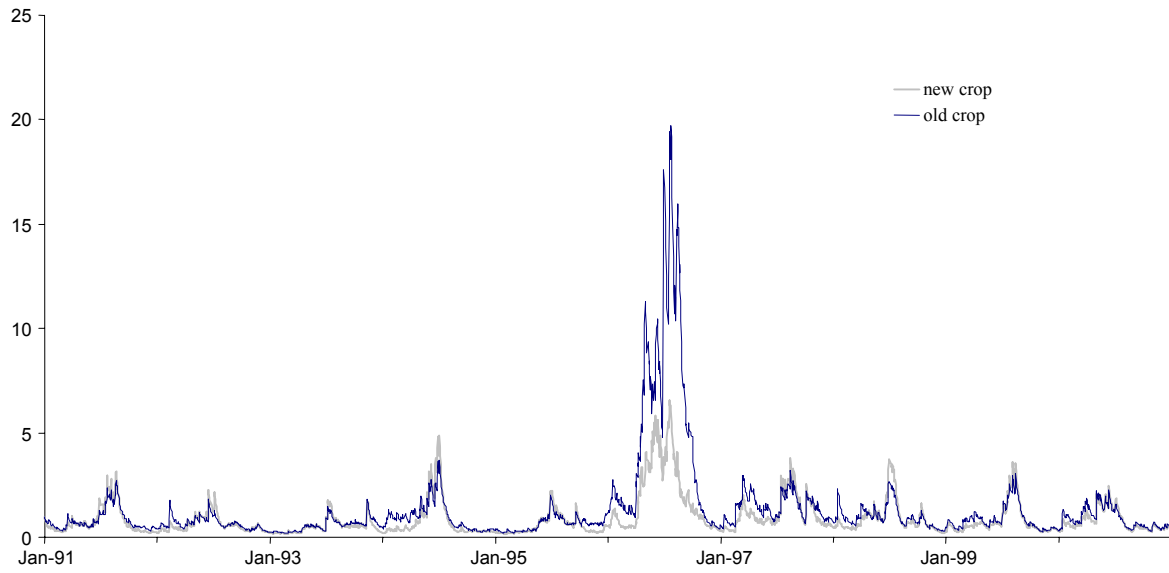
Note: The series are computed from the GARCH estimates of the 2-factor model (see Table 1).

Figure 6: Conditional Correlation and Relative Conditional Variance of Factors

Conditional Correlation

Relative Conditional Variance
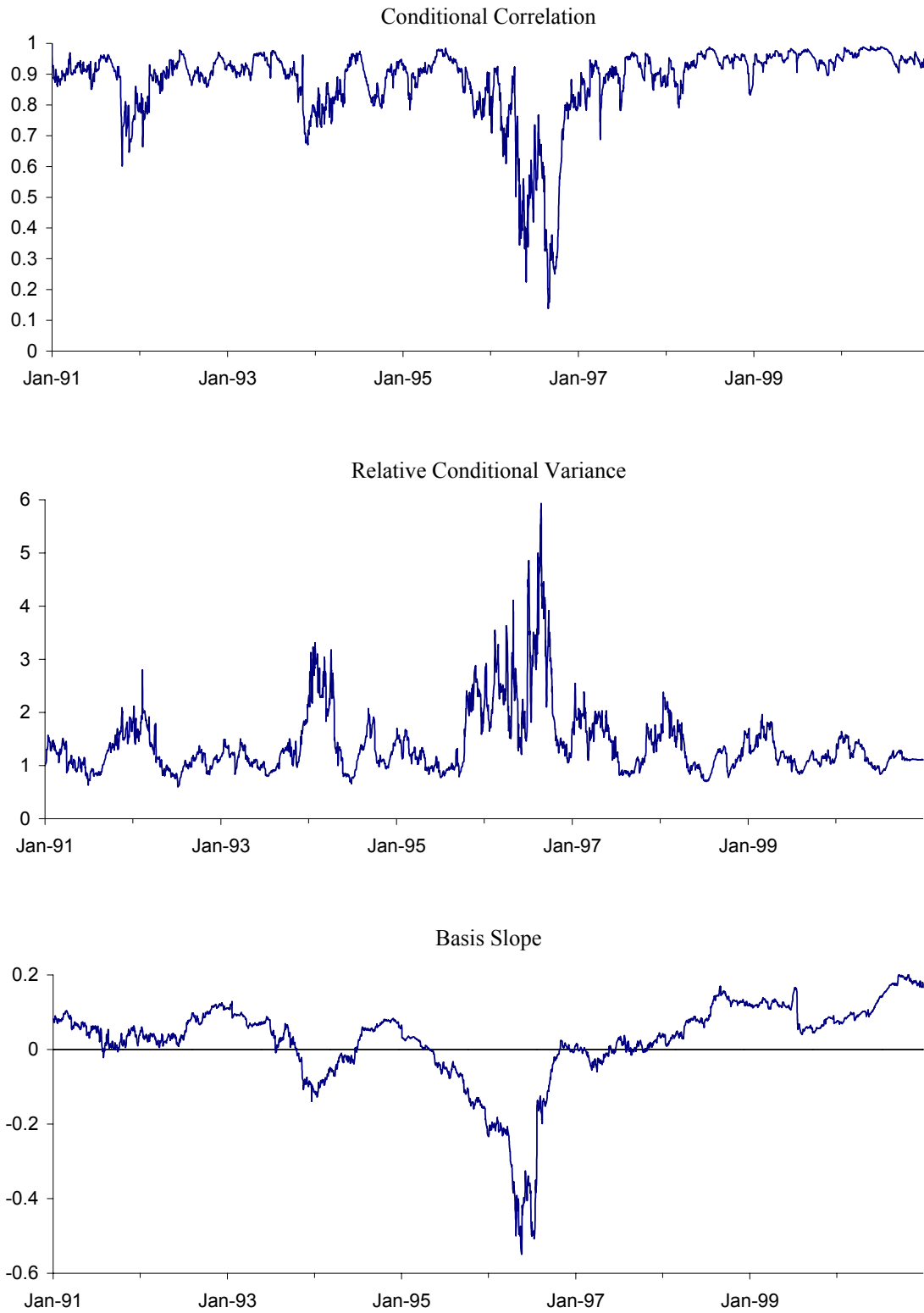
Basis Slope

Note: The conditional correlation and relative conditional variance series are computed from the GARCH estimates for the 2-factor model (see Table 1). The basis slope is computed from daily regressions of futures prices on number of days to delivery.

Figure 7: Unconditional Variance of the Nearby Contract
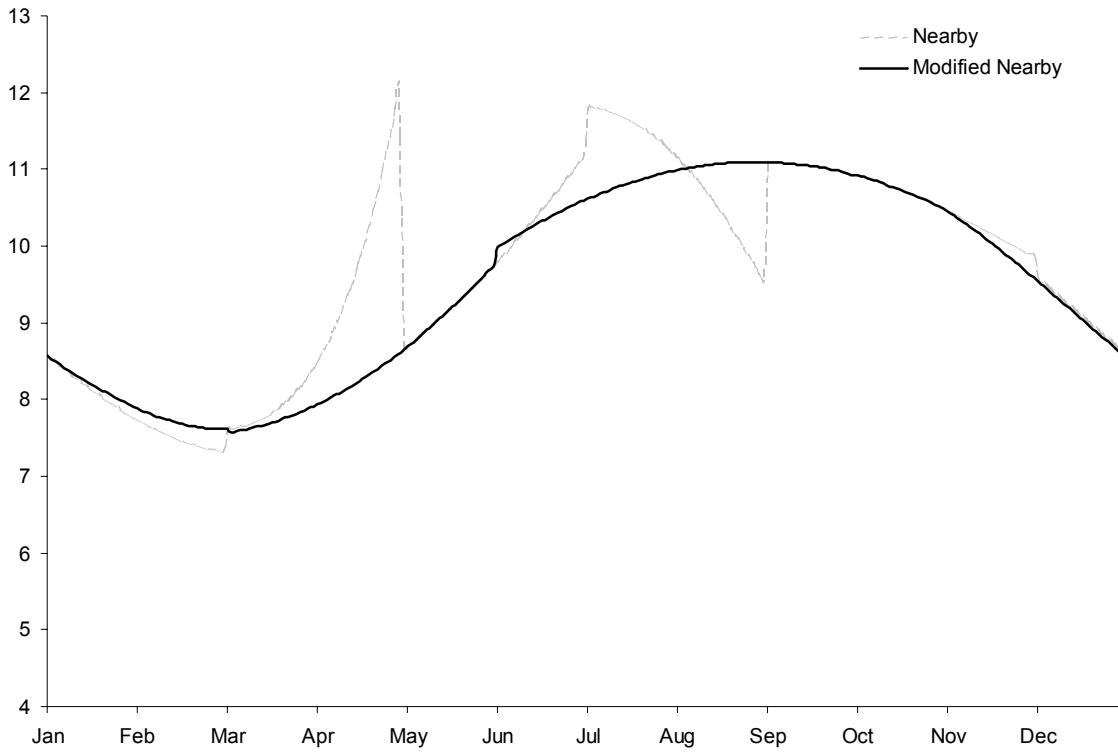
Note: Nearby contract rolls over on first day of delivery month. Modified nearby contract uses May contract for January and February, July contract for March though May, December contract for June through October, and March contract for November and December.