Estimation and Revision of a Sequential Auction Model for the Soybean Futures Current Contract

William D. O'Neill

I. INTRODUCTION

The Kyle (1985) sequential auction model was the first to explicitly account for the process by which market clearing prices converge to the liquidation value of the traded asset while simultaneously accounting for profit maximizing insider trading. His model posited inside traders as those with knowledge of the eventual liquidation asset price; market makers as those who match up buy and sell orders and trade as necessary to clear the market; and noise traders as those who take random, uniformed positions in the market regardless of present or past prices.

The unique feature of the Kyle model that motivates its application to futures trading is his grouping of multiple (sequential) auctions into distinct sessions. A given session prevails until the insider information is revealed, within a finite error, in a sequence of market clearing prices. The concept that a single sample of the insider information (the future liquidation price) requires multiple auctions to be transmitted by the market is unique in the literature. At the termination of a session the process starts anew with another sample of insider information.

The Kyle model theory supposes that insiders know the future liquidation value of the current contract but are constrained by considerations of market liquidity (the volume required to change price 1 unit) from taking extreme positions. If future market liquidity exceeds present liquidity, the insider saves his information and trades lightly. If he knows that liquidity will decline, he trades intensely now when profits are greater. Kyle shows this intuition to be optimal for securing maximum expected profits.

Futures markets in the U.S. are double auction markets with buyers and sellers freely exchanging bids and offers without the benefit of an auctioneer. Since trades occur when bids and offers match up, the market clears almost continuously. However, it has been observed, Forsythe (1982), Smith (1982), and commented upon by Monroe (1988), that experimental evidence implies multiple replications of the market are required, under the same set of supply and demand fundamentals, for price to settle to its Walrasian value.

This study was supported in part by the National Science Foundation under Grant No. IRI-8822346.

William D. O'Neill is a Research Associate, Office of Futures and Options, University of Illinois at Chicago.

The Journal of Futures Markets, Vol. 10, No. 4, 425–441 (1990)

CCC 0270-7314/90/040425-17$04.00

The replication requirement to achieve price equilibrium is well captured by the Kyle model. But there are no market makers (specialists) in a futures market. Therefore, though the cited literature agrees that efficiency is eventually achieved, the Kyle model appears to attribute this task to a nonexistent specialist. In an earlier article, Kyle (1984) shows that market makers do make profits because they have an informational advantage of being able to observe total orders before trading while the insiders and noise traders must take positions without this observation. The main conclusions of that article are that: uninformed speculation has a stabilizing effect on prices; that public generation of information increases price informativeness; and that markets do not depend critically on whether market makers trade profitably. Therefore, in his 1985 article, he replaces the profitable market makers with an auction between risk neutral market makers and the assumption that both observe total orders, the market efficiency condition he imposes drives their profits to zero.

It is unrealistic to assume that market makers would enter a market with prospects of zero profits. However, if inside and noise traders are considered as off-the-floor traders and market makers as floor traders then the Kyle (1985) model captures in an approximate way the observations about organized futures markets that: (1) Floor traders (market makers) are primarily concerned with order flow and less about supply-demand fundamentals and (2) Off-the-floor traders (inside and noise traders) are no better than floor traders at predicting which positions will clear the market.

On a more theoretical level, as shown by Telser (1978), the transient theory of market equilibrium is based on the hypothesis that the multiple supply and demand curves brought to the market by its many agents carry the information required for efficiency. In this sense, the market maker is a facilitator whose job it is to inflict losses on the uninformed and direct profits to the informed.

The central hypothesis tested here is that the Kyle (1985) sequential auction model explains the price and volume dynamics of the soybean futures markets by attributing inside information to a particular group of traders. The test consists of estimating the model structure parameters and comparing them against the predictions of the theory.

A secondary hypothesis tested here is that the information the insiders used in trading during the period studied was public weather data. This hypothesis is tested by comparing the time course of the estimated model parameters against their theoretical dependence on the length of the auction session. Session length is directly dependent on how long weather data influenced the liquidation price.

The final hypothesis tested is that inside traders take market positions which have a constant variance over the time course of the auction session. This hypothesis is suggested by the results of the central hypothesis test cited above. The constant variance hypothesis is tested by comparing the values of a constant parameter estimated from model structural parameters and a derived expected profit function of the inside traders.

Section II describes the data to be used and in Section III the quantitative form of the Kyle model is set out. Section IV reports empirical findings and Section V discusses those findings in the context of a modified Kyle which appears to conform better to the data and to the behavioral assumptions of the model. Section VI shows how the modified Kyle model can be used in situations where it is not clear that some specific public or private information is driving the market but rather a general forecasting model is used to predict the liquidation value. The results are summarized in Section VII.

II. THE DATA

The period March 1, 1983 to December 23, 1983 was chosen as the study period because this growing season for soybeans experienced a major drought throughout the entire soybean growing region of the United States.[1] The drought onset was sudden and persistent and weather data revealed the extent of drought damage. This allowed testing the hypothesis that major trading companies (inside traders) can forecast eventual crop production fairly accurately.

Weekly weather data (public information) was collected from 14 weather reporting stations covering the U.S. soybean growing region. The data consisted of deviations from normal rainfall and normal temperature. Decrements in rainfall R and increments in temperature T were considered unfavorable to crop production. These data were combined in a weather index $W = R + kT$ where $k = .1\ 21$ was derived by normalizing the sum of absolute rainfall deviations over the study period. This procedure provided equal weight for rainfall and temperature. $W > 0$ represents a negative impact on crop production.

Roll (1984) studied an orange juice futures price and weather relationship but related these variables by a regression model with weather as an explanatory variable. Weather data was used in this study simply to infer the length of the auction session and, as such, the data does not enter as an explanatory variable.

Knowledge of when the soybean crop is sensitive to weather is also important. One can see in Roll's study the price-weather connection. In the winter of 1981, severe freezes set off an orange juice futures price increase which required about 100 days to stabilize. If the Kyle model were applied to this data, it would be assumed that the freezes of January 11, 12 and 13, 1981 set off the observed price spiral and that insiders used the freeze data as a significant, but not sole, forecaster of an equilibrium price valid for 100 days hence. Knowing the equilibrium price and the number of auctions required to reach it, one would assume the insiders then traded to maximize expected profits over the 100 day session.

The daily close price of all current soybean contracts and the daily volume in those contracts were obtained from the Chicago Board of Trade. 1983 contracts for months 3, 5, 7, 8, 9, and 11, 1984 contracts for months 1, 3, 5, 7, 8, 9, and 11 and 1985 contracts for months 1 and 3 were all of the contracts traded on the CBOT over the study period. Daily close price in the data set was the average of close prices for all contracts traded on a given day. Volume was the sum of the volumes of these contracts.

III. THE TRADING MODEL

The Kyle sequential auction model is defined by the following variables and parameters: (1) x_n = Insider position taken at auction $n = 1, 2, \ldots, N$, where N is the end of the auction session; (2) u_n = Noise trader position at auction n, u_n ca. $N(0, \Delta t\sigma_u^2)$ independent of $v, \Delta t = 1/N$; (3) $\omega_n = x_n + u_n$ = Total order flow observed by market makers; (4) p_n = Market clearing auction price at day n; (5) v = Insider information, liquidation price of current contract at $n = N + 1$; v ca. $N(p_o, \Sigma_0)$. A sample of v is known and held fixed by the insiders for $n = 1, 2, \ldots, N$; and (6) $\pi_n = \Sigma_{k=n}^{N}(v - p_{k-k})x_k$ = Insider profits from auctions $n, n + 1, \ldots, N$ for $n = 1, 2, \ldots, N$.

Market equilibrium is defined for days $n = 1, 2, \ldots, N$ by the 2 conditions: (I) Profit maximization; for all positions $\underline{x}' = (x_1', x_2', \ldots, x_N')$, $\underline{x} = (x_1, x_2, \ldots, x_N)$, and $\underline{p} = (p_1, p_2, \ldots, p_N)$

$$E[\pi_n(\underline{x}, \underline{p}) \mid p_1, \ldots, p_{n-1}, v] \geq E[\pi_n(\underline{x}', \underline{p}) \mid p_1, \ldots, p_{n-1}, v]$$

(II) Market Efficiency;

$$p_n = E[v \mid \omega_1, \ldots \omega_n]$$

[1]Domestic soybean production fell by 31% from 1982 (USDA Crop Production Estimates, December, 1983).

The equilibrium is characterized by the following relationships in which $\Delta t = 1/N$ is the time increment (days) between auctions:

$$x_n = \beta_n \Delta t(v - p_{n-1}) \tag{1}$$

$$p_n = p_{n-1} + \lambda_n \omega_n \tag{2}$$

$$\sum_n = \text{var}[v \mid \omega_1, \ldots, \omega_n] \tag{3}$$

$$E[\pi_n \mid p_1, \ldots, p_{n-1}, v] = \alpha_{n-1}(v - p_{n-1})^2 + \delta_{n-1}. \tag{4}$$

The parameters β_n, λ_n, α_n and Σ_n are shown by Kyle to satisfy a third order, nonlinear, difference equation which has a unique solution for the parameters given the final values $\delta_N = \alpha_N = 0$, an initial value Σ_0 of Σ_n, and the second order profit maximization condition $\lambda_n(1 - \alpha_n \lambda_n) > 0$. Unfortunately, the solution is only exhibitable by numerical methods so that the intuitive economic properties of the parameters are difficult to see in general. However, as a futures model, the parameters β_n and λ_n are readily interpretable.

It is clear from the definition of β_n that $\beta_n \Delta t$ is the insider trading intensity in that it is the ratio of his current position to the deviation of the last auction price from the current contract liquidation price. The parameter λ_n is just the inverse of market liquidity, being the ratio of the current auction price difference to current, total market orders. These parameter descriptions simplify the overall conception of the model. The insiders determine their trading intensity so as to maximize their expected profits conditional on their information of past prices and the eventual liquidation price. Market makers, using current and past order flows, determine market liquidity such that an efficient auction price prevails in the sense of condition II in the definition of market equilibrium. The noise traders, by definition, take random positions in the current contract regardless of price.

The variance measure, Σ_n, indicates the information released by the insiders as the sequence of auctions unfolds. Since Σ_n can also be written as

$$\sum_n = E[(v - p_n)^2 \mid \omega_1, \ldots, \omega_n]$$

it is evident that $p_n \cong v$ certainly signals the complete revelation of insider information. As a trivial example, if $N = 1$, a single auction market, it is readily shown that

$$\lambda_1 = (\textstyle\sum_0 / \sigma_u^2)^{1/2}/2$$

$$\beta_1 = (\sigma_u^2 / \textstyle\sum_0)^{1/2}$$

so that $\Sigma_1 = \Sigma_0/2$ and ½ of the insider information is revealed.

The parameters α_n and δ_n are more difficult to interpret beyond their defining a quadratic profit function giving the value of trading opportunities at auction $n + 1$. Their more substantial roles are as follows: These parameters, through their nonlinear dependence on β_n, λ_n, and Σ_n, are intended to limit the position variance of the insider and thereby thwart opportunities of unbounded profits by market destabilization. Such a scheme would have the insider make unprofitable trades at the present auction, which destabilize price, and then cover such losses plus much more with profitable trades at future auctions.

IV. EMPIRICAL RESULTS

Figure 1 shows the close price data and the weather index W. Figure 2 is the daily volume during the same period. It is evident from these figures that the onset of unfavor-

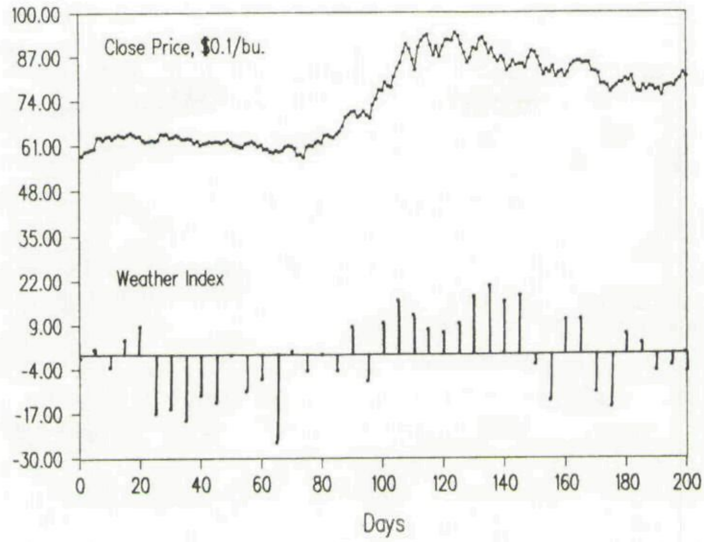

Figure 1

able weather and its persistence, weeks 18 through 29, is accompanied by a dramatic increase in the current contract price and the volume therein.

Technically, the weeks 18 through 29 encompass the critical growing period for soybeans over the entire growing region and by day 173, November 2, 1983, the resultant crop yield was known quite accurately. Thus, it is assumed that after day 173 the insiders were no longer trading the current contract based on an equilibrium price forecast dominated by weather, whose damage was done and known, but rather on some future liquidation price forecast based on other information. In effect, day 173 is taken as the end of the trading session.

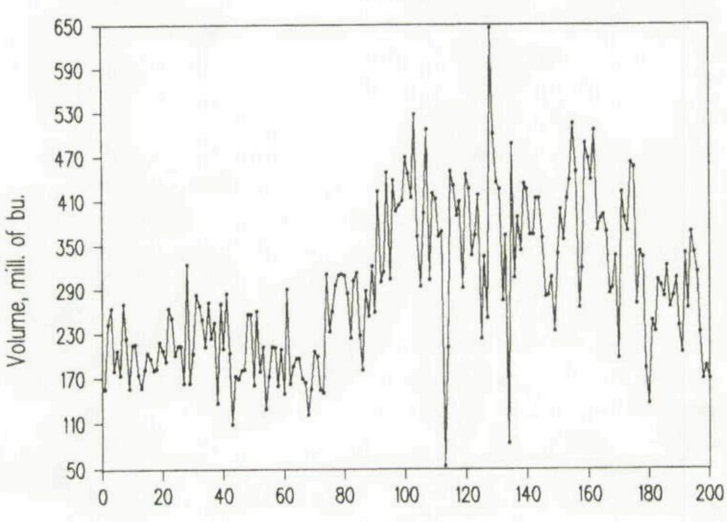

Figure 2

During the drought period, the daily price that appears to kick off the "drought-trading" is the local minimum daily price at day 74, June 14, 1983. It is evident from Figures 1 and 2 that both price and volume began upward spirals on this day which eventually led to a new asymptotic equilibrium some 100 days later. Daily weather data from all 14 reporting stations was not available and weekly averages contain information at least 5 days old, so daily weather data might show an earlier start of the drought than week 18. However, recall that weather is just one data set the insider uses to forecast the eventual liquidation price. Therefore, although weather data is used to estimate a trading session length (day 74 to day 173), this session does not have to correspond exactly to the drought period. In the Roll orange juice study, there were but 3 daily, consecutive, freezes that obstensibly set off a 100 day price movement. But damage from orange tree freezes and soybean droughts takes time to assess. Other fundamentals change contemporaneously and the market has its own inherent dynamics of information revelation to contend with. These other fundamentals impinging on the market are now addressed.

One must consider also how the demand side of the market for soybean meal and oil and exports reacted in the weeks studied. There are fundamentally 2 ways for the supply of soybeans to disappear, either through meal and oil, called crush, or by export. Weekly statistics from "Export Sales," USDA, and from the National Soybean Producers Association (NSPA) were used to estimate, respectively, export contractual agreements (metric tons) and conversion of soybeans into meal and oil (metric tons). These estimates were converted to millions of bushels and used to calculate a demand measure consisting of the change in export sales on a weekly basis added to the aggregate weekly NSPA crush. This measure averaged 18.36 million bushels for the weeks 1 through 18 and 24.3 million bushels for weeks 19 through 39. This 32 percent increase in demand in all probability contributed to the 48 percent increase in the current contract average price over the same period but falls short of explaining such a price rise. This observation is strengthened by noting that the largest demands, 48, 43, 45, 38, and 37 million bushels at respective weeks 22, 23, 27, 31, and 38 did not uniformly elicit significant price increases. Indeed, weeks 31 and 38 saw price falls and weeks 22 and 29 small price increments. By contrast, the monotonic price rise for the periods 18 to 27 was accompanied by 9 out of 10 detrimental weather indices.

Also, it must be kept in mind that weather is a crucial factor in future supply only during the critical growing period while demand for soybean products is influential throughout the year. Thus, the negative values of the weather index for the weeks immediately prior to week 39 are immaterial to future supply since the growing period damage of the drought is already history. However, if prices are significantly driven by deviations from normal demand, such price fluctuations should be apparent regardless of the time of the year. The latter was not the case. The hypothesis that the trading session of weather interest lasted from day 74 to day 173 is well supported.

In all events, it is theoretically irrelevant to this analysis whether demand contributed significantly to the observed price rise because the market model we used posits the existence of a forecasting model in the hands of the insiders which forecasts a liquidation price the insiders use to trade on for 100 days. Since demand and weather simultaneously turned in favor of a price rise it is unimportant if the forecastors used weather, demand, or any other price-synchronous set of variables to arrive at a liquidation price. What is important is the starting and ending time of the trading session.

The structural equations of the model are (1) and (2) plus the definition of total order flow, $\omega_n = x_n + u_n$. These can be put into estimable forms by using insider positions from (1) into (2) through the definition of ω_n. Thus

$$p_n = (1 - \lambda_n \beta_n \Delta t) p_{n-1} + \lambda_n \beta_n \Delta t v + \lambda_n u_n \tag{5}$$

$$\Delta p_n = p_n - p_{n-1} = \lambda_n \omega_n . \tag{6}$$

Figures 1 and 2 provide data on Δp_n and ω_n, so λ_n can be estimated from (6). An estimate of $(1 - \lambda_n \beta_n \Delta t)$ is in principle obtainable from (5) by writing it as

$$v - p_n = (1 - \lambda_n \beta_n \Delta t)(v - p_{n-1}) - \lambda_n u_n \,. \tag{7}$$

If the eventual liquidation price v in (7) were known could be used to estimate $1 - \lambda_n \beta_n \Delta t$ using the data series $v - p_n$ and $v - p_{n-1}$. With the benefit of hindsight, one can readily estimate v by iteratively estimating (7) for various assumed values of v and selecting that v which minimizes the mean square estimation error.

Equations (6) and (7) are the structural equations to be estimated and, as is apparent from the dependence of λ_n and β_n on the auction index n, a nonstationary estimation method is required. A natural choice is recursive least squares. A discussion of this method is presented in Appendix 2. A brief summary is presented here. If the parameter b is to be estimated from the linear model.

$$y_n = bx_n + z_n \,,$$

where $(x_n, y_n), n = 1, 2, \ldots, N$ are data pairs and z_n a model error term, then the criterion

$$V_N = \frac{1}{N} \sum_{n=1}^{N} [y_n - bx_n]^2$$

is minimized by the recursive estimates $\hat{b}_n$ satisfying

$$\hat{b}_n = \hat{b}_{n-1}(1 - x_n^2 A_n) + A_n x_n y_n \tag{8}$$

$$A_n = A_{n-1}/(1 + x_n^2 A_{n-1}) \,. \tag{9}$$

If, in fact, $b_n = b$ for all n and $b_1 = y_1/x_1$, $A_1 = 1/x_1^2$ then $\hat{b}_N$ is the OLS estimate of b. In the nonstationary parameter case, the estimates $\hat{b}_n$ track the sequence b_n. As explained in Appendix 2, the convergence of (8) and (9) is independent of the initial values $\hat{b}_0$ and A_0 as long as $A_0 > 0$ so arbitrary initial values $\hat{b}_0 = 0$, $A_0 = 1$ can be assumed. There is a brief transient period in the estimates from these arbitrary choices but for the estimates made for the data of Figures 1 and 2 the sensitivity of the estimates to initial values is very slight. As an example, in estimating $(1 - \lambda_n \beta_n \Delta t)$ in (7), the initial estimate ranges $0 \leq \hat{b}_0 \leq 10, 0.01 \leq A_0 \leq 100$ changed the estimates $\hat{b}_n$ by less than 0.6 percent. Since for all n, $0.96 < \hat{b}_n < 1.03, 0 < A_n < 1$, the sensitivity range checked was more than adequate.

From Figure 1, it is evident that the monotonic rise of price began at day 74 and, as noted above, by day 173 one can safely assume the liquidation price had been revealed. Therefore, parameter estimates cover the auction session from day 74 to day 173 inclusive so that $N = 100$ is the number of auctions required for the market to transmit the insider-known liquidation price. The sequence $\hat{\lambda}_n$, the inverse liquidity parameter, estimated from (6) is shown in Figure 3.

The sequence estimated for $(1 - \lambda_n \beta_n \Delta t)$, called the time constant, from (7) is shown in Figure 4. This sequence was estimated for $\hat{v} = 8522$, in units of 0.1 cents/bu, which minimized the mean square estimation error over v.

It is clear from Figure 4 that the estimate of $(1 - \lambda_n \beta_n \Delta t)$ is virtually constant over the entire estimation period. This observation is confirmed by the following OLS regression result in which this parameter is hypothesized to be a constant,

$$v - p_n = \underset{(-0.03)}{-0.62} + \underset{(48.22)}{0.95002}(v - p_{n-1}) \tag{10}$$

$$r^2 = 0.96 \,.$$

An estimate of the trading intensity sequence is obtainable from the estimates of λ_n and $1 - \lambda_n \beta_n \Delta t$ and this estimate, $\hat{\beta}_n \Delta t$, the trading intensity, is shown in Figure 5.

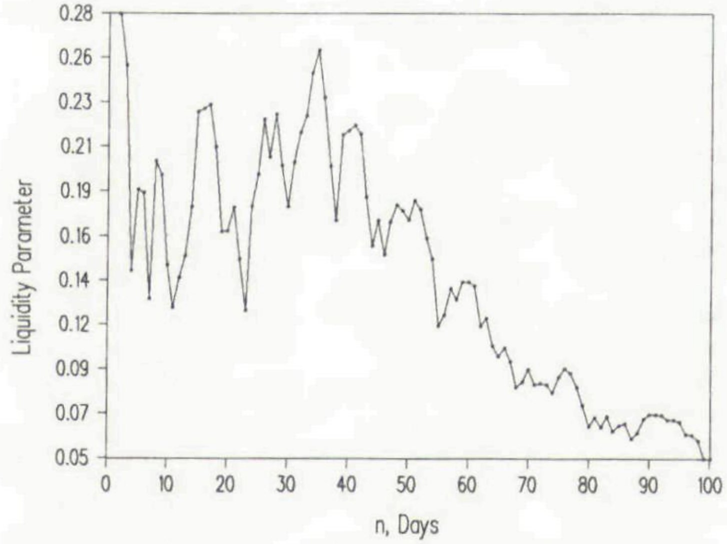

Figure 3

The estimates of the inverse liquidity parameter $\hat{\lambda}_n$ and the trading intensity parameter, $\hat{\beta}_n \Delta t$, qualitatively support the Kyle model theory in the following sense: As the session proceeds, market liquidity $(1/\lambda_n)$ increases in that small price differences are accompanied by large positions. The market can be said to gain depth through the session. Simultaneously, the auction price is approaching the liquidation price and since insider profits are proportional to this difference, the insider must take larger positions in a more liquid market. Thus, his trading intensity must increase through the session. Figures 3 and 5 illustrate the empirical agreement found for the soybean futures data but there are important caveats when comparing theory and empirical results.

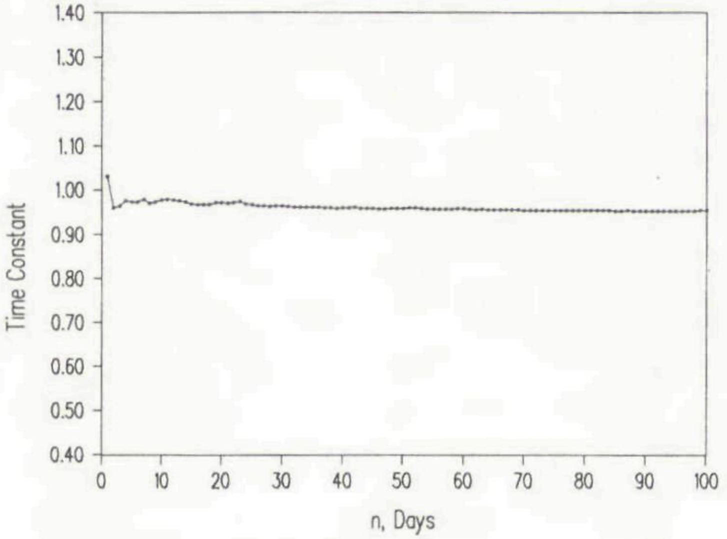

Figure 4

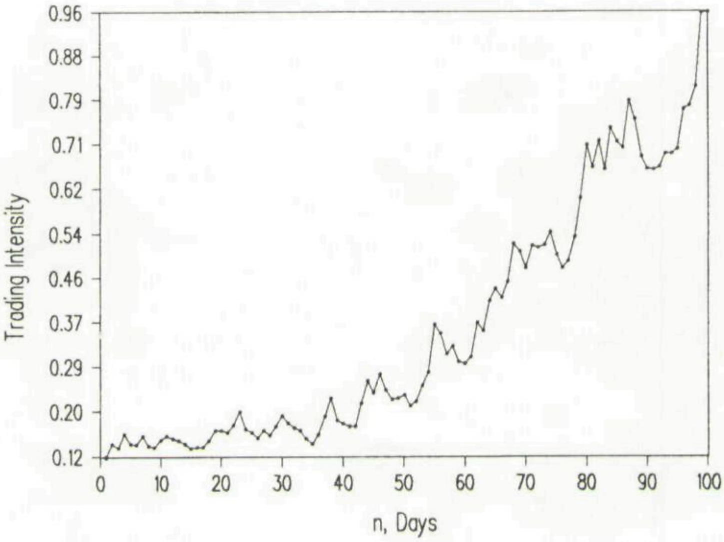

Figure 5

Kyle found from numerical solutions that $\beta_n \Delta t$ should increase through the session and λ_n should decrease but for the latter the decrease is monotonic with a negative *2nd* difference. Further, the theory indicates that $(1 - \lambda_n \beta_n \Delta t)$ should decrease also monotonically through the session, again with a negative *2nd* difference. He found, for example, for a 100 day session $1 - \lambda_1 \beta_1 \Delta t = 0.9905$ while $(1 - \lambda_{100} \beta_{100} \Delta t) = 0.5$. Clearly, Figure 4 gives no empirical support to these numerical results. This disagreement can be explained by the following argument. If one computes solutions of the Kyle model for the variance of the insider positions, x_n, one finds that this variance increases dramatically through the session. As an example, for $N = 100$, $\text{var}(x_{100})/\text{var}(x_1) = 103.6$ and even for a short, 4 auction session $\text{var}(x_4)/\text{var}(x_1) = 5.12$. Clearly, a 103-fold increase in insider trading variance is a strong signal to noninsiders. This represents a significant criticism of the Kyle model. However, the empirical result shown in Figure 4, and confirmed in the regression (10), obviates the need for an increasing insider position variance. The following section shows that when $1 - \lambda_n \beta_n \Delta t$ is not a function of n, then insiders trade with constant position variance.

V. A MODIFIED TRADING MODEL

The market efficiency definition of the market clearing price p_n is retained here; but assume the insider maximizes unconditional expected profits over the entire session subject to the constraint

$$\text{var}(x_n) = B$$

where B is a positive constant. Further, it is assumed that a linear equilibrium exists and λ_n and β_n in the equilibrium equations

$$x_n = \beta_n \Delta t(v - p_{n-1}) \tag{11}$$

$$p_n = p_{n-1} + \lambda_n \omega_n \tag{12}$$

must be determined such that: (1) The constraint $\text{var}(x_n) = B$ is met; (2) $p_n = E[v \mid \omega_1, \omega_2, \ldots, \omega_n]$; and (3) $E(\pi) = E[\sum_{n=1}^{N} (v - p_n)x_n]$ is maximized over positions $x_1, \ldots, x_N$.

Observe that if (11) and (12) are used in the expression for $E(\pi)$ there results

$$E(\pi) = \sum_{n=1}^{N} (1/\beta_n \Delta t - \lambda_n) \text{ var}(x_n)$$

and therefore if market liquidity is never 0, ($\lambda_n < \infty$), trading intensity is nonzero ($\beta_n \Delta t \neq 0$), and $\text{var}(x_n) = B$, then expected profits are bounded. In effect, the variance constraint eliminates the need for the α_n and δ_n parameters of the Kyle model.

To derive explicit expressions for λ_n and $\beta_n \Delta t$ we apply the market efficiency condition to the linear equilibrium and get from the projection theorem

$$\lambda_n = \beta_n \sum\nolimits_{n-1} / (\beta_n^2 \Delta t \sum\nolimits_{n-1} + \sigma_u^2) \tag{13}$$

$$\sum\nolimits_n = \sigma_u^2 \sum\nolimits_{n-1} / (\beta_n^2 \Delta t \sum\nolimits_{n-1} + \sigma_u^2) . \tag{14}$$

To show that $\lambda_n \beta_n \Delta t = k$, a constant independent of n, implies $\text{var}(x_n)$ is constant write (13) as

$$\lambda_n \beta_n \Delta t = k = \beta_n^2 \Delta t \sum\nolimits_{n-1} / (\beta_n^2 \Delta t \sum\nolimits_{n-1} + \sigma_u^2)$$

and solve for

$$\beta_n^2 \Delta t^2 \sum\nolimits_{n-1} = \Delta t k \sigma_u^2 / (1 - k) .$$

Now (11) implies

$$\text{var}(x_n) = \beta_n^2 \Delta t^2 \sum\nolimits_{n-1}$$

which is shown to be independent of n.

The implications of constant position variance are far-reaching, for if

$$\text{var}(x_n) \beta_n^2 \Delta t^2 E[v - p_{n-1}]^2 = \beta_n^2 \Delta t^2 \sum\nolimits_{n-1} = B ,$$

then writing (13) and (14) as $\beta_n = \sigma_u^2 \lambda_n / \sum_n$ allows a solution for $\sum_n$ as

$$\sum\nolimits_n = \frac{1}{s + 1} \sum\nolimits_{n-1}$$

where $s = B / \Delta t \sigma_u^2$ is defined as the market signal to noise ratio. It is now easy to solve (13), (14), and (15) simultaneously to yield

$$\beta_n = \sigma_u (sN / \sum\nolimits_0)^{1/2} (s + 1)^{(n-1)/2} \tag{16}$$

$$\lambda_n = (s \sum\nolimits_0 N)^{1/2} (s + 1)^{-(n+1)/2} / \sigma_u \tag{17}$$

$$\sum\nolimits_n = \sum\nolimits_0 / (s + 1)^n . \tag{18}$$

Expected profits have yet to be maximized. Writing $E(\pi)$ in terms of the solutions for λ_n and β_n:

$$E(\pi) = (\sigma_u^2 \textstyle\sum_0 s/N)^{1/2}(1 - (s + 1)^{-N/2}/(s + 1 - (s + 1)^{1/2}) \tag{19}$$

which the insider maximizes by his choice of s, or equivalently B. Figure 6 is a graph of $E(\pi)$ vs. s for $N = 100$ and $\sigma_u^2 \sum_0$ fixed at 1. It indicates that expected profits have a unique maximum at $s = 0.025$ for any $\sigma_u^2 \sum_0$.

The theoretical parameters from (16) and (17) are now compared to their estimates from the previous section. Calculating $1 - \lambda_n\beta_n\Delta t$ from (16) and (17) gives

$$(1 - \lambda_n\beta_n\Delta t) = (s + 1)^{-1}$$

which is well in accord with the almost constant sequence, Figure 4, estimated. (17) shows that λ_n decreases exponentially with n and one can judge its correspondence to $\hat{\lambda}_n$ of Figure 3 by performing the regression

$$\log \hat{\lambda}_n = \tfrac{1}{2} \log(s \textstyle\sum_0 N)/\sigma_u^2 + [(n + 1)/2] \log(s + 1)^{-1}.$$

The result is:

$$\log \hat{\lambda}_n = \underset{(-26.07)}{-1.30154} - \underset{(-17.36)}{0.029213}(n + 1)/2$$

$$r^2 = 0.757,$$

with the corresponding estimate $(s + 1)^{-1} = 0.9712$. This estimate of $(s + 1)^{-1}$ agrees well with the estimate 0.95 from (10). Further, since $E(\pi)$ is maximum for $s = 0.025$, profit maximization would yield the estimate $(s + 1)^{-1} = 0.976$.

Since these 3 estimates of $(s + 1)^{-1}$ differ at most by 2.7 percent (taking $(s + 1)^{-1} = 0.95$ as a base) it is evident that the hypothesis of constant insider signal variance has strong empirical support from the soybean futures data.

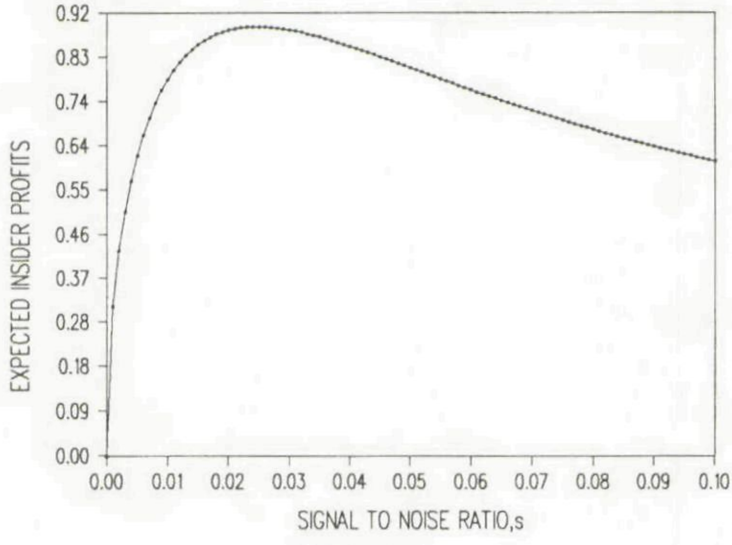

Figure 6

A check on an assumption of the modified auction model is a test of the hypothesis that insider positions x_n are independent of noise positions u_n. The time series for $x_n, n = 1, 2, \ldots, 100$ is estimated fom (11) using the estimates of $\beta_n \Delta t$, v, and the price series. The residuals from the estimated form of (7) are divided by the estimated series for λ_n to produce the estimated noise position time series $u_n, n = 1, 2, \ldots, 100$. The estimated, squared, correlation coefficient for these series was calculated to be 0.00751.

It is clear from (19) that $E(\pi)$ depends directly on the session length N but it depends also indirectly on N since the profit maximizing s, say $\hat{s}(N)$, is a function of N. It is numerically easy to find $\hat{s}(N)$ and compute $E(\pi)$ solely as a function of N. The result for $N = 1, 2, \ldots, 100$ is shown in Figure 7 for $\sigma_u^2 \Sigma_0 = 1$. It is also straightforward to use asymptotic analysis on (19) to show, O'Neill (1989), that $E(\pi) \rightarrow 0.902513$ as $N \rightarrow \infty$.

From Figure 7 one calculates that $E(\pi)$ increases less than 5% from $N = 30$ to $N = \infty$. There is little incentive, therefore, for the insider, or anyone for that matter, to engage in a long trading session. Indeed, this state of affairs challenges the implicit model assumption of zero transaction costs. In hindsight, it is evident that taking $N = 100$ was quite conservative, based as it was on the crop cycle of soybeans and weather data. In fact, Figure 1 shows that almost all of the price movement was over in 30 days after day 74. But, surely, a drought was not *certain* by day 104. It appears, then, that soybean insiders, taking 3 critical weeks of weather data, estimated crop yield conditional on a continuing drought and turned this into the eventual liquidation price v on which they traded. In 30 days after day 74, the hypothetical drought had been completely discounted by the insiders. What would have happened if, after day 120, 3 weeks of favorable weather had occurred? If the Kyle model is to be believed, prices would have fallen back to about "normal," say \$6.10/bu., demanded by "normal" weather conditions. This assertion is quite hypothetical but support for it is found in some of the Roll (1984) data on orange juice futures. On March 2, 1980, there was a single freezing night followed by normal temperatures thereafter. Prices increased quickly to 16.1% higher and in about 40 days were back to "normal." The point of discussion is that one does not need a long term future liquidation price fixed in concrete to make the Kyle model, as modified, work. As a trader, one only needs good forecasts of future liquidation prices to employ the model. This procedure is illustrated in the following section.

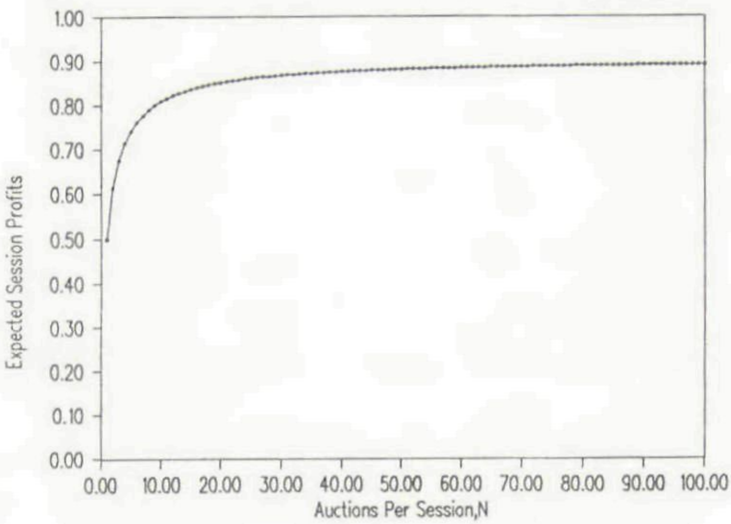

Figure 7

VI. PRACTICAL USE OF THE SEQUENTIAL AUCTION MODEL

The sequential auction model takes a certain set of information as known by the inside trader. Specifically, the insider has a forecasting model which produces the forecasts v with a forecast variance Σ_0. These forecasts embody all of his knowledge of past, current, and future events which impinge upon the future liquidation price v. Here one can envision an econometric model, perhaps a transfer function model with multiple explanatory variables, as his forecasting tool. He also must know the length of time for which a given forecast is valid. Given this time and the rate at which the market clears, allows the insider to estimate N, the number of auctions that will occur while his forecast v is valid.

Knowledge of N determines the optimal signal to noise ratio the insider should employ. To see this, maximize $E(\pi)$, eq. (19), for a range of N. The result is shown in Figure 8 and observe that the optimal s is independent of σ_u^2 and Σ_0.

It is clear from (16) that the insider trading rule is complete once he knows, or estimates, σ_u^2, the noise position variance. This variance could readily be estimated from estimates of the liquidity parameter λ_n since by (17) s, Σ_0, and N are known and λ_n is easily estimated recursively from price and volume data. His explicit trading rule is then given by (11) since $\beta_n \Delta t$, the trading intensity, is now known for all $n = 1, 2, \ldots, N$.

To illustrate the use of the model for the average trader, suppose his forecasting model produces 2 sequential forecasts; $v = v_1 = 10$ which will be valid by days $n = 0, 1, \ldots 10$. At $n = 10$ his forecasting model predicts $v = v_2 = 3$ which is valid for $n = 11, 12, \ldots, 40$. From Figure 8 he immediately knows that he should use $s = s_1 = 0.212$ when v_1 is valid and $s = s_2 = 0.0786$ when v_2 is valid. Suppose further that his forecasting model has error variance $\Sigma_0 = 1$ and that he estimates $\sigma_u^2 = 1$. It is assumed that at $n = 0, p_0 = 5$. From (16) and (17), the price sequence p_n given by (5) is now determined since λ_n, $\beta_n \Delta t$, and v have been specified for all n. A typical price sequence is shown in Figure 9 as generated by (5) and samples from u_n ca. $N(0, 1)$.

His optimal trading rule is as follows: (1) For $n = 1, 2, \ldots, 10$ he should take positions

$$x_n = \beta_n \Delta t(v_1 - p_{n-1}) = \left(\frac{0.212}{10}\right)^{1/2} (1.212^{n-1/2}(10 - p_{n-1})$$

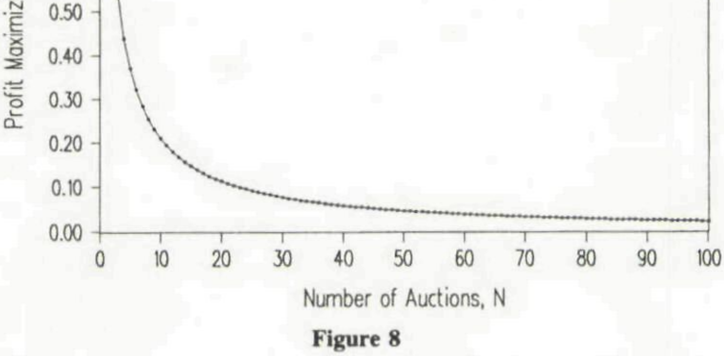

Figure 8

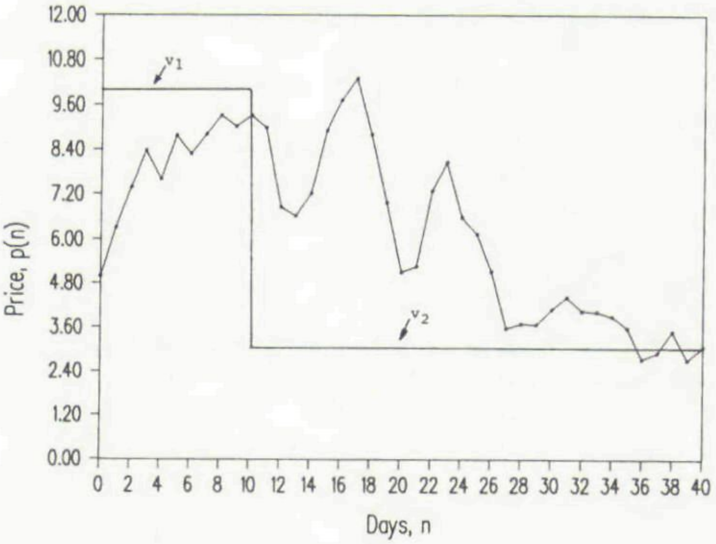

Figure 9

and (2) For $n = 11, 12, \ldots 40$ he should take positions

$$x_n = \left(\frac{0.0786}{30}\right)^{1/2} (1.0786)^{n-11/2}(3 - p_{n-1})\,.$$

Several comments are relevant to this simulation: (1) It is evident that the trading rule can be readily extended to future liquidation forecasts $v_3, v_4, \ldots$, and (2) The model and simulation are imbued with a "big trader" syndrome which suggests that insider positions, given v, drive price to v. While this is true, it does not limit the small, informed trader who could not, for example, afford to take the positions demanded by the model. Such a trader need merely scale back his positions by scaling down his trading intensity $\beta_n \Delta t$. As expression (19) for $E(\pi)$ indicates, by so doing he scales his profits proportionally.

VII. SUMMARY AND CONCLUSIONS

Kyle's 1985 sequential auction model with insider trading is applied in this study to the current soybeans futures contract of the CBOT. The length of the auction session is inferred from weather data which reflects a national drought that sharply cut soybean production. Model structural parameter estimates support the hypothesis that inside traders can forecast drought effect on price and that weather data is fundamental in that forecast.

Parameter estimates also suggest that a revised Kyle model, one containing the constraint that inside trader positions have a constant variance, better explains the soybean futures data. The modified model hypothesis is verified by comparing estimates derived from inverse market liquidity, insider trading intensity, and expected insider profits.

A simulation example is constructed to illustrate the use of the modified Kyle model in daily trading when a series of future liquidation price forecasts is available. The simulation illustrates that small traders can take advantage of information in the same manner as large, institutional traders do.

Appendix 1

WEATHER DATA

(A) Weather data reporting stations: (weekly deviation from normal rainfall and temperature) Little Rock, Arkansas, Peoria, Illinois, Quincy, Illinois, Springfield, Illinois, Fort Wayne, Indiana, Sioux City, Iowa, Waterloo, Iowa, Alexandria, Louisiana, Alexandria, Minnesota, Greenwood, Mississippi, Columbia, Missouri, Dayton, Ohio, Florence, South Carolina, and Memphis, Tennessee.

(B) Weather Index

Week, N	Index
1	2.2
2	−3.2
3	4.8
4	8.7
5	−16.8
6	−15.5
7	−18.8
8	−11.7
9	−13.7
10	0.2
11	−10.5
12	−6.8
13	−25.3
14	1.3
15	−4.4
16	0.4
17	−4.3
18	8.6
19	−7.4
20	9.7
21	16.4
22	12.0
23	7.9
24	6.9
25	9.5
26	17.4
27	20.7
28	1.6
29	17.8
30	−2.5
31	−13.1
32	10.7
33	10.8
34	−10.6
35	−15.1
36	6.5
37	3.8
38	−4.4
39	−2.8
40	−4.6

CBOT DATA

The "current contract" referred to in the text consists of 1983 contracts for the months 3, 5, 7, 8, 9, 11, 1984 contracts for the months 1, 3, 5, 7, 8, 9, 11, and 1985 contracts for months 1 and 3. These are all of the contracts traded on the CBOT over the study period March 1, 1983 to December 23, 1983. Volume referred to in the text is the sum of the volumes in such contracts traded on a given day. Daily close price is the average of close prices for all contracts traded on a given day.

Appendix 2

(1) The recursive theory is dealt with in depth by Ljung and Soderstrom (1983) and from an economic perspective in Chapter 10 of Johnston (1984). The method applies to a vector of parameters in, say, a k dimensional OLS model, but only a one parameter model is dealt with here.

For $n = 1, 2, \ldots, N$ data points (x_n, y_n) and the model

$$y_n = bx_n + z_n$$

The criterion is defined as:

$$V_N(b) = \frac{1}{N}\sum_{n=1}^{N}(y_n - bx_n)^2 \tag{1(A)}$$

and analytic minimization yields the OLS estimate

$$\hat{b}(N) = \sum_{n=1}^{N} x_n y_n \Big/ \sum_{n=1}^{N} x_n^2 .$$

This is written recursively by defining

$$R_n = \sum_{k=1}^{n} x_k^2$$

and writing (1(A)) as

$$\sum_{k=1}^{n-1} x_k y_k = R_{n-1}\hat{b}_{n-1}$$

Thus, $R_n = R_{n-1} + x_n^2$ and solving for $\hat{b}_n$ yields

$$\hat{b}_n = \hat{b}_{n-1} + R_n^{-1}x_n[y_n - \hat{b}_{n-1}x_n].$$

Now define A_n as R_n^{-1}/n and solve for $\hat{b}_n$ and A_n to get

$$\hat{b}_n = \hat{b}_{n-1}(1 - x_n^2)A_n + A_n x_n y_n$$

$$A_n = A_{n-1}/(1 + x_n^2 A_{n-1})$$

These are the "recursive residual" formulas found in Johnston (1984) and simplified for the 1 parameter model. Direct iteration shows that initial values $b_1 = y_1/x_1$, $A_1 = 1/x_1^2$ converge to the OLS estimate of b when $n = N$.

Bibliography

Forsythe, R., Palfrey, T. R., and Plott, C. R. (1982, May): "Asset Valuation in an Experimental Market," *Econometrica*, 50:537–567.

Johnston, J. (1984): *Econometric Methods, 3rd Edition,* McGraw-Hill Book Company, New York.

Kyle, Albert, S. (1985, November): "Continuous Auctions and Insider Trading," *Econometrica,* 53:1315–1335.

Kyle, Albert, S. (1984): "Market Structure, Information, Futures Markets, and Price Formation," in *International Agricultural Trade: Advanced Readings in Price Formation, Market Structure, and Price Stability,* Westview Press, Boulder and London, 45–65.

Ljung, Lennart and Soderstrom, Torsten (1983): *Theory and Practice of Recursive Identification,* The MIT Press, Cambridge, Massachusetts, London, England.

Monroe, Margaret, A. (1988, October): "Indeterminacy of Price and Quantity in Futures Markets," *Journal of Futures Markets,* 8:575–588.

O'Neill, William D. (1989, June): "Auction Markets as Communication Channels," *Summer Meeting of the North American Econometric Society,* Ann Arbor.

Roll, Richard (1984, December): "Orange Juice and Weather," *The American Economic Review,* 74, 3:861–880.

Smith, Vernon, L. (1982, December): "Microeconomic Systems as an Experimental Science," *American Economic Review,* 72:923–955.

Telser, Lester (1978): *Economic Theory and the Core,* University of Chicago Press, Chicago.