ANNOUNCEMENT EFFECTS AND THE THEORY OF STORAGE: AN EMPIRICAL STUDY OF LUMBER FUTURES

Berna Karali
Graduate Assistant
Department of Economics
North Carolina State University
Raleigh, NC 27695-8110
bkarali@ncsu.edu
(919) 413-3517

Walter N. Thurman
Professor
Department of Agricultural and Resource Economics
North Carolina State University
Raleigh, NC 27695-8109
wally_thurman@ncsu.edu
(919) 515-4545

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Portland, OR, July 29-August 1, 2007

Abstract

We investigate how lumber futures returns are affected by monthly housing starts announcements and analyze the dependence of the response on lumber inventories and time to delivery. We develop a Generalized Least Squares method to jointly analyze simultaneously traded contracts. We find that the unanticipated component of housing starts announcements increases returns on lumber futures contracts. Further, the effects of housing starts shocks decline with lumber inventories and time to delivery. Futures contracts up to four months out respond by a larger amount to the shocks than do more distant ones. For more distant delivery horizons, the effect of housing starts shocks declines linearly with time to delivery.

Key words: theory of storage, announcement effects, event study, futures markets, commodity futures

1. Introduction

Two concepts are key to understanding commodity futures price formation. One is market efficiency, which explains how well-functioning markets process information. The other is the theory of storage, which explains price dynamics when the commodity in question is storable.

According to the efficient markets hypothesis, asset prices move only when new information arrives to the market. Following the seminal paper by Fama, Fisher, Jensen, and Roll (1969) many economists have conducted event studies to measure the impact and information content of economic announcements or events and the speed of adjustment of prices following new information. Binder (1998) reviews the event study methodologies and previous work on this topic.

According to the theory of storage, the discounted expected price of a storable commodity should exceed its spot price by the cost of storage. See, for example, Brennan (1958), Telser (1978), Scheinkman and Schechtman (1983), Thurman (1988), and Williams and Wright (1991). For storable commodities, and unlike the situation in financial asset markets, physical inventories serve to stabilize price movements in response to shocks. Thus, the return from holding commodities and the variance of the return should depend on the level of physical inventories. In the physical asset markets represented by commodity futures, we expect smaller price responses to demand or supply shocks during high inventory regimes.

The implications of market efficiency and the theory of storage can be tested in commodity futures markets and futures price movements understood in their context. We empirically analyze the market for lumber futures contracts, both to test the implications of received theories and to contribute to an understanding of an important natural resource market.

Most past analysis of futures price movements (with the notable exception of Smith's analysis of corn futures (2005)) splices together time series of prices on multiple futures contracts—with different delivery dates—and studies the behavior of the resulting single time series. See, for example, Rucker, Thurman, and Yoder (2005) for an analysis of lumber futures. As Smith points out, this underutilizes the available information from futures exchanges, on which several contracts trade simultaneously. Multiple contracts provide more information about the market in consideration. Consequently, using all traded contracts increases statistical efficiency. For this purpose, and in the spirit of Smith's partially overlapping time series (POTS) model, we develop a Generalized Least Squares (GLS) method to jointly analyze all traded contracts for lumber. Our methods and results differ from Smith's but there are common themes. We compare and contrast them in the Conclusions section (see footnote 2) after presenting our results.

Our empirics measure the effects of U.S. Census Bureau housing starts announcements, one of the key indicators of the demand side of lumber markets, and their dependence on physical lumber inventories and time remaining to contract expiration. We analyze daily lumber futures prices from the Chicago Mercantile Exchange in Chicago, from 1992 to 2005, and define a housing starts shock as the difference between the released figure and a survey measure of the market's expectation. When there is an observable information flow, such as a housing starts announcement, lumber futures prices should move only if the announcement contains new information. The effect of the news component of housing starts announcements, or housing starts shock, should also depend on physical lumber inventories —see the theory of storage— and time remaining to contract expiration, due to the difference between contract delivery time and the perceived time horizon of the shock.

We find that housing starts announcements contain new information: the unexpected component of housing starts announcements increases expected returns on lumber futures contracts on announcement days. Further, the effects of housing starts shocks decline with inventories, as predicted by the theory of storage. We also find that housing starts shocks have effects that vary across delivery horizon. One to four-month out contracts respond by a larger magnitude to housing starts shocks than do five to eight-month out contracts, reflecting a greater elasticity of supply and demand response at more distant horizons. The response to news does not differ among the nearby contracts, that is, the housing starts shocks have the same impact on return for one to four-month out contracts. However, the response to news does vary among more distant contracts. We find that among five to eight-month out contracts, housing starts effects are more pronounced for contracts with the shortest time to delivery, and less pronounced for the ones with longest time to delivery.

2. Theoretical Implications

The efficient market hypothesis asserts that asset prices move in response to an announcement only if the announcement conveys new information. In most past studies, the price change of an asset is modeled as a function of the unexpected part of the announcement:

$$\ln P_t - \ln P_{t-1} = a + bA_t + \varepsilon_t,$$

where $\ln P_t$ is the natural logarithm of the asset price in question on day t and A_t is the unexpected part of the announcement made on day t. The theory suggests that $b \neq 0$, that is, the return on an asset is related to the news effect of an announcement.

The theory of storage suggests that the return from holding a commodity should depend on the level of physical inventories. Because inventories can be used to stabilize a supply or demand shock, we should expect to see a smaller price response due to a shock when inventories are large. Justification for this result can be seen from the theory developed in Thurman (1988), which is summarized graphically in figure 1.

Equilibrium in the markets for physical inventories is characterized by a Hotelling-like equation where the marginal return to holding inventories is just balanced by its marginal cost. Purchasing a unit of the commodity in the current period costs P_t, the current price. Holding and selling in the next period gives an expected return of $E(P_{t+1})/(1+r)-c-CY(S_t)$, where r and c are the (assumed constant) interest rate and cost of storage, respectively and $CY(S_t)$ is the marginal convenience yield enjoyed by stockholders. Convenience yield is large when aggregate inventories are small and declines with aggregate inventories. There are various interpretations of convenience yield, which are empirically equivalent for our purposes (see Brennan, Williams, and Wright (1997)). In equilibrium, then, inventory equilibrium requires:

$$P_t = E(P_{t+1})/(1+r) - c - CY(S_t). \tag{1}$$

The right-hand side of equation (1) can be interpreted as a demand for inventories, S_t, to carry out of period t. It is decreasing and convex in S_t, as depicted in figure 1(a). This decreasing and convex relationship is also shown in the three-period simulation study of Karali (2007), which is based on the analytical solutions of optimal storage rules.

Flow equilibrium derives from the material balance equation and the excess supply of the commodity:

$$
\begin{aligned}
S_t &= S_{t-1} + S(P_t) - D(P_t) \\
&= S_{t-1} + XS(P_t), \qquad (2)
\end{aligned}
$$

where $S(P_t)$ and $D(P_t)$ are the flow supply and demand functions for the commodity, $XS(P_t)$ is the excess supply function, and S_{t-1} is the level of inventories inherited from period $t-1$. The flow supply curve is upward sloping in $P_t - S_t$ space because excess supply is an increasing function of price. Flow supply shifts rightward for increases in P_{t-1}. Equilibrium price and inventories are determined by the intersection of the inventory demand and flow supply curves in figure 1(a).

Next consider the price effects observed in response to commodity supply and demand shocks. Refer to figure 1(b). There, the uppermost two supply functions are both conditioned on an initial inventory level of S^o_{t-1}. The top supply curve is subject to a negative excess supply shock of ϵ and the one immediately below it is subject to a positive excess supply shock of the same size. The difference between the price responses from the positive and negative shocks (as, say, forecast in the futures market) can be seen from the vertical distance between the two implied equilibrium prices.

The bottommost two supply functions replicate the excess supply shock experiment but for a higher beginning level of inventories, S^1_{t-1}. Because the demand for inventories is flatter in the region of the two equilibria for the second experiment, the implied difference in prices is smaller. When inventories are small, the rapidly declining marginal value of inventories, coupled with the large changes in expected future price with respect to changes in inventories, combine to make price changes large and inventory changes small. When inventories are large, more of the equilibrium adjustment is accommodated by inventory changes.

Beyond the effects of inventories, the futures prices response will depend upon the time between when the shock is observed and when the futures contract matures. The price response of a futures contract to a shock should be smaller when its delivery time is farther away due to greater elasticity of supply and demand curves over longer runs. We generically model these effects as:

$$\ln P_t - \ln P_{t-1} = a + b(S_t, TTD_t)A_t + \varepsilon_t,$$

where S_t is the physical inventory level on day t, TTD_t is the number of days remaining to contract expiration on day t, and $b(S_t, TTD_t)$ is a function of inventories and time to delivery. For an announcement that naturally increases equilibrium commodity price, we should expect to find:

$$\partial(\ln P_t - \ln P_{t-1})/\partial A_t = b(S_t, TTD_t) > 0$$

$$\partial^2(\ln P_t - \ln P_{t-1})/\partial A_t \partial S_t = \partial b(S_t, TTD_t)/\partial S_t < 0$$

$$\partial^2(\ln P_t - \ln P_{t-1})/\partial A_t \partial TTD_t = \partial b(S_t, TTD_t)/\partial TTD_t < 0.$$

3. Econometric Methodology for Multiple Contracts

Time series of futures contracts partially overlap. As time passes, some contracts expire and others begin trading. Figure 2 is a graphical depiction of overlapping futures contracts; each line represents the trading life of a contract. On any given day, we observe price data on multiple contracts. As an example, on September 15th 1992, the date that the September 1992 contract expires, three other futures are traded and so quotes exist for four contracts. As seen in the figure, no two contracts cover the same period. To combine all of these contracts, and account for contemporaneous correlations among the observations from the same trading day, we develop a GLS method, which is explained in detail in the appendix. The method corrects for contemporaneous correlation among contracts and also for delivery-horizon-specific heteroskedasticity.

4. Empirical Measures of Returns, Inventories, Announcements, and Market Expectations

4.a. Futures Returns

We analyze daily settlement prices of lumber futures from the Chicago Mercantile Exchange (CME) in Chicago, from 1992 to 2005. Trading takes place on Monday through Friday between 9:00 am and 1:05 pm (CT) in an open outcry trading pit. Lumber futures contracts expire every two months and the delivery months are January, March, May, July, September, and November. The last trading day for any contract is the last business day prior to the sixteenth calendar day of the contract's delivery month. A new contract is listed on the day after the front month expires, and a total of seven contracts are listed at any point in time, each with a different delivery date up to 14 months into the future. However, in the data we analyze there are price observations on a maximum of five contracts on any given day. Our sample period, in which total of 77 futures contracts are traded, covers the period from July 14, 1992 to November 15, 2005. There are 170 observations for each contract, resulting 13,090 observations in total (see appendix for details). The prices of each of these futures contracts are presented in figure 3.

4.b. Inventories

We employ inventory data from Monthly Wholesale Trade reports published by the U.S. Census Bureau from January 1992 to December 2005. Because they are not seasonally adjusted and stated in current dollars we convert inventory data into real dollar values by using not seasonally adjusted monthly lumber Producer Price Index (PPI) published by the Bureau of Labor Statistics. We interpolate the resulting monthly series, which are now in 1982 (PPI's base year) dollars, by a cubic spline method to obtain daily inventories. Figure 4 presents these daily inventory data.

4.c. Housing Starts Announcements

For announcement data we use housing starts estimates released by the U.S. Census Bureau in monthly New Residential Construction reports along with the estimates of building permits and housing completions. Housing starts is a measure of the initial construction of single-family and multi-family residential units. The start of construction is defined as the beginning of excavation for the footings or foundation of a building. New Residential Construction reports are released between the 15th and 20th of each month at 8:30 am (ET). Thus, the lumber futures market is closed when the announcement is made and opens one-and-a-half hours later. The announcements contain statistics for the previous month. For instance, the report released on August 16, 2005 contains information during the month of July 2005. Market participants know the release date and time of the current reports in advance because these are always listed in the prior release. We consider the news component of the announcements of housing starts estimates by measuring the difference between the actual numbers in the releases and the numbers expected by the market. The selection of market expectation data is discussed in the following section.

4.d. Market Expectations of Housing Starts

For market expectations data, we construct three different forecast series and then compare their forecasting accuracy by Root Mean Square Error (RMSE) measure. Money Market Services (MMS) surveys are conducted every Friday unless it is a holiday. Survey participants include economists from major financial institutions and universities, and are asked to predict the economic variables that will be released during the following week. Survey results are available commercially before the announcements. For more detailed information on MMS survey data, see Aggarwal, Mohanty, and Song (1995).

The first forecast series we consider is the median response of the MMS surveys. For the second forecast series, we assume that the best forecast of the housing starts that will be released in current month is the actual housing starts number released in the previous month. So, the forecast is the housing starts lagged by one month. The third forecast series is constructed by analyzing the released housing starts series through standard time series techniques. This analysis suggests a moving average of order one for the first difference of housing starts, which implies an exponential smoothing forecast. Therefore, the third forecast series takes the following ARIMA(0,1,1) form: $\widehat{HS}_m = HS_{m-1} - 0.57177(HS_{m-1} - \widehat{HS}_{m-1})$, where HS_m is the released housing starts in month m, and HS_{m-1} is the previous month's released housing starts.

Table 1(a) shows summary statistics of housing starts shocks, which are computed as the difference between the actual number and the forecasted number. In our sample period, there were 159 monthly announcements of housing starts. As seen in the table, the smallest RMSE is obtained from the forecast errors of the MMS survey. These median responses of market expectation data perform better than the other two forecast series in predicting housing starts.

The rationality of the MMS survey data has been widely studied in the literature. Studies that support the unbiasedness and efficiency of the MMS survey data include Pearce and Roley (1985), Pearce (1987), and Balduzzi, Elton, and Green (2001). To further analyze the performance of these survey data in forecasting housing starts announcements, the following regression is performed:

$$
\begin{aligned}
\ln F_t - \ln F_{t-1} &= \beta(HS_t - MMS_t) + \gamma MMS_t + \delta(HS_{t-1} - MMS_{t-1}) + \varepsilon_t \\
&\equiv \beta h_t + \gamma MMS_t + \delta h_{t-1} + \varepsilon_t, \qquad (3)
\end{aligned}
$$

where $\ln F_t$ is the natural logarithm of futures contract on day t, HS_t is the released housing starts on day t and MMS_t is the median response of the MMS survey for housing starts on day t. The variables HS_t and MMS_t are defined as zero on nonannouncement days. Efficient market hypothesis states that β should be nonzero, meaning that unanticipated information should cause a price movement. Further, γ and δ should be zero, implying that anticipated part of an announcement and the past forecast error should not cause a price movement. The results from equation (3) are presented in table 2(a). As seen, the estimate for β is positive and statistically significant, showing that the unexpected part of housing starts announcements is positively related to daily return on lumber futures contracts. The estimate for δ is negative and statistically insignificant. This shows that yesterday's forecast error does not affect today's price change. The negative and significant estimate of γ is somewhat puzzling. It seems that the expected component of the announcements has an explanatory power in explaining daily price changes. However, the magnitude of the coefficient is quite small compared to the magnitude of the shock coefficient (approximately 39 times smaller). The table also gives the F statistic and its p-value for the null hypothesis that both γ and δ are zero. We do not reject the null hypothesis with a p-value of 0.13, which is slightly higher than conventional 10% level due to the significant effect of γ. One possible reason that we observe a significant estimate for the market expectations might be the timing difference between the realization of survey results and the announcements.

The unbiasedness of the MMS data is tested in the following model:

$$HS_m = a + bMMS_m + e_m, \tag{4}$$

where HS_m is the released number of housing starts in month m and MMS_m is the median response of the MMS survey forecast for housing starts that will be released in month m. Unbiased expectations imply that $a = 0$ and $b = 1$. Results given in table 2(b) show that the intercept estimate is not statistically different from zero, with a t-statistics of 0.87. Also, the coefficient estimate for the MMS forecasts, b, is significantly different from zero but not significantly different from one. The F-statistic of the joint test of unbiasedness is 2.02 with a p-value of 0.14. Thus, we do not reject the null hypothesis and conclude that the MMS forecasts are unbiased estimators of housing starts.

Another issue that has attracted economists' attention is the possibility of an asymmetric response to positive and negative shocks. See, for example, Li and Engle (1998), Andersen, Bollerslev, Diebold, and Vega (2003), Pearce and Solakoglu (2006). We test whether or not there is asymmetric response to positive and negative housing starts shocks by estimating the following:

$$\ln F_t - \ln F_{t-1} = \psi^P p_t * h_t + \psi^N n_t * h_t + \varepsilon_t, \tag{5}$$

where p_t is a dummy variable that takes the value of one if the housing starts shock on day t, h_t, is positive, and zero otherwise. Similarly, n_t is a dummy variable that takes the value of one if housing starts shock on day t, h_t, is negative, and zero otherwise. Table 2(c) shows that both positive and negative shock coefficients are positive and statistically significant. The null hypothesis that the response to positive and negative shocks is the same is not rejected and has a p-value of 0.49. We conclude that lumber futures contracts respond by a similar magnitude to positive and negative housing starts shocks.[1]

Based on these results, we use the MMS survey data in the rest of the empirics to represent

[1] The equations (3) and (5) were estimated with intercepts as well. The intercept estimates were insignificant and therefore dropped from the equations.

expected housing starts by market participants. Figure 5(a) shows the actual housing starts and the MMS survey forecasts. Figure 5(b) shows the forecast errors, or news component of announcements, which are computed as the difference between the actual number and the forecasted number.

5. Housing Market Shocks and Expected Return: Measures of the Dependence on Time-to-Delivery and Inventories

In this section we explore the price effects of the unexpected component of housing starts announcements and their dependence on inventories and time to delivery. Housing starts shocks are computed as the difference between the released housing starts number and the median response of the Money Market Services (MMS) survey. These announcements are made between the 15th and 20th of each month and lumber contracts expire on the last business day prior to the 16th calendar day of the contract's delivery month. Therefore when the announcement is made in a contract's delivery month, that month's contract has already expired. Because lumber contracts expire every two months and announcements are made every month, the nearest contract on any announcement day will be either one month out or two months out. As a result, the distribution of the time to delivery (TTD) variable is such that it does not take values close to zero on announcement days. Figure 6 shows the distribution of TTD on announcement days. As the figure shows, there are eight TTD clusters on announcement days.

The clusters in figure 6 will be important to our empirical specifications. Figure 7 gives a graphical explanation of the clusters of the TTD variable on announcement days. In the data, at most four contracts were traded on announcement days. When an announcement is made in a delivery month, TTD takes values near 40, 80, 120, and 160 for the four contract traded. Because the contracts expire before the announcements, the nearest delivery contract is a two-month-out contract, and the next contract is four months out. When the announcement is not made in a delivery month, then TTD takes values near 20, 60, 100, and 140. So, the first two nearby contracts are one-month and three-month out contracts. Therefore, on any announcement day, the contracts that are closest to delivery will have TTD values near 20, 40, 60, and 80.

5.a. A Nonparametric Model of Time-to-Delivery Effects

Given the data structure just noted, a reasonable model for the conditional mean of lumber futures returns might include some function of inventories and a linear combination of eight dummy variables for the TTD categories, all interacted with housing starts shocks. Such a specification would look like

$$
\begin{aligned}
\ln F_t - \ln F_{t-1} = BS_t * h_t + b_1 T_{1,t} * h_t + b_2 T_{2,t} * h_t + b_3 T_{3,t} * h_t + b_4 T_{4,t} * h_t \\
+ b_5 T_{5,t} * h_t + b_6 T_{6,t} * h_t + b_7 T_{7,t} * h_t + b_8 T_{8,t} * h_t + \varepsilon_t,
\end{aligned}
\tag{6}
$$

where $\ln F_t$ is the natural logarithm of the futures price on day t, S_t is the lumber inventory level on day t, and h_t is the difference between released housing starts and the market expectation on day t. The dummy variables are defined as follows:

$$T_{1,t} = \begin{cases} 1, & \text{if } 15 \leq TTD_t \leq 21 \\ 0, & \text{otherwise} \end{cases} \quad T_{2,t} = \begin{cases} 1, & \text{if } 31 \leq TTD_t \leq 42 \\ 0, & \text{otherwise} \end{cases}$$

$$T_{3,t} = \begin{cases} 1, & \text{if } 57 \leq TTD_t \leq 64 \\ 0, & \text{otherwise} \end{cases} \quad T_{4,t} = \begin{cases} 1, & \text{if } 75 \leq TTD_t \leq 86 \\ 0, & \text{otherwise} \end{cases}$$

$$T_{5,t} = \begin{cases} 1, & \text{if } 98 \leq TTD_t \leq 106 \\ 0, & \text{otherwise} \end{cases} \quad T_{6,t} = \begin{cases} 1, & \text{if } 117 \leq TTD_t \leq 128 \\ 0, & \text{otherwise} \end{cases}$$

$$T_{7,t} = \begin{cases} 1, & \text{if } 140 \leq TTD_t \leq 150 \\ 0, & \text{otherwise} \end{cases} \quad T_{8,t} = \begin{cases} 1, & \text{if } 158 \leq TTD_t \leq 169 \\ 0, & \text{otherwise} \end{cases},$$

where TTD_t is the number of days remaining to delivery on day t. The dummy variables are constructed according to the minimum and maximum TTD observed on announcement days for each one of eight TTD categories. The h_t variable is defined as zero on nonannouncement days. Table 1(b) presents summary statistics of the key variables.

Estimates of specification (6) are presented in table 3(a). Both OLS and GLS estimates of all dummy coefficients are positive and statistically significant. The GLS estimate of B, which is the partial inventory effect of housing starts shocks, is negative and statistically significant at the 10% level (one-tailed). This implies that as inventory levels become smaller, housing starts shocks affect expected return on lumber contracts by a larger amount. As inventory levels increase, the effect of housing starts shocks on expected return is smaller. To see the overall effect of housing starts shocks on expected return, one needs to compute the derivative $\partial(\ln F_t - \ln F_{t-1})/\partial h_t$ for each one of eight TTD categories and evaluate the derivative at a particular value of inventories. Table 3(c) shows the marginal effects of housing starts shocks computed by using the GLS estimates of parameters. The marginal effect of housing starts shocks is positive for each TTD category in all cases. While these effects are significant for medium and low levels of inventories, they are statistically insignificant when the derivatives are evaluated at the maximum value of inventories.

Figure 8 shows predicted daily return from the table 3(a) GLS estimates for each TTD category separately. The predicted returns are computed using the mean values of inventories and housing starts shocks and so answer the following question: on a typical announcement day, with inventories equal to their sample, how does expected return on a futures contract vary with its time to delivery? From figure, it appears that after the first four categories ($TTD > 86$), the effect of housing starts on the expected return on lumber futures contracts starts to decline. The nearby contracts have a higher response to h_t than do the distant ones for a given shock and inventory level. This suggests a more parsimonious version of equation (6), with a parametric dependence on TTD. This, in turn, suggests performing F-tests of the equality of the parameters from equation (6). Table 3(b) shows such tests.

The hypothesis that all TTD dummy coefficients are zero is rejected at the 10% level with GLS estimates. The F-statistic for the hypothesis that the average of the first four TTD dummy coefficients is equal to the average of the last four dummy coefficients also is rejected at the 10% level. Although the hypothesis of the equality of all eight dummy coefficients is not rejected with a p-value of 0.16, the equality of averages is rejected. The clustered pattern of predicted price changes seen in figure 8 can be tested by the hypotheses:

$$H_0^1 : b_1 = b_2 = b_3 = b_4$$

$$H_0^2 : b_5 = b_6 = b_7 = b_8.$$

The first null hypothesis, H_0^1, is not rejected with a p-value of 0.88. The second null hypothesis, H_0^2, is also not rejected with a p-value of 0.80. Further, the joint test of these null hypotheses is not rejected with a p-value of 0.67. Given these test results and the visual pattern evident in figure 8, we define two dummy variables: one for the first four TTD categories and one for the last four TTD categories.

To this end we define the following new model:

$$\ln F_t - \ln F_{t-1} = BS_t * h_t + b_{1,4}T_{1,4,t} * h_t + b_{5,8}T_{5,8,t} * h_t + \varepsilon_t, \tag{7}$$

where $T_{1,4,t} = T_{1,t} + T_{2,t} + T_{3,t} + T_{4,t}$ and $T_{5,8,t} = T_{5,t} + T_{6,t} + T_{7,t} + T_{8,t}$. In this model, the effects of housing starts shocks are the same for the first four contracts and for the last four contracts. However, the effects can change between the two groups.

Estimation results for equation (7) are given in table 4(a). Once again, we obtain negative estimates for B, and positive estimates for the dummy coefficients. The estimate of B is statistically significant at the 10% level (one-tailed test). The estimates of $b_{1,4}$ and $b_{5,8}$ are also significant with t-values of 2.62 and 2.37. To see the overall marginal impact of housing starts on expected return, the derivatives $\partial(\ln F_t - \ln F_{t-1})/\partial h_t$ for both TTD categories are computed with GLS estimates. These are presented in table 4(b). As before, the marginal impact of housing starts is positive for each TTD category, however, insignificant at the maximum value of inventories. The estimated $\partial(\ln F_t - \ln F_{t-1})/\partial h_t$ for the first TTD category ($T_{1,4,t} = 1$) is larger in all cases. This reflects that housing starts shocks have larger impacts on nearby contracts than the distant ones. As evidenced by the sign of B, the effect of housing starts shocks on expected return declines with inventories.

When we consider the GLS estimates and evaluate the marginal effect of housing starts shocks at the mean value of inventories (second row of table 4(b)), a change in shocks from their minimum value, -273,000 housing units, to their maximum value, 200,000 housing units, causes a 3.6 percentage point increase in the expected returns of contracts that have 86 or less days to delivery. The same change in shocks causes a 2.9 percentage point increase in the expected returns of contracts that have 98 or more days to delivery. The respective increases in the expected return with the minimum value of inventories are 4.7 and 4.0 percentage points. These results show that supply and demand curves are inelastic in the short run, so that housing starts shocks, a demand shock, have a larger price impact on nearby contracts. On the other hand, supply and demand curves become more elastic in the long run, hence the price impact of housing starts shocks is smaller for more distant contracts. Also note that, as the theory of storage predicts, we obtain larger price response to shocks when inventories are small.

The results of F-tests performed on parameter estimates from the model with two dummy variables are given in the bottom part of table 4(a). The hypothesis that both dummy coefficients are equal to zero is rejected. The equality of $b_{1,4}$ and $b_{5,8}$ is also rejected with a p-value of 0.01.

5.b. A Piecewise Linear Model of Time-to-Delivery Effects

The results in the previous section suggest that the effect of housing starts shocks on expected return is different across contracts, and the effect declines with contract horizon.

To further analyze this time-to-delivery pattern we construct a new model that captures the time-horizon effects discussed above with a piecewise linear spline function of time to delivery. We define expected return as

$$E[\ln F_t - \ln F_{t-1}] = (A + BS_t) * h_t + g(TTD_t) * h_t, \tag{8}$$

where

$$g(TTD_t) = (1 - d_t) * (a + bTTD_t) + d_t * (c + eTTD_t)$$

and $d_t = 1$ if $TTD_t \geq T^*$, zero otherwise. A graphical depiction is given in figure 9. At T^*, $a + bT^* = c + eT^*$. So, we can solve for a and substitute in $g(TTD_t)$ to ensure that the two linear pieces are connected at T^*. Solving for a gives us $a = c + (e - b)T^*$. When we substitute the solution for a into the function $g(TTD_t)$, we obtain:

$$\begin{aligned} g(TTD_t) &= (1 - d_t)(c + (e - b)T^* + bTTD_t) + d_t(c + eTTD_t) \\ &= c + b[(1 - d_t)(TTD_t - T^*)] + e[(1 - d_t)T^* + d_tTTD_t]. \end{aligned}$$

Substituting this into equation (8) yields the new regression model:

$$\begin{aligned} \ln F_t - \ln F_{t-1} &= (A + BS_t) * h_t + \Big\{c + b[(1 - d_t)(TTD_t - T^*)] \\ &\qquad + e[(1 - d_t)T^* + d_tTTD_t]\Big\} * h_t \\ &= (A + c)h_t + BS_t * h_t + b(1 - d_t)(TTD_t - T^*) * h_t \\ &\qquad + e[(1 - d_t)T^* + d_tTTD_t] * h_t + \varepsilon_t. \end{aligned} \tag{9}$$

Given the results from the previous section, we set $T^* = 90$ because this splits the eight clusters of TTD variable observed on announcement days in half. Moreover, TTD never equals 90 on announcement days (see figure 6).

Regression results from the piecewise linear model are presented in table 5(a). The estimates of $A + c$ and b are positive but only the former is statistically significant. The insignificance of b (the slope before T^*) is in accordance with what we expect. For contracts that have less than 90 days to delivery, the effect of housing starts shocks does not change with TTD. The four nearby contracts respond to shocks by similar amounts. On the other hand, the estimate of e (the slope after T^*) is negative and statistically significant. This shows that when contracts have more than 90 days to delivery, the effect of housing starts shocks decreases with TTD. Among these contracts, the one with nearest delivery is affected by more than ones with more distant delivery. After some critical value of TTD, we observe a declining impact of housing starts shocks with time to delivery. The estimate of B is negative and statistically significant at the 10% level (one-tailed test). This confirms that the effect of housing starts shocks on expected return declines with inventories. When inventories are small, housing starts shocks have more impact on expected returns, whereas when inventories are large, shocks have less impact on returns.

Interest centers on the overall impact of housing starts shocks on expected returns. Once again, we need to compute the derivative $\partial(\ln F_t - \ln F_{t-1})/\partial h_t$ and evaluate it at some predetermined values of inventories and time to delivery. Table 5(b) shows the marginal effects computed at the minimum, mean, and maximum value of inventories for the cases when TTD takes its maximum value, the region beyond T^*, and its minimum value, the region before T^*, by using the GLS estimates. The second derivatives do not depend on any variable; they are simply the estimates of B, b, or e. We see that the expected return is positively related to housing starts shocks in each case. Once again, this marginal effect is insignificant at the maximum value of inventories. For the contracts that have more than 90 days to delivery, a change in the housing starts shocks from their minimum to maximum value results in a 2.3 percentage point increase in the expected return when the derivative is evaluated at the mean value of inventories. The same change in shocks causes a 3.5 percentage point increase in the expected return for contracts that have less than 90 days to delivery. When the derivative is evaluated at the minimum value of inventories, the respective changes are 3.4 and 4.6 percentage points. The second derivative $\partial^2(\ln F_t - \ln F_{t-1})/\partial h_t \partial S_t$ is equal to B. As discussed before, the effect of housing starts shocks on expected return decreases with inventories. The second derivative $\partial^2(\ln F_t - \ln F_{t-1})/\partial h_t \partial TTD_t$ is equal to b for the values of TTD that are less than 90. As seen in the table, it is positive and insignificant. The effect of housing starts shocks on expected return does not significantly change with time to delivery for $TTD < T^*$. The second derivative $\partial^2(\ln F_t - \ln F_{t-1})/\partial h_t \partial TTD_t$ is equal to e for the values of TTD that are equal or greater than 90. The estimate of e is negative and statistically significant. Therefore, the effect of housing starts shocks on expected return declines with TTD for more distant contracts. The marginal effect of housing starts shocks, $\partial(\ln F_t - \ln F_{t-1})/\partial h_t$, is plotted against TTD in figure 10 using the GLS estimates and evaluating the predicted effect at all sample points in the data.

The bottom part of table 5(a) gives results from the F-tests performed on parameter estimates. The hypothesis of no TTD effect on housing starts shocks, $b = e = 0$, is rejected with a p-value of 0.04. The hypothesis that the effect of housing starts shocks is the same for both TTD categories, $b = e$, is rejected with a p-value of 0.10.

6. Conclusions

The expected return on lumber futures is positively related to housing starts shocks. Housing starts convey important news to lumber markets. When market participants over- or underestimate housing starts, these forecast errors result in a price movement. In this analysis, we conclude that the effect of announcements on daily returns declines with physical inventories. As inventory levels become larger, these shocks cause a smaller change in return. This provides empirical evidence for a central implication from the theory of storage. It is the first such evidence, to our knowledge, that connects announcement effects in futures markets with measured commodity inventories.

The dependence of the effect of housing starts shocks on time to delivery is analyzed in different settings. All models we investigate lead to the conclusion that the effect of shocks is not the same for all contracts, and what differs across contracts is their delivery horizon. The most nearby (one to four-month out) contracts respond by a larger magnitude to housing starts shocks than do the more distant (five to eight-month out) contracts, which we interpret as reflecting a greater elasticity of supply and demand response at more distant horizons. Further, the decline is found not to be linear. The response to news does not differ among the nearby contracts (up to four months out), but does vary among more distant contracts. Among contracts more than four months out, housing starts effects are more pronounced for contracts with the shortest time to delivery, and less pronounced for those with longest time to delivery. These results from the lumber market are quite similar to those found by Smith in the corn market.[2]

[2] Our work is closely related in two ways to the work of Smith (2005) and his Partially Overlapping Time Series (POTS) model. Methodologically, we and he develop econometric models that exploit the information available from simultaneously traded futures contracts. Substantively, we both study the economics of futures

Our results suggest that the information embodied in housing starts announcements is short term. Other information sources may have different time profiles, suggesting further research using multiple futures contracts to characterize the term structure of information.

---

price movements and how they are affected by time-to-delivery and physical inventories.

There are, however, important differences between his work and ours. Smith studies time series of futures contract prices without reference to exogenous observables. He models unobservable information flows using a purely time series methodology—a sophisticated factor analytic model, incorporating the Kalman filter and cubic splines to specify the dependence of futures price movements on unobserved new- and old-crop information shocks. GARCH effects are included as well. To estimate the model, he uses an approximate EM algorithm in cooperation with more standard numerical methods to maximize the Gaussian likelihood function. In contrast, we relate futures price movements to observed information flows—Housing Starts announcements relative to survey forecasts of them—and condition the reaction to market news on observed measures of physical inventories. Because our measures of news and inventories are observable we are naturally led to a different empirical model, a Generalized Least Squares regression methodology based on the unbalanced panel structure of futures contracts. Our method has some advantages in its transparency and ready adaptability to other partially overlapping times series contexts with observable covariates.

With regards to economic results, the obvious contrast with Smith is that he studies corn and we study lumber. But despite the difference in commodities, and far from taking POTS shots at Smith's results, we find that our empirical results are quite consistent with his. We estimate the effect of inventories on the futures market's response to information and confirm directly the indirect inference that Smith draws from his time series factor model: that volatility (in our case, the size of the market response to news) is inversely related to inventories. We also find that volatility increases as time to delivery approaches (the Samuelson effect, also studied by Anderson and Danthine (1983), Anderson (1985), Milonas (1986), Leistikow (1989), and Black and Tonks (2000)). Also similar to Smith's findings, for contracts with more than about four months to delivery volatility is roughly constant with respect to time to delivery. We find this result both for observed announcement effects and for daily volatility due to unobserved factors.

APPENDIX

A. Econometric Methodology for Multiple Contracts

To combine simultaneously traded contracts we choose futures contracts according to the availability of lumber inventory data, which are available from January 1992 to December 2005. Therefore, we only consider contracts that started trading after January 31, 1992 and ones that matured before December 31, 2005. This results in a sample period of July 14, 1992-November 15, 2005. Within this time interval there are 77 contracts. Because each contract has a different number of active trading days, the maximum TTD varies somewhat among contracts. For consistency, we trimmed the data set to include only those observations for which TTD is equal to or less than 169, which is the shortest TTD in the sample. The resulting data set has 170 observations on each of 77 contracts (13,090 observations in total). During the sample period, at most five contracts were traded on a given day and on some days three or four were traded.

We organize the observations into five groups, according to the number of contracts traded on a given day. We denote the groups as:

$$\begin{aligned}
\mathbf{y}_1 &= \begin{bmatrix} y_{1,t_1^1} & y_{1,t_2^1} & \cdots & y_{1,t_{n_1}^1} \end{bmatrix}', \\
\mathbf{y}_2 &= \begin{bmatrix} y_{1,t_1^2} & y_{2,t_1^2} & y_{1,t_2^2} & y_{2,t_2^2} & \cdots & y_{1,t_{n_2}^2} & y_{2,t_{n_2}^2} \end{bmatrix}', \\
\mathbf{y}_3 &= \begin{bmatrix} y_{1,t_1^3} & y_{2,t_1^3} & y_{3,t_1^3} & y_{1,t_2^3} & y_{2,t_2^3} & y_{3,t_2^3} & \cdots & y_{1,t_{n_3}^3} & y_{2,t_{n_3}^3} & y_{3,t_{n_3}^3} \end{bmatrix}', \\
\mathbf{y}_4 &= \begin{bmatrix} y_{1,t_1^4} & y_{2,t_1^4} & y_{3,t_1^4} & y_{4,t_1^4} & y_{1,t_2^4} & y_{2,t_2^4} & y_{3,t_2^4} & y_{4,t_2^4} & \cdots & y_{1,t_{n_4}^4} & y_{2,t_{n_4}^4} & y_{3,t_{n_4}^4} & y_{4,t_{n_4}^4} \end{bmatrix}', \\
\mathbf{y}_5 &= \begin{bmatrix} y_{1,t_1^5} & y_{2,t_1^5} & y_{3,t_1^5} & y_{4,t_1^5} & y_{5,t_1^5} & y_{1,t_2^5} & y_{2,t_2^5} & y_{3,t_2^5} & y_{4,t_2^5} & y_{5,t_2^5} & \cdots & y_{1,t_{n_5}^5} & y_{2,t_{n_5}^5} & y_{3,t_{n_5}^5} & y_{4,t_{n_5}^5} & y_{5,t_{n_5}^5} \end{bmatrix}',
\end{aligned}$$

where t_k^j refers to the kth day in the j-contract group; for example t_2^1 refers to the second day during which only one contract was traded, t_5^2 refers to the fifth day during which exactly two contracts were traded. The variable y_{i,t_k^j} refers to the change in the price of the ith-nearest-delivery contract in the j-contracts-a-day group on day t_k^j; for example, y_{2,t_{17}^4} is the price change on the second-nearest-delivery contract (two to four months out) in the group with four contracts traded per day, on the 17th day in that group. The variable n_j, $j = 1, 2, 3, 4, 5$ is the number of trading days with j contracts traded in total. There are 3,365 trading days in the data and $n_1 + n_2 + n_3 + n_4 + n_5 = 3,365$. Specifically, $n_1 = 86$, $n_2 = 85$, $n_3 = 96$, $n_4 = 2,944$, $n_5 = 154$: $\mathbf{y}_1$ is $n_1 \times 1$, $\mathbf{y}_2$ is $2 \cdot n_2 \times 1$, $\mathbf{y}_3$ is $3 \cdot n_3 \times 1$, $\mathbf{y}_4$ is $4 \cdot n_4 \times 1$, and $\mathbf{y}_5$ is $5 \cdot n_5 \times 1$.

The five grouped vectors are stacked to compose a single data vector:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1{}' & \mathbf{y}_2{}' & \mathbf{y}_3{}' & \mathbf{y}_4{}' & \mathbf{y}_5{}' \end{bmatrix}',$$

Once the column vectors are stacked, the resulting vector is $13,090 \times 1$. We organize the entire data set in this way to have column vectors for each variable ordered according to the number of contracts traded on a given day.

Because information flows to the market affect, to some degree, all lumber contracts, price observations from the same calendar date will be correlated with each other. Simply pooling the time series and ignoring contemporaneous correlation would falsely imply that each observation provided an independent observation on the relationship between prices, TTD, and inventories. To deal with correlation among observations from the same day, we first run a regression for the model of interest like $\ln F_t - \ln F_{t-1} = \alpha + \delta h_t + \beta S_t * h_t + \gamma TTD_t * h_t + \varepsilon_t$. The variable $\ln F_t$ is the natural logarithm of futures price on day t, h_t is the difference between the released housing starts and market expectation on day t, S_t is the inventory level on day t, and TTD_t is the number of days remaining to delivery on day t. From any such regression we obtain the residual vector ordered as above. Then we construct submatrices from this residual vector. The residual submatrices are:

$$\mathbf{e}_1 = \begin{bmatrix} e_{1,t_1^1} & e_{1,t_2^1} & \cdots & e_{1,t_{n_1}^1} \end{bmatrix}', \quad \mathbf{e}_2 = \begin{bmatrix} e_{1,t_1^2} & e_{1,t_2^2} & \cdots & e_{1,t_{n_2}^2} \\ e_{2,t_1^2} & e_{2,t_2^2} & \cdots & e_{2,t_{n_2}^2} \end{bmatrix}', \quad \mathbf{e}_3 = \begin{bmatrix} e_{1,t_1^3} & e_{1,t_2^3} & \cdots & e_{1,t_{n_3}^3} \\ e_{2,t_1^3} & e_{2,t_2^3} & \cdots & e_{2,t_{n_3}^3} \\ e_{3,t_1^3} & e_{3,t_2^3} & \cdots & e_{3,t_{n_3}^3} \end{bmatrix}',$$

$$\mathbf{e}_4 = \begin{bmatrix} e_{1,t_1^4} & e_{1,t_2^4} & \cdots & e_{1,t_{n_4}^4} \\ e_{2,t_1^4} & e_{2,t_2^4} & \cdots & e_{2,t_{n_4}^4} \\ e_{3,t_1^4} & e_{3,t_2^4} & \cdots & e_{3,t_{n_4}^4} \\ e_{4,t_1^4} & e_{4,t_2^4} & \cdots & e_{4,t_{n_4}^4} \end{bmatrix}', \quad \mathbf{e}_5 = \begin{bmatrix} e_{1,t_1^5} & e_{1,t_2^5} & \cdots & e_{1,t_{n_5}^5} \\ e_{2,t_1^5} & e_{2,t_2^5} & \cdots & e_{2,t_{n_5}^5} \\ e_{3,t_1^5} & e_{3,t_2^5} & \cdots & e_{3,t_{n_5}^5} \\ e_{4,t_1^5} & e_{4,t_2^5} & \cdots & e_{4,t_{n_5}^5} \\ e_{5,t_1^5} & e_{5,t_2^5} & \cdots & e_{5,t_{n_5}^5} \end{bmatrix}'. \tag{A-1}$$

In the preceding matrices each row represents a trading day. The first column of each submatrix shows residuals associated with the first nearby contract. Similarly, the second columns are associated with the second nearby contract, the third columns with the third nearby contract, and so on. Using these submatrices, we compute the sample means of squared residuals by column to obtain variance estimates of first nearby contracts, second nearby contracts, etc. We also estimate covariances between first and second nearby contracts, between first and third nearby contracts, etc. by computing sample means of the products of related residuals. The variance estimate of the ith-nearest-delivery contract is computed as:

$$\hat{\sigma}_i^2 = \sum_{j=i}^{5} \frac{1}{n_j} \sum_{k=1}^{n_j} (e_{i,t_k^j})^2.$$

More specifically, the variance estimates are:

$$\hat{\sigma}_1^2 = \sum_{j=1}^{5} \frac{1}{n_j} \sum_{k=1}^{n_j} (e_{1,t_k^j})^2, \quad \hat{\sigma}_2^2 = \sum_{j=2}^{5} \frac{1}{n_j} \sum_{k=1}^{n_j} (e_{2,t_k^j})^2, \quad \hat{\sigma}_3^2 = \sum_{j=3}^{5} \frac{1}{n_j} \sum_{k=1}^{n_j} (e_{3,t_k^j})^2,$$

$$\hat{\sigma}_4^2 = \sum_{j=4}^{5} \frac{1}{n_j} \sum_{k=1}^{n_j} (e_{4,t_k^j})^2, \quad \hat{\sigma}_5^2 = \frac{1}{n_5} \sum_{k=1}^{n_5} (e_{5,t_k^5})^2.$$

The estimate of the covariance between the ith-nearest-delivery and ℓth-nearest-delivery contracts is computed as:

$$\hat{\sigma}_{i\ell} = \sum_{j=\max\{i,\ell\}}^{5} \frac{1}{n_j} \sum_{k=1}^{n_j} e_{i,t_k^j} e_{\ell,t_k^j}.$$

For example,

$$\hat{\sigma}_{23} = \sum_{j=3}^{5} \frac{1}{n_j} \sum_{k=1}^{n_j} e_{2,t_k^j} e_{3,t_k^j},$$

or

$$\hat{\sigma}_{23} = \frac{1}{n_3} \sum_{k=1}^{n_3} e_{2,t_k^3} e_{3,t_k^3} + \frac{1}{n_4} \sum_{k=1}^{n_4} e_{2,t_k^4} e_{3,t_k^4} + \frac{1}{n_5} \sum_{k=1}^{n_5} e_{2,t_k^5} e_{3,t_k^5}.$$

Using these estimates we construct a 5×5 variance-covariance matrix as

$$\mathbf{\Sigma} = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} & \hat{\sigma}_{13} & \hat{\sigma}_{14} & \hat{\sigma}_{15} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 & \hat{\sigma}_{23} & \hat{\sigma}_{24} & \hat{\sigma}_{25} \\ \hat{\sigma}_{13} & \hat{\sigma}_{23} & \hat{\sigma}_3^2 & \hat{\sigma}_{34} & \hat{\sigma}_{35} \\ \hat{\sigma}_{14} & \hat{\sigma}_{24} & \hat{\sigma}_{34} & \hat{\sigma}_4^2 & \hat{\sigma}_{45} \\ \hat{\sigma}_{15} & \hat{\sigma}_{25} & \hat{\sigma}_{35} & \hat{\sigma}_{45} & \hat{\sigma}_5^2 \end{bmatrix}.$$

This method allows variances to change for the five contract types and covariances to vary among contracts that have the same discrepancy in delivery month. Even though there are two-month delivery discrepancies between the first and second nearby contracts and between the second and third nearby contracts, the covariance between the first and second contracts ($\hat{\sigma}_{12}$) is not assumed to be the same as that between the second and third contracts ($\hat{\sigma}_{23}$). For all the models discussed in this paper, it is observed that variance estimates decline with time to delivery. That is, the variance estimates of distant contracts are smaller than those for nearby contracts. Further, the covariance estimates between contracts decrease as the discrepancy in contract delivery times becomes larger.

Once the variance-covariance matrix from a regression is obtained, its Cholesky decomposition can be used to apply a GLS transformation to the data to eliminate contemporaneous correlation among residuals and to adjust the observations to be homoskedastic. To accomplish this we construct five submatrices for each variable like the ones in equation (A-1).

We then compute the Cholesky factors, $\mathbf{C}_i$s, of the five submatrices of $\mathbf{\Sigma}$ using the following:

$$\Sigma_1 = C_1' C_1 = \hat{\sigma}_1^2, \quad \mathbf{\Sigma}_2 = \mathbf{C}_2'\mathbf{C}_2 = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 \end{bmatrix}, \quad \mathbf{\Sigma}_3 = \mathbf{C}_3'\mathbf{C}_3 = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} & \hat{\sigma}_{13} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 & \hat{\sigma}_{23} \\ \hat{\sigma}_{13} & \hat{\sigma}_{23} & \hat{\sigma}_3^2 \end{bmatrix},$$

$$\mathbf{\Sigma}_4 = \mathbf{C}_4'\mathbf{C}_4 = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} & \hat{\sigma}_{13} & \hat{\sigma}_{14} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 & \hat{\sigma}_{23} & \hat{\sigma}_{24} \\ \hat{\sigma}_{13} & \hat{\sigma}_{23} & \hat{\sigma}_3^2 & \hat{\sigma}_{34} \\ \hat{\sigma}_{14} & \hat{\sigma}_{24} & \hat{\sigma}_{34} & \hat{\sigma}_4^2 \end{bmatrix}, \quad \mathbf{\Sigma}_5 = \mathbf{C}_5'\mathbf{C}_5 = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} & \hat{\sigma}_{13} & \hat{\sigma}_{14} & \hat{\sigma}_{15} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 & \hat{\sigma}_{23} & \hat{\sigma}_{24} & \hat{\sigma}_{25} \\ \hat{\sigma}_{13} & \hat{\sigma}_{23} & \hat{\sigma}_3^2 & \hat{\sigma}_{34} & \hat{\sigma}_{35} \\ \hat{\sigma}_{14} & \hat{\sigma}_{24} & \hat{\sigma}_{34} & \hat{\sigma}_4^2 & \hat{\sigma}_{45} \\ \hat{\sigma}_{15} & \hat{\sigma}_{25} & \hat{\sigma}_{35} & \hat{\sigma}_{45} & \hat{\sigma}_5^2 \end{bmatrix}.$$

Then we premultiply the variable submatrices by the associated $(\mathbf{C}_i')^{-1}$, $i = 1, 2, 3, 4, 5$. For example, for the data group during which four contracts were traded on a given day, the GLS transformation to the variable $\mathbf{x}$ is performed as:

$$\mathbf{x}_4^{*\prime} = (\mathbf{C}_4')^{-1}\mathbf{x}_4' = (\mathbf{C}_4')^{-1} \begin{bmatrix} x_{1,t_1^4} & x_{1,t_2^4} & \cdots & x_{1,t_{n_4}^4} \\ x_{2,t_1^4} & x_{2,t_2^4} & \cdots & x_{2,t_{n_4}^4} \\ x_{3,t_1^4} & x_{3,t_2^4} & \cdots & x_{3,t_{n_4}^4} \\ x_{4,t_1^4} & x_{4,t_2^4} & \cdots & x_{4,t_{n_4}^4} \end{bmatrix},$$

where $(\mathbf{C}_4')^{-1}$ is 4×4 and $\mathbf{x}_4'$ is $4 \times n_4$.

With this method, we obtain new data that are corrected for the contemporaneous correlation among contracts and for delivery-horizon-specific heteroskedasticity.

References

Aggarwal, R., S. Mohanty, and F. Song. 1995. "Are Survey Forecasts of Macroeconomic Variables Rational?" *Journal of Business* 68:99–119.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and C. Vega. 2003. "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange." *The American Economic Review* 93:38–62.

Anderson, R.W. 1985. "Some Determinants of the Volatility of Futures Prices." *The 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 Prices." *The Review of Economic Studies* 50:249–266.

Balduzzi, P., E.J. Elton, and T.C. Green. 2001. "The Effects of Macroeconomic News on High Frequency Exchange Rate Behavior." *Journal of Financial and Quantitative Analysis* 36:523–543.

Binder, J.J. 1998. "The Event Study Methodology Since 1969." *Review of Quantitative Finance and Accounting* 11:111–137.

Black, J., and I. Tonks. 2000. "Time Series Volatility of Commodity Futures Prices." *The Journal of Futures Markets* 20:127–144.

Brennan, D., J.C. Williams, and B.D. Wright. 1997. "Convenience Yield without the Convenience: A Spatial-Temporal Interpretation of Storage Under Backwardation." *The Economic Journal* 107:1009–1022.

Brennan, M.J. 1958. "The Supply of Storage." *The American Economic Review* 48:50–72.

Fama, E.F., L. Fisher, M.C. Jensen, and R. Roll. 1969. "The Adjustment of Stock Prices to New Information." *International Economic Review* 10:1–21.

Karali, B. 2007. "Information and Price Response in Storable Commodity Futures Markets: An Application to Lumber Contracts." PhD dissertation, North Carolina State University.

Leistikow, D. 1989. "Announcements and Futures Price Variability." *The Journal of Futures Markets* 9:477–486.

Li, L., and R.F. Engle. 1998. "Macroeconomic Announcements and Volatility of Treasury Futures." Discussion paper 98-27, University of California, San Diego.

Milonas, N.T. 1986. "Price Variability and the Maturity Effect in Futures Markets." *The Journal of Futures Markets* 6:443–460.

Pearce, D.K. 1987. "Short-term Inflation Expectations: Evidence from a Monthly Survey." *Journal of Money, Credit, and Banking* 19:388–395.

Pearce, D.K., and V.V. Roley. 1985. "Stock Prices and Economic News." *Journal of Business* 58:49–67.

Pearce, D.K., and M.N. Solakoglu. 2006. "Macroeconomic News and Exchange Rates." *Journal of International Financial Markets, Institutions & Money* Forthcoming.

Rucker, R.R., W.N. Thurman, and J.K. Yoder. 2005. "Estimating the Structure of Market Reaction to News: Information Events and Lumber Futures Prices." *American Journal of Agricultural Economics* 87:482–500.

Scheinkman, J.A., and J. Schechtman. 1983. "A Simple Competitive Model with Production and Storage." *The Review of Economic Studies* 50:427–441.

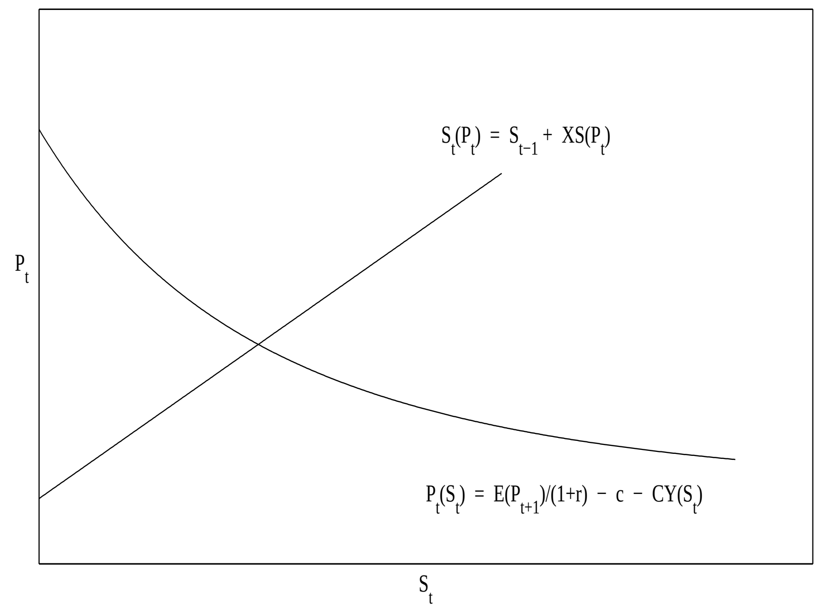
Smith, A. 2005. "Partially Overlapping Time Series: A New Model for Volatility Dynamics in Commodity Futures." *Journal of Applied Econometrics* 20:405–422.

Telser, L.G. 1978. *Economic Theory and the Core*, 3rd ed. The University of Chicago Press.

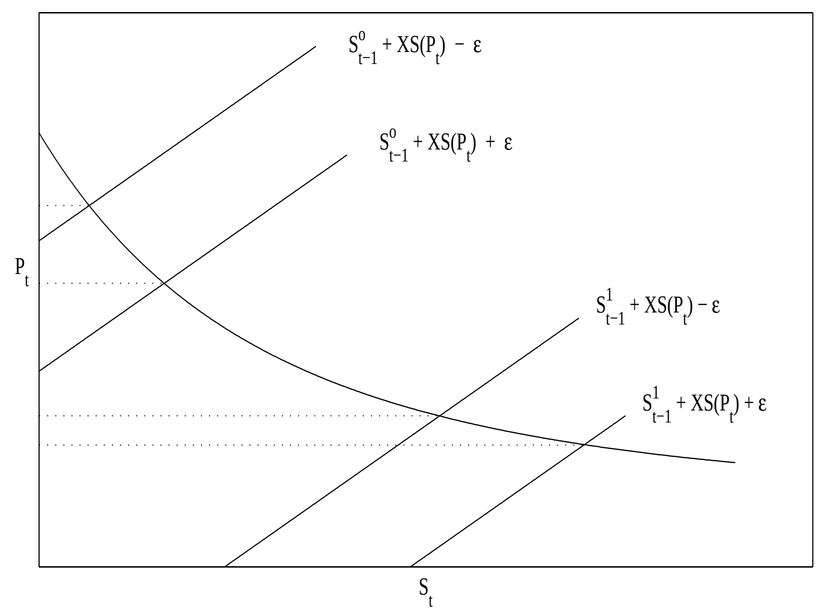
Thurman, W.N. 1988. "Speculative Carryover: An Empirical Examination of the U.S. Refined Copper Market." *The RAND Journal of Economics* 19:420–437.

Williams, J.C., and B.D. Wright. 1991. *Storage and Commodity Markets*. Cambridge University Press.

Figures



(a) Equilibrium Determination of Inventories



(b) Price Changes at Different Levels of S_{t-1}

Figure 1: Equilibrium in Storable Commodity Markets

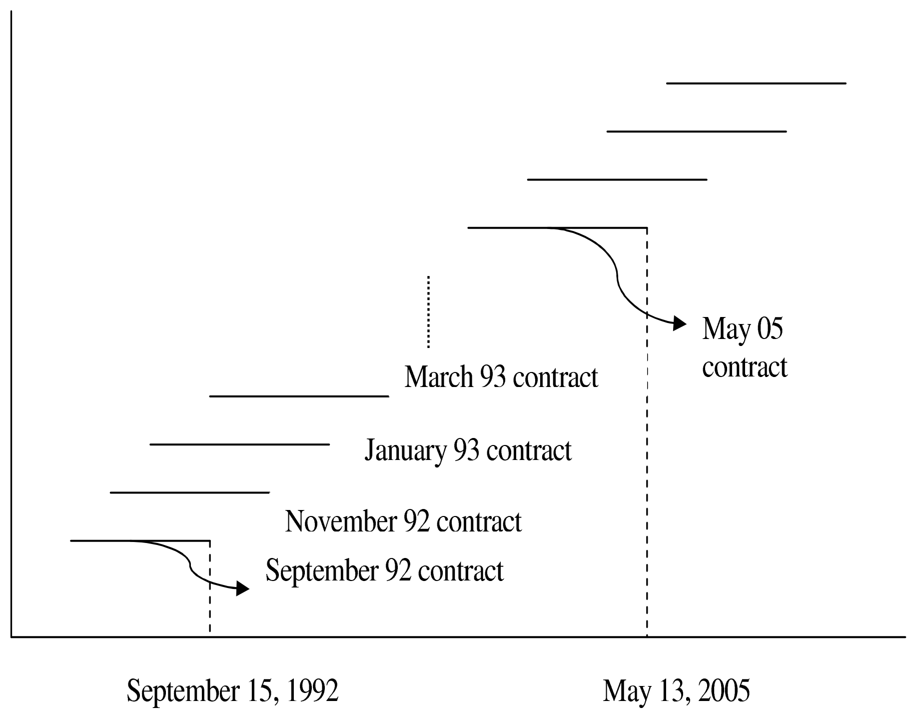


Figure 2: Overlapping Futures Contracts

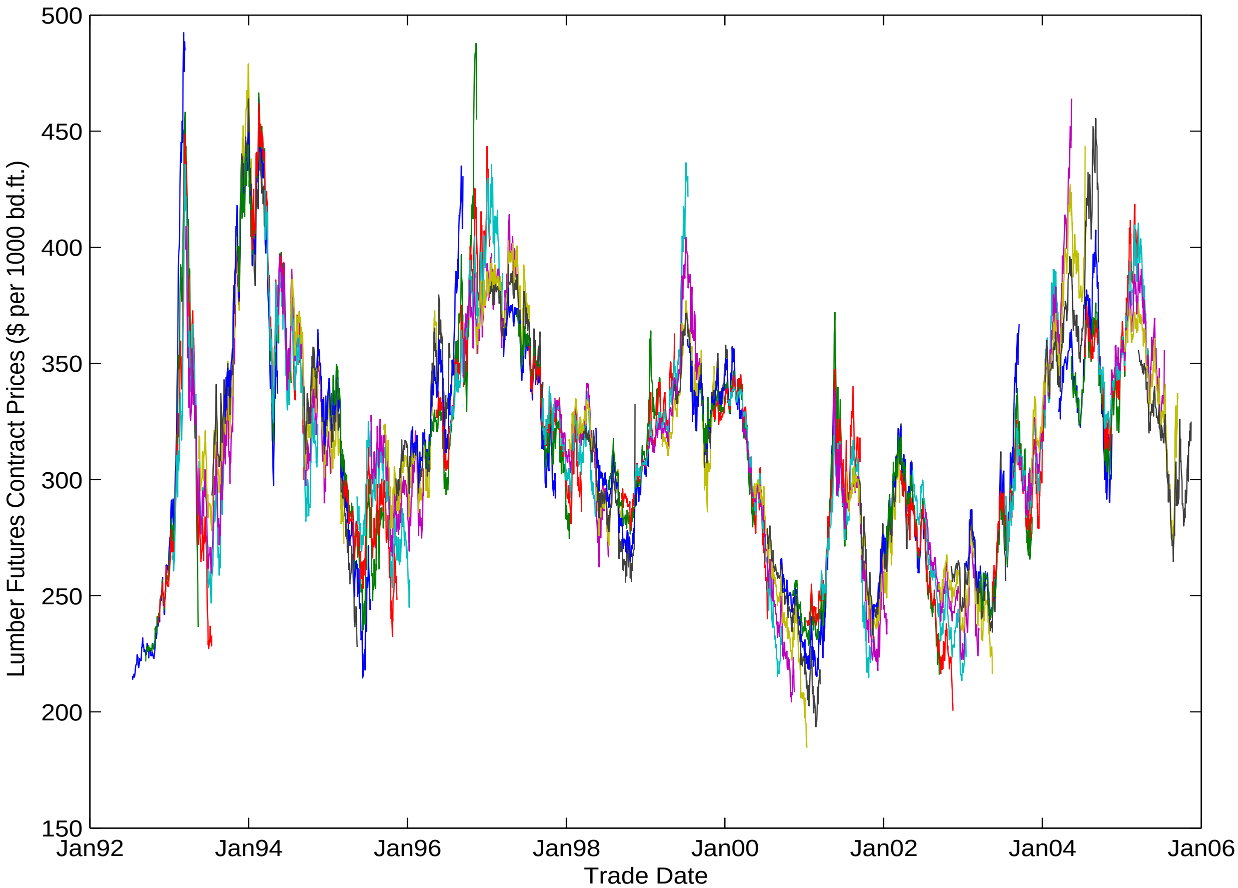


Figure 3: Lumber Futures Contract Prices ($ per 1,000 board feet)

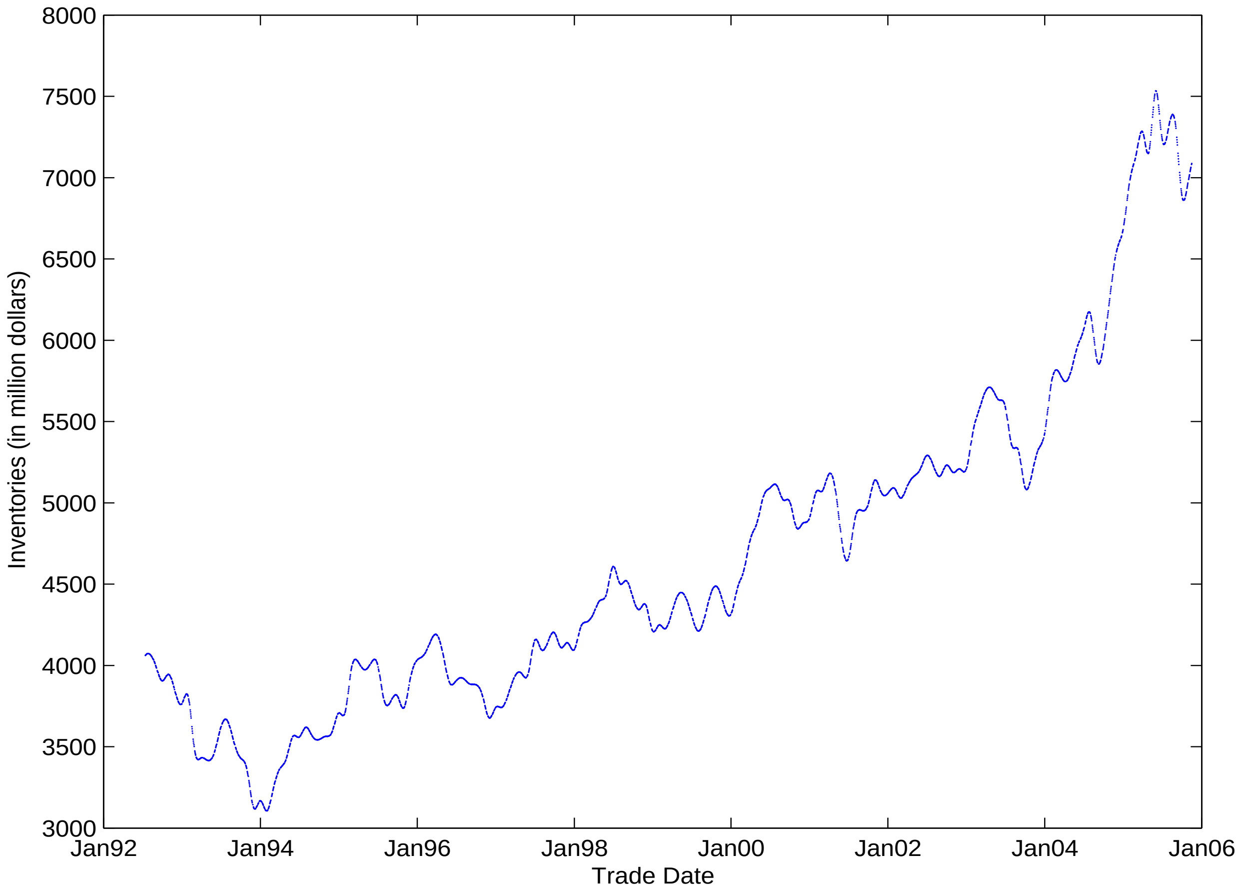
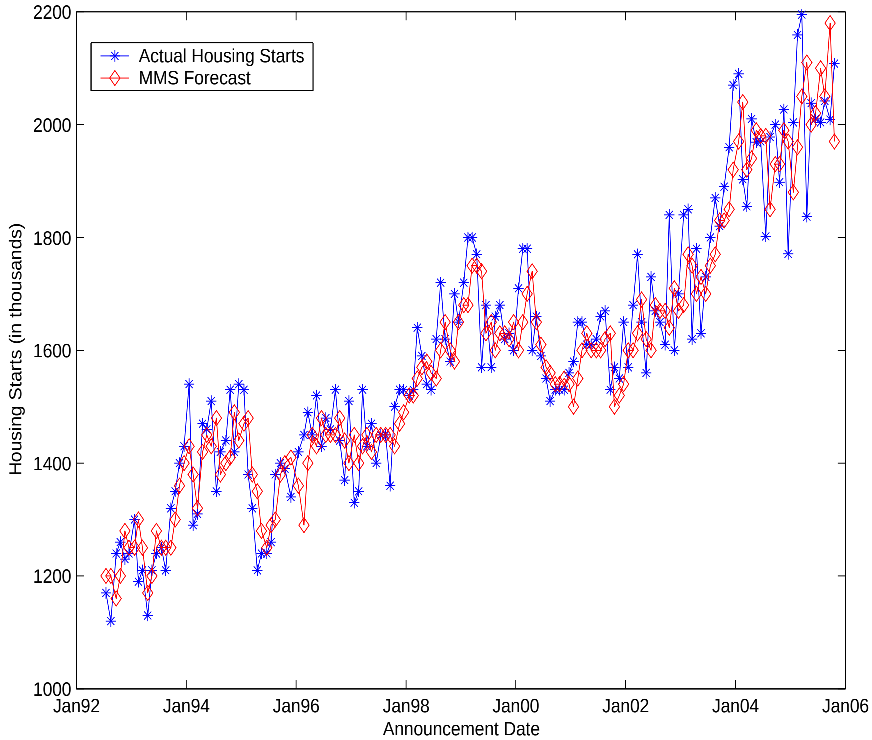
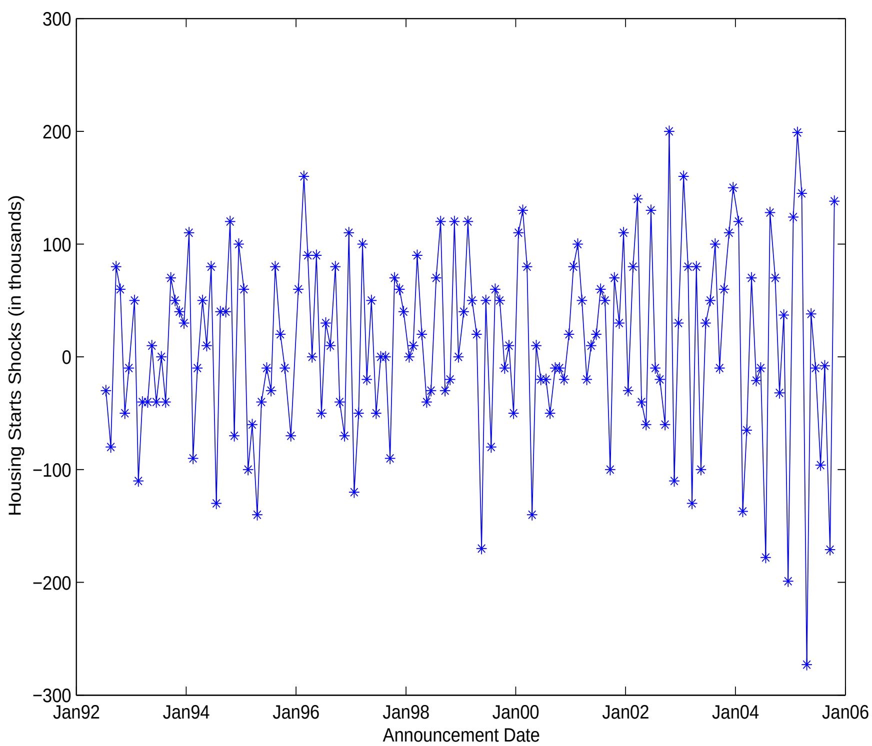


Figure 4: Lumber Inventories (millions of dollars)



(a) Housing Starts



(b) Housing Starts Shocks

Figure 5: Housing Starts Announcements and MMS Forecasts

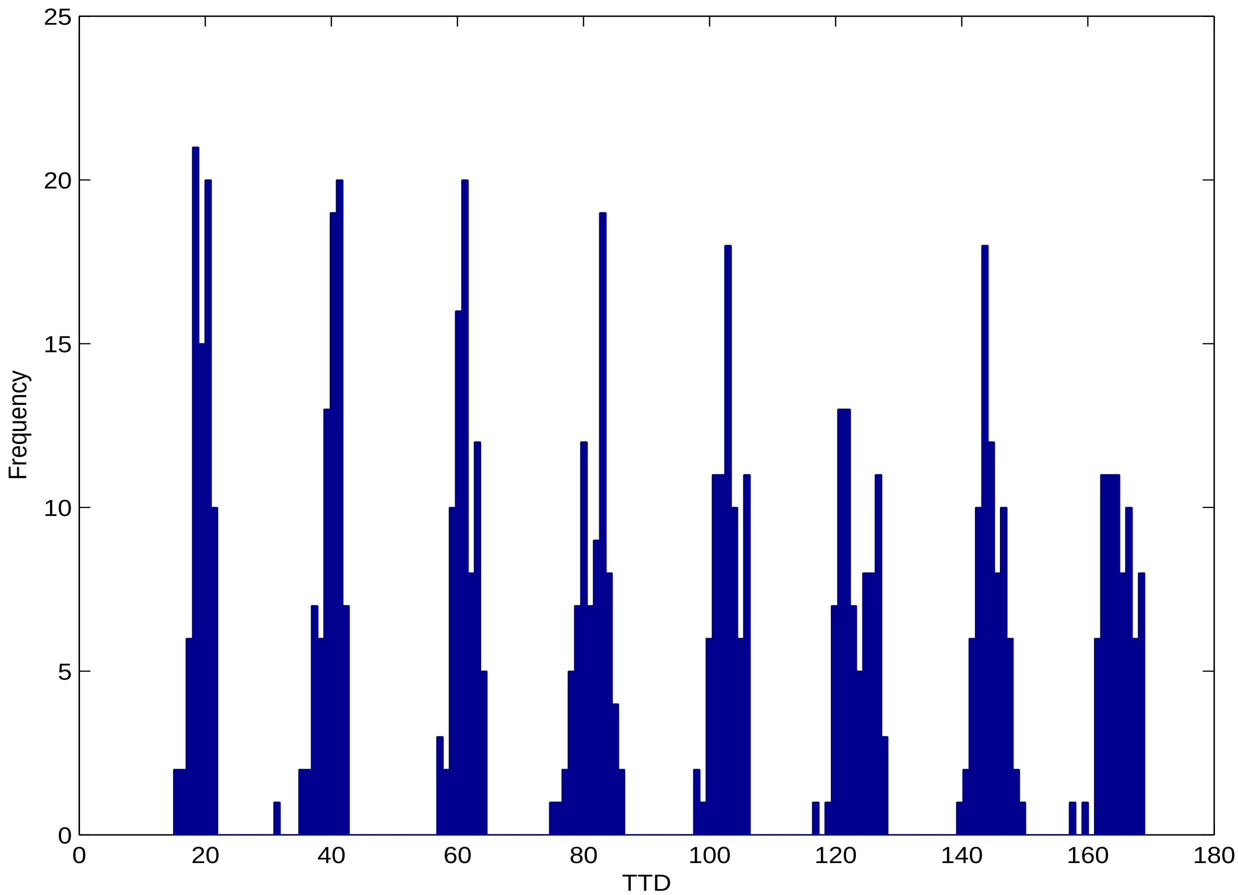


Figure 6: Time to Delivery on Announcement Days

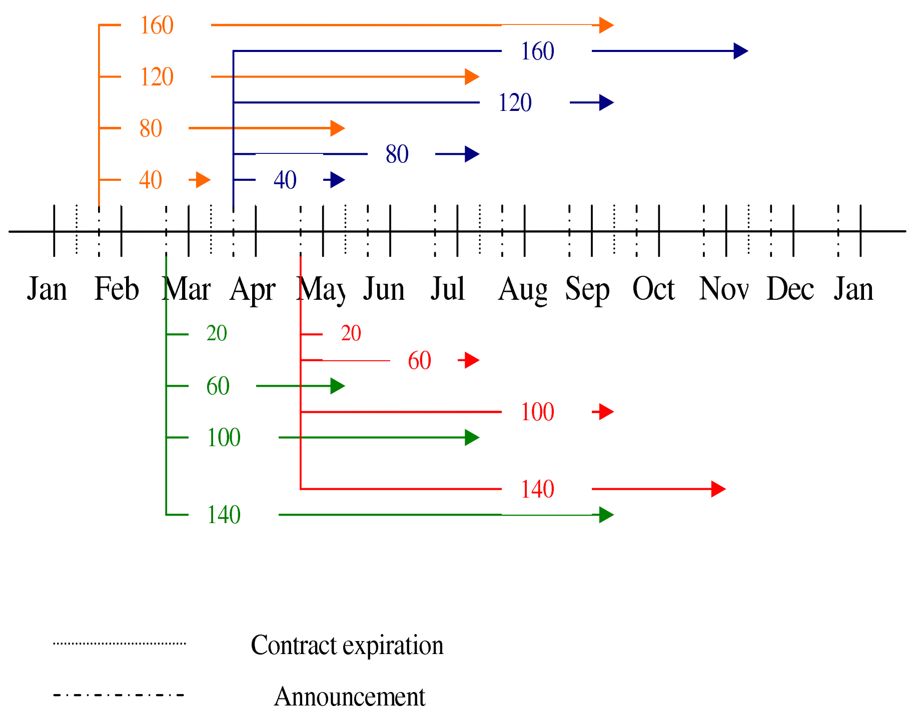


Figure 7: Contract Expiration and Time to Delivery on Announcement Days

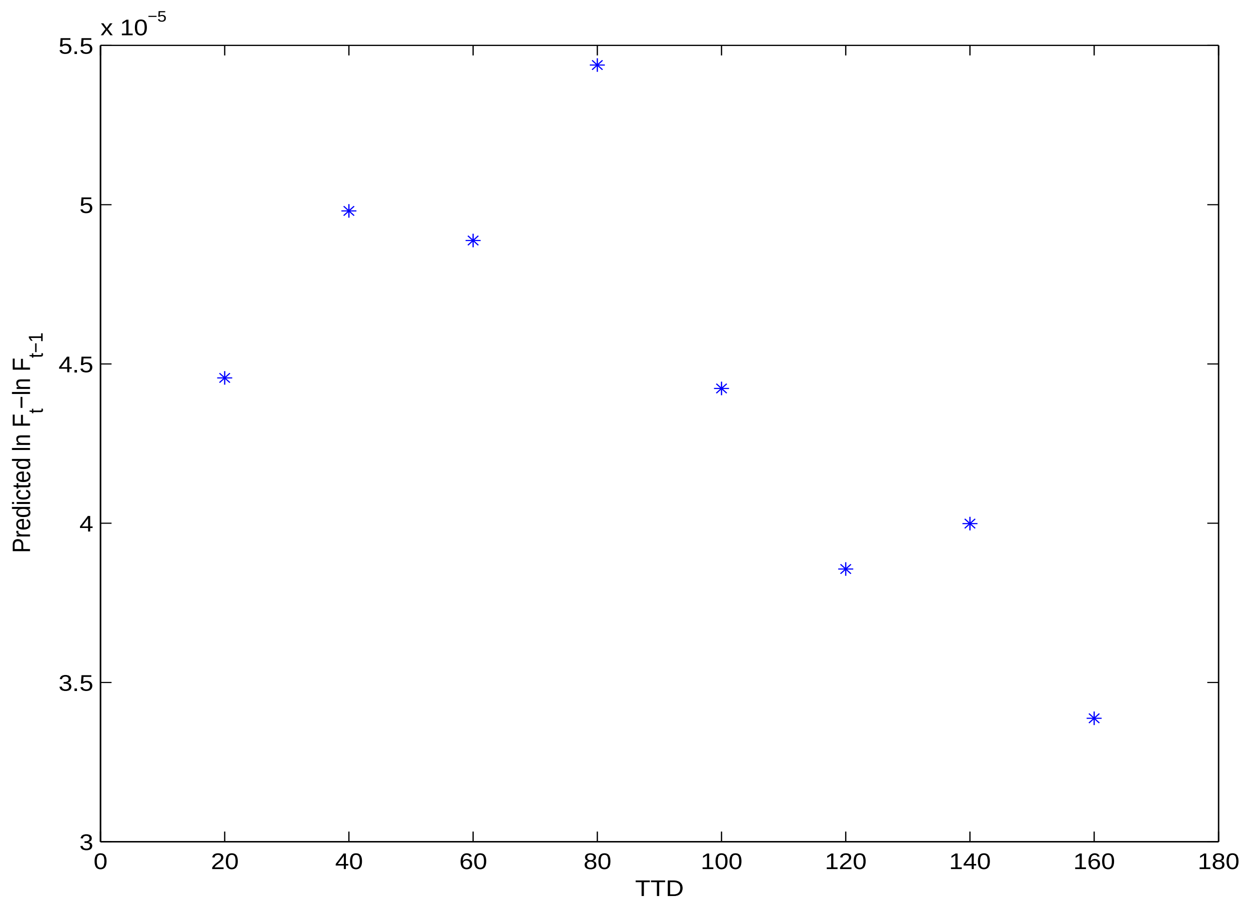


Figure 8: Predicted Daily Return with Eight TTD Categories

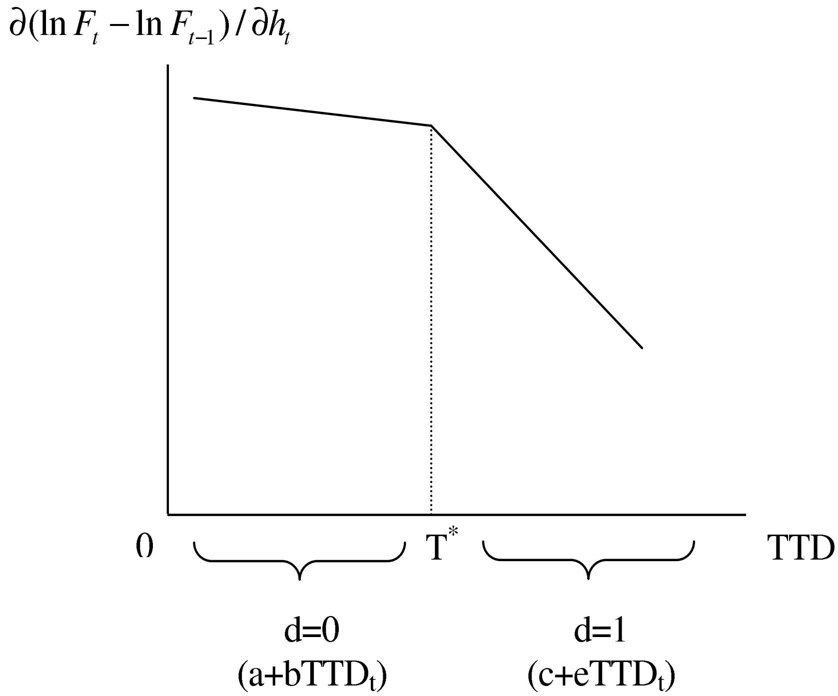


Figure 9: Piecewise Linear Function of Time to Delivery

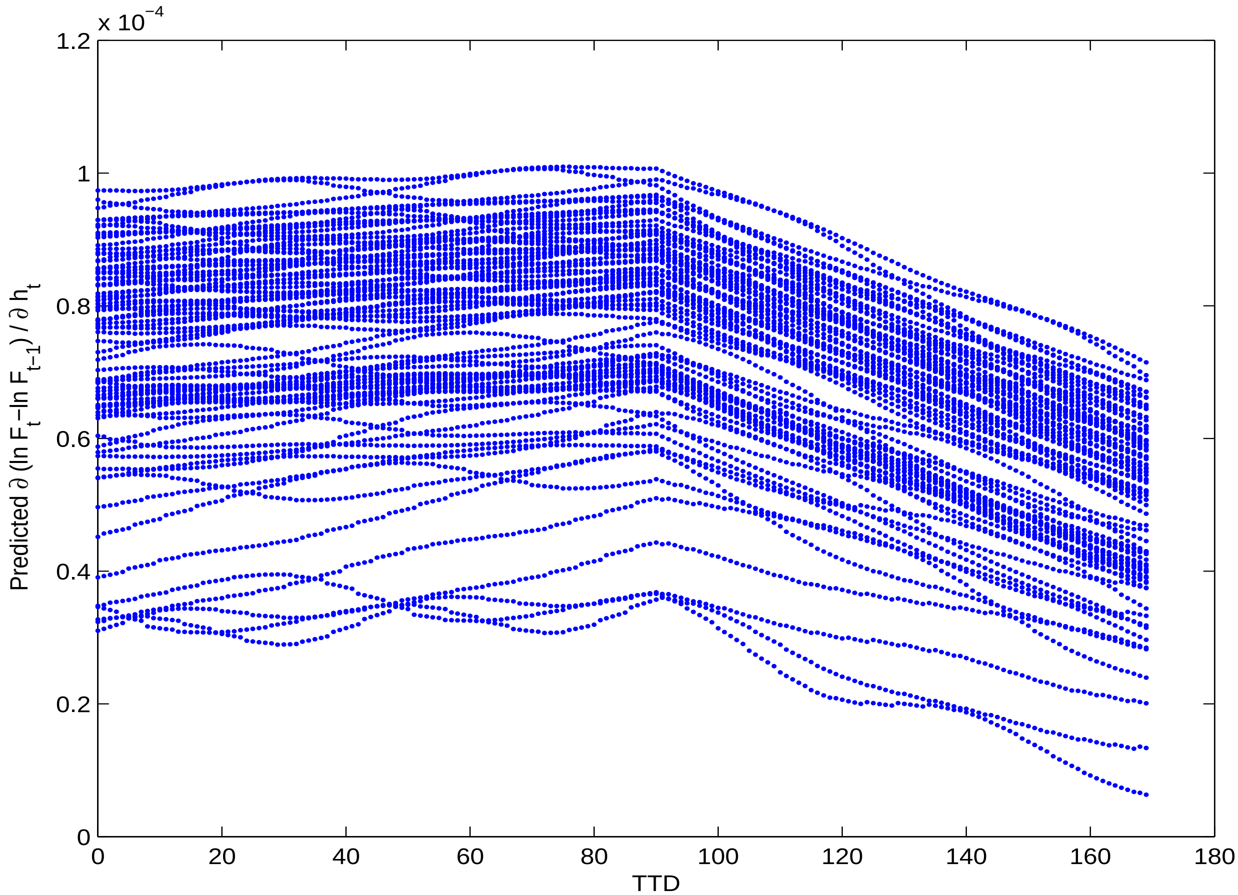


Figure 10: Predicted Marginal Effects of Housing Starts with Piecewise Linear TTD Function

Tables

Table 1: Summary Statistics

(a) Summary Statistics of Monthly Housing Starts Shocks

	$HS_m - MMS_m$	$HS_m - HS_{m-1}$	$HS_m - \widehat{HS}_m$
Mean	12.57	5.94	13.01
Median	10	10	19.35
Min	-273.00	-358.00	-264.85
Max	200	233	221.96
Std. Deviation	82.50	99.85	87.53
RMSE	83.20	99.71	88.22
N	159	158	159

(b) Summary Statistics of Daily Variables

N=13,090	$\ln F_t - \ln F_{t-1}$	$\lvert \ln F_t - \ln F_{t-1} \rvert$	Inventories	TTD	Housing Starts Shocks	|Housing Starts Shocks|
Mean	-0.00010534	0.01243	4,615.40	84.50	0.6226	3.0610
Median	0	0.00974	4,383.50	84.50	0	0
Min	-0.0786	0	3,105.40	0	-273	0
Max	0.1419	0.1419	7,533.70	169	200	273
Std. Deviation	0.0161	0.0102	954.15	49.08	17.8192	17.5654

Notes: Panel (a): HS_m is the released number of housing starts in thousands in month m, MMS_m is the median forecast of the Money Market Services (MMS) survey for housing starts that will be released in month m, HS_{m-1} is the previous month's released number for housing starts, $H\hat{S}_m$ is the forecast of housing starts constructed by an ARIMA(0,1,1) model $H\hat{S}_m = HS_{m-1} - 0.57177(HS_{m-1} - \widehat{HS}_{m-1})$. RMSE is the Root Mean Square Error measure of forecasting accuracy. Panel (b): Inventories are measured in millions of dollars. Inventory statistics are computed with all 13,090 observations. As a result, inventory levels repeat themselves when multiple contracts were traded on any day. The mean of unrepeated inventory series is 4,651.62 million dollars. Housing starts are measured in thousands. Shocks are computed as the difference between the released housing starts and the median response of the Money Market Services (MMS) survey. Housing starts shocks statistics are computed with zeros on nonannouncement days and with repeated shocks on announcement days if multiple contracts were traded.

Table 2: Market Efficiency Tests and Asymmetric Effects

(a) Market Efficiency

N=13,090	$\Delta y_t = \beta h_t + \gamma MMS_t + \delta h_{t-1} + \varepsilon_t$
β	5.09×10^{-5} $[1.26 \times 10^{-5}]$ (4.04)
γ	-1.31×10^{-6} $[6.55 \times 10^{-7}]$ (-2.00)
δ	-3.71×10^{-6} $[1.24 \times 10^{-5}]$ (-0.30)
$H_0 : \gamma = \delta = 0$	
F	2.0387
Pr>F	0.1302

(b) Unbiasedness

N=159	$HS_m = a + bMMS_m + e_m$
a	40.44 [46.29] (0.87)
b	0.98234 [0.02904] (33.83)
$H_0 : a = 0 \quad b = 1$	
F	2.0237
Pr>F	0.1356

(c) Asymmetric Effects

N=13,090	$\Delta y_t = \psi^P p_t * h_t + \psi^N n_t * h_t + \varepsilon_t$
ψ^P	3.97×10^{-5} $[1.64 \times 10^{-5}]$ (2.42)
ψ^N	5.72×10^{-5} $[1.92 \times 10^{-5}]$ (2.98)
$H_0 : \psi^P = \psi^N$	
F	0.4827
Pr>F	0.4872

Notes: In panel (a), the model is $\Delta y_t = \beta h_t + \gamma MMS_t + \delta h_{t-1} + \varepsilon_t$. Δy_t is the log price differences and computed as $\ln F_t - \ln F_{t-1}$, where $\ln F_t$ is the log price of futures contract on day t. h_t is the difference between released housing starts and the market expectation on day t and computed as $HS_t - MMS_t$, where HS_t is is the released number of housing starts in thousands on day t and MMS_t is the median forecast of the Money Market Services (MMS) Survey for housing starts on day t. In panel (b), the model is $HS_m = a + bMMS_m + e_m$, where HS_m is the released number of housing starts in thousands in month m and MMS_m is the median forecast of the MMS Survey for housing starts that will be released in month m. In panel (c), the model is $\Delta y_t = \psi^P p_t * h_t + \psi^N n_t * h_t + \varepsilon_t$. p_t is a dummy variable that equals one if h_t is positive on day t, zero otherwise. n_t is a dummy variable that equals one if h_t is negative on day t, zero otherwise. In panels (a) and (c), housing starts shock, h_t, is computed with zeros on nonannouncement days and with repeated shocks on announcement days if multiple contracts were traded. The parameter estimates in panels (a) and (c) are obtained using the Generalized Least Squares (GLS) method explained in the appendix. In panel (b), only data on announcement days are used and parameter estimates are obtained through Ordinary Least Squares (OLS) method. Standard errors and t-values of estimates are given in the brackets and parentheses, respectively.

Table 3: Housing Starts Effects on Expected Return

(a) Parameter Estimates

	Least Squares Estimates	Generalized Least Squares Estimates
B	-8.44×10^{-9}	-1.64×10^{-8}
	$[6.74 \times 10^{-9}]$	$[1.06 \times 10^{-8}]$
	(-1.25)	(-1.55)
b_1	1.07×10^{-4}	1.47×10^{-4}
	$[4.15 \times 10^{-5}]$	$[6.17 \times 10^{-5}]$
	(2.57)	(2.38)
b_2	1.18×10^{-4}	1.56×10^{-4}
	$[4.12 \times 10^{-5}]$	$[6.05 \times 10^{-5}]$
	(2.86)	(2.57)
b_3	1.08×10^{-4}	1.54×10^{-4}
	$[4.14 \times 10^{-5}]$	$[6.01 \times 10^{-5}]$
	(2.61)	(2.57)
b_4	1.32×10^{-4}	1.63×10^{-4}
	$[4.09 \times 10^{-5}]$	$[5.83 \times 10^{-5}]$
	(3.24)	(2.80)
b_5	1.01×10^{-4}	1.47×10^{-4}
	$[4.13 \times 10^{-5}]$	$[5.89 \times 10^{-5}]$
	(2.44)	(2.49)
b_6	9.88×10^{-5}	1.38×10^{-4}
	$[4.06 \times 10^{-5}]$	$[5.67 \times 10^{-5}]$
	(2.44)	(2.43)
b_7	9.35×10^{-5}	1.40×10^{-4}
	$[4.12 \times 10^{-5}]$	$[5.84 \times 10^{-5}]$
	(2.27)	(2.39)
b_8	9.15×10^{-5}	1.30×10^{-4}
	$[4.05 \times 10^{-5}]$	$[5.57 \times 10^{-5}]$
	(2.26)	(2.33)

(b) F-tests

	Least Squares Estimates				Generalized Least Squares Estimates			
H_0	F Value	D.f.(N)	D.f.(D)	Pr>F	F Value	D.f.(N)	D.f.(D)	Pr>F
$b_1 = b_2 = \cdots = b_8 = 0$	1.4567	8	13,081	0.1675	1.9118	8	13,081	0.0538
$b_1 = b_2 = \cdots = b_8$	0.3395	7	13,081	0.9361	1.4975	7	13,081	0.1629
$(b_1+b_2+b_3+b_4)/4 = (b_5+b_6+b_7+b_8)/4$	1.6041	1	13,081	0.2053	3.2878	1	13,081	0.0698
$b_1 = b_2 = b_3 = b_4$	0.2709	3	13,081	0.8464	0.2194	3	13,081	0.8830
$b_5 = b_6 = b_7 = b_8$	0.0379	3	13,081	0.9902	0.3327	3	13,081	0.8017
$b_1 = b_2 = b_3 = b_4$ and $b_5 = b_6 = b_7 = b_8$	0.1548	6	13,081	0.9882	0.6683	6	13,081	0.6753

(c) Marginal Effects

	$\frac{\partial y_t}{\partial h_t}\big\rvert_{T=20}$	$\frac{\partial y_t}{\partial h_t}\big\rvert_{T=40}$	$\frac{\partial y_t}{\partial h_t}\big\rvert_{T=60}$	$\frac{\partial y_t}{\partial h_t}\big\rvert_{T=80}$	$\frac{\partial y_t}{\partial h_t}\big\rvert_{T=100}$	$\frac{\partial y_t}{\partial h_t}\big\rvert_{T=120}$	$\frac{\partial y_t}{\partial h_t}\big\rvert_{T=140}$	$\frac{\partial y_t}{\partial h_t}\big\rvert_{T=160}$	$\frac{\partial^2 y_t}{\partial h_t \partial S_t}$
S_{min}	9.63×10^{-5} (2.77)	1.05×10^{-4} (2.99)	1.03×10^{-4} (3.24)	1.12×10^{-4} (3.56)	9.58×10^{-5} (3.23)	8.67×10^{-5} (3.02)	8.90×10^{-5} (3.09)	7.91×10^{-5} (2.88)	-1.64×10^{-8} (-1.55)
S_{mean}	7.16×10^{-5} (2.68)	8.00×10^{-5} (2.79)	7.85×10^{-5} (3.42)	8.73×10^{-5} (3.57)	7.10×10^{-5} (3.56)	6.19×10^{-5} (2.95)	6.42×10^{-5} (3.44)	5.44×10^{-5} (2.76)	-1.64×10^{-8} (-1.55)
S_{max}	2.38×10^{-5} (0.68)	3.22×10^{-5} (0.83)	3.07×10^{-5} (0.96)	3.96×10^{-5} (1.10)	2.33×10^{-5} (0.77)	1.41×10^{-5} (0.42)	1.64×10^{-5} (0.56)	6.62×10^{-5} (0.20)	-1.64×10^{-8} (-1.55)

Notes: In panel (a), the model is $\ln F_t - \ln F_{t-1} = BS_t * h_t + b_1 T_{1,t} * h_t + b_2 T_{2,t} * h_t + b_3 T_{3,t} * h_t + b_4 T_{4,t} * h_t + b_5 T_{5,t} * h_t + b_6 T_{6,t} * h_t + b_7 T_{7,t} * h_t + b_8 T_{8,t} * h_t + \varepsilon_t$. $\ln F_t$ is the log price of futures contract on day t, S_t is the lumber inventory level on day t, h_t is the difference between released housing starts and the market expectation on day t. $T_{1,t} = 1$ if $15 \leq TTD_t \leq 21$, 0 otherwise; $T_{2,t} = 1$ if $31 \leq TTD_t \leq 42$, 0 otherwise; $T_{3,t} = 1$ if $57 \leq TTD_t \leq 64$, 0 otherwise; $T_{4,t} = 1$ if $75 \leq TTD_t \leq 86$, 0 otherwise; $T_{5,t} = 1$ if $98 \leq TTD_t \leq 106$, 0 otherwise; $T_{6,t} = 1$ if $117 \leq TTD_t \leq 128$, 0 otherwise; $T_{7,t} = 1$ if $140 \leq TTD_t \leq 150$, 0 otherwise; $T_{8,t} = 1$ if $158 \leq TTD_t \leq 169$, 0 otherwise, where TTD_t is the number of days remaining to delivery on day t. Standard errors and t-values of estimates are given in the brackets and parentheses, respectively. In panel (b), F-tests for restrictions imposed on estimates are given. In panel (c), $y_t \equiv \ln F_t - \ln F_{t-1} = BS_t * h_t + b_1 T_{1,t} * h_t + b_2 T_{2,t} * h_t + b_3 T_{3,t} * h_t + b_4 T_{4,t} * h_t + b_5 T_{5,t} * h_t + b_6 T_{6,t} * h_t + b_7 T_{7,t} * h_t + b_8 T_{8,t} * h_t + \varepsilon_t$. Derivatives are computed by using the GLS estimates of parameters and evaluated at the minimum, mean, and maximum value of inventories. t-values of the derivative estimates are given in parentheses.

Table 4: Housing Starts Effects by Time to Delivery

(a) Parameter Estimates and F-tests

	Least Squares Estimates	Generalized Least Squares Estimates
B	-8.88×10^{-9}	-1.50×10^{-8}
	$[6.66 \times 10^{-9}]$	$[1.04 \times 10^{-8}]$
	(-1.33)	(-1.44)
$b_{1,4}$	1.17×10^{-4}	1.46×10^{-4}
	$[3.61 \times 10^{-5}]$	$[5.60 \times 10^{-5}]$
	(3.25)	(2.62)
$b_{5,8}$	9.85×10^{-5}	1.31×10^{-4}
	$[3.56 \times 10^{-5}]$	$[5.52 \times 10^{-5}]$
	(2.77)	(2.37)
$H_0 : b_{1,4} = b_{5,8} = 0$		
$F_{2,13087}$	5.3643	5.6432
$Pr > F$	0.0047	0.0035
$H_0 : b_{1,4} = b_{5,8}$		
$F_{1,13087}$	1.4483	6.4787
$Pr > F$	0.2288	0.0109

(b) Marginal Effects

	$\frac{\partial y_t}{\partial h_t}\Big\vert_{T_{1,4,t}=1}$	$\frac{\partial y_t}{\partial h_t}\Big\vert_{T_{5,8,t}=1}$	$\frac{\partial^2 y_t}{\partial h_t \partial S_t}$
S_{min}	9.99×10^{-5}	8.40×10^{-5}	-1.50×10^{-8}
	(3.80)	(3.39)	(-1.44)
S_{mean}	7.72×10^{-5}	6.14×10^{-5}	-1.50×10^{-8}
	(4.75)	(4.43)	(-1.44)
S_{max}	3.35×10^{-5}	1.76×10^{-5}	-1.50×10^{-8}
	(1.15)	(0.63)	(-1.44)

Notes: In panel (a), the model is $\ln F_t - \ln F_{t-1} = BS_t * h_t + b_{1,4}T_{1,4,t} * h_t + b_{5,8}T_{5,8,t} * h_t + \varepsilon_t$. $\ln F_t$ is the log price of futures contract on day t, S_t is the lumber inventory level on day t, h_t is the difference between released housing starts and the market expectation on day t. $T_{1,4,t} = 1$ if $15 \leq TTD_t \leq 21$, or if $31 \leq TTD_t \leq 42$, or if $57 \leq TTD_t \leq 64$, or if $75 \leq TTD_t \leq 86$; 0 otherwise. $T_{5,8,t} = 1$ if $98 \leq TTD_t \leq 106$, or if $117 \leq TTD_t \leq 128$, or if $140 \leq TTD_t \leq 150$, or if $158 \leq TTD_t \leq 169$; 0 otherwise. TTD_t is the number of days remaining to delivery on day t. Standard errors and t-values of estimates are given in the brackets and parentheses, respectively. In panel (b), $y_t \equiv \ln F_t - \ln F_{t-1} = BS_t * h_t + b_{1,4}T_{1,4,t} * h_t + b_{5,8}T_{5,8,t} * h_t + \varepsilon_t$. Derivatives are computed by using the GLS estimates of parameters and evaluated at the minimum, mean, and maximum value of inventories. t-values of the derivative estimates are given in parentheses.

Table 5: Housing Starts, Inventory, and Time-to-Delivery Effects on Expected Return

(a) Parameter Estimates and F-tests

	Least Squares Estimates	Generalized Least Squares Estimates
$A+c$	1.56×10^{-4} $[5.92 \times 10^{-5}]$ (2.64)	1.85×10^{-4} $[6.05 \times 10^{-5}]$ (3.05)
B	-8.98×10^{-9} $[6.67 \times 10^{-9}]$ (-1.35)	-1.59×10^{-8} $[1.05 \times 10^{-8}]$ (-1.52)
b	1.24×10^{-7} $[3.76 \times 10^{-7}]$ (0.33)	4.10×10^{-8} $[1.80 \times 10^{-7}]$ (0.23)
e	-4.07×10^{-7} $[3.75 \times 10^{-7}]$ (-1.08)	-3.74×10^{-7} $[1.47 \times 10^{-7}]$ (-2.54)
$H_0 : b = e = 0$		
$F_{2,13086}$	0.6806	3.2348
$Pr > F$	0.5063	0.0394
$H_0 : b = e$		
$F_{1,13086}$	0.6168	2.7532
$Pr > F$	0.4323	0.0971

(b) Marginal Effects

	$TTD_t = TTD_{max} = 169$			$TTD_t = TTD_{min} = 0$		
	$\partial y_t/\partial h_t$	$\partial^2 y_t/\partial h_t \partial TTD_t$	$\partial^2 y_t/\partial h_t \partial S_t$	$\partial y_t/\partial h_t$	$\partial^2 y_t/\partial h_t \partial TTD_t$	$\partial^2 y_t/\partial h_t \partial S_t$
S_{min}	7.21×10^{-5} (2.92)	-3.74×10^{-7} (-2.54)	-1.59×10^{-8} (-1.52)	9.79×10^{-5} (3.14)	4.10×10^{-8} (0.23)	-1.59×10^{-8} (-1.52)
S_{mean}	4.81×10^{-5} (3.46)	-3.74×10^{-7} (-2.54)	-1.59×10^{-8} (-1.52)	7.40×10^{-5} (3.20)	4.10×10^{-8} (0.23)	-1.59×10^{-8} (-1.52)
S_{max}	1.84×10^{-6} (0.06)	-3.74×10^{-7} (-2.54)	-1.59×10^{-8} (-1.52)	2.77×10^{-5} (0.84)	4.10×10^{-8} (0.23)	-1.59×10^{-8} (-1.52)

Notes: In panel (a), the model is $\ln F_t - \ln F_{t-1} = (A + BS_t) * h_t + [(1-d_t)(c + (e-b)T^* + bTTD_t) + d_t(c + eTTD_t)] * h_t + \varepsilon_t = (A+c)h_t + BS_t * h_t + b(1-d_t)(TTD_t - T^*) * h_t + e[(1-d_t)T^* + d_t TTD_t] * h_t + \varepsilon_t$. $\ln F_t$ is the log price of futures contract on day t, S_t is the lumber inventory level on day t, h_t is the difference between released housing starts and the market expectation on day t, TTD_t is the number of days remaining to delivery on day t, and d_t is a dummy variable, which takes the value of one when $TTD_t \geq T^*$, zero otherwise on day t. T^* is set equal to 90. Standard errors and t-values of estimates are given in the brackets and parentheses, respectively. In panel (b), $y_t \equiv \ln F_t - \ln F_{t-1} = (A + BS_t) * h_t + [(1-d_t)(c + (e-b)T^* + bTTD_t) + d_t(c + eTTD_t)] * h_t + \varepsilon_t = (A+c)h_t + BS_t * h_t + b(1-d_t)(TTD_t - T^*) * h_t + e[(1-d_t)T^* + d_t TTD_t] * h_t + \varepsilon_t$. Derivatives are computed by using the GLS estimates of parameters and evaluated at the minimum, mean, and maximum value of inventories. t-values of the derivative estimates are given in parentheses.