Title
Private Information and High-Frequency Stochastic Volatility

Authors
Kelly, David L.
Steigerwald, Douglas G

Publication Date
2003-08-01

Private Information and High-Frequency Stochastic Volatility[*]

David L. Kelly
Department of Economics
University of Miami
Box 248126
Coral Gables, FL 33134

Douglas G. Steigerwald
Department of Economics
University of California
Santa Barbara, CA 93106
FAX: 805.893.8830
telephone: 805.893.3151

August 1, 2003
JEL Classification Numbers: D82, G12, G14

Abstract

We study the effect of privately informed traders on measured high frequency price changes and trades in asset markets. We use a standard market microstructure framework where exogenous news is captured by signals that informed agents receive. We show that the entry and exit of informed traders following the arrival of news accounts for high-frequency serial correlation in squared price changes (stochastic volatility) and trades. Because the bid-ask spread of the market specialist tends to shrink as individuals trade and reveal their information, the model also accounts for the empirical observation that high-frequency serial correlation is more pronounced in trades than in squared price changes. A calibration test of the model shows that the features of the market microstructure, without serially correlated news, accounts qualitatively for the serial correlation in the data, but predicts less persistence than is present in the data.

[*]We thank Lawrence Harris, Bruce Lehman, Steve LeRoy, Charles Stuart, and seminar participants at University of Aarhus, University of Arizona, University of California at Santa Cruz, University of California at San Diego, Econometric Society Winter Meetings, Federal Reserve Bank of Kansas City, Federal Reserve Bank of San Fransisco, and University of Texas, for comments and suggestions. All errors are our own.

1 Introduction

The arrival of news is widely thought to have an important impact on asset prices. Despite such widespread belief, surprisingly little is known about the exact linkage between news and the intertemporal regularities that characterize many asset prices. Perhaps the most pronounced intertemporal regularity is positive serial correlation in squared price changes, detected via stochastic volatility (SV) and generalized auto-regressive conditional heteroskedasticity (GARCH) models, which has important implications for option pricing and conditional return forecasting. Empirical specification of SV and GARCH models vary widely in the literature, which suggests a need for theoretical guidance. We therefore derive the properties of transaction price changes from a standard microstructure model that incorporates the random arrival of news. In particular, the model replicates three features of the high frequency data: serial correlation in trades and squared price changes, and serial correlation in trades which is more persistent than serial correlation in squared price changes. We then test implications of the model. In particular, we derive that the market microstructure, without serial correlation in news, qualitatively accounts for the serial correlation in hourly squared IBM stock prices, albeit with less persistence. Our results therefore provide a theoretical explanation (and guidance) for much of the recent empirical results on the volatility of financial assets.[1]

We derive the properties of prices and trading behavior at the level of individual transactions from a repeated version of the asymmetric information model of Easley and O'Hara (1992). With some probability informed traders receive a private signal, or private news. Because uninformed (liquidity) traders are also in the market, private news is not immediately revealed by the trade decisions of the informed. The specialist, who clears trade, accounts for adverse selection when setting the bid and ask. As trade occurs, the specialist uses Bayes rule to update beliefs, and so the bid-ask spread declines as informed traders reveal their information through trade. We show that the bid-ask spread bounds the variance of transaction price changes. Because the bid-ask spread is dynamic in response to the specialist's learning, transaction price changes are neither independent nor identically distributed. In particular, transaction price changes have autocorrelated conditional heteroskedasticity (although not of GARCH form).

We assume that news arrivals are serially uncorrelated and so focus on the learning dynamics that result from information-based trade. Of course certain events may lead to serially correlated news; adding serial correlation into the exogenous news arrival process

[1] Bollerslev, Engle and Nelson (1993) provide a survey of GARCH models; Ghysels, Harvey and Renault (1996) provide a survey of SV models.

would augment the correlation that arises from the learning dynamics alone.[2] Perhaps surprisingly, we show that information-based trade alone (without serially correlated news) accounts for high frequency SV and two important related features of asset prices.[3]

The importance of private information as a determinant of asset price volatility is supported by French and Roll (1986), who conclude that revelation of private information (rather than public information or pricing errors) drives stock price changes. The entry and exit of informed traders after the arrival of private information is a key component of our explanation. First, the arrival of private news causes informed traders to enter the market, increasing the number of trades relative to calendar periods in which no private news exists. As the informed continue to trade until their information is fully revealed, informed traders enter and exit for stretches of calendar time. This behavior induces serial correlation in the number of calendar period trades (as well as trading volume), a feature documented by many authors (Harris, 1987; Andersen, 1996; Brock and LeBaron, 1996; Goodhart and O'Hara, 1997 page 96 provides a survey). Second, because the squared price change is determined by the number of trades in the calendar period and the variance of the price innovation for each trade, positive serial correlation in trades leads to SV. Because transaction prices have SV, the SV in calendar periods is not an artifact of discrete sampling. Third, because the bid-ask spread bounds the variance of trade-by-trade price innovations, the declining bid-ask spread reduces the serial correlation in squared price changes without affecting the serial correlation in trades. Thus serial correlation is more pronounced for trades than for squared price changes, also a well-known feature of the data (Harris, 1987; Andersen, 1996; Steigerwald 1997).[4] This third feature has proven to be a puzzle that is difficult to solve with traditional models that do not examine the properties of transaction price changes.[5]

We also derive other volatility related testable implications of the market microstructure model. In general, if the probability that the information advantage of informed traders is not eliminated between adjacent calendar periods increases, then informed traders are more likely to remain in the market in adjacent calendar periods. Thus the increase in trades and squared price changes resulting from the presence of informed traders is more likely to

[2] Engle et al. (1990) find some evidence of serial correlation in public news; although serial correlation in public news does not necessarily imply serial correlation in private news.

[3] Further, information-based trade can account for the positive contemporaneous relation between squared price changes and trading volume, which is the focus of the economic models of Epps (1975) and Tauchen and Pitts (1983).

[4] Similarly, Tauchen, Zhang, and Liu (1996) report that a price change has more persistent effects on volume than on squared price changes.

[5] In the analysis of Clark (1973), Gallant, Hsieh and Tauchen (1991), and Andersen (1996) the magnitude of stochastic volatility is determined by, and so is proportional to, the correlation of trades or trade volume.

remain in adjacent calendar periods, increases the magnitude and persistence of the serial correlation in trades and squared price changes. For example, we derive that the magnitude and persistence of the serial correlation in trades and squared price changes increases as the sampling frequency increases, because the information advantage of the informed is more likely to remain between adjacent hours than adjacent days. We also derive that the magnitude and persistence of the serial correlation in trades and squared price changes increases in markets where trade by the informed accounts for a relatively small proportion of the total trades.

We then test implications of the model using hourly IBM data (¯ltered of time of day and day of the week e®ects). The high frequency IBM data has all three empirical features of interest: serial correlation in trades and squared price changes, and the serial correlation in trades is more persistent. We calibrate the model to match certain moments of the IBM trade data. The ¯tted model has all three features of interest, although the persistence in trades is one day in the model rather three weeks in the data and the persistence in squared price changes is on the order of minutes in the model as opposed to one or two days in the data. An alternative calibration matches the persistence of the data, but then the magnitude is smaller than in the data.

As we focus on the properties of transaction price changes, we are implicitly modeling high-frequency calendar periods. Several researchers propose alternative explanations for stochastic volatility at lower frequencies. Timmerman (2001) shows rare structural breaks in the dividend process and incomplete learning generate ARCH and SV e®ects in an asset pricing model. Shorish and Spear (1996) show how moral hazard between the owner and manager of a ¯rm generates serial correlation in squared price changes in an asset pricing model. Den Haan and Spear (1998) show how agency costs and borrowing constraints give rise to wealth e®ects that yield serial correlation in squared interest rate changes. Serial correlation in such models does not arise from the trading process, since the \no trade" theorems hold.[6] While dividend-based models provide an important step by directly explaining stochastic volatility at low frequencies, these models cannot account for the stochastic volatility found in nearly all ¯nancial assets at high frequencies. In contrast we explain how news (say about the dividend process) generates high-frequency serial correlation through the trading process.

Section 2 presents an overview of the asymmetric information microstructure model. In Section 3 we derive basic properties of transaction price changes. Section 4 contains our

[6] Hu®man (1987) generates trade using an overlapping generations framework. However, Hu®man's model generates transitory negative serial correlation in both asset price and trading volume, which is inconsistent with the features described above.

results on the serial correlation of calendar period trades and price changes and Section 5 contains the empirical test of the model.

2 Model Overview

We work with the asymmetric information microstructure model of Easley and O'Hara (1992), which is derived in turn from Glosten and Milgrom (1985). In contrast to Easley and O'Hara, we assume that if news is not present, then the informed are inactive (Easley and O'Hara assume that the informed act as uninformed). The results are not qualitatively sensitive to the behavior of the informed when private news is absent. We also consider multiple information periods in which the news arrival process is independent and identically distributed. We use a standard model in order to show that standard models of the market microstructure can account for the persistence puzzle described in the introduction.

The information structure of the market is as follows. Informed traders learn the true share value with positive probability before trading starts, while the specialist and uninformed traders do not learn the true share value before trading starts. We de¯ne the interval of time over which asymmetric information is present to be an information period. At the beginning of each information period informed traders receive the signal S_m, where m indexes information periods. At the end of each information period the realization of the random dollar value per share, V_m, becomes public information and all traders agree upon the share value. We assume V_m takes one of two values $v_{L_m} < v_{H_m}$ with $P(V_m = v_{L_m}) = \pm$. We assume v_{L_m} and v_{H_m} are bounded from below by $v_{lb} > 0$ and above by $v_{ub} < 1$ for all m and are public information at the end of information period m ¡ 1:We also assume $0 < \pm < 1$ so that adverse selection is present in the market.

The signals received by informed traders at the start of an information period are independent across information periods and identically distributed. Therefore, serial correlation in trades and squared price changes generated by the model does not require serial correlation in the underlying news process. The signal S_m takes the value s_H if the informed receive the high signal and learn $V_m = v_{H_m}$, s_L if the informed receive the low signal and learn $V_m = v_{L_m}$, and s_0 if the informed receive the uninformative signal and hence, no private information. The probability that the informed learn the true value of the stock through the signal is µ, so the probability that S_m takes the value s_L is ±µ.

The signal completely determines the trading decisions of the informed. Conditional on receiving the uninformative signal, informed agents do not trade by assumption. If informed traders receive signal s_L, then informed traders always sell as long as the specialist is uncertain that the true value is v_{L_m}. If informed traders receive signal s_H, then informed traders always buy as long as the specialist is uncertain that the true value is v_{H_m}.

All traders and the market specialist, are risk neutral and rational. To induce uninformed rational traders to trade, some disparity of preferences or endowments across traders must exist. We let ω_i be the rate of time discount for the ith trader. As in Glosten and Milgrom each individual assigns random utility to shares of stock, s, and current consumption, c, as $\omega s V_m + c$.[7] We set $\omega = 1$ for the specialist and informed traders. Three types of uninformed traders exist, those with $\omega = 1$, who have identical preferences and do not trade, those with $\omega = 0$, who always sell the stock, and those with $\omega = 1$, who always buy the stock. Among the population of uninformed traders, the proportion with $\omega = 1$ is $1 - \epsilon$, the proportion with $\omega = 1$ is $(1 - \gamma)\epsilon$, and the proportion with $\omega = 0$ is $\gamma\epsilon$. The value of ω completely determines the trading decisions of the uninformed, which thus do not depend on the bid and ask.

Traders arrive randomly to the market one at a time, so we index traders by their order of arrival. The probability that an arriving trader is informed is $\alpha > 0$. A trader arrives, observes the bid and ask, and decides whether to buy, sell, or not trade. Let C_i be the random variable that corresponds to the trade decision of trader i. Then C_i takes one of three values: c_A if the ith trader buys one share at the ask, A_i; c_B if the ith trader sells one share at the bid, B_i; and c_N if the ith trader elects not to trade. The assumption that informed traders arrive randomly and trade at most one share is perhaps strong given the information advantage, but can be viewed as a simplification of a more complex model in which a pooling equilibrium exists where informed traders (or perhaps a single informed trader) mimic the both the timing of arrival and size of trades of the uninformed (see for example Laffont and Maskin, 1990 or Goodhart and O'Hara, 1997 page 94).

Because the specialist and the uninformed have the same information set, they have the same learning process. In what follows, we simply refer to the learning process for the specialist, noting that the same process applies to the uninformed. After the action of the trader, the specialist revises beliefs about the signal received by informed traders, and thence about the true value of a share. The sequence of trading decisions is public information. Let Z_i be the publicly available information set prior to the arrival of trader $i + 1$. After the ith trader has come to the market, the specialist's belief that informed traders received a high signal is $P(S_m = s_H | Z_i) = y_i$. Correspondingly, the specialist's belief that informed traders received a low signal is $P(S_m = s_L | Z_i) = x_i$. By construction, the specialist's belief that informed traders received an uninformative signal is $P(S_m = s_0 | Z_i) = 1 - x_i - y_i$. The action of each trader, even the decision not to trade, conveys information about the signal

[7] Because V_m is realized at the end of the information period, V_m is the random share value used to construct a trader's utility at the end of an information period.

received by informed traders.

The specialist sets a bid and ask, which are the prices at which he is willing to buy and sell, respectively, one share of stock. The bid and ask are determined so that the specialist earns zero expected profits from each trade. The zero expected profit condition is an equilibrium condition, which arises from the potential free entry of additional market specialists should the bid and ask lead to positive expected profits for the specialist. The quoted prices set the specialist's expected loss from trade with an informed trader equal to the specialist's expected gain from trade with an uninformed trader:

$$A_i = \frac{\alpha y_{i-1} v_{H_m} + (1-\alpha)\varepsilon(1-\gamma) E(V_m|Z_{i-1})}{\alpha y_{i-1} + (1-\alpha)\varepsilon(1-\gamma)};$$

where $E(V_m|Z_{i-1}) = x_{i-1} v_{L_m} + y_{i-1} v_{H_m} + (1 - x_{i-1} - y_{i-1}) EV_m$. In parallel fashion

$$B_i = \frac{\alpha x_{i-1} v_{L_m} + (1-\alpha)\varepsilon\gamma E(V_m|Z_{i-1})}{\alpha x_{i-1} + (1-\alpha)\varepsilon\gamma}.$$

It is straightforward to show that learning is consistent, that is, the bid and ask converge to the strong-form efficient value of a share, which reflects both public and private information. Hence the bid-ask spread, which reflects the specialist's uncertainty about private news, converges to zero as private information is revealed through trade. Because transaction prices are between the bid and ask, transaction prices also converge to the strong-form efficient value of a share.

3 Transaction Price Changes

To understand the behavior of transaction price changes implied by the model, we first present a simple expression for the price change associated with each possible trade decision. Following the decision of trader i, the price of the stock is its expected value conditional on public information. The resultant price change from the decision of trader i is

$$U_i = E(V_m|Z_i) - E(V_m|Z_{i-1}).$$

For example, if trader i elects to buy the stock at the ask, then the transaction price is $E(V_m|Z_i) = A_i$. (The equality between the conditional expected value of the stock and the ask is ensured by the equilibrium condition that governs quote setting, which implies

$A_i = E(V_m | Z_{i-1}; C_i = c_A))$. We refer to $\{U_i\}_{i \geq 1}$ as the sequence of transaction price changes, noting that a transaction occurs even if a trader elects not to trade.[8]

The information content of trade decisions, which depends on the history of trades and the parameter values, drives transaction price changes. To provide insight, we present simple expressions for each of the three possible values for U_i, one corresponding to each of the possible trade decisions. If $C_i = c_A$, then $E(V_m | Z_i) = A_i$, and

$$U_i = \frac{\alpha y_{i-1}}{P(C_i = c_A | Z_{i-1})} \left[v_{H_m} - E(V_m | Z_{i-1}) \right].$$

The price change that results from a trade at the ask is the price change that would result if the specialist knew the trader was informed $v_{H_m} - E(V_m | Z_{i-1})$, multiplied by the specialist's likelihood of such a trade with an informed trader $\frac{\alpha y_{i-1}}{P(C_i = c_A | Z_{i-1})}$. If $C_i = c_B$, then $E(V_m | Z_i) = B_i$ and

$$U_i = \frac{\alpha x_{i-1}}{P(C_i = c_B | Z_{i-1})} \left[v_{L_m} - E(V_m | Z_{i-1}) \right].$$

Finally, if $C_i = c_N$, then

$$E(V_m | Z_i) = \frac{\alpha (1 - x_{i-1} - y_{i-1}) EV_m + (1 - \alpha)(1 - \varepsilon) E(V_m | Z_{i-1})}{\alpha (1 - x_{i-1} - y_{i-1}) + (1 - \alpha)(1 - \varepsilon)}$$

and

$$U_i = \frac{\alpha (1 - x_{i-1} - y_{i-1})}{P(C_i = c_N | Z_{i-1})} \left[EV_m - E(V_m | Z_{i-1}) \right].$$

Even the decision not to trade conveys information and results in a transaction price change that is not zero.

In general, the expected value of the stock following a decision not to trade lies within the bid-ask spread. As a result, decisions to trade at the bid or the ask generally convey more information than do decisions not to trade. (For the first trade in an information period, trades at the bid or ask must convey more information, because $\delta y_0 = (1 - \delta) x_0$ which implies $B_1 < E(V_m | Z_0; C_1 = c_N) < A_1$). However, it is possible to have parameter

[8] In empirical work, U_i is not observed if either trader i or trader $i - 1$ elects not to trade. Econometricians therefore typically use the bid, ask, midpoint between the bid and ask, or last trade as a proxy for the unobserved transaction prices. Alternatively, estimates of the microstructure parameters could be used to construct a proxy. Our results on calendar period aggregates are virtually unchanged if a proxy replaces U_i on no trade decisions, because all measures respond to information in a similar fashion.

values and a trade history for which a decision not to trade conveys the most information. For example, if ε is nearly one and α is nearly zero, then no trade decisions are rare and are most often made by informed traders, which implies $E(V_m|Z_{i-1}; C_i = c_N) > A_i$ (if $E(V_m|Z_{i-1}) < EV_m$). For this reason we introduce the effective bid-ask spread

$$\tilde{A}_i - \tilde{B}_i = \max\{A_i; E[V_m|Z_{i-1}; C_i = c_N]\} - \min\{B_i; E[V_m|Z_{i-1}; C_i = c_N]\};$$

which is the difference between the maximum price change and the minimum price change.

We are now able to establish the statistical properties of transaction price changes $\{U_i\}_{i \geq 1}$.

Theorem 1: Transaction price changes satisfy:

1. $E(U_i|Z_{i-1}) = 0$ and $E(U_i|S_m \neq s_0) \neq 0$
2. $E(U_h U_i|Z_{i-1}) = 0$ for $h < i$
3. $c\left(\tilde{A}_i - \tilde{B}_i\right)^2 \leq E(U_i^2|Z_{i-1}) \leq \left(\tilde{A}_i - \tilde{B}_i\right)^2$ with $c \leq \frac{1}{4}$.
4. $\left(\tilde{A}_i - \tilde{B}_i\right) \stackrel{as}{\to} 0$ at an exponential rate.

Proof: See Appendix.

The first two parts of Theorem 1 deliver the traditional results that, with respect to public information, transaction price changes are mean zero and serially uncorrelated. Further, informed traders who are active anticipate transaction price changes that move in a systematic way in response to the flow of private information. Since transaction price changes have nonzero conditional variance, Parts 3 and 4 of Theorem 1 together imply that the effective bid-ask spread drives the variance in U_i and induces heteroskedasticity.[9] As the specialist becomes certain of the true value of the share, the bid and ask converge to the true value of the share and $E(U_i^2|Z_{i-1}) \to 0$ as $i \to \infty$.

The declining bid ask spread induces autocorrelated conditional heteroskedasticity and therefore serial correlation in squared transaction price changes. The difference in variance between information periods with and without news also induces serial correlation in squared transaction price changes, since transactions in which private information is present (and thus high variance) are most often followed by transactions in which private information is still present. In the next section, we derive the serial correlation properties of both transaction level and calendar period data.

[9]Hausman, Lo and MacKinlay (1992) find that the bid-ask spread is positively related to the variance of transaction price changes.

4 Serial Correlation Properties

To formally link the effects of individual trader decisions to the behavior of prices and trades measured at calendar period intervals, we first define how arrivals (economic time) are aggregated into calendar periods. Let each information period contain $k > 0$ calendar periods. For example, if an information period lasts one day, as in Easley, Kiefer and O'Hara (1993), then for data from the New York stock exchange (which is open for 6.5 hours) each information period contains thirteen 30 minute calendar periods. A calendar period, which is indexed by t, contains η trader arrivals, which as above are indexed by i. Transaction level properties are therefore the special case $\eta = 1$ and $k = \tau$. In general, we show serial correlation exists at both the transactions and calendar period data. A data sample, from which the serial correlation properties of calendar period quantities are estimated, consists of a large number of information periods. Because the information arrival process is independent over time, the k calendar measurements corresponding to one information period are independent of the k calendar measurements corresponding to any other information period. The sequence of calendar measurements is not itself generated by a stationary process.

Transaction Level and Calendar Period Trades

Let the number of trades in period t be I_t. Because η traders arrive each period, I_t takes integer values between 0 and η and so I_t is a binomial random variable. The parameters are η, the number of possible arrivals, and the probability of trade at each arrival. The probability of a trade depends on the signal, so:

$$I_t|(S_m \neq s_0) \sim B(\eta, \alpha + \varepsilon(1 - \alpha)),$$
$$I_t|(S_m = s_0) \sim B(\eta, \varepsilon(1 - \alpha)).$$

Unconditionally,

$$E[I_t] \equiv \mu = \theta\mu_1 + (1 - \theta)\mu_0 \quad (1)$$
$$Var[I_t] \equiv \sigma^2 = \theta\sigma_1^2 + (1 - \theta)\sigma_0^2 + \theta(1 - \theta)(\mu_1 - \mu_0)^2, \quad (2)$$

where the subscripts 0 and 1 indicate conditioning on $S_m = s_0$ and $S_m \neq s_0$, respectively.

Given this structure for the number of trades in a calendar period, we derive the serial correlation properties of $\{I_t\}_{t \geq 1}$.

Theorem 2: If $0 < r < k$ and $0 < \theta < 1$, then I_{t-r} and I_t are positively serially correlated. If $r \geq k$, then I_{t-r} and I_t are uncorrelated. Further for all $r > 0$, the correlation between $I_{t ¡ r}$ and I_t is given by:

$$\mathrm{Corr}(I_{t ¡ r}; I_t) = \frac{μ(1 ¡ μ)\,(®´)^2}{o2} \, " \frac{k ¡ \min(r;k)}{k} \# \qquad (3)$$

Proof: See Appendix.

Because of the nonstationary process generating trades, it may seem surprising that the correlation in I_t is not expressed as a function of time. To understand why, note that when connecting calendar period measurements to the data generating process, we do not know in which past calendar period the process began. Consider an information period that corresponds to one day for which news potentially arrives at the beginning of the day. As news could just as likely have arrived at any calendar period in the day, we do not want our calendar period implications to depend on an arbitrary assumption about news arrival. To avoid such dependence, we consider t to be randomly sampled, so that I_t is equally likely to correspond to any calendar period in the day. The serial correlation in I_t is then independent of time.

Many empirical studies focus on the correlation structure for one market at di®erent frequencies (e.g., comparing ¯ve minute intervals with hourly intervals). Because the data are gathered from the same market on the same asset, the number of trader arrivals in an information period, ¿ = k´, is constant even though both k and ´ depend on the sampling frequency. To understand the e®ect changing the frequency of observation has on the correlation, we substitute the formulas for the mean and variance of a binomial random variable and ´ = $\frac{¿}{k}$ into (3) to express the correlation (for r < k) as

$$\mathrm{Corr}(I_{t ¡ r}; I_t) = Ã \frac{k ¡ r}{k} ! \frac{(¿=k)μ(1 ¡ μ)®^2}{"(1 ¡ ®)[1 ¡ "(1 ¡ ®)] + μ®[(1 ¡ ®)(1 ¡ 2") + ®(¿=k)(1 ¡ μ)]};$$

and take the derivative with respect to k. As we decrease the frequency of observation we simultaneously decrease k and increase ´, yielding two countervailing e®ects on the correlation. The decrease in k reduces the serial correlation while the increase in ´ increases the serial correlation. Because the magnitude of the e®ect of a change in k on the correlation diminishes as r increases, it is for long lags that we would most likely see the serial correlation in trades decline as we move from ¯ve minute data to hourly data.

As the driving source of the serial correlation is the impact of trade by the informed, serial correlation persists only as long as the information advantage lasts. Liquidity parameters such as ´, the proportion of informed traders ®, and other parameters a®ect the magnitude of the serial correlation. To understand how the parameters individually in°uence the serial correlation in trades we calculate comparative static e®ects.

Corollary 3: For $0 < r < k$ and $0 < \theta < 1$ the correlation between I_{t-r} and I_t is decreasing in r, increasing in k, increasing in η and increasing in α.

Proof: The results follow from differentiation. ∎

The results in Corollary 3 imply certain patterns of serial correlation in trades across markets. Increasing the proportion of informed traders magnifies the impact of informed traders and increases the correlation.[10] In similar fashion, an asset for which the public realization of the news occurs rather slowly (a larger value of k), will be more impacted by the entry and exit of informed traders leading to more pronounced correlation. For a market with greater liquidity in the form of a larger value of η, the increased number of traders also magnifies the impact of informed traders and increases the correlation. An interesting implication is that serial correlation in trades exists in both liquid and illiquid markets, which provides a theoretical ground for the empirical serial correlation in illiquid markets found by Lange (1998).

Next, consider the relationship between ε (the fraction of uninformed who trade) and the serial correlation in trades. If ε is small (precisely, if $\varepsilon < \frac{1-2\alpha\theta}{2(1-\alpha)}$), then virtually all trades are by informed traders and increasing ε dilutes the informed traders and reduces the serial correlation in trades. If α is large (precisely if $\alpha\theta \geq \frac{1}{2}$), then increasing ε increases the variation in trades across information periods and increases the serial correlation in trades. In similar fashion, increasing the frequency of news θ increases the correlation if ε is large and θ is small (precisely $\varepsilon > \frac{1}{2}$ and $\theta < \frac{1}{2}$). Because good and bad news are symmetric with respect to the decision of whether or not to trade, the serial correlation is unaffected by changes to γ or δ.

Transaction Level and Calendar Period Squared Price Changes

Let P_t be the price at the end of period t. The period-t price change is

$$\Delta P_t = \sum_{i=(t-1)\eta+1}^{t\eta} U_i. \qquad (4)$$

Transaction price changes thus drive calendar price changes. Note that calendar price changes are equivalent to transaction price changes for $\eta = 1$. From Section 3, we know that the information content of a trade decision depends on the preceding sequence of trade

[10] Changes to each parameter affect both the covariance and the variance, so the relative effects determine the sign of each derivative. For example, as α increases there is a greater increase in trading in response to news, which increases the covariance. The variance may also increase, but because the informed trade identically at least one of the conditional variances $(\gamma_0; \gamma_1)$ that combine to form the variance must decrease, and the increase in the covariance dominates.

decisions. As a result, the conditional variance of each U_i depends on the path of trades and so analysis of the mixture process (4) does not yield straightforward analytical results for the correlation of $(\Delta P_t)^2$.

To make analysis of the correlation of $(\Delta P_t)^2$ tractable, we introduce an approximation to the mixture process. If period t is the first period following the arrival of news, then $E\left[(\Delta P_t)^2 | S_m \neq s_0\right] = \sigma_1$. Because trade decisions that occur shortly after the potential arrival of news contain more information than do later decisions, the expected squared price change for later periods declines, $\sigma_j > \sigma_{j+1}$ for $j = 1, \ldots, k-1$. If the informed are inactive, then the variance of calendar period price changes is driven by the random decisions of the uninformed and $E\left[(\Delta P_t)^2 | S_m = s_0\right] = \sigma_0$. Thus, we assume the information advantage of the informed persists until the information period ends, while for information periods without informed traders the uncertainty is quickly resolved, $\sigma_k > \sigma_0$. Because observation t is equally likely to correspond to any of the calendar periods in an information period, the unconditional expectation of calendar period squared price changes is

$$E\left[(\Delta P_t)^2\right] = \theta\bar{\sigma}_k + (1-\theta)\sigma_0;$$

where $\bar{\sigma}_k = \frac{1}{k}\sum_{j=1}^{k}\sigma_j$.

Given this structure for the squared price change in a calendar period, we derive the serial covariance of $\left\{(\Delta P_t)^2\right\}_{t \geq 1}$.

Theorem 4: For $0 < r < k$, the covariance between $(\Delta P_{t-r})^2$ and $(\Delta P_t)^2$ is

$$\frac{1}{k}\left\{\theta(1-\theta)\sum_{j=1}^{k-r}(\sigma_j - \sigma_0)(\sigma_{j+r} - \sigma_0) + \theta^2\sum_{j=1}^{k-r}(\bar{\sigma}_k - \sigma_j)(\bar{\sigma}_k - \sigma_{j+r}) + \theta^2\sum_{j=1}^{r}(\bar{\sigma}_k - \sigma_j)(\bar{\sigma}_k - \sigma_{k-r+j})\right\}.$$

For $r = 1, k-1$, if

$$\sigma_k > \theta\bar{\sigma}_k + (1-\theta)\sigma_0;$$

then

$$\mathrm{Cov}\left[(\Delta P_{t-r})^2, (\Delta P_t)^2\right] \geq 0.$$

Proof: See Appendix.

As in the covariance of calendar period trades, the covariance of calendar period squared price changes is zero if $r \geq k$. To determine the sign of the covariance at the longest lag, $r = k-1$, we compare the magnitude of the conditional covariances (which are positive and given by the first term in brackets) with the magnitude of the covariances of the conditional means (which are negative and given by the remaining two terms in brackets). The sufficient condition for positive serial covariance (and thus GARCH and SV) ensures that expected squared price changes are above their unconditional mean if the informed are active and below their unconditional mean if the informed are inactive. As a result, if $(\Delta P_t)^2$ is above the unconditional mean, then $(\Delta P_{t-r})^2$ also tends to be above the unconditional mean for $r = k - 1$, and so prices have stochastic volatility.

Persistence Puzzle

If U_i^2 is assumed to be homoskedastic, then the covariance of calendar period squared price changes is driven exclusively by the covariance in calendar period trades, and the persistence in the covariance in trades should be matched by the persistence in the covariance in squared price changes. The heteroskedasticity in U_i^2 that arises from the movements in the expected bid-ask spread breaks this persistence link. The variance of U_i declines in response to the information revealed through trade, causing the serial covariance in squared price changes to decline, but not affecting the serial covariance in the number of trades. Hence stochastic volatility is less persistent than is the serial correlation in trades.

We first obtain an analytic result for the simplified structure of Theorem 4. Because the persistence of both the stochastic volatility and the serial correlation increases with k, the relative persistence depends on k.

Proposition 5: Let $\sigma_k > \theta\bar{\sigma}_k + (1 - \theta)\sigma_0$ and $\sigma_j - \sigma_{j+1} = \phi$ for all $j = 1 \ldots k - 1$. Then for $0 < r < k$,

$$2(k - 2)(k + 3) > 3\theta(20 - 11k + k^2) \tag{5}$$

implies the covariance, and hence the correlation, of calendar period squared price changes decays more rapidly than the covariance of calendar period trades.

Proof: See Appendix.

If information persistence is moderate (precisely if $k \le 32$), then (5) is satisfied for all θ. Alternatively, if the news is not too frequent (precisely $\theta \le \frac{2}{3}$), then (5) is satisfied for all k. Part 4 of Theorem 1 implies the decline in the variance of U_i is exponential, hence Proposition 5, which assumes linear decline in calendar period squared price changes, likely understates the difference in persistence between trades and squared price changes.

5 Empirical Results

To see how well the predictions of the microstructure model accord with the data, we turn to analyses of transaction data for IBM from the New York Stock Exchange (NYSE). From the NYSE Trades and Quotes (TAQ) database, we study IBM transactions on the NYSE for the year 2000. We filter the trade data to remove trades that were recorded out of sequence, canceled, executed with special conditions, or recorded with some other anomaly. Because of certain institutional details, occasionally large trades are broken up into a sequence of smaller trades, all at the same price (see Hasbrouck (1988)). In order to avoid misidentifying these sequences of same sided trades as bursts of informed trades, we aggregate all trades recorded within five seconds of each other without an intervening price change or quote revision.[11]

The data are further filtered to remove time stamps outside of the official trading hours of the NYSE (9:30 AM to 4:00 PM). Finally, the first half-hour of each trading day is removed in order to avoid modeling the market opening of the NYSE, which is characterized by heavy activity following the morning call auction. As Harris (1987), Engle and Russell (1998), and many other authors have noted, the first half-hour of trade exhibits substantially different properties than the rest of the day.

We analyze hourly totals for each of the 252 trading days in the year. The (hourly) average number of trades is 331 with an average squared price change of 0.91. As noted by previous authors (e.g. Harris (1987)) exchange data exhibits periodic features, which we remove as these features likely arise from sources of trade not captured by the model. In addition to day-of-the-week effects, we must remove any diurnal pattern. The hourly data exhibit a U-shaped pattern, with higher transaction activity and volatility at the start and end of the day. In addition, the number of trades on the NYSE exhibits a significant decline during the lunch period. We capture the U-shaped diurnal pattern for squared price changes with a quadratic function in hours. To capture the lunch effect in the number of trades, we replace the quadratic function for hours with a linear spline, the middle part of which captures the slow period of trade around the lunch hour. The periodic features are estimated to be (parentheses enclose the t-statistics)

$$\begin{aligned} T_t^P = \; & \underset{(63.4)}{429.2} + \underset{(3.6)}{19.0}\, Mo_t + \underset{(2.2)}{11.3}\, Tu_t + \underset{(4.1)}{20.6}\, We_t + \underset{(2.4)}{12.3}\, Th_t + \underset{(2.0)}{19.3}\, Ho_t \\ & - \underset{(20.7)}{56.2}\, H_t + \underset{(12.4)}{57.1}\,(H_t - 3)\cdot L_t + \underset{(8.8)}{60.0}\,(H_t - 4)\cdot AL_t; \end{aligned}$$

[11] We use quotes only from the NYSE (Blume and Goldstein (1997) find that the NYSE quote determines or matches the national best quote about 95 percent of the time). We also filter the quote data to remove recording anomalies.

$$
\begin{aligned}
\left(\Delta P_t^2\right)^P &= \underset{(7.4)}{2.0} + \underset{(0.9)}{0.2}\, Mo_t + \underset{(0.5)}{0.1}\, Tu_t + \underset{(0.7)}{0.1}\, We_t - \underset{(0.3)}{0.1}\, Th_t + \underset{(0.1)}{0.1}\, Ho_t \\
&\quad - \underset{(5.4)}{0.9}\, H_t + \underset{(5.5)}{0.1}\, H_t^2,
\end{aligned}
$$

where superscript P indicates predicted value, Mo_t, Tu_t, We_t, and Th_t, are day-of-the-week indicator variables, and Ho_t is an indicator variable that takes the value 1 if the succeeding trading day is a holiday or if the market closes early (the days prior to July 4 and after Thanksgiving end at noon). Next, H_t takes the integer value corresponding to the hour of the day (1 for the first hour, 6 for the last hour) and L_t and AL_t are indicator variables equal to 1 for all hours after 12 p.m. and 1 p.m., respectively. To see how the lunch effect is captured, hourly trades decline by 56 each hour until 1 p.m., hourly trades from 1 to 3 p.m. are roughly unchanged from the noon hour, and hourly trades rise by about 120 from the previous hour during the last hour of trading. As is immediately apparent, the diurnal effects are more substantial for this data than the daily effects. In what follows we work with the adjusted series $\left[T_t - T_t^P + 429.2,\ (\Delta P_t^2) - (\Delta P_t^2)^P\right]$.

Figure 1 contains the autocorrelation functions for adjusted hourly trades and squared price changes.[12] The trade correlation remains significant for more than three weeks (ninety trading hours). The squared price change correlation, which appears to die away within one or two days (although there are several significant correlations at longer lags), does not appear to be proportional to the trade correlation. These results are certainly consistent with the literature and the implications of the microstructure model.

Another common way to capture the correlation in squared price changes is to model the volatility with a GARCH model. For hourly squared price changes, the estimates of the GARCH model are (standard errors in parentheses)

$$
\begin{aligned}
\Delta P_t &= \underset{(0.02)}{0.02} + \left(H_t^P\right)^{1/2} U_t^P, \\
H_t^P &= \underset{(0.01)}{0.05} + \underset{(0.01)}{0.06}\left[\left(H_{t-1}^P\right)^{1/2} U_{t-1}^P\right]^2 + \underset{(0.02)}{0.88}\, H_{t-1}^P.
\end{aligned}
$$

The significant coefficient on the ARCH term (the estimated coefficient of 0.88 in the equation for the scale) indicates that stochastic volatility is a statistically significant feature of the hourly data.

To obtain predicted serial correlation properties of the model, we must assign parameter values to the model. The standard method in the literature is to use maximum likelihood estimation (ML) to obtain model parameters from the probabilities of trade (for example

[12] Results are unchanged if squared returns are used in place of squared price changes.

Easley, Kiefer, O'Hara and Paperman (1996)). However, the ML estimator is constructed with an assumed information period length, a critical parameter for the persistence of the serial correlation. Ideally, a method of moments estimator could be constructed to estimate the length of an information period using the trade auto-covariance moments. However, construction of such an estimator involves a number of difficulties and is thus beyond the scope of this paper. Therefore, we instead calibrate the model to match certain moments in the trade data. In particular, we set θ so that the mean number of trades in the model match the mean number of trades in the data and ϵ to maximize the first order serial correlation in trades (which is approximately equal to the first order serial correlation in the data). The number of trader arrivals η is then set so that the variance of trades in the model matches the variance of trades in the data (the ML estimator also significantly underestimates the variance of trades), given $\alpha = 0.05$. In general, the model is consistent either with a large number of trader arrivals and low probability of trade (α and ϵ) or the reverse. As is commonly done with ML, we set the probability that an uninformed trader trades at the ask and the probability of good news equal to 0.5. We examine both $k = 6$, in which an information period is one day, as is commonly assumed in ML, and $k = 90$, so that the persistence in serial correlation in trades matches the data (note that increasing k does not change the time required for private information to be incorporated into the share price through informed trade, but instead increases the time between the arrival of private news and the public announcement of private news).

k	ϵ	θ	α	η	δ	γ
6	0.17	0.48	0.05	2304	0.5	0.5
90	0.17	0.48	0.01	2304	0.5	0.5

Table 1: Calibrated Parameters and alternative specification.

Given the estimated parameter values, the predicted trade moments may be computed according to Theorem 2 (see the proof of Theorem 2). Table 2 compares a variety of trade moments predicted by the model to the data.

The model, given an information period of $k = 6$ hours, does a reasonable job matching the magnitude of serial correlation in trades. However, the model predicts the correlation in trades lasts for only 6 hours, which is inconsistent with persistence of 90 hours observed in the data. Hence it may be the case that the time between the arrival of private information and the public announcement of the private information, the length of an information period, is

moment	$E\,I_t$	$\mathrm{Var}\,I_t$	$\mathrm{Corr}\,(I_{t-1}, I_t)$	$\mathrm{Corr}\,(I_{t-2}, I_t)$	$\mathrm{Corr}\,(I_{t-50}, I_t)$
Sample Moment	429.00	3658.03	0.76	0.51	0.20
Model, k = 6	429.24	3661.98	0.75	0.60	0
Model, k = 90	400.42	463.34	0.283	0.280	0.127

Table 2: Moments of the Market Microstructure Model and IBM Data.

longer than commonly assumed. While misspecifying the length of an information period has no bias on estimates of uninformed behavior, it does bias estimates of informed behavior. In particular, if information periods are assumed to be one day, when in practice information periods last longer than one day, then ML overestimates the impact of informed trade relative to uninformed trade. As depicted in Table 2, increasing k to 90 and decreasing α to 0.01 matches the persistence in trades but underestimates the variance in trades and overestimates the magnitude of the serial correlation at lower lags, since the model predicts a linear decline when in fact the decline in the IBM data appears more geometric.

To approach the persistence puzzle for the general mixture model, we simulate the model using the parameters from the calibration. Figure 2 depicts the simulated model (with k = 6) serial correlation in hourly trades and squared price changes. The calendar period price change is calculated with the last price associated with a trade in the calendar period. Because η is large, all information is resolved in one calendar period with probability very close to one. Hence the expected first squared price change (σ_1) is positive while the next five squared calendar price changes are zero. Clearly then the model must predict negative serial correlation for lags 1-5 and the positive serial correlation for lag 6 (it is straightforward to calculate these moments analytically). Although the model does not match the hourly squared price data very well, the model does predict quite a bit of positive serial correlation at 5 minute and transaction level data. Thus the model predicts positive serial correlation in squared price changes on the order of minutes and positive serial correlation in trades for a few hours. Figure 3 shows the autocorrelation function for the simulated model with k = 90. As noted above, the model captures the persistence in trades, but predicts a linear decline. Squared price change correlations remain positive for about three or four hours, which is close to the persistence of the data, although again the magnitude of first order serial correlation is smaller.

Although the calibrated model qualitatively matches the three key features of the data either the persistence or the magnitude is quantitatively less in the model than in the data. The persistence in the data for trades is a few weeks and one or two days for squared price changes, whereas in the calibrated model the persistence is one day and a few minutes respectively. Conversely, if the model is calibrated to match the persistence, the serial correlation in trades and squared price changes have about half the magnitude in the model as in the data.

6 Conclusions

The possible presence of private information in an asset market leads to transaction price changes that, while uncorrelated, are dependent and heterogeneous. The heterogeneity is present in the conditional variance, which moves in accord with the bid-ask spread. As trading reveals private information, the conditional variance of transaction prices declines with the spread. As a result, transaction price changes have stochastic volatility.

Serial correlation in calendar period quantities, for trades and squared price changes, as well as the persistence puzzle can also be explained by the arrival of private information. Given that informed traders are trading in the current period, informed traders will most likely trade in the following period, which generates serial correlation in trades. The serial correlation in trades is positive and persistent. Serial correlation in trades generates serial correlation in squared price changes. Given that the informed traders are trading, more variance exists in squared price changes simply because more trades occur in a calendar period. More trades implies that the price change is the sum of more random transaction price changes, which in turn implies that price changes have greater variance. Because serial correlation exists in trades, serial correlation exists in squared price changes. However, there is an additional e®ect on the serial correlation in squared price changes, the decline in the bid-ask spread. All trades are at the bid or ask, hence expected price changes are bounded by the bid-ask spread. The bid-ask spread declines as learning proceeds, which reduces the variance and the persistence of the serial correlation in squared price changes. Given more trades occur in a calendar period, most likely more trades occur in the next calendar period, which implies higher variance in both periods. However, the trades in the second calendar period are from a random variable with a smaller variance, due to the smaller bid-ask spread. Hence the serial correlation is smaller and less persistent. We thus replicate the observed empirical features of the data and explain the serial correlation through the entry and exit of informed traders and the associated revelation of information in prices.

The correlation in calendar period quantities is not an artifact of aggregation; as transaction price changes themselves have stochastic volatility. Further, the stochastic volatility in calendar period data arises without correlated news; the news arrival process we consider is independent over time. Instead, the endogenous news revelation process over the information period generates a persistent information advantage for the informed, leading to di®erences in the number of trades on news versus no news periods. When information periods are aggregated together, serial correlation results. Because we presume no serial correlation in the news arrival process, obtaining serial correlation at lower frequencies requires a long information period. As a long information period may not be plausible for all news arrivals, our results provide an explanation for high-frequency serial correlation and indicate that other factors must play a role in low-frequency serial correlation.

We calibrated the model to obtain parameters for the model and compare the serial correlation in trades and squared price changes in the fitted model with that of high frequency IBM data. We find that the fitted model qualitatively predicts all three features of interest, although either the persistence or the magnitude is less than in the data. The assumed length of an information period plays an important role in the results, however. Therefore, a fruitful direction of future research might be to estimate the length of an information period by exploiting the autocovariance moments in trades perhaps with a method of moments estimator.

What information set should be used to form conditional expectations of $(\Delta P_t)^2$? The above results indicate that prediction of the variance of price changes depends on prediction of the entry and exit of informed traders. Specifically, the conditional variance of stock prices depends on the previous number of trades, but does so in a nonlinear way. This finding underpins recent models of stochastic volatility that are based on jump-diffusion processes. Many of these models have a jump arrival rate that is constant through time. Our work suggests that future models of stochastic volatility include a jump arrival rate that varies through time, in response to innovations in the number of trades.

References

[1] Andersen, T., 1996, \Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility" Journal of Finance 51, 169-204.

[2] Bollerslev, T, R. Engle and D. Nelson, 1993, \ARCH Models" in Handbook of Econometrics, Volume 4, Amseterdam: North-Holland.

[3] Blume, M. and M. Goldstein, 1997, \Quotes, Order Flow, and Price Discovery" Journal of Finance 52, 221-44.

[4] Brock, W. and B. LeBaron, 1996, \A Dynamic Structural Model for Stock Return Volatility and Trading Volume" Review of Economics and Statistics 78, 94-110.

[5] Clark, P., 1973, \A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices" Econometrica 41, 135-159.

[6] Den Haan, W. and Spear, S., 1998, \Volatility Clustering in Real Interest Rates: Theory and Evidence" Journal of Monetary Economics 41, 431-53.

[7] Easley, D., and M. O'Hara, 1992, \Time and the Process of Security Price Adjustment" Journal of Finance 47, 577-605.

[8] Easley, D., N. Kiefer, and M. O'Hara, 1993, \One Day in the Life of a Very Common Stock" Review of Financial Studies 10, 805-35.

[9] Easley, D., N. Kiefer, M. O'Hara, and J. Paperman, 1996, \Liquidity, Information, and Infrequently Traded Stocks" Journal of Finance 51, 1405-1436.

[10] Engle, R., and Ng, V., and Rothschild, M., 1990, \Asset Pricing With a Factor-ARCH Covariance Structure: Empirical Estimates For Treasury Bills" Journal of Econometrics 45, 213-37.

[11] Engle, R. and Russell, J., 1998, \Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data" Econometrica 66, 1127-62.

[12] Epps, T., 1975, \Security Price Changes and Transaction Volumes: Theory and Evidence" American Economic Review 65, 586-597.

[13] French, D. and R. Roll, 1986, \Stock Return Variances: The Arrival of Information and the Reaction of Traders" Journal of Financial Economics 17, 5-26.

[14] Gallant, R., D. Hsieh, and G. Tauchen, 1991, \On Fitting a Recalcitrant Series: The pound/dollar Exchange Rate, 1974-1983" in Nonparametric and Semiparametric Methods in Econometrics and Statistics, W. Barnett, J. Powell, and G. Tauchen eds., Cambridge: Cambridge University.

[15] Ghysels, E., A. Harvey and E. Renault, 1996, \Stochastic Volatility" in Handbook of Statistics, Volume 14: Statistical Methods in Finance, G. Maddala and C. Rao eds., Amsterdam: North-Holland.

[16] Goodhart, C. and M. O'Hara, 1997, \High Frequency Data in Financial Markets: Issues and Applications" Journal of Empirical Finance 4, 73-114.

[17] Glosten, L. and P. Milgrom, 1985, \Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders" Journal of Financial Economics 14, 71-100.

[18] Harris, L, 1986, \A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns" Journal of Financial Economics 16, 99-117.

[19] Hausman, J., A. Lo and A. MacKinlay, 1992, \An Ordered Probit Analysis of Transaction Stock Price Changes" Journal of Financial Economics 31, 319-379.

[20] Hasbrouck, J., 1988, \Trades, Quotes, and Information," Journal of Financial Economics 22, 229-52.

[21] Hu®man, G., 1987, \A Dynamic Equilibrium Model of Asset Prices and Transaction Volume" Journal of Political Economy 95, 138-159.

[22] Kelly, D. and D. Steigerwald, 2001, \Private Information and High-Frequency Stochastic Volatility", web manuscript.

[23] La®ont, J. and E. Maskin, 1990, \The E±cient Market Hypothesis and Insider Trading on the Stock Market" Journal of Political Economy, 98, 70-93.

[24] Lange, S., 1998. \Modeling Asset Market Volatility in a Small Market: Accounting for Non-synchronous Trading E®ects," Journal of International Financial Markets, Institutions and Money, 9, 1-18.

[25] Owens, J. and D. Steigerwald, 2003, \Inferring Information Frequency and Quality" manuscript, U.C. Santa Barbara.

[26] Shorish, J. and Spear, S., 1996 \Shaking the Tree: An Agency-Theoretic Model of Asset Pricing" Technical Report 1996-E1, Graduate School of Industrial Administration, Carnegie Mellon University.

[27] Steigerwald, D., 1997, \Mixture Models and Conditional Heteroskedasticity" manuscript, U.C. Santa Barbara.

[28] Tauchen, G. and M. Pitts, 1983, \The Price Variability-Volume Relationship on Speculative Markets" Econometrica 51, 485-505.

[29] Tauchen, G., H. Zhang, and M. Liu, \Volume, Volatility, and Leverage: A Dynamic Analysis" Journal of Econometrics 74, 177-208.

[30] Timmermann, A., 2001, \Structural Breaks, Incomplete Learning, and Stock Prices " Journal of Business and Economic Statistics 19, 299-314.

7 Appendix

Proof of Theorem 1

For the proof of Theorem 1, let C_N represent $C_i = c_N$ in the conditioning information set.

Part 1. The expected price change from trader i, conditional on the public information set Z_{i-1} is

$$E(U_i|Z_{i-1}) = \sum_{j=A,B,N} P(C_i = c_j|Z_{i-1})\, U_i(C_i = c_j);$$

which equals

$$\alpha y_{i-1} v_{H_m} + \alpha x_{i-1} v_{L_m} + \alpha(1 - x_{i-1} - y_{i-1}) EV_m - \alpha E(V_m|Z_{i-1}) = 0.$$

Because

$$P(C_i = c_A|S_m \neq s_0) \neq P(C_i = c_A|Z_i)$$

for any finite i, price changes are not mean zero with respect to the information set of the informed.

Part 2. Let h and i be distinct values with $h < i$,

$$E(U_h U_i|Z_{i-1}) = U_h \cdot [E(V_m|Z_{i-1}) - E(V_m|Z_{i-1})] = 0.$$

Part 3. Recall $E(U_i^2|Z_{i-1})$ equals

$$P(C_i = c_A)(A_i - E(V_m|Z_{i-1}))^2 + P(C_i = c_B)(B_i - E(V_m|Z_{i-1}))^2$$
$$+P(C_i = c_N)(E[V_m|Z_{i-1}; C_N] - E(V_m|Z_{i-1}))^2.$$

The upper bound for the conditional variance is

$$\begin{aligned}
E\left(U_i^2|Z_{i-1}\right) &\leq P(C_i = c_A)(\bar{A}_i - E(V_m|Z_{i-1}))^2 + P(C_i = c_B)(\bar{B}_i - E(V_m|Z_{i-1}))^2 \\
&\quad + P(C_i = c_N)(E[V_m|Z_{i-1}; C_N]) - E(V_m|Z_{i-1}))^2 \\
&\leq [P(C_i = c_A) + P(C_i = c_N)](\bar{A}_i - E(V_m|Z_{i-1}))^2 \\
&\quad + [P(C_i = c_B) + P(C_i = c_N)](\bar{B}_i - E(V_m|Z_{i-1}))^2 \\
&\leq (\bar{A}_i - E(V_m|Z_{i-1}))^2 + (\bar{B}_i - E(V_m|Z_{i-1}))^2 \\
&\leq \left[(\bar{A}_i - E(V_m|Z_{i-1})) - (\bar{B}_i - E(V_m|Z_{i-1}))\right]^2 \\
&= \left[\bar{A}_i - \bar{B}_i\right]^2;
\end{aligned}$$

where the first inequality follows from the definition of $\bar{A}_i$ and $\bar{B}_i$ and the fourth inequality follows from $B_i \le E[V_m|Z_{i-1}] \le A_i$. Note that the unconditional variance is immediately obtained from Jensen's inequality

$$EU_i^2 \le E\left(\bar{A}_i - \bar{B}_i\right)^2 \le \left(E\bar{A}_i - E\bar{B}_i\right)^2.$$

To obtain the lower bound for the conditional variance we consider three cases. For each case we consider the set T_i, which has three elements: $|A_i - E(V_m|Z_{i-1})|$, $|B_i - E(V_m|Z_{i-1})|$ and $|E[V_m|Z_{i-1}, C_N]) - E(V_m|Z_{i-1})|$. Let $P_j = P(C_i = c_j)$. If $\min T_i = |A_i - E(V_m|Z_{i-1})|$, then

$$\begin{aligned} E\left(U_i^2|Z_{i-1}\right) &\ge (P_A + P_N)(A_i - E(V_m|Z_{i-1}))^2 + P_B(B_i - E(V_m|Z_{i-1}))^2 \\ &\ge P_B(P_A + P_N)\left(\bar{A}_i - \bar{B}_i\right)^2, \end{aligned}$$

where the second inequality follows from Lemma 1.1, which is proven below. If $\min T_i = |B_i - E(V_m|Z_{i-1})|$, then

$$\begin{aligned} E\left(U_i^2|Z_{i-1}\right) &\ge P_A(A_i - E(V_m|Z_{i-1}))^2 + (P_B + P_N)(B_i - E(V_m|Z_{i-1}))^2 \\ &\ge P_A(P_B + P_N)\left(\bar{A}_i - \bar{B}_i\right)^2, \end{aligned}$$

where the second inequality follows from Lemma 1.1. If $\min T_i = |E[V_m|Z_{i-1}, C_N]) - E(V_m|Z_{i-1})|$, then

$$\begin{aligned} E\left(U_i^2|Z_{i-1}\right) &\ge P_A(A_i - E(V_m|Z_{i-1}))^2 + (P_B + P_N)(E[V_m|Z_{i-1}, C_N] - E(V_m|Z_{i-1}))^2 \\ &\ge P_A(P_B + P_N)\left(\bar{A}_i - \bar{B}_i\right)^2, \end{aligned}$$

where the second inequality follows from Lemma 1.1.

The unconditional variance thus satisfies:

$$\min\{P_A(P_B + P_N), P_B(P_A + P_N)\}E\left(\bar{A}_i - \bar{B}_i\right)^2 \le E\left(U_i^2\right).$$

Hence $c = \min\{P_A(P_B + P_N), P_B(P_A + P_N)\}$, which by direct analysis is maximized at $P_A = P_B = \frac{1}{2}$.

Part 4. The proof follows logic in Easley and O'Hara (1992). Full details are contained in Kelly and Steigerwald (2001). ■

Lemma 1.1: Let $c \in [0, 1]$. For any pair of real numbers a and b

$$c(1 - c)(a + b)^2 \le ca^2 + (1 - c)b^2.$$

Proof. The left side of the inequality is $c(1-c)(a^2+b^2+2ab)$, which when subtracted from both sides converts the inequality to

$$0 \le c^2a^2 + (1-c)^2 b^2 - 2c(1-c)ab = [ca - (1-c)b]^2 .$$

∎

Proof of Theorem 2

The proof is a straightforward calculation of the correlation. By definition, the covariance is

$$\mathrm{Cov}(I_{t-r}, I_t) = E(I_{t-r}I_t) - EI_{t-r} \cdot EI_t .$$

If $r \ge k$, then the independence of the signal process implies that I_{t-r} is independent of I_t, so $E(I_{t-r}I_t) = EI_{t-r} \cdot EI_t$ and the covariance is zero.

For all calendar periods on information period m

$$\begin{aligned} E[I_t|S_m \neq s_0] &= \mu_1 = \delta(\alpha + \varepsilon(1-\alpha)); \\ E[I_t|S_m = s_0] &= \mu_0 = \delta\varepsilon(1-\alpha); \end{aligned}$$

$$\begin{aligned} Var[I_t|S_m \neq s_0] &= \sigma_1^2 = \delta[\alpha + \varepsilon(1-\alpha)](1-\alpha)(1-\varepsilon); \\ Var[I_t|S_m = s_0] &= \sigma_0^2 = \delta\varepsilon(1-\alpha)[1-\varepsilon(1-\alpha)]. \end{aligned}$$

If $r < k$ and I_{t-r} and I_t are measured on the same information period the conditional expectation of $(I_{t-r}I_t)$ is

$$\theta\mu_1^2 + (1-\theta)\mu_0^2,$$

which occurs with probability $\frac{k-r}{k}$. Second, if I_{t-r} and I_t are measured on consecutive information periods then the covariance is zero since information events are independent across trading days. Because the process for calendar period trades is stationary, EI_{t-r} equals EI_t. As noted in the text

$$EI_t = \theta\mu_1 + (1-\theta)\mu_0,$$

so

$$\begin{aligned} \mathrm{Cov}(I_{t-r}, I_t) &= \frac{k-r}{k}\theta(1-\theta)(\mu_1 - \mu_0)^2 \\ &= \frac{k-r}{k}\theta(1-\theta)(\alpha\delta)^2 . \end{aligned}$$

Combining the two possible cases for r relative to k yields

$$\mathrm{Cov}(I_{t-r}, I_t) = \begin{cases} \theta(1-\theta)(\alpha\delta)^2 \left[\frac{k-r}{k}\right] & r < k \\ 0 & r \geq k \end{cases} \tag{6}$$

Combining the covariance and variance of I_t given by (2) gives the desired correlation. Because all terms are positive for $r < k$, the correlation is positive. ■

Proof of Theorem 4

We derive $\mathrm{Cov}\left[(\Delta P_{t-r})^2, (\Delta P_t)^2\right]$ for $r = 1$ and $k = 3$. Derivation of the covariance for general r and k follows similar logic. Let $N = j$ if $t-1$ is the j^{th} calendar period in an information period. Then for $j = 1, 2, 3$:

$$E\left[(\Delta P_{t-1})^2 | N = j\right] = \theta\sigma_j + (1-\theta)\sigma_0$$

and

$$E\left[(\Delta P_t)^2 | N = j\right] = \begin{cases} E\left[(\Delta P_{t-1})^2 | N = j+1\right] & \text{for } j = 1, 2 \\ E\left[(\Delta P_{t-1})^2 | N = 1\right] & \text{for } j = 3 \end{cases}.$$

Because N is equally likely to take each value,

$$E\left[(\Delta P_t)^2\right] = \theta\bar{\sigma}_3 + (1-\theta)\sigma_0.$$

The covariance equals the conditional covariance plus the covariance of the conditional means. The conditional covariance is

$$\frac{1}{3}\sum_{j=1}^{3}\left\{E\left[(\Delta P_{t-1})^2(\Delta P_t)^2 | N = j\right] - E\left[(\Delta P_{t-1})^2 | N = j\right] E\left[(\Delta P_t)^2 | N = j\right]\right\},$$

which from the formulae for the expected calendar period squared price change given the value of N equals

$$\frac{1}{3}\left\{\begin{array}{c} [\theta\sigma_1\sigma_2 + (1-\theta)\sigma_0^2] - (\theta\sigma_1 + (1-\theta)\sigma_0)(\theta\sigma_2 + (1-\theta)\sigma_0) + \\ [\theta\sigma_2\sigma_3 + (1-\theta)\sigma_0^2] - (\theta\sigma_2 + (1-\theta)\sigma_0)(\theta\sigma_3 + (1-\theta)\sigma_0) + \\ \theta[\theta\sigma_3\sigma_1 + (1-\theta)\sigma_3\sigma_0] + (1-\theta)[\theta\sigma_0\sigma_1 + (1-\theta)\sigma_0^2] - \\ -(\theta\sigma_3 + (1-\theta)\sigma_0)(\theta\sigma_1 + (1-\theta)\sigma_0) \end{array}\right\}.$$

This expression simplifies to

$$\frac{1}{3}\theta(1-\theta)\left[(\sigma_1 - \sigma_0)(\sigma_2 - \sigma_0) + (\sigma_2 - \sigma_0)(\sigma_3 - \sigma_0)\right]. \tag{A4.2}$$

The covariance of the conditional means is

$$E\left\{\left(E(\Delta P_{t-1})^2 - E\left[(\Delta P_{t-1})^2|N\right]\right)\left(E(\Delta P_t)^2 - E\left[(\Delta P_t)^2|N\right]\right)\right\};$$

which equals

$$\sum_{j=1}^{3} P(N=j)\left\{\left(E(\Delta P_{t-1})^2 - E\left[(\Delta P_{t-1})^2|N=j\right]\right)\left(E(\Delta P_t)^2 - E\left[(\Delta P_t)^2|N=j\right]\right)\right\}:$$

Note

$$E(\Delta P_{t-1})^2 - E\left[(\Delta P_{t-1})^2|N=j\right] = \theta(\bar\sigma_3 - \sigma_j);$$

and

$$E(\Delta P_t)^2 - E\left[(\Delta P_t)^2|N=j\right] = \begin{cases} \theta(\bar\sigma_3 - \sigma_{j+1}) & \text{for } j = 1;2 \\ \theta(\bar\sigma_3 - \sigma_1) & \text{for } j = 3 \end{cases}.$$

Thus, the covariance of the conditional means is

$$\frac{\theta^2}{3}\left[(\bar\sigma_3 - \sigma_1)(\bar\sigma_3 - \sigma_2) + (\bar\sigma_3 - \sigma_2)(\bar\sigma_3 - \sigma_3) + (\bar\sigma_3 - \sigma_1)(\bar\sigma_3 - \sigma_3)\right]: \qquad \text{(A4.3)}$$

Combining (A4.2) with (A4.3) yields the formula.

Proof of second assertion: From the condition in the theorem, for all j,

$$\sigma_j > \theta\bar\sigma_k + (1-\theta)\sigma_0;$$

or:

$$(1-\theta)(\sigma_j - \sigma_0) > \theta(\bar\sigma_k - \sigma_j)$$

Hence from the covariance formula:

$$\begin{aligned} \operatorname{cov}(\Delta P_t^2; \Delta P_{t-r}^2) &> \theta\sum_{j=1}^{k-r}(\bar\sigma_k - \sigma_{j+r})(\sigma_j - \sigma_0) + \theta\sum_{j=1}^{k-r}(\bar\sigma_k - \sigma_{j+r})(\bar\sigma_k - \sigma_j) + \\ &\quad \theta\sum_{j=1}^{r}(\bar\sigma_k - \sigma_j)(\bar\sigma_k - \sigma_{k-r+j}) \\ &= \theta(\bar\sigma_k - \sigma_0)\sum_{j=1}^{k-r}(\bar\sigma_k - \sigma_{j+r}) + \theta\sum_{j=1}^{r}(\bar\sigma_k - \sigma_j)(\bar\sigma_k - \sigma_{k-r+j}) \end{aligned}$$

$$
\begin{aligned}
&= \theta(\bar{\sigma}_k - \sigma_0)\left(\sum_{j=1}^{k}(\bar{\sigma}_k - \sigma_j) - \sum_{j=1}^{r}(\bar{\sigma}_k - \sigma_j)\right) + \theta\sum_{j=1}^{r}(\bar{\sigma}_k - \sigma_j)(\bar{\sigma}_k - \sigma_{k-r+j}) \\
&= -\theta(\bar{\sigma}_k - \sigma_0)\sum_{j=1}^{r}(\bar{\sigma}_k - \sigma_j) + \theta\sum_{j=1}^{r}(\bar{\sigma}_k - \sigma_j)(\bar{\sigma}_k - \sigma_{k-r+j}) \\
&= \theta\sum_{j=1}^{r}(\sigma_j - \bar{\sigma}_k)(\sigma_{k-r+j} - \sigma_0);
\end{aligned}
$$

which is clearly positive for $r = 1$. The covariance is positive for all k for $r = k - 1$, because:

$$
\theta\sum_{j=1}^{k-1}(\sigma_j - \bar{\sigma}_k)(\sigma_{j+1} - \sigma_0) = \theta\sum_{j=1}^{i}(\sigma_j - \bar{\sigma}_k)(\sigma_{j+1} - \sigma_0) + \theta\sum_{j=i+1}^{k-1}(\sigma_j - \bar{\sigma}_k)(\sigma_{j+1} - \sigma_0);
$$

in which i is the largest integer for which $\sigma_j > \bar{\sigma}_k$. Hence:

$$
\theta\sum_{j=1}^{i}(\sigma_j - \bar{\sigma}_k)(\sigma_{j+1} - \sigma_0) + \theta\sum_{j=i+1}^{k-1}(\sigma_j - \bar{\sigma}_k)(\sigma_{j+1} - \sigma_0) >
$$

$$
\theta\sum_{j=1}^{i}(\sigma_j - \bar{\sigma}_k)(\sigma_{i+1} - \sigma_0) + \theta\sum_{j=i+1}^{k-1}(\sigma_j - \bar{\sigma}_k)(\sigma_{i+1} - \sigma_0)
$$

$$
= \theta(\sigma_{i+1} - \sigma_0)\sum_{j=1}^{k-1}(\sigma_j - \bar{\sigma}_k) = \theta(\sigma_{i+1} - \sigma_0)(\bar{\sigma}_k - \sigma_k) > 0 \blacksquare
$$

Proof of Proposition 5

For calendar period trades, let $\phi = \theta(1-\theta)(\alpha\delta)^2$, then:

$$
\frac{\mathrm{Cov}[I_{t-1}; I_t] - \mathrm{Cov}[I_{t-2}; I_t]}{\mathrm{Cov}[I_{t-1}; I_t]} = \frac{\phi\frac{k-1}{k} - \phi\frac{k-2}{k}}{\phi\frac{k-1}{k}} = \frac{1}{k-1};
$$

so the covariance in trades declines by a factor of $\frac{1}{k-1}$. Thus for the covariance in squared price changes to decline faster, we must show:

$$
\frac{\mathrm{Cov}\left[(\Delta P_{t-1})^2; (\Delta P_t)^2\right] - \mathrm{Cov}\left[(\Delta P_{t-2})^2; (\Delta P_t)^2\right]}{\mathrm{Cov}\left[(\Delta P_{t-1})^2; (\Delta P_t)^2\right]} > \frac{1}{k-1}: \tag{7}
$$

Because Condition 1 holds for period k, Proposition 5 implies the covariance is positive for $r = 1$. Hence the Equation (7) holds if and only if:

$$(k-2)\,\mathrm{Cov}\left[(\Delta P_{t-1})^2, (\Delta P_t)^2\right] > (k-1)\,\mathrm{Cov}\left[(\Delta P_{t-2})^2, (\Delta P_t)^2\right]. \tag{8}$$

Since $\sigma_j = \sigma_{j+1} + \phi$, we have:

$$\sigma_j = \sigma_k + (k-j)\phi,$$

which in turn implies:

$$\bar{\sigma}_k = \sigma_k + \phi\sum_{j=1}^{k}(k-j) = \sigma_k + \frac{\phi}{2}(k-1)$$

Substituting these facts and the formula for the covariance into (8) and performing some tedious algebra, we see that (8) holds if and only if

$$2(k-2)(k+3) > 3\theta(20 - 11k + k^2).$$

Or:

$$2(k-2)(k+3) > 3\theta(k-2.3)(k-8.7). \blacksquare$$

8 Figures

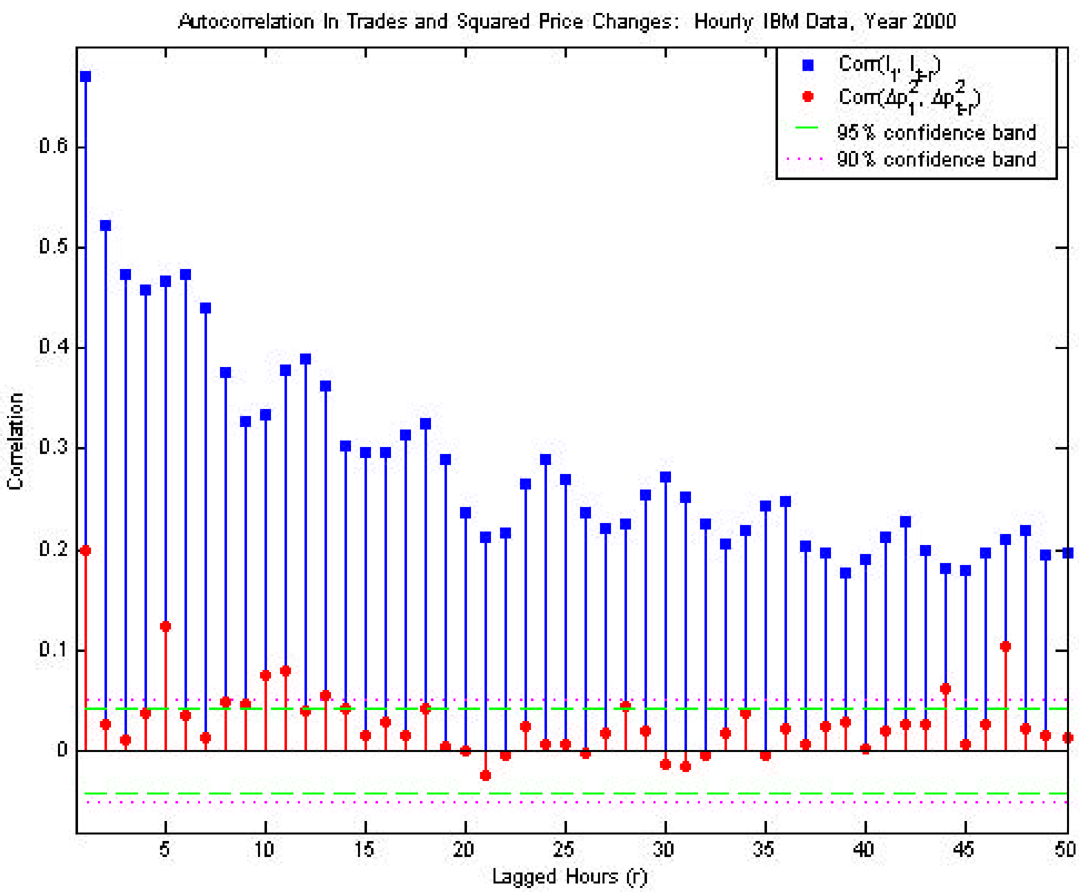

Figure 1: Serial Correlation Properties of 2000 IBM Data.

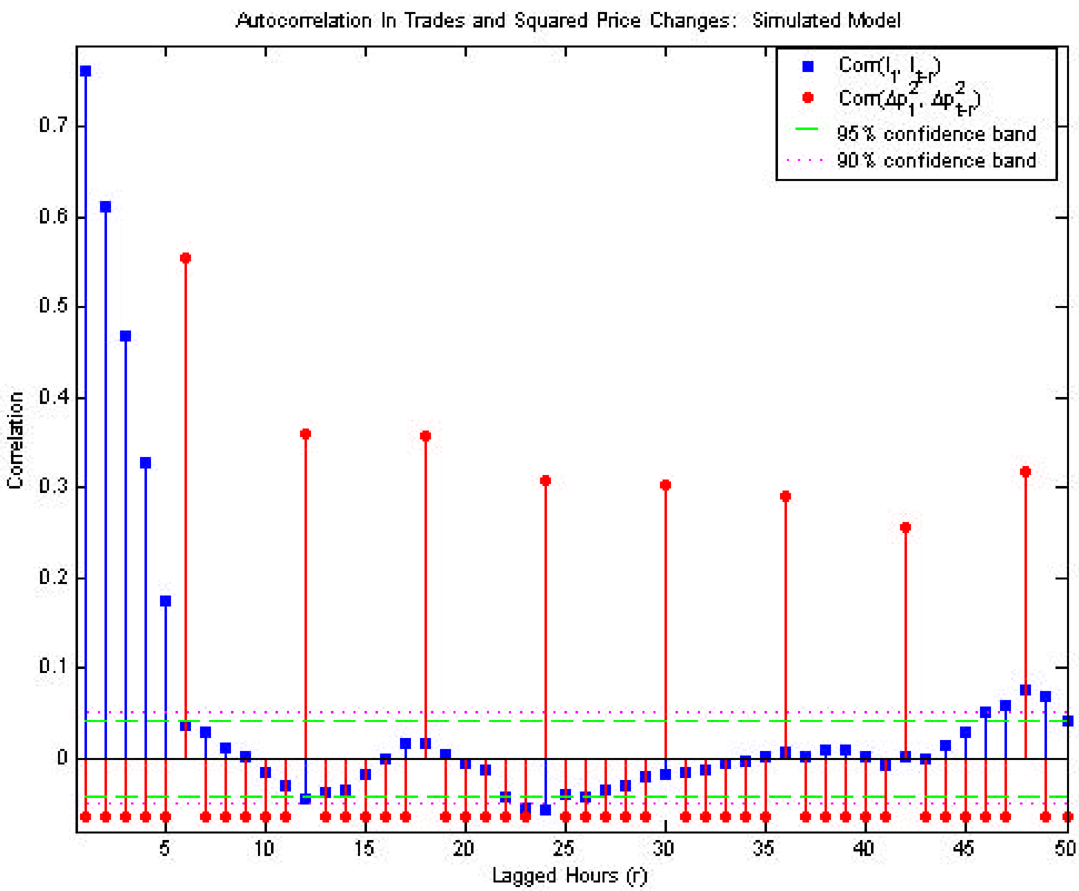

Figure 2: Calibrated model, k = 6.

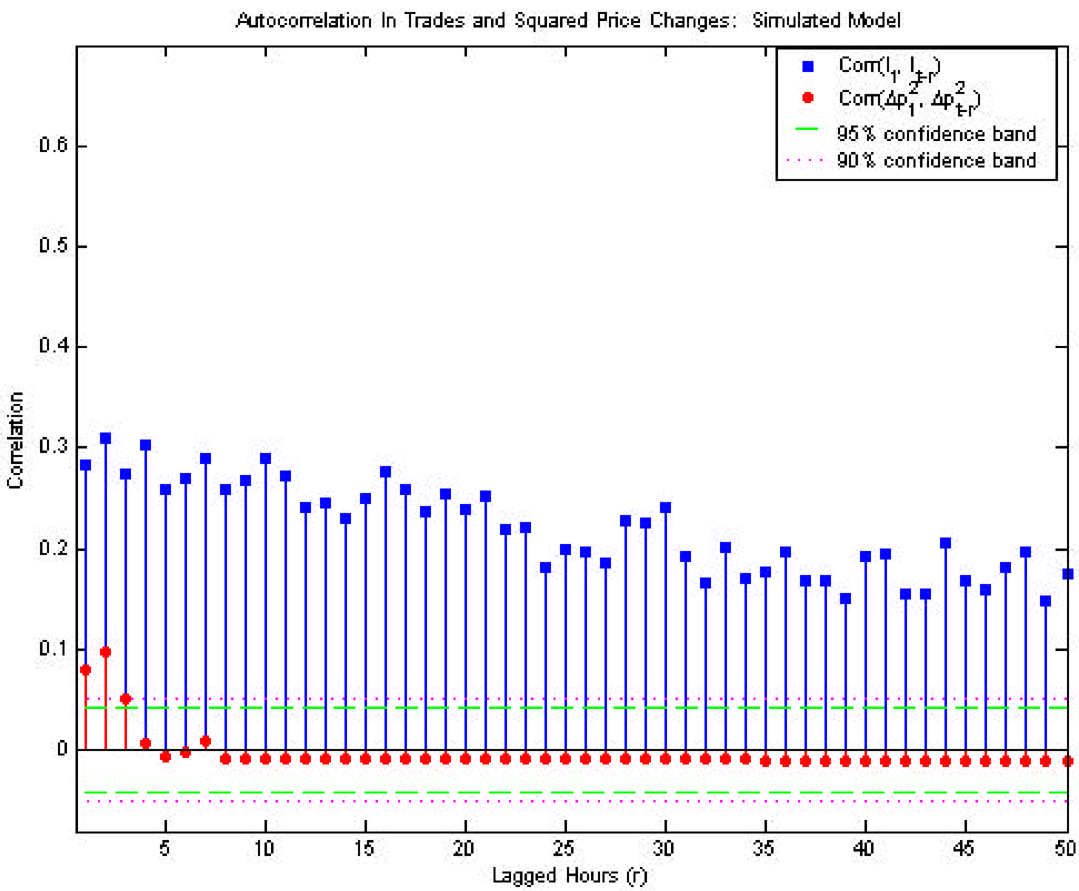

Figure 3: Calibrated Model, k = 90.