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

September 12, 2001

Abstract

We study the effect of privately informed traders on measured price changes and trades in asset markets. In the model exogenous news is captured by signals that informed agents receive. Agents trade anonymously through a market specialist, who does not receive a signal. 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.

*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 Francisco, 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 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, termed stochastic volatility, which has important implications for option pricing and conditional return forecasting. We provide an economic model that links the behavior of traders in a financial market following the arrival of news, to stochastic volatility in the asset prices. As such, our model provides an economic underpinning for the statistical models, such as GARCH and SV models, that form the basis of much recent empirical work.¹

There are two important related features of asset prices that our model must account for. First, there is extensive positive serial correlation in the number of trades (see for example Harris, 1987; Goodhart and O’Hara, 1997 page 96 provides a survey) and in trading volume (Harris, 1987; Andersen, 1996; Brock and LeBaron, 1996). Second, high-frequency serial correlation in trades is more persistent than serial correlation in squared price changes (Harris, 1987; Andersen, 1996; Steigerwald 1997).² To model the temporal features of asset prices, one must consider trades that occur sequentially. In Huffman (1987), the sequential nature of trades arises from the creation of a new generation each period. However, Huffman’s model appears to generate transitory negative serial correlation in both asset price and trading volume, which is inconsistent with the features described above.³

In our approach, the sequential nature of trades is derived through a model of the market’s microstructure. Traders, who arrive consecutively, differ in the information that they receive: Informed traders receive private news about the firm’s price that is unknown to other traders. Trade occurs anonymously with a market specialist who does not receive private news. The market specialist faces an adverse selection problem because the specialist trades with more informed agents with positive probability.

Private news arrives randomly and has two important effects. First, informed traders enter the market, increasing the number of trades relative to trading periods in which there is no private news. Because uninformed (liquidity) traders are also in the market, private

---

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

²Similarly, Tauchen, Zhang, and Liu (1996) report that a price change has more persistent effects on volume than on squared price changes.

³Specifically, the autocorrelation functions for squared price changes and trading volume derived from the figures that Huffman presents are insignificant for the first ten lags and the point estimates are generally negative.
news is not immediately revealed by the trade decisions of the informed. Thus the entry and exit of informed traders in response to the random arrival of private news implies that trades are serially correlated. Second, the market specialist widens the bid-ask spread in response to the possible adverse selection problem. As trade occurs, the market specialist uses Bayes rule to update beliefs and hence the bid and ask. As informed traders trade and reveal their information, the bid-ask spread declines. Because the squared (calendar period) price change is determined by the number of trades in the period and the variance of the price innovation for each trade, positive serial correlation in trades leads to positive serial correlation in squared price changes. 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.4

The entry and exit of informed traders after the arrival of private information is a key component of our model. The importance of private information as a determinant of stock 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. Our model is based on the market microstructure model of Easley and O’Hara (1992) which models the news arrival process. Market microstructure models which do not model the news arrival process generally do not exhibit serial correlation in trades. Glosten and Milgrom (1985) consider only a single news event, so trades are constant and thus serially uncorrelated. Sargent (1993) and Brock and LeBaron (1996) model traders who receive noisy signals. Because traders do not decide to leave the market, trades are serially uncorrelated, although volume generally declines through time.

Several researchers propose alternative explanations for serial correlation in squared price changes. Timmerman (1996) combines rare structural breaks in the dividend process with incomplete learning. Shorish and Spear (1996) show how moral hazard between the owner and manager of a firm generates serial correlation in squared price changes in a Lucas asset pricing model. Den Haan and Spear (1997) show how agency costs and borrowing constraints give rise to wealth effects that yield serial correlation in squared interest rate changes. Dividend based models provide an important first step by directly explaining serial correlation in squared price changes at low frequencies. Serial correlation in such models does not arise from the trading process, since the “no trade” theorems hold. In contrast our model explains how news (say about the dividend process) generates high frequency serial correlation through the trading process.

4Our model also accounts 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).
2 Market Microstructure Model

We consider a pure dealership market. In this way we rule out brokerage services provided by the specialist, implying that all orders are market orders. 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. Thus, as in Glosten and Milgrom (1985) and Easley and O’Hara (1992), we assume a Bertrand-style market.

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 define 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 signal is revealed to uninformed traders and to the specialist, and all traders agree upon the share value.

On each information period the random dollar value per share, $V_m$, takes one of two values $v_{Lm} < v_{Hm}$ with $P(V_m = v_{Lm}) = \delta$. To ensure the continuity of prices over information periods, $EV_m = v_{m-1}$ if the informed learn the true value of the stock on information period $m - 1$. If the informed do not learn the true value of the stock on information period $m - 1$, then we presume the possible share values are unchanged and $v_{Lm} = v_{Lm-1}$ and $v_{Hm} = v_{Hm-1}$.

The signals received by informed traders at the start of an information period are independent across information periods and identically distributed. Therefore, the 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 $V_m = v_{Hm}$, $s_L$ if the informed receive the low signal and learn $V_m = v_{Lm}$, 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 $\theta$, so the probability that $S_m$ takes the value $s_L$ is $\delta\theta$.

The signal completely determines the trading decisions of the informed. Conditional on receiving the uninformative signal, informed agents do not trade because of identical preferences. If informed traders receive signal $s_L$, then informed traders always sell as long as the specialist is uncertain that the true value is $v_{Lm}$. If informed traders receive signal $s_H$, then informed traders always buy as long as the specialist is uncertain that the true value is $v_{Hm}$.

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 $\omega_i$ be the rate of time discount for the $i$th trader. As in Glosten and Milgrom each individual assigns random utility to shares of stock, $s$, and current consumption, $c$, as $\omega s V_m + c$.\(^5\) The larger the value of $\omega$ the greater is the desire to invest and forego current consumption. We set $\omega = 1$ for the specialist and informed traders. There are three types of uninformed traders, those with $\omega = 1$, who have identical preferences and do not trade, those with $\omega = 0$, who always sell the stock, and those with $\omega = \infty$, who always buy the stock. Among the population of uninformed traders, the proportion with $\omega = 1$ is $1 - \varepsilon$, the proportion with $\omega = \infty$ is $(1 - \gamma)\varepsilon$, and the proportion with $\omega = 0$ is $\gamma \varepsilon$. The trading decisions of the uninformed are determined completely by the value of $\omega$ and do not depend on the bid and ask.

As in Easley and O’Hara (1992) and Glosten and Milgrom (1985), traders arrive randomly to the market one at a time, so we index traders by their order of arrival. The probability that the arriving trader is informed is $\alpha > 0$. The arrival of an informed trader within an information period is serially uncorrelated. The serial correlation properties of trades and squared price changes instead will follow from the difference in mean trades by the informed on news and no news days. 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 $i$th trader buys one share at the ask, $A_i$; $c_B$ if the $i$th trader sells one share at the bid, $B_i$; and $c_N$ if the $i$th 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). The sequence of trading decisions is public information. Let $Z_i$ be the publicly available information set prior to the arrival of trader $i + 1$. The information set available to the specialist and the uninformed is $Z_i$.

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. After the $i$th trader has come to the market, the specialist’s

---

\(^5\)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.
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 specialist’s beliefs about \( S_m \) translate directly into beliefs about the value of a share. If the specialist believes \( S_m = s_H \), then the accuracy of the signal implies that the specialist believes \( V_m = v_{Hm} \). Similarly, if the specialist believes \( S_m = s_L \), then the specialist also believes \( V_m = v_{Lm} \). If the specialist believes \( S_m = s_0 \), then the specialist assigns the unconditional probabilities to the possible values for \( V_m \). To summarize, after the \( i \)th trader has come to the market, the specialist’s conditional probability that \( V_m = v_{Hm} \) is

\[ P(V_m = v_{Hm}|Z_i) = y_i + (1 - x_i - y_i)(1 - \delta), \]

while \( P(V_m = v_{Lm}|Z_i) = 1 - P(V_m = v_{Hm}|Z_i) \). The action of each trader, even the decision not to trade, conveys information about the signal received by informed traders.

### 2.1 Determination of Ask and Bid

At the beginning of each information period, \( x_0 = \theta\delta \) and \( y_0 = \theta(1 - \delta) \). Let \( A_1 \) and \( B_1 \) be the initial ask and bid, respectively. (Thus \( A_1 \) is the ask that the first trader faces.) The equilibrium condition that the specialist earn zero expected profit from each trade provides the equations that determine the quoted prices \( (B_1, A_1) \). In essence, 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. We explicitly derive \( A_1 \) (derivation of \( B_1 \) follows similar logic). If the first trader trades at the ask, then the specialist’s expected loss from trade with an informed trader is

\[ \alpha \cdot y_0(A_1 - v_{Hm}), \]

where \( y_0(A_1 - v_{Hm}) \) is the expected loss if the first trader trades at the ask, given that the first trader is informed. Similarly, if the first trader trades at the ask, then the specialist’s expected gain from trade with an uninformed trader is

\[ (1 - \alpha) \varepsilon (1 - \gamma) \{(x_0 + \delta(1 - x_0 - y_0))(A_1 - v_{Lm}) + [y_0 + (1 - \delta)(1 - x_0 - y_0)](A_1 - v_{Hm})\}. \]
If expected profits equal zero, then
\[ A_1 = \frac{\alpha y_0 v_{Hm} + (1 - \alpha) \varepsilon (1 - \gamma) E(V_m | Z_0)}{\alpha y_0 + (1 - \alpha) \varepsilon (1 - \gamma)}, \]
where \( E(V_m | Z_0) = x_0 v_{Lm} + y_0 v_{Hm} + (1 - x_0 - y_0) E V_m. \) In parallel fashion
\[ B_1 = \frac{\alpha x_0 v_{Lm} + (1 - \alpha) \varepsilon \gamma E(V_m | Z_0)}{\alpha x_0 + (1 - \alpha) \varepsilon \gamma}. \]
The equations for \((B_i, A_i)\) are simply the equations for \((B_1, A_1)\) with \(y_0\) replaced by \(y_{i-1}\) and \(x_0\) replaced by \(x_{i-1}\) (which implies \(E(V_m | Z_0)\) is replaced by \(E(V_m | Z_{i-1})\)). As one would expect, both the bid and ask increase with \(y_{i-1}\) and decrease with \(x_{i-1}\).

It is easy to see that \(v_{Lm} \leq B_i \leq A_i \leq v_{Hm}\), with strict inequality unless the specialist is certain the informed learned the true value of \(V_m\) (no adverse selection). Mathematically, the specialist is certain the informed learned the true value of \(V_m\) if \(x_{i-1} = 1\) or \(y_{i-1} = 1\). It is also easy to see that \(B_i \leq E(V_m | Z_{i-1}) \leq A_i\), which follows directly from \(v_{Lm} \leq E(V_m | Z_{i-1}) \leq v_{Hm}\).

### 2.2 Learning Rules

As trading occurs, information accrues to the specialist. In response, the specialist updates the probabilities \((x_i, y_i)\). We begin by examining how the specialist learns from the action of the first trader and explicitly discuss only updating of \(y_1\) (updating of \(x_1\) follows similar logic). The key parameters that govern the speed of learning are \(\alpha\) and \(\varepsilon\). If the first trader trades at the ask
\[ y_1 = y_0 \frac{\alpha + (1 - \alpha) \varepsilon (1 - \gamma)}{\alpha y_0 + (1 - \alpha) \varepsilon (1 - \gamma)}. \]
As long as \(y_0 < 1\), a trade at the ask increases \(y_1\). If \(\alpha = 1\) or \(\varepsilon = 0\) only informed traders trade, so learning is immediate and \(y_1 = 1\). If the first trader trades at the bid
\[ y_1 = y_0 \frac{(1 - \alpha) \varepsilon \gamma}{\alpha x_0 + (1 - \alpha) \varepsilon \gamma}. \]
As long as \(x_0 > 0\), a trade at the bid decreases \(y_1\). If \(\alpha = 1\) or \(\varepsilon = 0\) again learning is immediate, so \(y_1 = 0\) and \(x_1 = 1\). Finally, if the first trader does not trade
\[ y_1 = y_0 \frac{(1 - \alpha) (1 - \varepsilon)}{\alpha (1 - x_0 - y_0) + (1 - \alpha) (1 - \varepsilon)}. \]

---

6 The updating, or learning formulae are derived from Bayes rule in the Appendix.
As long as \((1 - x_0 - y_0) > 0\), a decision not to trade decreases \(y_1\). If \(\alpha = 1\), or if \(\varepsilon = 1\) in which case all uninformed traders trade, then learning is immediate with \(y_1 = 0\) and \(x_1 = 0\).

The learning formulae for \(y_i\) are simply the learning formulae for \(y_1\) with \(y_0\) replaced by \(y_{i-1}\). The learning formulae for \(x_i\) are:

\[
x_i = x_{i-1} \frac{(1 - \alpha) \varepsilon (1 - \gamma)}{\alpha y_{i-1} + (1 - \alpha) \varepsilon (1 - \gamma)},
\]

if trader \(i\) trades at the ask;

\[
x_i = x_{i-1} \frac{\alpha + (1 - \alpha) \varepsilon \gamma}{\alpha x_{i-1} + (1 - \alpha) \varepsilon \gamma},
\]

if trader \(i\) trades at the bid; and

\[
x_i = x_{i-1} \frac{(1 - \alpha)(1 - \varepsilon)}{\alpha (1 - x_{i-1} - y_{i-1}) + (1 - \alpha)(1 - \varepsilon)},
\]

if trader \(i\) does not trade.

### 2.3 Consistency of Learning

We have posited that the signal is revealed at the end of a information period, which consists of a finite number of trader arrivals. To ensure that the learning formulae we described above are useful, we establish that if there were an infinite number of trader arrivals, the specialist would learn the value of \(S_m\). As a result, the bid and ask converge to the strong-form efficient value of a share, in which the bid and ask reflect both the public and private information. Because transaction prices are determined by the bid and ask, transaction prices also converge to the strong-form efficient value of a share.

Three sets of beliefs capture the specialist’s uncertainty about the value of \(S_m\). The first is the specialist’s belief that \(S_m = s_H\), which is expressed as the sequence of conditional probabilities \(\{y_i\}_{i=1}^{\infty}\). The second is the specialist’s belief that \(S_m = s_L\), which is expressed as the sequence of conditional probabilities \(\{x_i\}_{i=1}^{\infty}\), and finally the third is the belief that \(S_m = s_0\), which is expressed as the sequence \(\{1 - x_i - y_i\}_{i=1}^{\infty}\).
Theorem 1: The sequence of bids and asks, and hence the sequence of transaction prices, converge almost surely to their strong-form efficient values at an exponential rate. Formally, as \( i \to \infty \):

If \( S_m = s_H \), then \( x_i \overset{a.s.}{\to} 0 \), \( y_i \overset{a.s.}{\to} 1 \) and \( A_i \overset{a.s.}{\to} v_{Hm} \), \( B_i \overset{a.s.}{\to} v_{Hm} \).

If \( S_m = s_L \), then \( x_i \overset{a.s.}{\to} 1 \), \( y_i \overset{a.s.}{\to} 0 \) and \( A_i \overset{a.s.}{\to} v_{Lm} \), \( B_i \overset{a.s.}{\to} v_{Lm} \).

If \( S_m = s_0 \), then \( x_i \overset{a.s.}{\to} 0 \), \( y_i \overset{a.s.}{\to} 0 \) and \( A_i \overset{a.s.}{\to} EV_m \), \( B_i \overset{a.s.}{\to} EV_m \).

Proof: See Appendix.

Although the asymptotic behavior of prices is straightforward to determine, calculating the serial correlation properties requires knowledge of the distribution of share prices in each time period, a more difficult task which we turn to next.

3 Calendar Period Implications

With the learning rules established, we now show that the model accounts for the main empirical findings described in the introduction. We first show that one implication of the model is that the number of trades in a calendar period is serially correlated. As described in the introduction, such serial correlation leads to serial correlation in squared price changes. We then show that the serial correlation in the number of trades per calendar period is more persistent than is the serial correlation in squared price changes.

To derive calendar period implications, we must be clear about how trading opportunities are aggregated. A information period contains \( k \) calendar periods (such as an hour). A calendar period, which is indexed by \( t \), contains \( \eta \) trader arrivals, which as above are indexed by \( i \). (We can think of a trader arrival, or trading opportunity, as a unit of economic time.) The sample period consists of a large sample of information periods.

Calendar Period Trades

First we examine the covariance structure of the number of trades per calendar period. Let the number of trades in (calendar) period \( t \) be \( I_t \). Because \( \eta \) traders arrive each period, \( I_t \) takes integer values between 0 and \( \eta \). In fact \( I_t \) is distributed as a binomial random variable where the number of trades corresponds to the number of “successes” in \( \eta \) “trials”. For all periods on information period \( m \) the probability that a trader decides to trade is

\[
\begin{align*}
P(C_i \neq c_N | S_m \neq s_0) &= \alpha + (1 - \alpha) \varepsilon, \\
P(C_i \neq c_N | S_m = s_0) &= (1 - \alpha) \varepsilon,
\end{align*}
\]
so the distribution of \( I_t \) conditional on the value of \( S_m \) is
\[
I_t | (S_m \neq s_0) \sim B(\eta, \alpha + \varepsilon (1 - \alpha)) , \\
I_t | (S_m = s_0) \sim B(\eta, \varepsilon (1 - \alpha)) .
\]
Thus, for all calendar periods on information period \( m \)
\[
E[I_t | S_m \neq s_0] = \mu_1 = \eta(\alpha + \varepsilon (1 - \alpha)) ,
E[I_t | S_m = s_0] = \mu_0 = \eta \varepsilon (1 - \alpha),
\]
\[
Var[I_t | S_m \neq s_0] = \sigma_1^2 = \eta [\alpha + \varepsilon (1 - \alpha)] (1 - \alpha) (1 - \varepsilon) ,
Var[I_t | S_m = s_0] = \sigma_0^2 = \eta \varepsilon (1 - \alpha) [1 - \varepsilon (1 - \alpha)] .
\]

Unconditionally, we have:
\[
E[I_t] = \mu = \theta \mu_1 + (1 - \theta) \mu_0 \\
Var[I_t] = \sigma^2 = \theta \sigma_1^2 + (1 - \theta) \sigma_0^2 + \theta (1 - \theta) (\mu_1 - \mu_0)^2
\]

Given the above structure for the number of trades in a calendar period, we can derive
the serial correlation properties of the number of trades.

**Theorem 2:** Let \( r > 0 \). If \( r < k \), 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 \( r \), the correlation between \( I_{t-r} \) and \( I_t \) is given by:
\[
Cor(I_{t-r}, I_t) = \frac{\theta (1 - \theta) (\alpha \eta)^2}{\sigma^2} k - \min(r,k)^\# \]

**Proof:** See Appendix.

The underlying random process generates sets of \( k \) elements, with each set corresponding
to a information period. For example if \( k = 2 \), then there are two calendar measurements on
each information period. Because the information arrival process is independent over time,
the pair of calendar measurements corresponding to one information period is independent
of the pair of calendar measurements corresponding to any other information period. That is, the process generating the calendar measurements is pairwise (weakly) stationary. The sequence of calendar measurements is not itself generated by a stationary process. In light of the nonstationarity, 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 period in the past the process began. Specifically, for ease of exposition we begin the process with information arrivals in the morning. As information arrivals could just as likely have started in the afternoon, we do not want our calendar period implications to depend on such an arbitrary assumption. To avoid such an implication, we assume that $t$ is randomly sampled, so that $I_t$ is equally likely to be a morning or an afternoon observation. The correlation in $I_t$ is then independent of time.

According to Theorem 2, positive serial correlation persists for $k$ calendar periods. Our definition of a calendar period is only in terms of an information period, which is arbitrary. Hence, serial correlation of any persistence and frequency of data is possible for an appropriately defined information period. However, low frequency serial correlation implies a long information period, so our model may best thought of as explaining high frequency serial correlation.

Theorem 2 gives the exact formula for the correlation. Therefore, it is straightforward to establish comparative statics, which we summarize in the following corollary.

**Corollary 3:** If $r < k$, then the correlation between $I_{t-r}$ and $I_t$ is decreasing in $r$, increasing in $k$, increasing in $\eta$ and increasing in $\alpha$.

**Proof:**
Substituting in the definitions of $k$ and $\sigma^2$ into equation (3) and assuming $r < k$, results in:

$$\text{Cor}(I_{t-r}, I_t) = \frac{\mu \tau - r\eta}{\tau} \frac{\eta \theta (1 - \theta) \alpha^2}{\varepsilon (1 - \alpha)[1 - \varepsilon (1 - \alpha)] + \theta \alpha [(1 - \alpha) (1 - 2\varepsilon) + \alpha \eta (1 - \theta)]}.$$  

The results then follow by taking the appropriate derivatives.

The comparative static calculations in Corollary 3 imply certain patterns of serial correlation in trades across markets. We first study how the serial correlation in trades is affected by changes in the parameters characterizing aggregation over time. As the number of periods in a information period, $k$, increases, the impact of the entry and exit of informed traders grows and the serial correlation increases. As the number of trader arrivals in a calendar period, $\eta$, increases, the difference between the mean number of trades on news and no news
days widens, thus increasing the covariance quadratically. Increasing \( \eta \) also increases the variance linearly as more trades generate more variance. Because the covariance effect is stronger, the correlation is increasing in \( \eta \). Thus one may expect to see more pronounced serial correlation in asset markets in which the revelation of private information takes a relatively longer period of time. Similarly, one may expect to see more pronounced serial correlation in markets with more trader activity.

Both of the preceding calculations allow only one parameter to change; implicit in our comparison of market thickness is the assumption that the length of the information period is fixed. Yet for many comparisons, both \( k \) and \( \eta \) are changing. A leading case would be comparison of information gathered at two different calendar period frequencies, say 5 minute intervals versus hourly intervals. Because the data are gathered for the same asset, the number of trader arrivals in a information period, \( \tau = k \eta \), is constant for both frequencies. To understand the effect on the correlation caused by changing from 5 minute intervals to hourly intervals, we substitute \( \frac{\tau}{k} \) for \( \eta \) and take the derivative with respect to \( k \). As the change from 5 minute intervals to hourly intervals simultaneously decreases \( k \) and increases \( \eta \), we have two countervailing effects on the correlation. In general, the serial correlation can either increase or decrease with a change in calendar period and, perhaps most interestingly, the change is not constant across \( r \). Because the magnitude of the effect of a change in \( k \) on the correlation depends on \( r \), it is for long lags that we would most likely see the serial correlation in trades decline as we move from 5 minute data to hourly data.

To understand how the serial correlation in trades depends upon the underlying parameters of the market microstructure model, we can decompose the correlation into three terms. The first term is the difference between the number of trades on a information period with news and on a information period without news, which is \((\mu_1 - \mu_0)^2\). The remaining two terms are the conditional variances on a information period with news \((\sigma_1^2)\) and a information period without news \((\sigma_0^2)\), respectively. An increase in \((\mu_1 - \mu_0)^2\) increases both the covariance of trades and the variance of trades, where \((\mu_1 - \mu_0)^2\) enters the variance through the component for the variance of the conditional means, so the overall impact on the serial correlation in trades must be calculated. An increase to the conditional variances leads only to an increase in the variance, so the overall impact is to reduce the serial correlation in trades.

To understand why the serial correlation is an increasing function of \( \alpha \), observe that increasing \( \alpha \) has two effects on the correlation. First, with a larger number of informed traders, there is a wider difference between the number of trades on a information period with news and on a information period without news. Second, because the informed traders all make the same trading decision, an increase in the number of informed traders decreases at least one of the conditional variances. As a result, the positive impact on the covariance
outweighs the positive impact on the variance. Because the serial correlation in trades is an increasing function of $\alpha$, a market with many informed traders has more serial correlation in trades than does a market with fewer informed traders.

Next, consider the relationship between $\varepsilon$ (the fraction of uninformed who do not trade) and the serial correlation in trades. For specific parameter values the partial derivative is definitively signed. If $\varepsilon$ is small (precisely, if $\varepsilon < \frac{1-2\alpha\varepsilon^2}{2(1-\alpha)}$), then virtually all trades are by informed traders and increasing $\varepsilon$ dilutes the informed traders and reduces the serial correlation in trades. If $\alpha$ is large (precisely if $\alpha \theta \geq \frac{1}{2}$), then increasing $\varepsilon$ increases the variation in trades across days and increases the serial correlation in trades.

In similar fashion, increasing $\theta$ increases the correlation if $\varepsilon$ is large and $\theta$ is small (precisely $\varepsilon > \frac{1}{2}$ and $\theta < \frac{1}{2}$). Because good and bad news are symmetric in the model, the serial correlation is unaffected by changes to $\gamma$ or $\delta$. An interesting implication is that our model predicts correlation in a variety of markets. For example, there is serial correlation in trades in both liquid and illiquid markets. Our findings of serial correlation even in illiquid markets is also supported empirically by Lange (1998).

Of course, serial correlation in the number of trades could be artificially imposed by creating serial correlation in the private information arrival process. Engle et al. (1990) find some evidence of serial correlation in public news; although serial correlation in public news does not imply serial correlation in private news. Appealing to serial correlation in private news does not really provide an economic cause for serial correlation in squared price changes as it begs the question as to what causes serial correlation in private news.
Behavior of Individual Trader Price Changes

To understand the serial correlation in squared price changes per calendar period, we first study the behavior of the price changes that follow the arrival of each trader. The price change that results from the action of trader $i$ is $U_i = E(V_m|Z_i) - E(V_m|Z_{i-1})$, where $E(V_m|Z_i) = x_i v_{Lm} + y_i v_{Hm} + (1 - x_i - y_i) E(V_m)$.\(^7\) The definition of $U_i$ incorporates the arrival of public information after the decision of trader $i - 1$ but before the decision of trader $i$. To relate decisions in economic time given by our model to the calendar period measurements, we write calendar period price changes as

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

Price changes in economic time thus drive calendar price changes. In turn, the information content of trades (or no trades) drive price changes in economic time. The information content of a trade or no trade depends on the history of trades and the parameter values. For example, if $\varepsilon$ is large, no trades convey relatively more information. If $\gamma$ is large, a trade at the ask conveys relatively more information. Trades or no trades at early economic time periods convey more information than trades at later time intervals. In this way, squared price changes are serially correlated.

To provide insight, we study in detail the price change associated with the arrival of the first trader on information period $m$. There are three possible values for $U_1$, one corresponding to each of the possible trade decisions. If $C_1 = c_A$, then $E(V_m|Z_1) = A_1$, and

$$U_1 = \frac{\alpha y_0 [v_{Hm} - E(V_m|Z_0)]}{P(C_1 = c_A|Z_0)}.$$ If $C_1 = c_B$, then $E(V_m|Z_1) = B_1$ and

$$U_1 = \frac{\alpha x_0 [v_{Lm} - E(V_m|Z_0)]}{P(C_1 = c_B|Z_0)}.$$ Finally, if $C_1 = c_N$, then

$$E(V_m|Z_1) = \frac{\alpha (1 - x_0 - y_0) EV_m + (1 - \alpha) (1 - \varepsilon) E(V_m|Z_0)}{\alpha (1 - x_0 - y_0) + (1 - \alpha) (1 - \varepsilon)}$$

\(^7\)The price is conditional on public information and is hence theoretically observable to the econometrician. In reality the set of parameters must be estimated, resulting in an estimate of the price based on the estimated parameters. However, in most empirical studies of serial correlation in squared price changes, econometricians use the bid, ask, or last trade, which may have different properties from the price.
and
\[ U_1 = \frac{\alpha (1 - x_0 - y_0) [EV_m - E (V_m|Z_0)]}{P (C_1 = c_N|Z_0)}. \]

Because \( \delta y_0 = (1 - \delta) x_0 \) updating from the first trader is more informative if a trade occurs then if a trade does not occur, so \( B_1 < E (V_m|Z_0, C_1 = c_N) < A_1 \). While the inequality is generally satisfied for the remaining traders, it is possible for \( E (V_m|Z_{i-1}, C_i = c_N) \) to fall outside \( (B_i, A_i) \).

The mean price change from trader 1 is
\[ E (U_1|Z_0) = \sum_{j=A,B,N} P (C_1 = c_j|Z_0) U_1 (C_1 = c_j), \]
which equals
\[ \alpha y_0 v_{Hm} + \alpha x_0 v_{Lm} + \alpha (1 - x_0 - y_0) EV_m - \alpha E (V_m|Z_0) = 0. \]

Because
\[ P (C_i = c_A|s_m) \neq P (C_i = c_A|Z_i) \]
for any finite \( i \), price changes are not mean zero with respect to the information set of the informed.

The variance of the price change from trader 1 is
\[ E U_1^2|Z_0 = \sum_{j=A,B,N} P (C_1 = c_j|Z_0) U_1^2 (C_1 = c_j), \]
which equals
\[ \frac{(\alpha y_0)^2 [v_{Hm} - E (V_m|Z_0)]^2}{P (C_1 = c_A|Z_0)} + \frac{(\alpha x_0)^2 [v_{Lm} - E (V_m|Z_0)]^2}{P (C_1 = c_B|Z_0)} + \frac{\alpha^2 (1 - x_0 - y_0)^2 [EV_m - E (V_m|Z_0)]^2}{P (C_1 = c_N|Z_0)}. \]

To understand the impact of informed traders on the behavior of calendar period squared price changes, we must compare the variance of \( U_1 \) for \( S_m = s_0 \) with the variance of \( U_1 \) for \( S_m \neq s_0 \). (In general, the comparison will depend on whether the low or high signal was received. If \( \gamma = \delta = .5 \), then \( x_0 = y_0 \) and the variance of \( U_1 \) is identical for the low and high signals. In the remainder of the section we assume \( \gamma = \delta = .5 \) and so we do not need to

---

\(^8\)For example, if \( \varepsilon \) is very large and \( \alpha \) is very small (so that the rare no trade decisions are most often made by informed traders), then it is possible that \( E (V_m|Z_{i-1}, C_i = c_N) > A_i \).
distinguish between the low and high signals.) The addition of the signal alters the variance
of $U_1$ only through the impact on the probability with which each trade outcome is observed

$$E \ U_1^2 | S_m = s_m, Z_0 = \prod_{j=A,B,N} P(C_1 = c_j | S_m = s_m) U_1^2 (C_1 = c_j).$$

We compare the probabilities of each trade outcome for $S_m = s_H$ with $S_m = s_0$:

$$P(C_1 = c_A | S_m = s_H) = P(C_1 = c_A | S_m = s_0) + \alpha$$
$$P(C_1 = c_N | S_m = s_H) = P(C_1 = c_N | S_m = s_0) - \alpha,$$

where the probability that $C_1 = c_B$ is the same for the two values of $S_m$. Thus $E(U_1^2 | S_m = s_H, Z_0) - E(U_1^2 | S_m = s_0, Z_0)$ equals

$$\alpha \left( (\alpha^2 y_0)^2 [v_{Hm} - E(V_m | Z_0)]^2 - \alpha^2 (1 - x_0 - y_0)^2 [EV_m - E(V_m | Z_0)]^2 \right) \left[ P(C_1 = c_A | Z_0) \right]^2 - \left[ P(C_1 = c_N | Z_0) \right]^2,$$

which is greater than zero because $EV_m = E(V_m | Z_0)$. Because the term is positive, the
price uncertainty from the first trader is higher on a day with news than on a day without
news. The impact of trader 2 and following traders is not immediately signed because
$EV_m \neq E(V_m | Z_i)$ for $i > 0$. To determine the sign of the difference we study the behavior
of $U_i$ for general $i$.

For general $i$ there are $3^i$ possible values for $U_i$, so direct calculation of the moments
of $U_i$ is tedious. Rather, we construct analytic bounds to the moments that describe the
behavior of the distribution of $U_i$. Let

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

be the “spread” or the difference between the maximum price change and the minimum
price change. For most parameter values, the spread is equal to the familiar bid-ask spread.
However, as noted earlier, for some parameter values a no trade may induce larger or smaller
price changes than a trade at the ask or bid, respectively.

Let $\{U_i\}_{i=1}^{kn}$ be the sequence of trader price changes (price changes in economic time)
for a information period. With respect to the public information set, the elements of the
sequence are uncorrelated but are dependent and not identically distributed. Specifically,
the trader price changes are heteroskedastic and the heteroskedasticity is autoregressive.
Theorem 4: Price changes in economic time satisfy:

1. \( E(U_i|Z_{i-1}) = 0 \)
2. \( E(U_hU_i|Z_{i-1}) = 0 \) for \( h < i \)
3. \( (P(C_i = c_A)P(C_i = c_B) + P(C_i = c_N)\min\{P(C_i = c_A)P(C_i = c_B)\})^3 \mathcal{E}_i - \mathcal{B}_i^2 \leq E(U_i^2|Z_{i-1}) \leq (\alpha)^2 (1 - x_{i-1} - y_{i-1})^2 \left\{ EV_m - E(V_m|Z_{i-1}) \right\}^2 \left\{ P(C_1 = c_N|Z_{i-1}) \right\}^2 \)

Proof: See Appendix.

The first two parts of Theorem 4 deliver the traditional results that \( EU_i = 0 \) and that \( E(U_iU_j) = 0 \) if \( i \) does not equal \( j \). The spread drives the variance in \( U_i \) and induces heteroskedasticity. Theorem 4 and Proposition 1 together imply that \( E(U_i^2|Z_{i-1}) \to 0 \) as \( i \to \infty \). As the market maker becomes certain of the true value of the share, the bid and ask converge to the true value of the share and squared price changes go to zero.

To determine how the properties of the distribution of \( U_i \) are affected by the signal received by informed traders, we compare the variance of \( U_i \) if \( S_m = s_0 \) with the variance of \( U_i \) if \( S_m \neq s_0 \). Parallel to the case for trader 1, \( E(U_i^2|S_m = s_H, Z_{i-1}) - E(U_i^2|S_m = s_0, Z_{i-1}) \) equals

\[
\frac{\alpha (\alpha y_{i-1})^2 [v_{Hm} - E(V_m|Z_{i-1})]^2}{[P(C_1 = c_A|Z_{i-1})]^2} - \frac{\alpha^2 (1 - x_{i-1} - y_{i-1})^2 [EV_m - E(V_m|Z_{i-1})]^2}{[P(C_1 = c_N|Z_{i-1})]^2}.
\]

The difference (5) depends on \( E(V_m|Z_{i-1}) \), which in turn depends on \( (x_{i-1}, y_{i-1}) \).

If the proportion of informed traders is high enough, then learning takes place quickly and the entire distribution of \( U_i \) can be directly calculated. For example, if \( \alpha = .9 \), then the bid-ask spread is reduced very close to zero in only 10 trades. For \( \alpha = .9 \) the columns of Table 1 contain the values of (5) corresponding to trader 1 through trader 10 and the rows of Table 1 correspond to different values of \( \varepsilon \).

\footnote{For the calculations in Table 1 and the simulations in Figure 1 we set \( \theta = .4, v_L = 10, v_H = 20, \) and \( \gamma = \delta = .5 \).}
The entries in Table 1 reveal two important features. First, all entries are positive, reflecting the greater likelihood of large squared price changes on days with news than on days without news. (Also note that as the value of $\epsilon$ declines the information content of a trade increases, and so the first squared price change increases.) Second, as learning accumulates (moving across a row) the difference in squared price changes tends toward zero.

For smaller values of $\alpha$ learning occurs more slowly, so reduction of the bid-ask spread to zero takes many more trades and calculation of the exact distribution is cumbersome. To understand the behavior of (5) with a smaller value of $\alpha$, we approximate the exact distribution with simulations. We let $\alpha$ decline dramatically and find that in each case the variance of $U_i$ is higher on a news day than on a day without news, which confirms the fundamental insight of Table 1.\textsuperscript{10} In Figure 1 we report the simulated distribution for $\alpha = .2$ and $\epsilon = .5$. The vertical axis indexes $E(U_i^2|S_m = s_H, Z_{i-1})$, for the curve labeled News Days, and $E(U_i^2|S_m = s_0, Z_{i-1})$, for the curve labeled No News Days. Thus the quantity in (5) is given by difference between the two lines. The horizontal axis indexes trader arrivals. As is clear from the vertical scale, learning is essentially complete after 100 trader arrivals. As the figure reveals, the quantity in (5) is positive and, because both lines converge to zero as the number of traders increases, the quantity (5) shrinks to zero as the number of traders increases.

\textsuperscript{10}Specifically, we let $\alpha \in (.2, .4, .75)$ and set $\epsilon = .5$. Each experiment consists of 3000 simulations.
Calendar Period Squared Price Changes

From the results above, we make assumptions to simplify the structure for the expectation of the calendar period squared price change. The analytic and simulation results indicate that for information periods in which the informed do not trade, the specialist’s initial uncertainty about the signal is resolved quickly. As a result, the variance in calendar period price changes, which is driven by the random decisions of the uninformed, is constant over much of the information period. In contrast, for information periods in which the informed do trade, the specialist’s uncertainty is not resolved quickly, but rather declines over the course of the information period. To capture these phenomena mathematically, let $j$ index the calendar periods in a information period. For $t = 1, \ldots, j$ we assume

$$
E (\Delta P_t)^2 | S_m = s_0 = \sigma_0
$$

$$
E (\Delta P_t)^2 | S_m \neq s_0 = \sum_{j=1}^{\infty} \sigma_j 1(t = j).
$$
Further, we assume $\sigma_1 > \sigma_2 > \cdots > \sigma_k > \sigma_0$. If calendar period $t$ is the first period of information period $m$ on which $S_m \neq s_0$, the expected squared price change is $\sigma_1$. The inequality $\sigma_k > \sigma_0$ arises from the assumption that the information advantage of the informed persists until the information period ends, while for information periods without informed traders the uncertainty is quickly resolved. Given the above structure, the unconditional expectation of calendar period squared price changes is

$$E (\Delta P_t)^2 = \theta \bar{\sigma}_k + (1 - \theta) \sigma_0,$$

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

We begin our analytic derivations with a information period in which there are two periods, so $k = 2$. For this case an important condition emerges that is needed to ensure the covariance is positive.

*Condition 1* is said to hold for period $j^*$, with $1 < j^* \leq k$, if $j^*$ is the largest value of $j$ for which

$$\sigma_j > \theta \bar{\sigma}_k + (1 - \theta) \sigma_0.$$

Condition 1 is perhaps most intuitive for the case $k = 2$. (If a information period corresponded to a calendar day, then empirical study of mornings versus afternoons would yield $k = 2$.) Recall that positive covariance between two random variables, with the same unconditional mean, implies that if one random variable is below the unconditional mean, then the other random variable tends to be below the unconditional mean. Correspondingly, if one random variable is above the unconditional mean, then the other random variable tends to be above the unconditional mean. From the structure for the expectation of calendar period squared price changes it follows that $\sigma_1$ lies above the unconditional mean and $\sigma_0$ lies below the unconditional mean. Let $(\Delta P_{t-1})^2$ be the first period of information period $m$. If $S_m = s_0$, then in expectation $(\Delta P_{t-1})^2$ equals $\sigma_0$ and so tends to be below the unconditional mean. Yet if $S_m = s_0$, then in expectation $(\Delta P_t)^2$ also equals $\sigma_0$ and so also tends to be below the unconditional mean. If $S_m \neq s_0$, then in expectation $(\Delta P_{t-1})^2$ equals $\sigma_1$ and so tends to be above the unconditional mean. With $S_m \neq s_0$, then in expectation $(\Delta P_t)^2$ equals $\sigma_2$ and the behavior of the covariance depends on the relative magnitude of $\sigma_2$ and the conditional mean. If Condition 1 holds, then $\sigma_2$ is larger than the unconditional mean, so if $(\Delta P_{t-1})^2$ tends to be above the unconditional mean, then $(\Delta P_t)^2$ also tends to be above the unconditional mean.
**Proposition 5:** Let \( r > 0 \). The covariance of calendar period squared price changes is

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

where the addition is wrapped at \( k \). That is, if \( j + r > k \), then replace \( j + r \) with \( j + r - k \).

If \( r < k = 2 \) and Condition 1 holds for period 2, then

\[
\text{Cov} \left( (\Delta P_{t-r})^2, (\Delta P_t)^2 \right) = \frac{k - r}{k} \left\{ \theta (1 - \theta) (\sigma_1 - \sigma_0) (\sigma_2 - \sigma_0) + \frac{\theta^2}{2} (\sigma_1 - \sigma_2)^2 \right\} \geq 0.
\]

**Proof:** See Appendix.

As for 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 if \( r < k \), we must examine each term in detail. The first term in brackets is the sum of the conditional covariances, which is positive. The second term in brackets is the sum of the covariances of the conditional means, which is generally negative. Determination of the sign of the covariance depends on the relative magnitudes of the two terms.

To shed further light on the behavior of the covariance of calendar period squared price changes, we derive analytic results for a information period with 3 periods. If \( k = 3 \), then the second term in the formula for \( \text{Cov} \left( (\Delta P_{t-r})^2, (\Delta P_t)^2 \right) \), which corresponds to the covariance of the conditional means, is identical for \( r = 1 \) and \( r = 2 \). As the conditional covariance for \( r = 1 \) exceeds the conditional covariance for \( r = 2 \),

\[
\text{Cov} \left( (\Delta P_{t-1})^2, (\Delta P_t)^2 \right) > \text{Cov} \left( (\Delta P_{t-2})^2, (\Delta P_t)^2 \right).
\]

We begin by establishing under what condition \( \text{Cov} \left( (\Delta P_{t-r})^2, (\Delta P_t)^2 \right) \) is positive.

**Proposition 6:** Let Condition 1 hold for period 3 with \( k = 3 \). For \( r < k \) the covariance of calendar period squared price changes is positive.

**Proof:** See Appendix.

While Condition 1 holds naturally for period 2, there is no such natural intuition for extending Condition 1 to hold for period 3. If there are 3 periods in a information period, then in the last period of a information period enough of the information of the traders may have been revealed that the expected squared price change for that period need not exceed the unconditional expected squared price change for a period. If Condition 1 holds only for period 2, then the decline of the covariance in calendar period squared price changes can be dramatically rapid.
More Rapid Decay of Covariance for Calendar Period Squared Price Changes

As we shall see, the heteroskedasticity in $U_i^2$ that arises from the movements in the expected bid-ask spread plays an important role in explaining the persistence puzzle. If $U_i^2$ is assumed to be homoskedastic, as in Gallant, Hsieh, and Tauchen (1991), 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. Our model breaks the persistence link because one prediction of our model is that the variance of $U_i$ is not constant. In fact, the variance of $U_i$ declines as trades occur because information is revealed and the bid-ask spread declines over time. If the variance of $U_i$ declines, then the covariance in squared price changes will eventually be less than the covariance in the number of trades. We show that even during the period in which all traders are willing to trade, the variance of $U_i$ declines so that the news arrival has a more persistent effect on the number of trades than on squared price changes.

Close study of the case in which $k = 3$ reveals much about the relative decay of the correlation in calendar period trades and squared price changes. As the variance of either quantity is constant as the lag of the correlation changes, the decay of the correlation is driven by the decay of the covariance.

Suppose Condition 1 holds for period 3 so that $\text{Cov} \left( \Delta P_{t-r}, \Delta P_t^2 \right)$ is positive for $r = 1, 2$.

**Proposition 7:** Let Condition 1 hold for period 3 with $k = 3$. The covariance, and hence the correlation, of calendar period squared price changes decays more rapidly than the covariance of calendar period trades.

**Proof:** The proportional decay rates are revealed by direct calculation from the covariances. For calendar period trades

$$\frac{\text{Cov} \left[ I_{t-1}, I_t \right] - \text{Cov} \left[ I_{t-2}, I_t \right]}{\text{Cov} \left[ I_{t-1}, I_t \right]} = \frac{1}{2},$$

so the covariance declines by fifty percent.

For calendar period squared price changes the corresponding quantity is

$$\frac{1}{2} + \frac{1}{\text{Cov} \left( \Delta P_{t-1}^2, \Delta P_t^2 \right) h} \frac{1}{3(1 - \theta)} \left[ (\sigma_1 - \sigma_0)(\sigma_2 - \sigma_3) + (\sigma_2 - \sigma_0)(\sigma_3 - \sigma_0) \right]^{\frac{3}{4}}.$$
Because Condition 1 holds for period 3, the covariance between \((\Delta P_{t-1})^2\) and \((\Delta P_t)^2\) is positive. By definition \(\sigma_1 > \sigma_2 > \sigma_3 > \sigma_0\), so the second term is positive and the covariance declines by more than fifty percent.

If Condition 1 holds for period 2, it is possible that the covariance of calendar period squared price changes at lag 2 is negative. As the covariance of calendar period trades is always nonnegative, such a finding further enforces the more rapid decay of the covariance of calendar period squared price changes.

**Corollary 8:** Let Condition 1 hold for period 2 with \(k = 3\). There is an open subset of parameter values for which \(\text{Cov} \ (\Delta P_{t-2})^2, (\Delta P_t)^2\) is negative.

**Proof:** We need only establish that there exist parameter values for which the covariance is negative. Consider the set \(\{\sigma_j\}_{j=1}^3 = \{20, 7, 3\}\) so \(\bar{\sigma}_3 = 10\). Let \(\bar{\sigma}_0 = 1\). From the definition of the covariance of calendar period squared price changes, \(\text{Cov} \ (\Delta P_{t-2})^2, (\Delta P_t)^2\) is negative if \(\theta > .33\). Let \(\theta = .4\), so Condition 1 holds for period 2. As Condition 1 continues to hold for period 2 for an open set of values of \(\theta\) above .4, Corollary 8 is established.

Note that for the set of parameter values that establish Corollary 8, Condition 1 does not hold for period 3, as must be the case from Proposition 7.

To provide an idea of the pattern of serial correlation that is implied by our model for general \(k\), we simulate sequences of trades and the associated price changes over a period of many information periods. We assume that a trading opportunity arises every five minutes and that information is revealed at the end of each week. We follow the New York stock exchange and allow 6.5 trading hours in a day, so there are \(\tau = 390\) trading opportunities in an information period.\(^{11}\) Our simulated sample consists of 3000 information periods, each of which has probability of news \(\theta = .4\). Given news, the probability of good news is \(\delta = .5\). To ensure that asymmetries in the model are not driving our results we set \(\gamma = \epsilon = .5\), so that the uninformed are equally likely to buy or sell. A key parameter that remains is the proportion of traders with private information. In accord with Figure 1, we set \(\alpha = .2\).\(^{12}\)

Figure 2 depicts the serial correlation in trades and squared price changes, if the calendar period measurements correspond to 30 minute intervals \((\eta = 6, k = 65)\). The calendar period

\(^{11}\)Easley, Kiefer and O’Hara (1993) also assume that a trading opportunity arises every five minutes in their study of Ashland Oil. As the maximum number of trades in the stock during the period under study is 73, they assume that a trading day is one calendar day. Our parameter values could also be interpreted as assuming that a trading day corresponded to a calendar day if we assume that information is fully revealed over the course of one day rather than one week.

\(^{12}\)Easley, Kiefer and O’Hara (1993) estimate \(\alpha = .17\).
price change is calculated with the last price associated with a trade in the calendar period. The graph captures the essence of the persistence puzzle, the serial correlation in trades is much more persistent than is the serial correlation in squared price changes. The magnitude of the serial correlation in trades is much larger and more persistent, barely declining over 9 lags. Serial correlation in squared price changes is much smaller and declines to nearly zero after 7 lags.

![Graph of autocorrelation in trades and squared price changes](image)

**Figure 2**

## 4 Conclusions

In this paper we provide an economic model that generates serial correlation in trades and serial correlation in squared price changes. Further, serial correlation in trades is more persistent than serial correlation in squared price changes. We propose that serial correlation in trades arises simply from the entry and exit of informed traders, who receive a private signal. 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.

In our model serial correlation in trades generates serial correlation in squared price changes. Given that the informed traders are trading, there is more variance in squared
price changes simply because there are more trades in a calendar period. More trades implies that the price change is the sum of more random trades, which in turn implies that the price change has a larger variance. Because there is serial correlation in trades, there is serial correlation in squared price changes. However, there is an additional effect on the serial correlation in squared price changes, the decline in the bid-ask spread. All trades are at the bid-ask spread, 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 there are more trades in a calendar period, there are most likely more trades 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. Our model thus replicates the observed empirical features of the data and explains serial correlation through the entry and exit of informed traders and the associated revelation of information in prices.

Our model has no serial correlation in the news arrival process, nor in the arrival of informed traders within an information period. Instead, the endogenous news revelation process over the information period generates a persistent information advantage for the informed, leading to differences in the number of trades on news versus no news periods. When information periods are aggregated together, serial correlation results. Because our model has 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.

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. The probability of information arriving, $\theta$, plays an intriguing role. Serial correlation is highest in markets in which news arrives infrequently, and where the arrival of news dramatically impacts trading.
References


5 Appendix

Derivation of Learning Formulae

We explicitly derive the learning formula for \( y_i \) given that trader \( i \) trades at the ask. All other learning formulae follow the same logic. From Bayes rule

\[
P(S_m = s_H|Z_{i-1}, C_i = c_A) = \frac{P(S_m = s_H|Z_{i-1})P(C_i = c_A|S_m = s_H)}{P(S_m = j|Z_{i-1})P(C_i = c_A|S_m = j)}.
\]

We must calculate \( P(C_i = c_A|S_m = j) \) for \( j = s_L, s_H, s_0 \). If the informed receive the signal \( s_H \), then the informed will trade at the ask. Further, the fraction \( \varepsilon (1 - \gamma) \) of the uninformed will also trade at the ask. Hence

\[
P(C_i = c_A|S_m = s_H) = \alpha + (1 - \alpha) \varepsilon (1 - \gamma).
\]

If the informed receive the signal \( s_L \) or do not receive a signal, then the informed will not trade at the ask. Because only the uninformed trade at the ask if \( S_m \) equals \( s_L \) or \( s_0 \), both \( P(C_i = c_A|S_m = s_L) \) and \( P(C_i = c_A|S_m = s_0) \) equal \( (1 - \alpha) \varepsilon (1 - \gamma) \).

Proof of Theorem 1

The learning formulae for \( x_i \) and \( y_i \) are nonlinear in \((x_{i-1}, y_{i-1})\) and are not recursive, which make it difficult to determine the asymptotic behavior of \( x_i \) and \( y_i \). Because the denominator of the learning formula, conditional on the decision of trader \( i \), is identical for \( x_i \), \( y_i \) and \( 1 - x_i - y_i \), the learning formulae for ratios of \( x_i \) and \( y_i \) are linear in ratios of \((x_{i-1}, y_{i-1})\) and recursive. We work with ratios \( x_i \) and \( y_i \) and begin with the case \( S_m = s_H \), for which the relevant ratios are \( \frac{\hat{x}}{\hat{y}} \) and \( \frac{1 - \hat{x} - \hat{y}}{\hat{y}} \). Consider \( \frac{\hat{x}}{\hat{y}} \). If trader \( i \) trades at the ask

\[
\frac{x_i}{y_i} = \frac{x_{i-1}}{y_{i-1}} \cdot \frac{P(C_i = c_A|S_m = s_L)}{P(C_i = c_A|S_m = s_H)}.
\]

If trader \( i \) trades at the bid, then the expression for \( \frac{\hat{x}}{\hat{y}} \) is as above with \( C_i = c_A \) replaced by \( C_i = c_B \). If trader \( i \) does not trade

\[
\frac{x_i}{y_i} = \frac{x_{i-1}}{y_{i-1}},
\]

because \( P(C_i = c_N|S_m = s_L) \) equals \( P(C_i = c_N|S_m = s_H) \). We have

\[
\ln \frac{x_i}{y_i} = \ln \frac{x_0}{y_0} + n_A \ln \frac{P(C_i = c_A|S_m = s_L)}{P(C_i = c_A|S_m = s_H)} + n_B \ln \frac{P(C_i = c_B|S_m = s_L)}{P(C_i = c_B|S_m = s_H)}
\]

28
where \( n_A \) is the number of the first \( i \) trading opportunities for which there was a trade at the ask, \( n_B \) is the number of the first \( i \) trading opportunities for which there was a trade at the bid, and \( n_N \) is the number of the first \( i \) trading opportunities for which there was no trade.

Because the trader arrival process is i.i.d.,

\[
\frac{1}{i} \ln \frac{A_i!}{y_i^{x_i}} \xrightarrow{i \to \infty} \prod_{j=s_A,s_B} P(C_i = j|S_m = s_H) \ln \frac{P(C_i = j|S_m = s_L)}{P(C_i = j|S_m = s_H)} \tag{6}
\]

as \( i \to \infty \). The right-hand side of (6), multiplied by minus one, is a measure of distance between the probability measure \( P(\cdot|S_m = s_H) \) and the probability measure \( P(\cdot|S_m = s_L) \), which is termed the entropy of \( P(\cdot|S_m = s_H) \) relative to \( P(\cdot|S_m = s_L) \) and is denoted \( H(s_H,s_L) \). By construction the entropy is nonnegative and equals zero only if the probability measures differ solely on a set with measure zero. Hence

\[
\frac{1}{i} \ln \frac{A_i!}{y_i^{x_i}} \xrightarrow{i \to \infty} -H(s_H,s_L) < 0
\]

as \( i \to \infty \), so that \( \frac{A_i}{y_i} \) behaves as \( e^{-iH(s_H,s_L)} \). Thus \( \frac{A_i}{y_i} \) converges almost surely to zero at the exponential rate \( iH(s_H,s_L) \). A similar argument shows that \( \frac{1-x_i-y_i}{y_i} \) converges almost surely to zero at the exponential rate \( iH(s_H,s_N) \).

If \( S_m = s_H \), then \( \frac{x_i}{y_i} \xrightarrow{a.s.} 0 \) and \( \frac{1-x_i-y_i}{y_i} \xrightarrow{a.s.} 0 \) as \( i \to \infty \).

If \( S_m = s_L \), then the relevant ratios are \( \frac{y_i}{x_i} \) and \( \frac{1-x_i-y_i}{x_i} \). If \( S_m = s_0 \), then the relevant ratios are \( \frac{x_i}{1-x_i-y_i} \) and \( \frac{y_i}{1-x_i-y_i} \). Again the fact that the trader arrival process is i.i.d. is sufficient to establish that

- if \( S_m = s_L \), then \( \frac{y_i}{x_i} \xrightarrow{a.s.} 0 \) and \( \frac{1-x_i-y_i}{x_i} \xrightarrow{a.s.} 0 \),
- if \( S_m = s_0 \), then \( \frac{x_i}{1-x_i-y_i} \xrightarrow{a.s.} 0 \) and \( \frac{y_i}{1-x_i-y_i} \xrightarrow{a.s.} 0 \),

as \( i \to \infty \).

From the convergence properties of the ratios, we can easily deduce the convergence properties of \( x_i \) and \( y_i \). We continue with the case \( S_m = s_H \) and note that similar arguments hold for \( S_m = s_L \) and \( S_m = s_0 \). The statement \( \frac{1-x_i-y_i}{y_i} \xrightarrow{a.s.} 0 \) is equivalently written as

\[
\frac{1}{y_i} - \frac{x_i}{y_i} \xrightarrow{a.s.} 0.
\]

Because \( \frac{x_i}{y_i} \xrightarrow{a.s.} 0 \), the statement (7) is equivalent to

\[
\frac{1}{y_i} \xrightarrow{a.s.} 0.
\]
which directly implies \( y_i \overset{a.s.}{\to} 1 \). If \( y_i \overset{a.s.}{\to} 1 \), then the statement \( x_i \overset{a.s.}{\to} 0 \) implies \( x_i \overset{a.s.}{\to} 0 \). From the definition of \( A_i \) and \( B_i \), if \( x_i \overset{a.s.}{\to} 0 \) and \( y_i \overset{a.s.}{\to} 1 \), then \( A_i \overset{a.s.}{\to} v_H \) and \( B_i \overset{a.s.}{\to} v_H \).

Proof of Theorem 2

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

\[
\text{Cov}(I_{t-r}, I_t) = \text{E}(I_{t-r}I_t) - \text{E}I_{t-r} \cdot \text{E}I_t.
\]

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

If \( r < k \), then there are three possible conditional expectations of \( I_{t-r}I_t \). First, if \( I_{t-r} \) 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 and \( S_{m+1} \neq s_0 \), the conditional expectation of \( I_{t-r}I_t \) is

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

which occurs with probability \( \frac{k}{k} \theta \). Third, if \( I_{t-r} \) and \( I_t \) are measured on consecutive information periods and \( S_{m+1} = s_0 \), the conditional expectation of \( I_{t-r}I_t \) is

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

which occurs with probability \( \frac{k}{k} (1 - \theta) \). We combine the three conditional expectations to yield

\[
\text{E}(I_{t-r}I_t) = \frac{k-r}{k} \theta \mu_1^2 + (1 - \theta) \mu_0^2 + \frac{r}{k} \left[ \theta \mu_1 + (1 - \theta) \mu_0 \right]^2.
\]

Because the process for calendar period trades is stationary, \( \text{E}I_{t-r} = \text{E}I_t \). As noted in the text

\[
\text{E}I_t = \theta \mu_1 + (1 - \theta) \mu_0,
\]

so

\[
\text{Cov}(I_{t-r}, I_t) = \frac{k-r}{k} \theta (1 - \theta) (\mu_1 - \mu_0)^2 = \frac{k-r}{k} \theta (1 - \theta) (\alpha \eta)^2.
\]
Combining the two possible cases for \( r \) relative to \( k \) yields
\[
\text{Cov}(I_{l-r}, I_0) = \begin{cases} 
\theta(1 - \theta) (\alpha \eta)^2 \frac{h_i}{k} & r < k \\
0 & r \geq k 
\end{cases}
\] (8)

Combining the covariance and variance of \( I_t \) given by (1) gives the desired correlation. Because all terms are positive for \( r < k \), the correlation is positive.

**Proof of Theorem 4**

For the proof of Theorem 4, let \( C_N \) represent \( C_i = c_N \) in the conditioning information set. We have
\[
E(U_i|Z_{i-1}) = P(C_i = c_A)(A_i - E(V_m|Z_{i-1})) + P(C_i = c_B)(B_i - E(V_m|Z_{i-1})) \\
+ P(C_i = c_N)(E[V_m|Z_{i-1}, C_N] - E(V_m|Z_{i-1}))
\]
\[
= P(C_i = c_A)A_i + P(C_i = c_B)B_i + P(C_i = c_N)E[V_m|Z_{i-1}, C_N] - E(V_m|Z_{i-1})
\]
In similar fashion we find that \( U_i \) is a serially uncorrelated random variable. Let \( h \) and \( i \) be distinct values with \( h < i \),
\[
E(U_h U_i|Z_{i-1}) = U_h [E(V_m|Z_{i-1}) - E(V_m|Z_{i-1})] = 0.
\]
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
\[
E(U_i^2|Z_{i-1}) \leq P(C_i = c_A)(\tilde{A}_i - E(V_m|Z_{i-1}))^2 + P(C_i = c_B)(\tilde{B}_i - E(V_m|Z_{i-1}))^2 \\
+ P(C_i = c_N)(E[V_m|Z_{i-1}, C_N] - E(V_m|Z_{i-1}))^2
\]
\[
\leq [P(C_i = c_A) + P(C_i = c_N)] (\tilde{A}_i - E(V_m|Z_{i-1}))^2 \\
+ [P(C_i = c_B) + P(C_i = c_N)] (\tilde{B}_i - E(V_m|Z_{i-1}))^2
\]
\[
\leq (\tilde{A}_i - E(V_m|Z_{i-1}))^2 + (\tilde{B}_i - E(V_m|Z_{i-1}))^2
\]
\[
= h_{\tilde{A}_i - \tilde{B}_i}^2,
\]
where the first inequality follows from the definition of \( \tilde{A}_i \) and \( \tilde{B}_i \) and the fourth inequality follows from \( \tilde{B}_i \leq E[V_m|Z_{i-1}] \leq A_i \). Note that the unconditional variance is immediately obtained from Jensen’s inequality
\[
EU_i^2 \leq E^3 \tilde{A}_i - \tilde{B}_i^2 \leq E\tilde{A}_i - E\tilde{B}_i^2.
\]

31
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:

\[
\left| A_i - E(V_m|Z_{i-1}) \right|, \left| B_i - E(V_m|Z_{i-1}) \right| \text{ and } \left| E[V_m|Z_{i-1}, C_N] - E(V_m|Z_{i-1}) \right|.
\]

Let \( P_j = P(C_i = c_j) \). If \( \min T_i = |A_i - E(V_m|Z_{i-1})| \), then

\[
\mathbb{E} U_i^2 | Z_{i-1} \geq (P_A + P_N)(A_i - E(V_m|Z_{i-1}))^2 + P_B(B_i - E(V_m|Z_{i-1}))^2 \\
\geq P_B(P_A + P_N) \hat{A}_i - \hat{B}_i^2,
\]

where the second inequality follows from Lemma 4.1, which is proven below. If \( \min T_i = |B_i - E(V_m|Z_{i-1})| \), then

\[
\mathbb{E} U_i^2 | Z_{i-1} \geq P_A(A_i - E(V_m|Z_{i-1}))^2 + (P_B + P_N)(B_i - E(V_m|Z_{i-1}))^2 \\
\geq P_A(P_B + P_N) \hat{A}_i - \hat{B}_i^2,
\]

where the second inequality follows from Lemma 4.1. If \( \min T_i = |E[V_m|Z_{i-1}, C_N] - E(V_m|Z_{i-1})| \), then

\[
\mathbb{E} U_i^2 | Z_{i-1} \geq P_A(A_i - E(V_m|Z_{i-1}))^2 + (P_B + P_N)(E[V_m|Z_{i-1}, C_N] - E(V_m|Z_{i-1}))^2 \\
\geq P_A(P_B + P_N) \hat{A}_i - \hat{B}_i^2,
\]

where the second inequality follows from Lemma 4.1.

The unconditional variance thus satisfies:

\[
\min \{P_A(P_B + P_N), P_B(P_A + P_N)\} \mathbb{E} \hat{A}_i - \hat{B}_i^2 \leq \mathbb{E} U_i^2.
\]

**Lemma 4.1** Let \( c \in [0, 1] \). For any pair of real numbers \( a \) and \( b \)

\[
c(1-c)(a+b)^2 \leq ca^2 + (1-c)b^2.
\]

**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

\[
0 \leq c^2a^2 + (1-c)^2b^2 - 2c(1-c)ab = [ca - (1-c)b]^2.
\]
Proof of Proposition 5

We first carefully derive \( \text{Cov} (\Delta P_{t-1})^2, (\Delta P_t)^2 \) for \( k = 2 \). Let \( N = 1 \) if \( t - 1 \) is the first period of an information period and let \( N = 2 \) if \( t - 1 \) is the second period. Then

\[
E_h (\Delta P_{t-1})^2 | N = 1 \quad \text{if} 
\]

\[
E_h (\Delta P_{t-1})^2 | N = 2 \quad \text{if} 
\]

\[
E (\Delta P_t)^2 = \theta \sigma_2^2 + \frac{\sigma_2^2}{2} + (1 - \theta) \sigma_0.
\]

Because \( N \) is equally likely to take the values 1 or 2, the conditional covariance is

\[
\frac{1}{2} \left\{ E_h (\Delta P_{t-1})^2 (\Delta P_t)^2 | N = 1 \right\} i
\]

\[
+ \frac{1}{2} \left\{ E_h (\Delta P_{t-1})^2 (\Delta P_t)^2 | N = 2 \right\} i
\]  

\[
\text{(4.1)}
\]

From the formulae for the expected calendar period squared price change given the value of \( N \), (4.1) equals

\[
\frac{1}{2} \left\{ (\theta \sigma_1 \sigma_2 + (1 - \theta) \sigma_0^2) - (\theta \sigma_1 + (1 - \theta) \sigma_0) (\theta \sigma_2 + (1 - \theta) \sigma_0) + \theta (\sigma_2 \sigma_1 + (1 - \theta) \sigma_2) \right\}
\]

\[
+ (1 - \theta) \left[ \theta \sigma_0 \sigma_1 + (1 - \theta) \sigma_0^2 \right] - (\theta \sigma_2 + (1 - \theta) \sigma_0) (\theta \sigma_1 + (1 - \theta) \sigma_0)
\]

which is simplified as

\[
\frac{1}{2} \theta (1 - \theta) (\sigma_1 - \sigma_0) (\sigma_2 - \sigma_0).
\]

The covariance of the conditional means is

\[
E \{ (\Delta P_{t-1})^2 - E_h (\Delta P_{t-1})^2 | N = 1 \quad E (\Delta P_t)^2 - E (\Delta P_t)^2 | N = 1 \}
\]

which equals

\[
P (N = 1) \{ E (\Delta P_{t-1})^2 - E_h (\Delta P_{t-1})^2 | N = 1 \quad E (\Delta P_t)^2 - E (\Delta P_t)^2 | N = 1 \}
\]

\[
+ P (N = 2) \{ E (\Delta P_{t-1})^2 - E_h (\Delta P_{t-1})^2 | N = 2 \quad E (\Delta P_t)^2 - E (\Delta P_t)^2 | N = 2 \}
\]

Note

\[
E (\Delta P_{t-1})^2 - E_h (\Delta P_{t-1})^2 | N = 1 = \theta (\sigma_2 - \sigma_0),
\]

\[
E (\Delta P_t)^2 - E (\Delta P_t)^2 | N = 1 = \theta (\sigma_2 - \sigma_0),
\]

\[
E (\Delta P_{t-1})^2 - E_h (\Delta P_{t-1})^2 | N = 2 = \theta (\sigma_2 - \sigma_0),
\]

\[
E (\Delta P_t)^2 - E (\Delta P_t)^2 | N = 2 = \theta (\sigma_2 - \sigma_0).
\]

Thus, the covariance of the conditional means is
From (A4.2) and (A4.3) the Cov \((\Delta P_{t-1})^2, (\Delta P_t)^2\) equals
\[
\frac{1}{2}\theta (1 - \theta) (\sigma_1 - \sigma_0) (\sigma_2 - \sigma_0) + \theta^2 \frac{\sigma_1}{2} - \frac{\sigma_2}{2} \cdot \left( \frac{3}{2} - \frac{\sigma_1}{2} - \frac{\sigma_2}{2} \right).
\]

Derivation of the general covariance expression follows identical logic. Because \(\sigma_1 > \sigma_2\), the second term of the covariance is negative (while the first term is positive) and the covariance is positive if
\[
(1 - \theta) (\sigma_1 - \sigma_0) (\sigma_2 - \sigma_0) > \frac{\theta}{2} (\sigma_1 - \sigma_2)^2.
\]

First, by inspection
\[
(\sigma_1 - \sigma_0) > (\sigma_1 - \sigma_2).
\]

Thus to verify (A4.4), we need only show
\[
(1 - \theta) (\sigma_2 - \sigma_0) > \frac{\theta}{2} (\sigma_1 - \sigma_2).
\]

Because \(\frac{\theta}{2} (\sigma_1 - \sigma_2) = \theta (\bar{\sigma}_2 - \sigma_2)\), to verify the preceding inequality, we must show
\[
(1 - \theta) (\sigma_2 - \sigma_0) - \theta (\bar{\sigma}_2 - \sigma_2) > 0.
\]

Condition 1 implies
\[
(1 - \theta) (\sigma_2 - \sigma_0) > \theta (1 - \theta) (\bar{\sigma}_2 - \sigma_0).
\]

Hence
\[
(1 - \theta) (\sigma_2 - \sigma_0) - \theta (\bar{\sigma}_2 - \sigma_2) > \theta (1 - \theta) (\sigma_2 - \sigma_0) - \theta (\bar{\sigma}_2 - \sigma_2).
\]

The right-hand side of the preceding inequality equals
\[
\theta [(\sigma_2 - \sigma_0) - \theta (\bar{\sigma}_2 - \sigma_0)],
\]
and Condition 1 implies
\[
(\sigma_2 - \sigma_0) - \theta (\bar{\sigma}_2 - \sigma_0) > 0.
\]

**Proof of Proposition 6**

If \(k = 3\), then the covariance of calendar period squared price changes is larger for \(r = 1\) than for \(r = 2\), so Proposition 6 is established if \(\text{Cov} \ (\Delta P_{t-2})^2, (\Delta P_t)^2\) is positive. We have \(\text{Cov} \ (\Delta P_{t-2})^2, (\Delta P_t)^2\) equals.
\[
\frac{1}{3} \theta \{ (1 - \theta) (\sigma_1 - \sigma_0) (\sigma_3 - \sigma_0) + \theta (\sigma_3 - \sigma_1) (\sigma_3 - \sigma_2) + \theta (\sigma_3 - \sigma_2) (\sigma_3 - \sigma_1) \}.
\]

The first term is positive, the second negative, and the remaining two terms are opposite in sign and depend on the sign of \((\bar{\sigma}_3 - \sigma_2)\). We consider each of the three cases: \((\bar{\sigma}_3 - \sigma_2) > 0\), \((\bar{\sigma}_3 - \sigma_2) < 0\), and \((\bar{\sigma}_3 - \sigma_2) = 0\) in turn.

**Case 1: \((\bar{\sigma}_3 - \sigma_2) > 0\)**

If \((\bar{\sigma}_3 - \sigma_2) > 0\), then \(\sigma_1 > \bar{\sigma}_3 > \sigma_2 \Rightarrow \bar{\sigma}_3 > \sigma_0\). Define \(d_1 = \sigma_1 - \bar{\sigma}_3\), \(d_2 = \bar{\sigma}_3 - \sigma_2\), \(d_3 = \sigma_2 - \bar{\sigma}_3\), and \(d_4 = \bar{\sigma}_3 - \sigma_0\). The \(\text{Cov} \ (\Delta P_{t-2})^2, (\Delta P_1)^2\) is positive if

\[
(1 - \theta) (\sigma_1 - \sigma_0) (\sigma_3 - \sigma_0) + \theta (\bar{\sigma}_3 - \sigma_3) (\bar{\sigma}_3 - \sigma_2) > |\theta (\bar{\sigma}_3 - \sigma_1) (\bar{\sigma}_3 - \sigma_3) + \theta (\bar{\sigma}_3 - \sigma_2) (\bar{\sigma}_3 - \sigma_1)|,
\]

which is equivalently expressed as

\[
(1 - \theta) \prod_{j=1}^{4} d_j d_4 + \theta (d_2 + d_3) d_2 > \theta d_1 (2d_2 + d_3).
\]

(A5.1)

Rewrite (A5.1) as

\[
d_1 (1 - \theta) d_4 + (1 - \theta) \prod_{j=2}^{4} d_j d_4 + \theta (d_2 + d_3) d_2 > \theta d_1 (2d_2 + d_3) + \theta d_1 d_2.
\]

If Condition 1 holds for period 3, then

\[
(1 - \theta) d_4 > \theta (d_2 + d_3),
\]

and (A5.1) is satisfied if

\[
d_2 (1 - \theta) d_4 + \theta d_2 (d_2 + d_3) + (1 - \theta) \prod_{j=3}^{4} d_j d_4 > \theta d_1 d_2.
\]

(A5.2)

If Condition 1 holds for period 3, then

\[
(1 - \theta) d_4 > \theta d_2,
\]

and (A5.2) is satisfied if

\[
\theta d_2 (2d_2 + d_3) - \theta d_1 d_2 = \theta d_2 (2d_2 + d_3 - d_1) = 0
\]

From the definition of \(\bar{\sigma}_3\), \(\prod_{j=1}^{3} (\sigma_j - \bar{\sigma}_3) = d_1 - 2d_2 - d_3\), so

\[
(2d_2 + d_3 - d_1) = 0.
\]

**Case 2: \((\bar{\sigma}_3 - \sigma_2) < 0\)**

35
If \((\bar{\sigma}_3 - \sigma_2) < 0\), then \(\sigma_1 > \sigma_2 > \bar{\sigma}_3 \bar{\sigma}_3 > \sigma_0\). Define \(d_1 = \sigma_1 - \sigma_2\), \(d_2 = \sigma_2 - \bar{\sigma}_3\), \(d_3 = \bar{\sigma}_3 - \bar{\sigma}_3\), and \(d_4 = \sigma_3 - \sigma_0\). The \(\text{Cov} \ (\Delta P_{t-2})^2, (\Delta P_t)^2\) is positive if

\[
(1 - \theta) (\sigma_1 - \sigma_0) (\sigma_3 - \sigma_0) + \theta (\bar{\sigma}_3 - \sigma_1) (\bar{\sigma}_3 - \sigma_2) > \theta (\bar{\sigma}_3 - \sigma_1) (\bar{\sigma}_3 - \sigma_3) + \theta (\bar{\sigma}_3 - \sigma_2) (\bar{\sigma}_3 - \sigma_3),
\]

which is equivalently expressed as

\[
(1 - \theta) \prod_{j=1}^4 d_j \quad d_4 + \theta (d_1 + d_2) d_2 > \theta d_3 (2d_2 + d_1).
\] (A5.3)

From the definition of \(\bar{\sigma}_3\),

\[
2d_2 + d_1 = d_3,
\]

so (A5.3) is satisfied if

\[
(1 - \theta) d_3 d_4 - \theta d_3^2 > 0.
\]

Note \((1 - \theta) d_3 d_4 - \theta d_3^2 = d_3 (d_4 - \theta (d_3 + d_4))\). If Condition 1 holds for period 3

\[
\sigma_3 - \sigma_0 > \theta (\bar{\sigma}_3 - \sigma_0),
\]

which is equivalently expressed as

\[
d_4 > \theta (d_3 + d_4).
\]

**Case 3: \((\bar{\sigma}_3 - \sigma_2) = 0\)**

The \(\text{Cov} \ (\Delta P_{t-2})^2, (\Delta P_t)^2\) is positive if

\[
(1 - \theta) (\sigma_1 - \sigma_0) (\sigma_3 - \sigma_0) > \theta (\bar{\sigma}_3 - \sigma_1) (\bar{\sigma}_3 - \sigma_3).\]

First, by inspection

\[
(\sigma_1 - \sigma_0) > (\sigma_1 - \bar{\sigma}_3).
\]

Second, if Condition 1 holds for period 3

\[
(1 - \theta) (\sigma_3 - \sigma_0) > \theta (\bar{\sigma}_3 - \sigma_3).
\]