Option Market Microstructure and Stochastic Volatility*

Douglas G. Steigerwald and Richard J. Vagnoni
Department of Economics
University of California
Santa Barbara, CA 93106

July 27, 2001

Abstract

Our primary goal is to develop and analyze a dynamic economic model that takes into account several sources of information-based trade—the markets for a stock and options on that stock—and that ultimately accounts for salient features of stock price data, including serial correlation in stock trades, serial correlation in squared stock price changes (stochastic volatility), and more persistent serial correlation in stock trades than in squared stock price changes. We derive the dynamic relationships among the stock, the call option, and the put option and capture the leverage effect offered by options. We derive standard consistency results for the learning process and the convergence of an asset's quotes to the asset's true value. We also derive closed-form analytic results for expected calendar period price changes and trades, and we examine calendar period serial correlation properties of squared price changes and trades.

1. Introduction

Much recent attention has focused on modeling high-frequency stock price behavior. On the theoretical side, the blossoming area of market microstructure

*We thank Steve LeRoy and John Owens for helpful comments.

is providing valuable insights into the trade-by-trade stock price process.[1] On the empirical side, a wealth of research has focused on capturing salient features of calendar period stock data, including conditional heteroskedasticity in calendar period price changes, or stochastic volatility. We provide a theory-based link among asymmetric information, the behavior of market participants, and stochastic volatility through a market microstructure model of securities markets.

Our work follows on from Kelly and Steigerwald (2000), in which the stochastic properties of calendar period trades and squared price changes are derived from a market microstructure model. In the current paper, we make two principle contributions. First, we consider a model in which trade occurs in an option market as well as the stock market. Working from the microstructure model of Easley, O'Hara and Srinivas (1998) we derive the dynamic pattern of trade across markets as well as the stochastic properties of trades and squared price changes for each market. Second, we obtain analytic expressions for the serial correlation in calendar period squared price changes and so can directly relate stochastic volatility to the parameters of the underlying model.

Market microstructure, broadly speaking, is the area of economics that deals with the evolution of prices by focusing on the actual trading process [Demsetz (1968), Garman (1976), Amihud and Mendelson (1980), Kyle (1985), Glosten and Milgrom (1985), Easley and O'Hara (1987), Easley and O'Hara (1992), and Harris and Raviv (1993), among many others]. While market microstructure models come in a wide variety of styles—from inventory- and information-based models, to batch-order and sequential-trade models, to strategic behavior and game-theoretic models—we focus on information-based, sequential-trade models [Glosten and Milgrom and Easley and O'Hara (1992)]. In particular, information-based, sequential-trade market microstructure models capture the link between asset prices and informational asymmetries among traders and model the bid-ask spread as an adverse selection problem. The *raison d'être* of these models rests in the assertion that trades in a stock are correlated with private information regarding the value of that stock.

In modeling the market microstructure of a stock market, Glosten and Milgrom allow for fully informed traders, uninformed (liquidity) traders, and a market maker, all of whom are risk neutral and competitive. The market consists of a single stock and the market maker is the asset dealer. The risk neutrality of the market maker eliminates inventory effects. Both the informed and uninformed trade with the market maker and are chosen to transact randomly. The informed

[1] Cohen *et al* (1986) and O'Hara (1997) are notable references.

receive a signal indicating the true value of the asset, but the uninformed and the market maker do not—thus, there is information asymmetry. The information asymmetry poses an adverse selection problem for the market maker and is the reason for the bid-ask spread. Glosten and Milgrom show that the bid-ask spread bounds the expected value of the asset via a Bayesian learning process. Easley and O'Hara (1992) expand the basic model to take into account "event uncertainty" by allowing for the possibility that there is no signal regarding the true value of the asset—thus in some periods all traders are uninformed. This event uncertainty gives rise to the importance of time in the stock price process: if an information event is not a certainty, then a given trader's decision to trade or not to trade may provide information to the market. They examine the time process of the stock price in this less restrictive framework and show that the sequence of bid and ask prices converges to the true value of the asset for a given information event.

We posit that private information is a driving factor in the stock price process, as suggested by French and Roll (1986). It seems reasonable, therefore, that when modeling the stock price the sources of information-based trade related to the stock should be considered. Heretofore most market microstructure models considered only the market for a stock—but the stock market is not the only medium for information-based trade in the stock. Derivative instruments based on the underlying stock—such as call and put options—may provide another vehicle through which informed traders can profit from their information. The introduction of the options market is a natural extension of asymmetric information market microstructure models that allows for further sources of information-based trade, recognizing the important insight of Black (1975) that the options market may provide a better venue for informed traders and the fact that, when investigating insider trading cases, the Securities and Exchange Commission carefully examines trades in options. Easley, O'Hara, and Srinivas (1998) expand the standard market microstructure model to include both a stock market and an options market. Their sequential-trade, asymmetric information model of the stock and options markets demonstrates that both markets can host information-based trade. The model is analogous to that of Easley and O'Hara (1987), a sequential trade, asymmetric information model of the stock market that allows for multiple trade sizes. In Easley, O'Hara, and Srinivas, rather than choosing between stock trade sizes, traders choose between transacting in the stock and options markets. Via equilibrium arguments, they show that informed traders will transact in the options market under certain conditions regarding the depths of the stock and options markets as well as the available leverage offered by the options relative to the stock.

Easley, O'Hara, and Srinivas note that the inclusion of sources other than the stock market of information-based trade can blur the linkages between stock market transactions and information. They also point out the seemingly paradoxical implication that information-based transactions in derivatives on an underlying stock may have on the stock price itself [see Biais and Hillion (1994) and Back (1993)]. The prices of derivative instruments such as options are supposed to be determined unilaterally by the underlying stock price [for example, see Black and Scholes (1975)]; if, however, these derivatives are not redundant assets then this relation does not necessarily hold. Given that the trading process itself contains information about the stock price, transactions in derivatives on the underlying stock can have information for the stock price and thus affect the stock price. In particular, if derivatives markets are more attractive to informed traders, then the trading process in these derivatives may contain new information about the stock price before it is reflected in the stock or derivatives prices. Hence, these derivatives are not redundant assets.

Empirically, calendar period stock data tend to exhibit numerous regularities. There is strong evidence of serial correlation in calendar period squared price changes and in the number of trades across calendar periods, and the serial correlation in the number of trades tends to be more persistent than serial correlation in squared price changes [Harris (1987), Andersen (1996) and Steigerwald (1997)]. For example, we consider the behavior of the stock of Alcoa from January 2, 1996 through June 28, 1996. Using trade-by-trade data from the New York Stock Exchange Trades and Quotes database, we aggregate the price and trade data to half hour intervals, resulting in 1625 caldnear period observations. Because we observe cyclic patterns in the intraday data, on average the number of trades tends to be greater at the open and close of the market while prices are more volatile near the open, we remove the half-hourly means. The autocorrelation functions of the demeaned calendar period data in Figure 1.1 reveal that both squared price changes and trades are serially correllated, with trade correlations much higher and flatter than squared price change correlations.

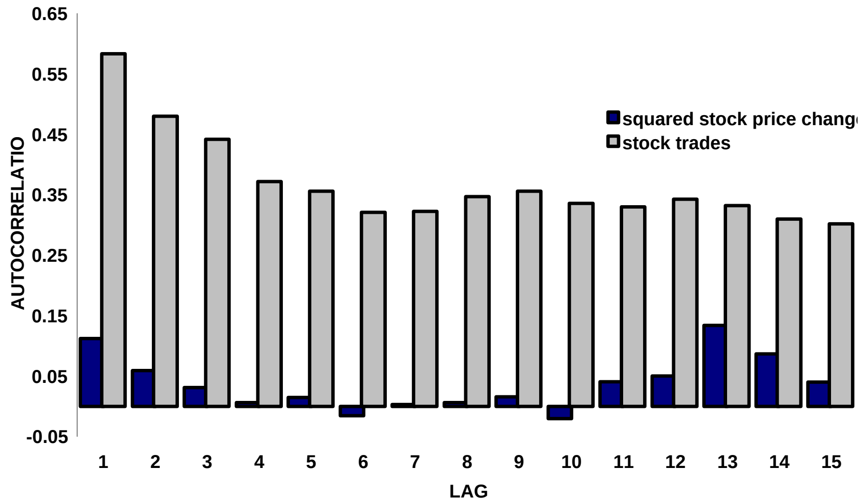

Figure 1.1: Autocorrelations in the number of trades and squared price changes for Alcoa stock using trade-by-trade data from January 2, 1996 through June 28, 1996. The data are sampled every half-hour from 9:30 AM until 4:00 PM for a total of 1625 observations.

The aforementioned empirical regularities have important implications for economic modeling of the stock price process. One very popular—and successful—avenue for capturing serial correlation in squared price changes has been through the employment of sundry conditional heteroskedasticity models [see Bollerslev, Engle, and Nelson (1993) for a survey]. Such models are purely statistical, however, and are not based on economic theory. Steigerwald provides a step toward linking economic theory and conditional heteroskedasticity models through the formulation and analysis of mixture models [Clark (1973) and Gallant, Hsieh, and Tauchen (1991)]. Economic theory suggests that the evolution of a stock price is not necessarily concordant with calendar time but rather is based on random transactions between buyers and sellers. These transactions can be driven by the random arrival of information, independent of calendar time. In this way, the information contained in a price can vary from one calendar period to the next, as can the number of transactions between buyers and sellers.

Easley, O'Hara, and Srinivas develop an information-based, sequential-trade market microstructure model of both the stock and options markets, but they do not allow for no-trade intervals and do not attempt to capture the important role of time in the price process. We generalize the model of Easley, O'Hara, and Srinivas, allowing for no-trade intervals, and consider the dynamic properties of the price and trade processes over calendar period intervals. We derive the dynamic relationships among the stock, the call option, and the put option and capture the leverage effect offered by options. We derive consistency results for the learning process and the convergence of an asset's quotes to the asset's true value. We also derive closed-form analytic results for expected calendar period price changes and trades, and we examine calendar period serial correlation properties of squared price changes and trades. While we model the leverage effect of options, we do not address the timing difference between option and stock ownership.

2. A Market Microstructure Model of the Stock and Options Markets

We consider a model with markets for a stock and for call and put options on the stock. We base our dual-market, sequential-trade, asymmetric information model on the market microstructure models of Easley and O'Hara (1992) and Easley, O'Hara, and Srinivas (1998). There is a market maker in the stock market and a market maker in the options market, and there are an infinite number of traders who can trade in either market. The market makers and the traders are assumed to be risk-neutral.

The model consists of pure dealership markets in which orders are solely of the market type—thus the market makers neither keep order books nor provide brokerage services. The market maker in the stock market sets an ask and a bid, collectively termed the quotes, for one share of the stock. These are the prices at which the market maker is willing to sell one share of the stock and buy one share of the stock, respectively. Analogously, the market maker in the options market sets quotes for the call and put option contracts. Each option contract grants rights to a fixed number of shares of the underlying stock, λ, with $\lambda \geq 1$. Thus, each transaction in the options market involves either buying or writing λ options.

Trade in the stock and options markets occurs over a sequence of trading days, indexed by m. On trading day m, the stock realizes some per share dollar value, given by the random variable V_m. The stock can take one of two known values, $V_m \in \{v_{L_\mathsf{m}}, v_{H_\mathsf{m}}\}$, with $v_{L_\mathsf{m}} < v_{H_\mathsf{m}}$. The stock takes the lower value, v_{L_m}, with positive probability δ. Prior to the commencement of trading on day m, informed traders receive a randomly determined signal, S_m, about the value of the stock on m. This signal is meant to capture private information and can take one of three values, $S_m \in \{s_L, s_H, s_O\}$. The informative signals, s_L and s_H, reveal the true value of the stock. The low signal, s_L, indicates bad news such that $V_m = v_{L_\mathsf{m}}$, the high signal, s_H, indicates good news such that $V_m = v_{H_\mathsf{m}}$, but the uninformative signal, s_O, provides no information regarding the true value of the stock. The probability that the informed traders learn the true value of a share of the stock through the signal is θ, with $\theta > 0$, so the probability that they receive the high signal is $\theta(1-\delta)$. Proportion α of the traders receives the signal, characterizing the informed universe of traders. The proportion of traders that does not receive the signal characterizes the uninformed universe of traders. Neither market maker is privy to the signal. At the end of each trading day, the signal is revealed to the market makers and uninformed traders and, hence, all agree on the value of a share of the stock.[2] The general trading day construct is designed to capture the interval over which asymmetric information due to a particular signal persists in the markets. This interval is not necessarily coincident with a calendar day.

Each option is of the European type—precluding the possibility of exercise prior to the end of the trading day—and expires upon revelation of the signal. Each call option provides the owner with the right to buy one share of the stock for a specified strike price, κ_{C_m}, with $\kappa_{C_\mathsf{m}} \in [v_{L_\mathsf{m}}, v_{H_\mathsf{m}}]$, from the call option writer at the end of the trading day. The value of the call option, V_{C_m}, is $\max(V_m - \kappa_{C_\mathsf{m}}, 0)$. If $S_m = s_L$, the call option is "out-of-the-money" and expires worthless. If $S_m = s_H$, the call option is "in-the-money" and expires worth $v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}$. If the signal is uninformative, however, $V_{C_\mathsf{m}} = \max(EV_m - \kappa_{C_\mathsf{m}}, 0)$. Each put option provides the owner with the right to sell one share of the stock for a specified strike price, κ_{P_m}, with $\kappa_{P_\mathsf{m}} \in [v_{L_\mathsf{m}}, v_{H_\mathsf{m}}]$, to the put option writer at the end of the trading day. The value of the put option, V_{P_m}, is $\max(\kappa_{P_\mathsf{m}} - V_m, 0)$. If $S_m = s_L$, the put option is in-the-money and expires worth $\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}$. If $S_m = s_H$, the put option is out-of-the-money and expires worthless. If the signal is uninformative, however, $V_{P_\mathsf{m}} = \max(\kappa_{P_\mathsf{m}} - EV_m, 0)$.

Traders randomly arrive to the markets one at a time, so we index them by their order of arrival, i.[3] Let ω_i denote the rate of time discount for consumption at the end of a trading day for the ith trader. We define W_{i_m} as a random variable representing the value of trader i's investment at the end of trading day m. For example, if the ith trader buys λ call option contracts, then $W_{i_\mathsf{m}} = \lambda V_{C_\mathsf{m}}$.

[2] To ensure the continuity of prices over trading days, $EV_\mathsf{m} = v_{\mathsf{m}-1}$ if the informed traders received an informative signal on trading day $m-1$. If the informed traders received an uninformative signal on trading day $m-1$, then we presume the possible share values are unchanged.

[3] Due to the infinite number of traders, with probability one no trader will arrive to the markets more than once.

Each market participant assigns random utility to his investment and current consumption, c, as $\omega W + c$. The larger the value of ω, the greater is the desire to forego current consumption.

We set $\omega = 1$ for the market makers and informed traders. Conditional on receiving $S_m = s_O$, informed traders do not trade because of identical preferences. Conditional on receiving an informative signal, informed traders trade as long as the market makers are uncertain of the true value of the stock. If $S_m = s_L$, then an informed trader implements one of three possible "bearish" strategies, selling short one share of the stock with probability ϵ_{IB}, writing λ call options with probability ϵ_{IBC}, or buying λ put options with probability $\epsilon_{IAP} = 1 - \epsilon_{IB} - \epsilon_{IBC}$. If $S_m = s_H$, then an informed trader implements one of three possible "bullish" strategies, buying one share of the stock with probability ϵ_{IA}, buying λ call options with probability ϵ_{IAC}, or writing λ put options with probability $\epsilon_{IBP} = 1 - \epsilon_{IA} - \epsilon_{IAC}$. Conditional on receiving an informative signal, the informed trader employs the strategy that provides the largest net gain.

Given currently available public information, a trader who buys an asset at the market maker's ask pays more than the expected value of the asset, and a trader who sells an asset at the market maker's bid receives less than the expected value of the asset. Thus, the trader who buys the stock if $V_m = v_{L_\mathsf{m}}$ or sells short the stock if $V_m = v_{H_\mathsf{m}}$ loses on the trade. Given this knowledge, the rational uninformed are assumed to trade for liquidity reasons and not speculation. To induce them to trade, we let ω characterize three types of uninformed trader. Of the uninformed traders with $\omega = 0$, proportion ϵ_{UB} potentially sells the stock short, proportion ϵ_{UBC} potentially writes λ call options, and proportion ϵ_{UBP} potentially writes λ put options. Proportion $1 - \epsilon$ of the uninformed traders has $\omega = 1$ and never makes any trade. Of the uninformed traders with $\omega = \infty$, proportion ϵ_{UA} potentially buys the stock, proportion ϵ_{UAC} potentially buys λ call options, and proportion ϵ_{UAP} potentially buys λ put options.

The ith trader arrives, observes the quotes, and makes a trade decision, D_i. The random variable, D_i, takes one of seven values. If trader i buys the stock at the ask, A_i, then $D_i = d_A$. If trader i sells the stock short at the bid, B_i, then $D_i = d_B$. If trader i buys λ call options at the ask, A_{C_i}, then $D_i = d_{AC}$. If trader i writes λ call options at the bid, B_{C_i}, then $D_i = d_{BC}$. If trader i buys λ put options at the ask, A_{P_i}, then $D_i = d_{AP}$. If trader i writes λ put options at the bid, B_{P_i}, then $D_i = d_{BP}$. If trader i elects not to trade, then $D_i = d_N$. We define the sequence of trading decsions on m as $\{D_k\}_{k=1}^{i}$. Given all publicly avalable information prior to the commencement of trade on m, Z_0, we specify the publicly available information set prior to the arrival of trader $i+1$ as Z_i, with $Z_i = \{Z_0, D_k\}_{k=1}^{i}$.

The information set, Z_i, is shared by the market makers and all traders. We assume that the market makers and uninformed traders also have the same Bayesian updating process by which they learn the signal received by the informed. We refer to the learning process of the market makers, noting that the same process applies to the uninformed traders. After witnessing the ith trading decision, the market makers' beliefs regarding the signal that the informed traders received are

$$P\left(S_m = s_L | Z_i\right) = x_i,$$

$$P\left(S_m = s_H | Z_i\right) = y_i,$$

and, by construction,

$$P\left(S_m = s_O | Z_i\right) = 1 - x_i - y_i.$$

Each trading decision—even if the decsion is not to trade—conveys information about the signal received by the informed traders.

2.1. Quote Determination

Quotes are determined by two equilibrium conditions. The first condition is that a market maker earns zero expected profit from each trade. The zero expected profit condition arises from the potential free entry of additional market makers. From the zero expected profit condition it follows that the quotes are equal to the expected value of the asset conditional on the trade. The second condition involves comparing the net gains from the various trading strategies available to the informed. Because the informed will make the trade in the asset that offers the highest net gain, in equilibrium it must be the case that the quotes are set so that an informed trader earns an equal net gain from each possible trade.

To understand the interplay of the two conditions, we begin by studying the opening ask in the call market, A_{C_1}. The zero expected profit condition sets the market maker's expected gain from trade with an uninformed trader equal to the expected loss from trade with an informed trader. The market maker's belief that an uninformed trader will buy λ call options on the first trade is $(1-\alpha)\,\epsilon_{UAC}$.

Because the uninformed have the same information as the market maker, the market maker's expected gain from writing λ call options to an uninformed trader is

$$(1-\alpha)\,\epsilon_{UAC}\cdot\lambda\left[A_{C_1}-E\left(V_{C_{\mathsf{m}}}|Z_0\right)\right].$$

The market maker's belief that an informed trader will buy λ call options on the first trade is $\alpha\epsilon_{IAC_1}y_0$. Because each of the λ call options purchased by an informed trader will expire in-the-money, worth $v_{H_{\mathsf{m}}}-\kappa_{C_{\mathsf{m}}}$ per share, the market maker's expected loss from writing λ call options to an informed trader is

$$\alpha\epsilon_{IAC_1}y_0\cdot\lambda\left[A_{C_1}-\left(v_{H_{\mathsf{m}}}-\kappa_{C_{\mathsf{m}}}\right)\right].$$

If net profits are zero, then

$$A_{C_1}=\frac{\alpha\epsilon_{IAC_1}y_0\left(v_{H_{\mathsf{m}}}-\kappa_{C_{\mathsf{m}}}\right)+(1-\alpha)\,\epsilon_{UAC}E\left(V_{C_{\mathsf{m}}}|Z_0\right)}{\alpha\epsilon_{IAC_1}y_0+(1-\alpha)\,\epsilon_{UAC}}.$$

The opening ask depends on ϵ_{IAC_1}, which in turn depends on the potential net gains available to the informed trader. To uniquely determine the probability that an informed trader chooses to buy λ call options at the market opening, we first calculate the net gain from each of the three possible bullish trades. We define the net gain from a particular trade as the ultimate proceeds to the trader less the cost of the trade. The informed trader can buy one share of the stock at the ask, A_1, and then sell it at the end of the trading day for $v_{H_{\mathsf{m}}}$ for a net gain of $v_{H_{\mathsf{m}}}-A_1$. The informed trader can buy λ call options at the ask, A_{C_1}, each of which will expire in-the-money at the end of the trading day for a net gain of $\lambda\left[\left(v_{H_{\mathsf{m}}}-\kappa_{C_{\mathsf{m}}}\right)-A_{C_1}\right]$. The informed trader can write λ put options at the bid, B_{P_1}, each of which will expire out–of-the-money, or worthless, at the end of the trading day for a net gain of λB_{P_1}.[4]

[4] As A_{C_1} depends on $\epsilon_{\mathsf{IAC}_1}$, so to do the other quotes depend on informed trade frequencies. For example, B_1 depends on the informed trade frequency at the bid in the stock market, ϵ_{IB_1}. To determine ϵ_{IB_1}, we equate the gains from each of the three bearish trades. The informed trader can sell short one share of the stock at the bid, B_1, and then buy it back at the end of the trading day for $v_{\mathsf{L}_\mathsf{m}}$ for a net gain of $B_1-v_{\mathsf{L}_\mathsf{m}}$. The informed trader can write λ call options at the bid, B_{C_1}, each of which will expire out-of-the-money at the end of the trading day for a net gain of λB_{C_1}. The informed trader can buy λ put options at the ask, A_{P_1}, each of which will expire in-the-money at the end of the trading day for a net gain of $\lambda\left[\left(\kappa_{\mathsf{P}_\mathsf{m}}-v_{\mathsf{L}_\mathsf{m}}\right)-A_{\mathsf{P}_1}\right]$.

In equilibrium, the market makers set A_1, A_{C_1}, and B_{P_1} so that the net gains from the three bullish strategies are equal,

$$v_{H_m} - A_1 = \lambda\left[(v_{H_m} - \kappa_{C_m}) - A_{C_1}\right] = \lambda B_{P_1}.$$

Solution of this uniquely determines ϵ_{IAC_1}, ϵ_{IA_1}, and ϵ_{IBP_1}. We find that

$$\epsilon_{IAC_1} = \epsilon_{UAC}\frac{\left\{\begin{array}{c}\left[\alpha y_0 + (1-\alpha)(\epsilon_{UA} + \epsilon_{UBP})\right]\lambda(v_{H_m} - \kappa_{C_m}) - \\ (1-\alpha)\left[(v_{H_m} - v_{L_m})\epsilon_{UA} + \lambda(\kappa_{P_m} - v_{L_m})\epsilon_{UBP}\right]\end{array}\right\}}{\alpha y_0\left[(v_{H_m} - v_{L_m})\epsilon_{UA} + \lambda(v_{H_m} - \kappa_{C_m})\epsilon_{UAC} + \lambda(\kappa_{P_m} - v_{L_m})\epsilon_{UBP}\right]}.$$

The opening ask in the call market is obtained as the solution to the zero profit condition with the value of ϵ_{IAC_1} determined above. With $\varphi_S = v_{H_m} - v_{L_m}$, $\varphi_C = v_{H_m} - \kappa_{C_m}$, and $\varphi_P = \kappa_{P_m} - v_{L_m}$, we have

$$A_{C_1} = \varphi_C - \frac{(1-\alpha)\left[\varphi_C - E(V_{C_m}|Z_0)\right](\varphi_S\epsilon_{UA} + \lambda\varphi_C\epsilon_{UAC} + \lambda\varphi_P\epsilon_{UBP})}{\left[\alpha y_0 + (1-\alpha)(\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP})\right]\lambda\varphi_C}$$

while the opening bid in the call market is

$$B_{C_1} = \frac{(1-\alpha)(\varphi_S\epsilon_{UB} + \lambda\varphi_C\epsilon_{UBC} + \lambda\varphi_P\epsilon_{UAP})E(V_{C_m}|Z_0)}{\left[\alpha x_0 + (1-\alpha)(\epsilon_{UB} + \epsilon_{UBC} + \epsilon_{UAP})\right]\lambda\varphi_C}.$$

2.2. Intra-Trading Day Dynamics

The market makers adjust, or update, their beliefs about the signal received by the informed traders—the conditional probabilities, x_i and y_i—to reflect the information revealed through the sequence of trading decisions. We detail how the market makers learn from the first trader; learning from all successive trade decisions follows similar logic. If $D_1 = d_A$, then from Bayes' Rule

$$x_1 = x_0\frac{(1-\alpha)\epsilon_{UA}}{\alpha\epsilon_{IA_1}y_0 + (1-\alpha)\epsilon_{UA}} \text{ and } y_1 = y_0\frac{\alpha\epsilon_{IA_1} + (1-\alpha)\epsilon_{UA}}{\alpha\epsilon_{IA_1}y_0 + (1-\alpha)\epsilon_{UA}}.$$

For $D_1 = d_{AC}$, we simply replace ϵ_{UA} with ϵ_{UAC} and replace ϵ_{IA_1} with ϵ_{IAC_1}. For $D_1 = d_{BP}$, we replace ϵ_{UA} with ϵ_{UBP} and replace ϵ_{IA_1} with ϵ_{IBP_1}. If $D_1 = d_B$, then

$$x_1 = x_0\frac{\alpha\epsilon_{IB_1} + (1-\alpha)\epsilon_{UB}}{\alpha\epsilon_{IB_1}y_0 + (1-\alpha)\epsilon_{UB}} \text{ and } y_1 = y_0\frac{(1-\alpha)\epsilon_{UB}}{\alpha\epsilon_{IB_1}y_0 + (1-\alpha)\epsilon_{UB}}.$$

For $D_1 = d_{BC}$, we replace ϵ_{UB} with ϵ_{UBC} and replace ϵ_{IB_1} with ϵ_{IBC_1}. For $D_1 = d_{AP}$, we replace ϵ_{UBC} with ϵ_{UAP} and replace ϵ_{IBC_1} with ϵ_{IAP_1}.

Consider the case in which the first trade is at the ask in the stock market. If $\epsilon_{IA_1} = 0$, then learning occurs only from trade in the options market. Suppose that $\epsilon_{IA_1} \neq 0$, so that learning occurs from trade in both the stock and options markets. If $y_0 < 1$, then a trade at the ask in the stock or the call option or a trade at the bid in the put option increases y_1 relative to y_0. Similarly, if $x_0 > 0$, then a trade at the ask in the stock or the call option or a trade at the bid in the put option decreases x_1 relative to x_0. We also find that revisions in x_1 and y_1 are not symmetric in that an increase in y_1 is accompanied by both a decrease in x_1 and a change in $1 - x_1 - y_1$. If $\alpha = 1$ or $\epsilon = 0$ so that only the informed trade, then learning is immediate with $x_1 = 0$ and $y_1 = 1$.

A decision not to trade also reveals information. If $D_1 = d_N$, then

$$x_1 = x_0 \frac{(1-\alpha)(1-\epsilon)}{\alpha(1 - x_0 - y_0) + (1-\alpha)(1-\epsilon)}$$

and

$$y_1 = y_0 \frac{(1-\alpha)(1-\epsilon)}{\alpha(1 - x_0 - y_0) + (1-\alpha)(1-\epsilon)}.$$

If $1 - x_0 - y_0 > 0$, then a decision not to trade decreases both x_1 and y_1. If $\alpha = 1$, or if $\epsilon = 1$ so that all uninformed traders trade, then learning is immediate and $x_1 = y_1 = 0$.

The market makers adjust, or update, their beliefs about the signal received by the informed traders—the conditional probabilities, x_i and y_i—to reflect the information revealed through the sequence of trading decisions. We detail how the market makers learn from the first trader; learning from all successive trade decisions follows similar logic. If $D_1 = d_A$, then from Bayes' Rule

$$x_1 = x_0 \frac{(1-\alpha)\epsilon_{UA}}{\alpha\epsilon_{IA_1} y_0 + (1-\alpha)\epsilon_{UA}} \text{ and } y_1 = y_0 \frac{\alpha\epsilon_{IA_1} + (1-\alpha)\epsilon_{UA}}{\alpha\epsilon_{IA_1} y_0 + (1-\alpha)\epsilon_{UA}}.$$

For $D_1 = d_{AC}$, we simply replace ϵ_{UA} with ϵ_{UAC} and replace ϵ_{IA_1} with ϵ_{IAC_1}. For $D_1 = d_{BP}$, we replace ϵ_{UA} with ϵ_{UBP} and replace ϵ_{IA_1} with ϵ_{IBP_1}. If $D_1 = d_B$, then

$$x_1 = x_0 \frac{\alpha\epsilon_{IB_1} + (1-\alpha)\epsilon_{UB}}{\alpha\epsilon_{IB_1} y_0 + (1-\alpha)\epsilon_{UB}} \text{ and } y_1 = y_0 \frac{(1-\alpha)\epsilon_{UB}}{\alpha\epsilon_{IB_1} y_0 + (1-\alpha)\epsilon_{UB}}.$$

For $D_1 = d_{BC}$, we replace ϵ_{UB} with ϵ_{UBC} and replace ϵ_{IB_1} with ϵ_{IBC_1}. For $D_1 = d_{AP}$, we replace ϵ_{UBC} with ϵ_{UAP} and replace ϵ_{IBC_1} with ϵ_{IAP_1}.

The key parameters that govern the speed of learning are α, ϵ, and the frequencies of informed trade. As α increases, the information content of a given trade increases and learning is more rapid. Increasing the propensity of the uninformed to trade, ϵ, decreases the information content of a given trade and slows learning. Increasing the frequency with which an informed trader makes a particular trade also affects learning. Learning occurs most rapidly from trade in the market with the least depth, that is the market for which the ratio of informed trade frequency to uninformed trade frequency is highest.[5] Consider the case in which the strike prices are the limit values and $\lambda > 1$, so the informed trade frequency is higher in the options market than in the stock market. If the uninformed trade frequency is equal in all assets, then the market depth is lower and so trade is more informative in the options market. Further, as the number of shares of the underlying stock controlled by an option contract, λ, increases, then informed trade frequencies in the options market increase. As a result, the information content of an option trade increases and, because trade is more concentrated in the options market, learning quickens. From Figure 2.1, which contains the average beliefs over 1000 simulations with $S_m = s_H$, one can readily confirm that as either ϵ declines or λ increases, learning is more rapid.

The formula for ϵ_{IAC_1} generalizes to the ith trader in a straightforward fashion; simply replace y_0 with y_{i-1}. We present the remaining frequencies of informed trade in the Appendix.[6] Analysis of the equilibrium frequencies of informed trade reveals the

Equal Payoff Condition: The options leverage and strike prices satisfy

$$v_{H_m} - v_{L_m} = \lambda \left(v_{H_m} - \kappa_{C_m} \right) = \lambda \left(\kappa_{P_m} - v_{L_m} \right).$$

If the equal payoff condition is satisfied, then in equilibrium the informed trade

[5] The depth of a market is the proportion of trades in a market made by uninformed traders. For example, the depth of the stock market at the bid is

$$\frac{\epsilon_{UB}}{\epsilon_{IB} + \epsilon_{UB}}.$$

[6] We simplify the algebra underlying the informed trade frequencies by assuming, as do Easley, O'Hara and Srinivas (1998), that: if $S_m = s_O$ then $V_{C_m} = (1-\delta)\left(v_{H_m} - \kappa_{C_m}\right)$ and $V_{P_m} = \delta\left(\kappa_{P_m} - v_{L_m}\right)$. While such an assumption implies revelation of V_m rather than S_m, the two are equivalent if the strike prices are at their limits and $\delta = .5$.

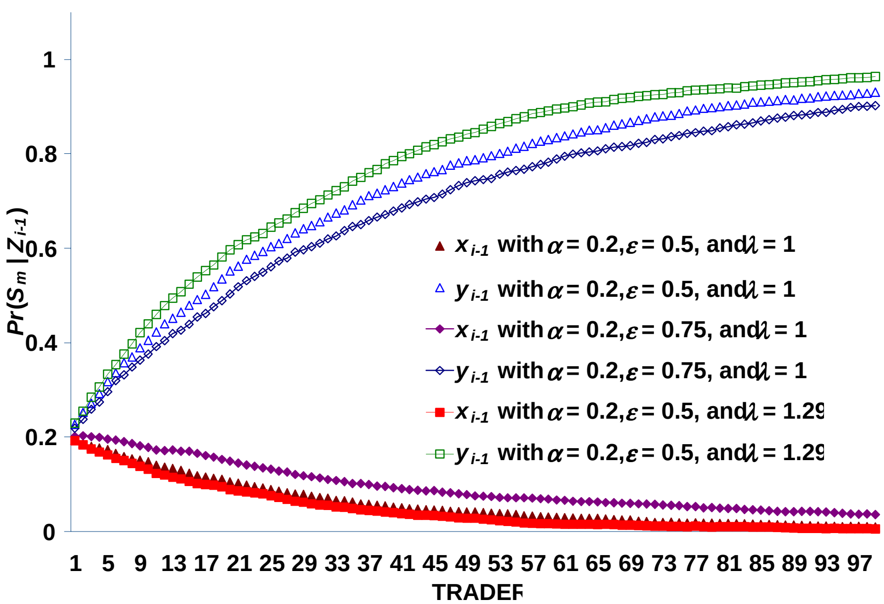

Figure 2.1: The Learning Process

with *constant and equal* frequency in each asset:

$$\epsilon_{IA_{\mathsf{i}}} = \frac{\epsilon_{UA}}{\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}},$$

$$\epsilon_{IAC_{\mathsf{i}}} = \frac{\epsilon_{UAC}}{\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}},$$

and

$$\epsilon_{IBP_{\mathsf{i}}} = \frac{\epsilon_{UBP}}{\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}}.$$

In essence, the informed traders mirror the behavior of uninformed traders in that they trade with identical relative frequency. Because the uninformed have six potential trades, while the informed have only three, the frequency of informed trade is always at least as great as the frequency of uninformed trade. If the uninformed are equally likely to make each of the six possible uninformed trades, then the informed are equally likely to make each of the three possible informed trades. For example, if $S_m = s_H$ then $\epsilon_{IA_{\mathsf{i}}} = \epsilon_{IAC_{\mathsf{i}}} = \epsilon_{IBP_{\mathsf{i}}} = \frac{1}{3}$ for all i.

If the equal payoff condition is not satisfied, then in equilibrium the informed trade with *variable and unequal* frequency in each asset throughout the trading day.

For variable informed trade frequencies, we study how the frequencies depend on the underlying parameters. To make the analysis concise, we focus on an empirically relevant case in which the options offer greater leverage (over stocks, so $\lambda\left(\kappa_{P_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right) > v_{H_{\mathsf{m}}} - v_{L_{\mathsf{m}}}$) and the option payoffs are symmetric ($\lambda\left(\kappa_{P_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right) = \lambda\left(v_{H_{\mathsf{m}}} - \kappa_{C_{\mathsf{m}}}\right) \equiv \lambda\beta$). Our results are contained in

Theorem 1: *If the options offer greater leverage and have symmetric payoffs, then the informed trade frequencies behave in the following ways.*

(a) As λ increases, the informed are less likely to trade in the stock market. As α increases, the informed are more likely to trade in the stock market.

(b) As learning evolves, the informed flow from the options market to the stock market. The rate of flow declines over the course of a trading day. The rate of flow also declines as α increases.

(c) Informed trade frequencies in the option market are always positive. If the uninformed trade each asset with equal frequency, then $\epsilon_{IAC_{\mathsf{i}}} = \epsilon_{IBP_{\mathsf{i}}} > \epsilon_{IA_{\mathsf{i}}}$ and $\epsilon_{IBC_{\mathsf{i}}} = \epsilon_{IAP_{\mathsf{i}}} > \epsilon_{IB_{\mathsf{i}}}$.

(d) The ith informed trade frequencies in the stock market are positive if, for $j = H, L$,

$$\lambda < \frac{(v_{H_\mathsf{m}} - v_{L_\mathsf{m}})}{\beta}\left(1 + \frac{\alpha}{1-\alpha}\frac{1}{\varepsilon_j} b_{j,i-1}\right),$$

with $\varepsilon_H = \varepsilon_{UAC} + \varepsilon_{UBP}$, $\varepsilon_L = \varepsilon_{UBC} + \varepsilon_{UAP}$, $b_{H,i-1} = y_{i-1}$ and $b_{L,i-1} = x_{i-1}$.

Proof. *See the Appendix.*

Increasing the payoff of an option causes the informed to flow into the options market, as detailed in (a). Further, an increase in the proportion of informed traders reduces the potential depth of the options market, which in turn makes the stock market more attractive to informed traders. To understand the dynamic pattern revealed in (b), consider a day on which $S_m = s_H$. As the informed trade and reveal their information, y_i increases. As y_i increases, the gains to trade on information shrink, as does the advantage from trading in the options market. Hence, over the course of a trading day the informed flow from the options market to the stock market. As the updating of y_i slows over the course of a trading day to reflect the reduced information content of trades, so too does the rate of flow of informed traders. In similar fashion, as α increases, the information gain from each trader increases, so higher values of α lead to faster learning and greater attenuation of the rate of flow of informed between markets over the course of a trading day. While the informed flow from the options market to the stock market over the course of a trading day, if the uninformed are equally likely to trade in each market then the informed trade frequency is higher in the options market uniformly over the trading day, as stated in (c).

As noted in (a), leverage attracts informed traders to the options market. If λ is large enough, then the frequency of informed trade in the stock market is zero and the equilibrium separates the markets in which the informed trade. As either α decreases or ε_j increases, the *separating* bound in (d) decreases and informed trade is more likely to occur only in the options market. Such a result is intuitive in that decreases in the proportion of informed traders or increases in the proportion of uninformed traders in the (relevant) options allow the informed to more easily hide in the options market.

Because λ is fixed over the course of a trading day while b_i evolves with the trade flow, it will generally not be the case that a separating equilibrium exists in all periods. Consider a trading day on which $S_m = s_H$. The relevant informed trade frequencies correspond to the bullish trades and the corresponding bound for λ is based on y_{i-1}. As the market makers learn that the high signal is increasingly likely, y_{i-1} increases toward 1 thereby increasing the bound. As the bound increases the informed enter the stock market. As an aside, at each trader arrival the equilibrium also requires calculation of the irrelevant bearish informed trade frequencies, whose bound is based on x_{i-1}. As x_i declines to 0 the bound for the bearish informed trade frequencies is violated and these irrelevant informed trade frequencies can move outside $[0, 1]$. In essence, the market maker is attempting to determine the frequency with which the informed are making bearish trades when, in fact, the informed are making only bullish trades. The import is potentially to slow down learning, as uninformed trade at one of the three bearish quotes could lead to erratic updating if the informed trade frequencies are not on $[0, 1]$. To eliminate the erratic behavior, if the informed trade frequencies fall outside $[0, 1]$, we fix them at their last values in $[0, 1]$.

In parallel to the opening quotes in the call market, the ith-trade quotes for each asset are obtained as the solution t the zero profit condition with the relevant informed trade frequency (given in the Appendix). The ith-trade quotes for one share of the stock are

$$A_i = v_{H_\mathsf{m}} - \frac{(1-\alpha)\left[v_{H_\mathsf{m}} - E\left(V_m|Z_{i-1}\right)\right]\left(\varphi_S\epsilon_{UA} + \lambda\varphi_C\epsilon_{UAC} + \lambda\varphi_P\epsilon_{UBP}\right)}{\left[\alpha y_{i-1} + (1-\alpha)\left(\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}\right)\right]\varphi_S}$$

and

$$B_i = v_{L_\mathsf{m}} + \frac{(1-\alpha)\left[E\left(V_m|Z_{i-1}\right) - v_{L_\mathsf{m}}\right]\left(\varphi_S\epsilon_{UB} + \lambda\varphi_C\epsilon_{UBC} + \lambda\varphi_P\epsilon_{UAP}\right)}{\left[\alpha x_{i-1} + (1-\alpha)\left(\epsilon_{UB} + \epsilon_{UBC} + \epsilon_{UAP}\right)\right]\varphi_S},$$

and the ith-trade quotes for the put option are

$$A_{P_\mathsf{i}} = \varphi_P - \frac{(1-\alpha)\left[\varphi_P - E\left(V_{P_\mathsf{m}}|Z_{i-1}\right)\right]\left(\varphi_S\epsilon_{UB} + \lambda\varphi_C\epsilon_{UBC} + \lambda\varphi_P\epsilon_{UAP}\right)}{\left[\alpha x_{i-1} + (1-\alpha)\left(\epsilon_{UB} + \epsilon_{UBC} + \epsilon_{UAP}\right)\right]\lambda\varphi_P}$$

and

$$B_{P_\mathsf{i}} = \frac{(1-\alpha)\left[\varphi_S\epsilon_{UA} + \lambda\varphi_C\epsilon_{UAC} + \lambda\varphi_P\epsilon_{UBP}\right]E\left(V_{P_\mathsf{m}}|Z_{i-1}\right)}{\left[\alpha y_{i-1} + (1-\alpha)\left(\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}\right)\right]\lambda\varphi_P}.$$

From these equations it is easy to see that each set of quotes is bounded by the respective limit values of the asset, with strict inequality unless the specialist is certain the informed learn the true value of V_m (no adverse selection). We also find that the quotes for the stock and the options bound the respective expected values of the assets, which illustrates the spread generated by the market makers in an effort to offset expected losses to traders with superior information. Thus $B_i \leq E(V_m|Z_{i-1}) \leq A_i$, $B_{C_i} \leq E(V_{C_m}|Z_{i-1}) \leq A_{C_i}$, and $B_{P_i} \leq E(V_{P_m}|Z_{i-1}) \leq A_{P_i}$, each of which follows directly from the condition that the informed trade frequencies lie on $[0,1]$.

To understand the key influences of the bid-ask spread, suppose that $x_{i-1} = y_{i-1}$ and that the uninformed are equally likely to make each of the six possible trades. If the strike prices are fixed at their limit values, we find that

$$A_i - B_i = (v_{H_m} - v_{L_m}) \frac{\alpha y_{i-1} + (1-\alpha)(1-\lambda)\frac{\epsilon}{3}}{\alpha y_{i-1} + (1-\alpha)\frac{\epsilon}{2}},$$

and

$$A_{C_i} - B_{C_i} = A_{P_i} - B_{P_i} = (v_{H_m} - v_{L_m}) \frac{\alpha y_{i-1}\lambda + (1-\alpha)(\lambda-1)\frac{\epsilon}{6}}{\left[\alpha y_{i-1} + (1-\alpha)\frac{\epsilon}{2}\right]\lambda}.$$

While the spread in the options market is clearly positive, the sign of the spread in the stock market depends on the magnitude of λ. Because the bid-ask spread arises from potential trade with the informed, the separating bound that ensures positive informed trade in the stock market also ensures a postive bid-ask spread. If λ equals the separating bound, then the frequency of informed trade in the stock market is zero and the bid-ask spread is zero. With $v_{H_m} - v_{L_m}$ as a measure of the variance of the stock price, we find that, as $v_{H_m} - v_{L_m}$ increases, the bid-ask spreads for all of the assets increase. As ϵ increases, so does the market depth, and the bid-ask spread declines. Increasing the probability of informed trade, α, widens the bid-ask spread in all assets.

To illustrate the dynamics of informed trade frequencies implied by our analytic results, we simulate the arrival of traders over the course of 1000 trading days on which $S_m = s_H$. We set the information advantage of the informed at five percent of the initial value of the asset, so $v_{H_m} = 105$, $v_{L_m} = 95$ and $\delta = .5$. To ensure that option payoffs are symmetric, we set $\kappa_{C_m} = v_{L_m}$ and $\kappa_{P_m} = v_{H_m}$. (The greater leverage afforded by options is then captured by $\lambda > 1$.[7]) We further suppose that the uninformed are equally likely to trade each asset, so that the informed trade frequencies, and hence the spreads, are identical for the two options. Finally, we suppose that $\alpha = 0.2$ and $\epsilon = 0.75$, noting that the essential

[7] For the given parameter values, the separating bound is 1.2.

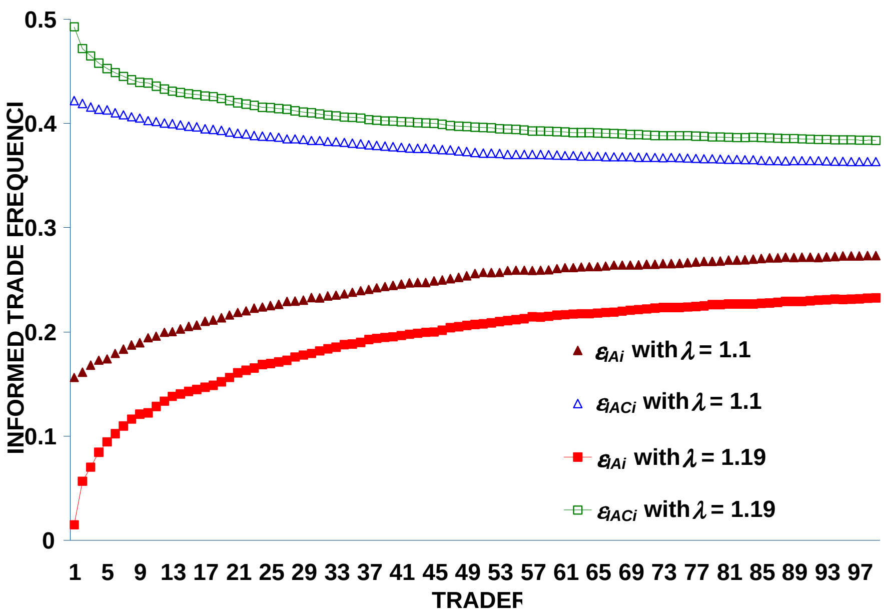

Figure 2.2: Informed Trader Flow

features we report are obtained with the other parameter values as well.[8]

In Figure 2.2 we present the average frequency of informed trade in each asset over the course of a trading day. The analytic results are clearly revealed. First, as λ increases ϵ_{IAC_i} and ϵ_{IBP_i} increase toward .5 and ϵ_{IA_i} decreases toward zero. Second, as the market makers learn the signal we find that informed traders flow into the stock market and do so at a decreasing rate over the course of a trading day.

In Figure 2.3 we present the average bid-ask spread over the course of a trading day. First, as λ increases the adverse selection problem in the options is exacerbated and forces the market maker to widen the bid-ask spreads for the call and put options, while the adverse selection problem in the stock is mitigated and allows the market maker to reduce the spread for the stock. This is particularly

[8] We conduct a series of a series of simulations with $\alpha \in \{0.2, 0.4, 0.6, 0.8\}$ and $\epsilon \in \{0.25, 0.5, 0.75, 0.9\}$. Our decision to focus on the results of simulations with $\alpha = 0.2$ and $\epsilon = 0.75$ is based on the fact that these parameters produce slower learning and allow us to more easily observe the dynamic process of the model.

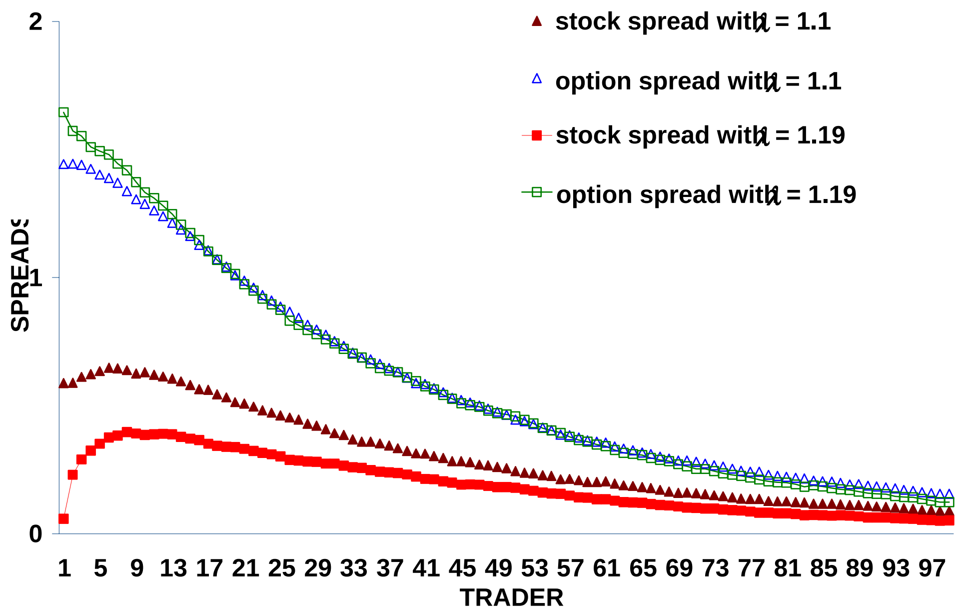

Figure 2.3: Bid-Ask Spreads

apparent initially, but over the course of the trading day diminishes as informed traders flow out of the options and into the stock.

2.3. Behavior of Individual Trader Price Changes

Price changes reflect public information after the decision of trader i but before the arrival of trader $i+1$. The stock price change associated with a specific trade decision for trader i is $U_i\left(D_i = d_j\right) = E\left(V_m | Z_{i-1}, D_i = d_j\right) - E\left(V_m | Z_{i-1}\right)$. The form of the price change is intuitive and mirrors the logic of the learning rules. To understand the logic for a specific trade, at the ask in the stock, note first that because $E\left(V_m | Z_i\right) = x_i v_{L_\mathrm{m}} + y_i v_{H_\mathrm{m}} + \left(1 - x_i - y_i\right) EV_m$, the stock price change is

$$U_i\left(D_i = d_A\right) = \left[v_{H_\mathrm{m}} - E\left(V_m | Z_{i-1}\right)\right] \frac{\alpha \epsilon_{IA_\mathrm{i}} y_{i-1}}{P\left(D_i = d_A | Z_{i-1}\right)}.$$

The price change reflects expected learning from the informed; if the market maker knows that the trader is uninformed, there is no learning from the trade and the price change is zero. The expected learning from the informed is the price change that would occur if the specialist knows that the trader is informed $v_{H_\mathsf{m}} - E(V_m|Z_{i-1})$ multiplied by the frequency of trade with an informed trader. The market maker's belief that the trader is informed is simply the probability that an informed trader trades at the ask in the stock market, $\alpha\epsilon_{IA_\mathsf{i}} y_{i-1}$, divided by the probability of a trade at the ask in the stock market, $\alpha\epsilon_{IA_\mathsf{i}} y_{i-1} + (1-\alpha)\epsilon_{UA}$.

As expected, a trade in the stock also affects the prices of the options. Given a trade at the ask in the stock, the corresponding price changes for the options are

$$U_{C_\mathsf{i}}(D_i = d_A) = [(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}) - E(V_{C_\mathsf{m}}|Z_{i-1})] \frac{\alpha\epsilon_{IA_\mathsf{i}} y_{i-1}}{P(C_1 = d_A|Z_{i-1})}$$

and

$$U_{P_\mathsf{i}}(D_i = d_A) = [0 - E(V_{P_\mathsf{m}}|Z_{i-1})] \frac{\alpha\epsilon_{IA_\mathsf{i}} y_{i-1}}{P(D_i = d_A|Z_{i-1})}.$$

With a bullish trade the value of the stock and the call option rise while the value of the put option falls.

Another important feature of the model is that a trade in an option affects the price of the stock. If trader i elects to buy the call option contract rather than the stock, the three price change formulae differ only in the frequency of informed trade. For example,

$$U_i(D_i = d_{AC}) = [v_{H_\mathsf{m}} - E(V_m|Z_{i-1})] \frac{\alpha\epsilon_{IAC_\mathsf{i}} y_{i-1}}{P(D_i = d_{AC}|Z_{i-1})}.$$

Because the decision not to trade also conveys information to the market makers,

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

We first establish that prices are unpredictable with respect to public information. The price change expected by market makers and uninformed traders resulting from the decision of the ith trader is

$$E(U_i|Z_{i-1}) = \sum_{j=A,B,AC,BC,AP,BP,N} P(D_i = d_j|Z_{i-1}) U_i(D_i = d_j)$$

or

$$\begin{aligned} E(U_i|Z_{i-1}) &= \alpha y_{i-1}(\epsilon_{IA_\mathsf{i}} + \epsilon_{IAC_\mathsf{i}} + \epsilon_{IBP_\mathsf{i}})[v_{H_\mathsf{m}} - E(V_m|Z_{i-1})] + \\ &\quad \alpha x_{i-1}(\epsilon_{IB_\mathsf{i}} + \epsilon_{IBC_\mathsf{i}} + \epsilon_{IAP_\mathsf{i}})[v_{L_\mathsf{m}} - E(V_m|Z_{i-1})] + \\ &\quad \alpha(1 - x_{i-1} - y_{i-1})[EV_m - E(V_m|Z_{i-1})] \\ &= 0, \end{aligned}$$

where the final line follows from $\epsilon_{IA_{\mathsf{i}}} + \epsilon_{IAC_{\mathsf{i}}} + \epsilon_{IBP_{\mathsf{i}}} = \epsilon_{IB_{\mathsf{i}}} + \epsilon_{IBC_{\mathsf{i}}} + \epsilon_{IAP_{\mathsf{i}}} = 1$. By identical logic, the expected price changes with respect to the public information set are zero in the options market.

As one would expect, price changes are predicatable on the basis of the signal. Consider the case in which $S_m = s_H$. The expected stock price change for trader i, given knowledge of the signal, is

$$E\left(U_i | S_m = s_H, Z_{i-1}\right) = \sum_{j=A,B,AC,BC,AP,BP,N} P\left(D_i = d_j | S_m = s_H, Z_{i-1}\right) U_i\left(D_i = d_j\right).$$

Because the market maker does not know the signal, only the probability of trade is affected by the signal. For the three bullish trades, $j = A$, AC, and BP, we have

$$P\left(D_i = d_j | S_m = s_H, Z_{i-1}\right) > P\left(D_i = d_j | Z_{i-1}\right),$$

while for the three bearish trades, $j = B$, BC, and AP, we have

$$P\left(D_i = d_j | S_m = s_H, Z_{i-1}\right) < P\left(D_i = d_j | Z_{i-1}\right).$$

The differences in probabilities are summarized in the positive constants a_i and b_i, so

$$\begin{aligned} E\left(U_i | S_m = s_H, Z_{i-1}\right) &= \alpha y_{i-1}\left(1 + a_i\right)\left[v_{H_{\mathsf{m}}} - E\left(V_m | Z_{i-1}\right)\right] + \\ &\quad \alpha x_{i-1}\left(1 - b_i\right)\left[v_{L_{\mathsf{m}}} - E\left(V_m | Z_{i-1}\right)\right] + \\ &\quad \alpha\left(1 - x_{i-1} - y_{i-1}\right)\left[EV_m - E\left(V_m | Z_{i-1}\right)\right] \\ &> 0. \end{aligned}$$

In similar fashion, $E\left(U_i | S_m = s_L, Z_{i-1}\right) < 0$.

We next establish that price changes are serially uncorrelated with respect to public information. Let h and i be distinct, with $h < i$:

$$E\left(U_h U_i | Z_{i-1}\right) = U_h E\left(U_i | Z_{i-1}\right) = 0.$$

Price changes are serially correlated on the basis of the signal. As above, consider the case in which $S_m = s_H$. The serial correlation between the price changes resulting from the decisions of traders h and i, given knowledge of the signal, is

$$E\left(U_h U_i | S_m = s_H, Z_{i-1}\right) = U_h E\left(U_i | S_m = s_H, Z_{i-1}\right) \neq 0.$$

The variance of price changes varies over the course of the trading day, with

$$E\left(U_i^2|Z_{i-1}\right)=\sum_{j=A,B,AC,BC,AP,BP,N} P\left(D_i=d_j|Z_{i-1}\right)U_i^2\left(D_i=d_j\right).$$

To determine the pattern of variation over the course of a trading day, we construct analytic bounds. These bounds depend on the effective bid-ask spread. The effective bid-ask spread captures the maximum revision in the price resulting from a trade. If $\epsilon_{IAC_i}>\epsilon_{IA_i}$ and $\epsilon_{IBP_i}>\epsilon_{IA_i}$, or if $\epsilon_{IBC_i}>\epsilon_{IB_i}$ and $\epsilon_{IAP_i}>\epsilon_{IB_i}$, then the information content of a trade in the options market is greater than that of a trade in the stock market. If a decision not to trade is quite rare and generally made by informed traders (when ϵ is very large and α is very small) then a decision not to trade can yield a larger price change than a decision to trade. With this in mind, we define the effective bid-ask spread for the stock as

$$\hat{A}_i-\hat{B}_i=\max_{j\in\{AC,BP,N\}}\left[A_i,E\left(V_m|Z_{i-1},D_i=d_j\right)\right]-\min_{j\in\{BC,AP,N\}}\left[B_i,E\left(V_m|Z_{i-1},D_i=d_j\right)\right].$$

We define the effective bid-ask spreads for the call option and the put option, $\hat{A}_{C_i}-\hat{B}_{C_i}$ and $\hat{A}_{P_i}-\hat{B}_{P_i}$, in the same way. As in Kelly and Steigerwald, we find that, with respect to public information, although price changes of an asset are mean zero and uncorrelated, they are *dependent and not identically distributed.* Thus, an asset's bid-ask spread drives the variance of its price changes, introducing autoregressive heteroskedasticity.

Theorem 2: *Price changes in economic time for each asset are mean zero and serially uncorrelated with respect to the public information set. In addition*

$$E\left(U_i^2|Z_{i-1}\right)\leq\left(\hat{A}_i-\hat{B}_i\right)^2,$$

$$E\left(U_{C_i}^2|Z_{i-1}\right)\leq\left(\hat{A}_{C_i}-\hat{B}_{C_i}\right)^2,$$

and

$$E\left(U_{P_i}^2|Z_{i-1}\right)\leq\left(\hat{A}_{P_i}-\hat{B}_{P_i}\right)^2.$$

Proof: *See the Appendix.*

The fact that the price change variance is bounded by the effective bid-ask spread is an important component of the model. The result suggests that price change behavior is systematically different on days for which the signal is informative than on days for which the signal is s_O. In particular, the price uncertainty associated with informed trading should widen the effective bid-ask spread, leading to greater price variance on days for which $S_m \neq s_O$. We examine the price uncertainty on a trading day with $S_m = s_H$ relative to the price uncertainty on a trading day with $S_m = s_O$.[9] We study $E\left(U_i^2|S_m = s_H, Z_{i-1}\right) - E\left(U_i^2|S_m = s_O, Z_{i-1}\right)$, which equals

$$\sum_{j=A,B,AC,BC,AP,BP,N} \left[\begin{array}{c} P\left(D_i = d_j|S_m = s_H, Z_{i-1}\right) - \\ P\left(D_i = d_j|S_m = s_O, Z_{i-1}\right) \end{array} \right] U_i^2\left(D_i = d_j\right).$$

While the probability of a bearish trade is identical under either signal, the probabilities of the remaining trades differ under the two signals as

$$P\left(D_i = d_A|S_m = s_H, Z_{i-1}\right) = P\left(D_i = d_A|S_m = s_O, Z_{i-1}\right) + \alpha\epsilon_{IA_i},$$

$$P\left(D_i = d_{AC}|S_m = s_H, Z_{i-1}\right) = P\left(D_i = d_{AC}|S_m = s_O, Z_{i-1}\right) + \alpha\epsilon_{IAC_i},$$

$$P\left(D_i = d_{BP}|S_m = s_H, Z_{i-1}\right) = P\left(D_i = d_{BP}|S_m = s_O, Z_{i-1}\right) + \alpha\epsilon_{IBP_i},$$

and

$$P\left(D_i = d_N|S_m = s_H, Z_{i-1}\right) = P\left(D_i = d_N|S_m = s_O, Z_{i-1}\right) - \alpha.$$

Hence, we have

$$\begin{aligned} E\left(U_i^2|S_m = s_H, Z_{i-1}\right) - E\left(U_i^2|S_m = s_O, Z_{i-1}\right) = \\ \alpha\epsilon_{IA_i}U_i^2\left(D_i = d_A\right) + \alpha\epsilon_{IAC_i}U_i^2\left(D_i = d_{AC}\right) + \\ \alpha\epsilon_{IBP_i}U_i^2\left(D_i = d_{BP}\right) - \alpha U_i^2\left(D_i = d_N\right). \end{aligned}$$

This difference is clearly positive for the price change associated with the first trader, U_1, because $\delta = \frac{1}{2}$ implies that $EV_m = E\left(V_m|Z_0\right)$ so that $U_1\left(C_1 = d_N\right) =$

[9] If the low and high signals are equally likely ($\delta = \frac{1}{2}$) and the uninformed trade with equal frequency in all markets, then the behavior of price changes is identical on low and high signal days and it is enough to study high signal days.

Table 2.1: Value of $E\left(U_i^2|S_m = s_H, Z_{i-1}\right) - E\left(U_i^2|S_m = s_O, Z_{i-1}\right)$, with $\alpha = 0.9$, $\epsilon = 0.75$, $v_{Hm} = 100$, $v_{Lm} = 90$, $\kappa_{Cm} = 90$, $\kappa_{Pm} = 100$, and $\lambda = 1$.

trader	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$\epsilon = 0.9$	14.40	1.12	1.80	0.25	0.23	0.05	0.04	0.01	0.01	0.00
$\epsilon = 0.8$	15.06	1.21	1.81	0.20	0.20	0.04	0.03	0.01	0.00	0.00
$\epsilon = 0.7$	15.77	1.24	1.78	0.21	0.18	0.04	0.02	0.01	0.00	0.00
$\epsilon = 0.6$	16.53	1.22	1.65	0.23	0.15	0.04	0.02	0.00	0.00	0.00
$\epsilon = 0.5$	17.34	1.17	1.43	0.24	0.10	0.03	0.01	0.00	0.00	0.00
$\epsilon = 0.4$	18.23	1.11	1.15	0.27	0.06	0.04	0.01	0.00	0.00	0.00
$\epsilon = 0.3$	19.17	1.10	0.82	0.30	0.05	0.03	0.01	0.00	0.00	0.00
$\epsilon = 0.2$	20.19	1.20	0.52	0.31	0.04	0.01	0.01	0.00	0.00	0.00
$\epsilon = 0.1$	21.30	1.52	0.29	0.22	0.05	0.00	0.00	0.00	0.00	0.00

0. In general, however, $EV_m \neq E\left(V_m|Z_{i-1}\right)$ for $i > 1$, so the sign of the difference is not clear.

To determine the sign of $E\left(U_i^2|S_m = s_H, Z_{i-1}\right) - E\left(U_i^2|S_m = s_O, Z_{i-1}\right)$, we study the behavior of U_i^2 for general i. We assume that the equal payoff condition is satisfied, with $\lambda = 1$. For trader i, there are 7^i possible values for U_i, so calculation of the distribution of U_i^2 is cumbersome for large i. If α is large, then learning is rapid and largely occurs within the first 10 traders. For illustration, in Table 2.1 we directly calculate $E\left(U_i^2|S_m = s_H, Z_{i-1}\right) - E\left(U_i^2|S_m = s_O, Z_{i-1}\right)$ for $\alpha = .9$, from the exact distributions for U_i^2. We first note that the price uncertainty during a trading day on which $S_m \neq s_O$ is always larger than the price uncertainty during a day on which $S_m = s_O$. We also note that, as traders arrive to the markets, the market makers learn and the relative price uncertainty for an asset decreases as the ask and bid converge to the true value of the asset. Convergence depends on the speed of learning, and the speed of learning increases as the proportion of informed traders, α, increases and the proportion of uninformed traders who trade, ϵ, decreases. The magnitude of each difference reflects the information content of a trade: as ϵ decreases, the magnitude of the expected price change resulting from the first trader arrival increases.

For smaller values of α, learning is slowed and reduction of an asset's bid-ask spread to zero requires many more trader arrivals. With direct calculation of the exact distribution cumbersome, we approximate the exact distribution with simulations. In Figure 2.5 we calculate $E\left(U_i^2|S_m = s_H, Z_{i-1}\right) - E\left(U_i^2|S_m = s_O, Z_{i-1}\right)$ for $\alpha = .2$, from 1000 simulations. We again find that the variance of U_i is higher, uniformly, on a day with $S_m \neq s_O$ than it is on a day with $S_m = s_O$, confirming the results of Table 2.1. As the proportion of the uninformed who trade decreases, the information content of a trade increases and learning speeds up, resulting in more rapid convergence of an asset's bid-ask spread and faster dissipation of price uncertainty.

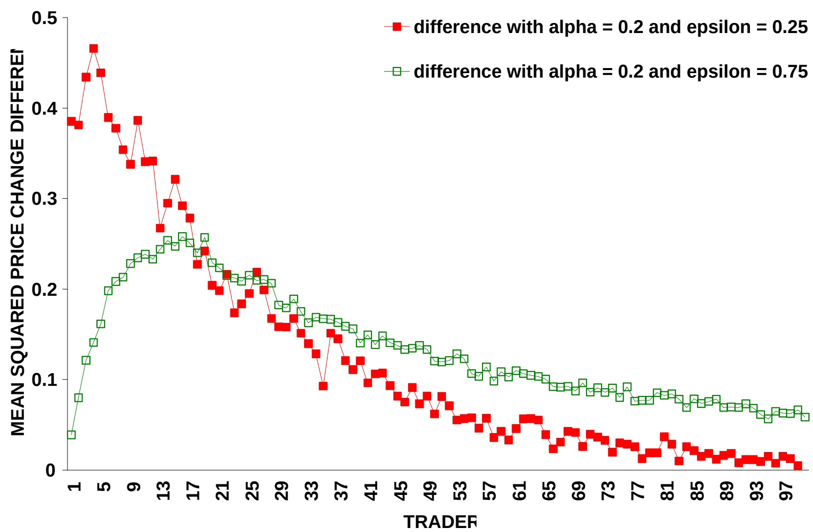

Figure 2.5: Price uncertainty difference, $E\left(U_i^2|Z_{i-1}, S_m = s_H\right) - E\left(U_i^2|Z_{i-1}, S_m = s_O\right)$, with $\alpha = 0.2$, $v_{H_\mathrm{m}} = 100$, $v_{L_\mathrm{m}} = 90$, $\kappa_{C_\mathrm{m}} = 90$, $\kappa_{P_\mathrm{m}} = 100$, and $\lambda = 1$ with two values of ϵ.

2.4. Consistency of Learning

To confirm consistency of the learning rules, we must establish the limiting values of x_i and y_i as the number of trader arrivals on a trading day grows without bound.

We establish that if there were an infinite number of trader arrivals on m, then market makers would learn the signal, S_m. As a result, the quotes for each asset converge to the strong-form efficient value of that asset, reflecting both public and private information. As transaction prices are determined by the quotes, these prices also converge to the respective strong-form efficient values of the assets.

Theorem 3: *The sequence of quotes and, hence, the sequence of transaction prices for each asset converge almost surely to the strong-form efficient value of* *that asset at an exponential rate. Specifically, the following results obtain as* $i \longrightarrow \infty$.

If $S_m = s_L$ then $x_i \xrightarrow{as} 1$, $y_i \xrightarrow{as} 0$, so $A_i \xrightarrow{as} v_{L_\mathsf{m}}$, $B_i \xrightarrow{as} v_{L_\mathsf{m}}$, $A_{C_\mathsf{i}} \xrightarrow{as} 0$, $B_{C_\mathsf{i}} \xrightarrow{as} 0$, $A_{P_\mathsf{i}} \xrightarrow{as} \kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}$ and $B_{P_\mathsf{i}} \xrightarrow{as} \kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}$.

If $S_m = s_H$ then $x_i \xrightarrow{as} 0$, $y_i \xrightarrow{as} 1$, so $A_i \xrightarrow{as} v_{H_\mathsf{m}}$, $B_i \xrightarrow{as} v_{H_\mathsf{m}}$, $A_{C_\mathsf{i}} \xrightarrow{as} v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}$, $B_{C_\mathsf{i}} \xrightarrow{as} v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}$, $A_{P_\mathsf{i}} \xrightarrow{as} 0$ and $B_{P_\mathsf{i}} \xrightarrow{as} 0$.

If $S_m = s_O$ then $x_i \xrightarrow{as} 0$, $y_i \xrightarrow{as} 0$, so $A_i \xrightarrow{as} EV_m$, $B_i \xrightarrow{as} EV_m$, $A_{C_\mathsf{i}} \xrightarrow{as} EV_{C_\mathsf{m}}$, $B_{C_\mathsf{i}} \xrightarrow{as} EV_{C_\mathsf{m}}$, $A_{P_\mathsf{i}} \xrightarrow{as} EV_{P_\mathsf{m}}$ and $B_{P_\mathsf{i}} \xrightarrow{as} EV_{P_\mathsf{m}}$.

Proof: *See the Appendix.*

Convergence of the beliefs $\{x_i\}_{i\geq 0}$ and $\{y_i\}_{i\geq 0}$ immediately implies that $U_i \xrightarrow{as} 0$, so that individual trader price volatility converges to zero.

3. Calendar Period Implications

To determine the serial correlation properties of trades and squared price changes for calendar periods, such as thirty-minute intervals, we divide each trading day into k calendar periods. Each calendar period contains η trader arrivals, and each trader arrival can be thought of as a unit of economic time. For a given trading day, m, we have $\tau = k\eta$ trading opportunities.

3.1. Calendar Period Trades

First we focus on the covariance structure of the number of trades in the stock and the call and put options in a calendar period. We let calendar periods be indexed by t, so I_{S_t}, I_{C_t}, and I_{P_t} represent the respective number of trades in the stock, the call option, and the put option in t, and we let I_t represent the total number of trades. Given η trader arrivals in t, I_t (as well as the number of trades in each market) takes integer values between 0 and η. Each trade variable is a binomial random variable for which the number of trades in t corresponds to the number of successes in η trials.

As Kelly and Steigerwald consider only a stock market, their result on serial correlation in trades applies directly to our total trades variable I_t. Because news arrivals are independent across trading days, if $r > k$ then I_{t-r} and I_t are uncorrelated. For $0 < r < k$, Kelly and Steigerwald prove that trades are postively correlated as

$$Cor\left(I_{t-r}, I_t\right) = \frac{\theta\left(1-\theta\right)\left(\alpha\eta\right)^2}{\sigma^2}\left(\frac{k-r}{k}\right).$$

They further show that the positive correlation between I_t and I_{t-r} is increasing in α, increasing in k, increasing in η, but decreasing in r.

For correlation in trades in a specific asset, we focus on trades in the call option. (Analogous results hold for the stock and the put option.) For the call option, in each period on trading day m we have

$$E\left(I_{C_t}|S_m \neq s_O\right) = \eta\left(1-\alpha\right)\epsilon_{UC} + \sum_{i=\eta(t-1)+1}^{\eta t} \alpha\left[\delta\epsilon_{IBC_i} + \left(1-\delta\right)\epsilon_{IAC_i}\right] = \mu_{C1}$$

and

$$E\left(I_{C_t}|S_m = s_O\right) = \eta\left(1-\alpha\right)\epsilon_{UC} = \mu_{C0}.$$

In general, derivation of calendar period trades is quite complicated, as the informed trade frequencies are not constant. To begin, we assume that the equal payoff condition holds so that the informed trade frequencies are constant throughout the trading day. For simplicity, we assume that the uninformed trade frequencies are equal across assets, so that the each informed trade frequency is $\frac{1}{3}$. We then have that

$$E\left(I_{C_t}|S_m \neq s_O\right) = \eta\left[\frac{\alpha + \left(1-\alpha\right)\epsilon}{3}\right] = \mu_{C1}$$

and

$$E\left(I_{C_t}|S_m = s_O\right) = \eta\left[\frac{\left(1-\alpha\right)\epsilon}{3}\right] = \mu_{C0}.$$

We have

$$Var\left(I_{C_t}|S_m \neq s_O\right) = \eta\left[\frac{\alpha + \left(1-\alpha\right)\epsilon}{3}\right]\left[1 - \frac{\alpha + \left(1-\alpha\right)\epsilon}{3}\right] = \sigma^2_{C1}$$

and

$$Var\left(I_{C_t}|S_m = s_O\right) = \eta\left[\frac{\left(1-\alpha\right)\epsilon}{3}\right]\left[1 - \frac{\left(1-\alpha\right)\epsilon}{3}\right] = \sigma^2_{C0}.$$

Unconditionally, we have

$$EI_{C_t} = \theta\mu_{C1} + (1-\theta)\,\mu_{C0} = \mu_C$$

and

$$Var\left(I_{C_t}\right) = \theta\sigma_{C1}^2 + (1-\theta)\,\sigma_{C0}^2 + \theta\,(1-\theta)\,(\mu_{C1}-\mu_{C0})^2 = \sigma_C^2.$$

We arrive at the following theorem and corollary.

Theorem 4: *Let $r > 0$. If $r < k$, then $I_{C_{t-r}}$ and I_{C_t} are positively serially correlated. If $r \geq k$, then $I_{C_{t-r}}$ and I_{C_t} are uncorrelated. For all r, we have*

$$Cor\left(I_{C_{t-r}}, I_{C_t}\right) = \frac{\theta\,(1-\theta)\left(\frac{\alpha}{3}\eta\right)^2}{\sigma_C^2}\left[\frac{k-\min(k,r)}{k}\right].$$

Proof: *See the Appendix.*

Because the probability of success for a binomial random variable is not the scale of the random variable, the variance of trades in a specific asset is not a scale transformation of the variance of total trades. Thus the correlation in trades for a specific asset is not the same as the correlation in total trades. Because the covariance in trades for a given asset is simply one-ninth of the covariance of total trades, while the variance of trades in an asset is larger than one-ninth of the variance of total trades, the correlation in trades for a specific asset is smaller than is the correlation in total trades.

With our specification of k calendar periods per trading day, the underlying random process generates sets of k observations. For example, if $k = 2$ then there are two calendar period measurements in a trading day. In this case we can heuristically segment the trading day into a morning period and an afternoon period. As we assume that the information arrival process is independent over time, the pair of calendar period measurements corresponding to one trading day is independent of the pair of calendar period measurements corresponding to any other trading day. We do not know in which calendar period—either the morning or the afternoon—an information event occurs, and we assume that t is randomly sampled so that I_{C_t} is equally likely to be a morning or afternoon observation. In this way, the correlation in I_{C_t} is independent of time.

Corollary 5: *Let $r < k$. The positive correlation between $I_{C_{t-r}}$ and I_{C_t} is increasing in α, increasing in k and decreasing in r. The effects of changing the market parameters, ϵ and θ, and the trade aggregation parameter, η, on the positive correlation between $I_{C_{t-r}}$ and I_{C_t} are ambiguous.*

Proof: *Using $\tau = k\eta$, we have*

$$Cor\left(I_{C_{t-r}}, I_{C_t}\right) = \frac{\theta(1-\theta)\,\alpha^2\eta\left(\frac{\tau - r\eta}{\tau}\right)}{\begin{array}{c}(1-\alpha)\,\epsilon\,[1-(1-\alpha)\,\epsilon] + \theta\alpha\,[\alpha\eta(1-\theta) + (1-\alpha)(1-2\epsilon)] + \\ 2\theta\,[\alpha + (1-\alpha)\,\epsilon] + 2(1-\theta)(1-\alpha)\,\epsilon\end{array}}.$$

The comparative static results follow from differentiation.

To explore the comparative static results for the underlying parameters of the market microstructure model—the proportion of informed traders, α, the proportion of uninformed traders who trade, ϵ, and the likelihood of an informative signal, θ—we study the three key components of the $Cor\left(I_{C_{t-r}}, I_{C_t}\right)$. The three components are: $(\mu_{C1} - \mu_{C0})^2$, which reflects the difference in the expected number of call option trades in a given calendar period on a day with $S_m \neq s_O$ compared to a day with $S_m = s_O$, and the conditional variances $(\sigma^2_{C1}, \sigma^2_{C0})$. An increase in $(\mu_{C1} - \mu_{C0})^2$ increases both $Cov\left(I_{C_{t-r}}, I_{C_t}\right)$ and $Var\left(I_{C_t}\right)$, so the overall effect is ambiguous. An increase in either of the conditional variances, σ^2_{C1} or σ^2_{C0}, only leads to an increase in $Var\left(I_{C_t}\right)$, and so it weakens the serial correlation in call option trades. Increasing α leads to a larger difference in the expected number of trades in a given calendar period on a day with $S_m \neq s_O$ compared to a day with $S_m = s_O$, and thus $(\mu_{C1} - \mu_{C0})^2$ increases. In general, it is unclear whether the positive impact on $Cov\left(I_{C_{t-r}}, I_{C_t}\right)$ outweighs the positive impact on $Var\left(I_{C_t}\right)$, and thus increasing α may or may not increase $Cor\left(I_{C_{t-r}}, I_{C_t}\right)$. We do find that $Cor\left(I_{C_{t-r}}, I_{C_t}\right)$ tends to increase as α increases if the proportion of uninformed traders who trade, ϵ, is relatively low and the likelihood of an informative signal, θ, is high. Increasing ϵ has ambiguous effects: if the proportion of uninformed traders who trade is relatively small so that virtually all trades are made by informed traders (if $S_m \neq s_O$), increasing the proportion of uninformed traders who trade has a dilutive effect on the serial correlation, while if the proportion of informed traders is relatively large, then increasing ϵ increases the serial correlation. Increasing the probability of an informative signal, θ, also has ambiguous effects, but it tends to increase the serial correlation in call option trades if ϵ is relatively large and θ is relatively small.

We next analyze the effects on $Cor\left(I_{C_{t-r}}, I_{C_t}\right)$ of the trade aggregation parameters, k and η. As the number of calendar periods in a trading day, k, increases, the serial correlation in trades increases through the heightened impact of the entry and exit of informed traders. However, increasing the number of trader arrivals during a trading day, η, has ambiguous effects. In particular, we find that increasing η increases $Cor\left(I_{C_{t-r}}, I_{C_t}\right)$ when $Cor\left(I_{C_{t-r}}, I_{C_t}\right) \leq 1 - \frac{r}{k-r}$. It is important to note, however, that the maximum number of trades on a trading day is $\tau = k\eta$. If we increase the frequency of calendar period measurement, we necessarily decrease the number of trader arrivals during each calendar period. For example, let us suppose that a trading day coincides with 32.5 calendar hours, or five NYSE trading days each with 6.5 hours per day. We assume that a trading opportunity arises once every minute, so there are 1950 trader arrivals. Initially, we focus on five-minute calendar periods and gather data every five minutes; in this case $\eta = 5$ and $k = 390$. If we instead decide to focus on thirty-minute calendar periods and hence gather data for the same asset every thirty minutes, then $\eta = 30$ and $k = 65$. The overall effect on the serial correlation in the number of calendar period call option trades is ambiguous due to the countervailing pressures and, moreover, is not constant across r.

Finally, we turn our attention to calendar period trades when the frequency of informed trade varies over the trading day. In what follows we do not assume that the equal payoff condition holds nor do we assume that the frequency of uninformed trade is constant across markets. We study the serial correlation properties of stock trades and note that the correlation properties of option trades follow from similar logic. From the analytic and simulation results of Section 2, we have that the frequency of informed trade in the stock market rises as the trading day evolves. To capture this mathematically, for each calendar period in a trading day $j = 1, ..., k$ we have

$$E\left(I_{S_{\mathrm{j}}} | S_m = s_O\right) = \mu_{S0}$$

and

$$E\left(I_{S_{\mathrm{j}}} | S_m \neq s_O\right) = \mu_{Sj},$$

with $0 < \mu_{S0} < \mu_{S1} < \mu_{S2} < ... < \mu_{Sk}$.

The preceding displayed equations provide the conditional expectation of each calendar period in the trading day. As news does not always arrive at the same point in each day, such as the beginning of a day, the random arrival of news is an important feature of actual stock prices. To capture this, we consider a given calendar period t to be drawn at random over the course of the trading day. As such, the unconditional mean of calendar period trades is

$$EI_{S_{\mathrm{t}}} = \theta\overline{\mu_{Sk}} + (1-\theta)\,\mu_{S0},$$

in which

$$\overline{\mu_{Sk}} = \frac{1}{k} \sum_{j=1}^{k} \mu_{Sj}.$$

In deriving the serial correlation properties of $\{I_{S_t}\}_{t \geq 1}$, an important condition emerges that ensures the correlation is positive.

Positive Trade Covariance Condition: *The positive trade covariance condition is said to hold for period $\underline{j}$, with $1 \leq \underline{j} \leq k$, if $\underline{j}$ is the smallest value of j for which*

$$\mu_{Sj} > \theta \overline{\mu_{Sk}} + (1-\theta)\, \mu_{S0}.$$

The positive trade covariance condition is most intuitive for the case $k = 2$. From the structure for the expectation of calendar period trades it follows that μ_{S0} lies below the unconditional mean and μ_{S2} lies above the unconditional mean. Suppose that $t-1$ corresponds to the first calendar interval—the morning—of the trading day. For days without private news, we have $E\left(I_{S_{t-1}} | S_m = s_O\right) = \mu_{S0}$ and $E\left(I_{S_t} | S_m = s_O\right) = \mu_{S0}$. Thus for days on which the morning observation tends to be below the unconditional mean, the afternoon observation also tends to be below the unconditional mean. For days with private news, we have $E\left(I_{S_{t-1}} | S_m \neq s_O\right) = \mu_{S1}$ and $E\left(I_{S_t} | S_m \neq s_O\right) = \mu_{S2}$. While it is clear that the afternoon observation tends to be above the unconditional mean, it is not clear whether $EI_{S_t} < \mu_{S1}$. If the positive trade covariance condition holds (for period 1), then $EI_{S_t} < \mu_{S1}$. As a result, on days with private news both the morning and afternoon observations tend to lie above the unconditional mean and positive serial correlation is assured.

Proposition 6: *Let $r > 0$. The covariance of calendar period stock trades is*

$$\left[\frac{k - \min(k,r)}{k}\right] \left[\begin{array}{c} \theta(1-\theta) \sum_{j=1}^{k-r} \left(\mu_{Sj} - \mu_{S0}\right)\left(\mu_{Sj+r} - \mu_{S0}\right) + \\ \theta^2 \sum_{j=1}^{k} \left(\overline{\mu_{Sk}} - \mu_{Sj}\right)\left(\overline{\mu_{Sk}} - \mu_{Sj+r}\right) \end{array} \right],$$

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 the Positive Trade Covariance Condition holds for period one, then

$$Cov\left(I_{S_{t-r}}, I_{S_t}\right) = \left[\frac{2-r}{2}\right] \left[\begin{array}{c} \theta(1-\theta)\left(\mu_{S1} - \mu_{S0}\right)\left(\mu_{S2} - \mu_{S0}\right) + \\ \theta^2 \left(\overline{\mu_{S2}} - \mu_{S1}\right)\left(\overline{\mu_{S2}} - \mu_{S2}\right) \end{array} \right] \geq 0.$$

3.2. Calendar Period Price Changes

Serial correlation in an asset's squared price changes stems from the information content of trades. The information content of a trade depends on the history of trades and the parameter values. Trade decisions in early economic time contain more information than later trade decisions. For larger α, the information content of a trade decision increases, while for larger ϵ, a decision not to trade carries relatively more information. We attempt to analytically model the covariance structure of calendar period squared price changes and begin with the simplified setting of Kelly and Steigerwald in which there is only a stock market. We define the calendar period price change of the stock as

$$\Delta P_t = \sum_{i=(t-1)\eta+1}^{t\eta} U_i = E\left(V|Z_{t\eta}\right) - E\left[V|Z_{(t-1)\eta}\right],$$

and we note that the number of possible values that ΔP_t can take is the number of possible values that $E\left(V|Z_{t\eta}\right) - E\left[V|Z_{(t-1)\eta}\right]$ can take. While there are 3^η possible trade sequences in this setting, the recursive structure of the model reveals that there are far fewer possible values for ΔP_t—namely, there are $3 + \sum_{j=2}^{\eta}(j+1)$ possible values. For example, for the third trader in the period there are 27 possible trade sequences but only ten possible values for ΔP_t.

We begin with several recursions from Bayes' Rule,

$$a_i = \frac{y_i}{1-x_i-y_i} = \frac{y_{i-1}}{1-x_{i-1}-y_{i-1}} \cdot \begin{cases} f_A \text{ if } D_i = d_A \\ f_N \text{ if } D_i = d_N \\ 1 \text{ if } D_i = d_B \end{cases}$$

and

$$b_i = \frac{x_i}{1-x_i-y_i} = \frac{x_{i-1}}{1-x_{i-1}-y_{i-1}} \cdot \begin{cases} 1 \text{ if } D_i = d_A \\ f_N \text{ if } D_i = d_N \\ f_B \text{ if } D_i = d_B \end{cases},$$

in which

$$f_A = 1 + \frac{\alpha}{(1-\alpha)\,\epsilon_{UA}},$$

$$f_B = 1 + \frac{\alpha}{(1-\alpha)\,\epsilon_{UB}},$$

and

$$f_N = 1 - \frac{\alpha}{\alpha + (1-\alpha)(1-\epsilon)}.$$

These allow us to determine x_i and y_i recursively as

$$y_i = \frac{a_i}{1 + a_i + b_i}$$

and

$$x_i = \frac{b_i}{1 + a_i + b_i}.$$

In general, expressing P_t as a function of all of the underlying parameters is difficult. We simplify the task by noting that

$$a_\eta = a_0 f_A^{\eta_A} \ f_N^{\eta_N}$$

and

$$b_\eta = b_0 f_B^{\eta_B} \ f_N^{\eta_N},$$

in which η_A and η_B are the number of trades at the ask and the bid, respectively, of the first η trades, and $\eta_N = \eta - \eta_A - \eta_B$. If $EV \neq 0$ we can remove the mean from the price and obtain

$$E(V^*|Z_i) = x_i (v_L - EV) + y_i (v_H - EV),$$

from which we can reconstruct actual prices as

$$E(V|Z_i) = E(V^*|Z_i) + EV.$$

If we assume that $\delta = \frac{1}{2}$, then $-(v_L - EV) = (v_H - EV)$. Further, if $v_H - EV \neq 1$ then we can remove the scale from the price and find

$$E(V^{**}|Z_i) = -x_i + y_i,$$

from which we can reconstruct correctly scaled prices as

$$E(V^*|Z_i) = E(V^{**}|Z_i) \cdot (v_H - EV).$$

Thus, we need only study

$$E\left(V^{**}|Z_i\right) = \frac{a_i - b_i}{1 + a_i + b_i}.$$

We consider a simple case as an example. Suppose that all of the parameters are set equal to $\frac{1}{2}$ so that

$$E\left(V^{**}|\eta_A = n, \eta_B = m\right) = \frac{3^{n+m}\left(5^n - 5^m\right)}{2\left(3^\eta\right) + 3^{n+m}\left(5^n + 5^m\right)}.$$

To determine $E\left(V^{**}|Z_i\right)$ we need to weight each possible value by the corresponding probability. We consider the case in which $\eta = 2$ so that there are six unique possibilities. Each trader arrival results in one of three possible outcomes, d_A, d_B, or d_N, and trader arrivals are independent and identically distributed. Thus the distribution of the number of outcomes of each type is multinomial with the density function

$$f\left(n, m, \eta - n - m\right) =$$
$$\frac{n!}{\left(n!\right)\left(m!\right)\left[\left(\eta - n - m\right)!\right]} \cdot \left[P\left(D_i = d_A\right)\right]^n \left[P\left(D_i = d_B\right)\right]^m \left[P\left(D_i = d_N\right)\right]^{\eta - n - m}.$$

The possible outcomes, possible values, and associated probabilities are ennumerated in Table 3.1.

Table 3.1: Values and Probabilities

$(n, m, \eta - n - m)$	**probability**	$\mathsf{E}\left(V^{**}\vert\eta_A = n, \eta_B = m\right)$
$(2, 0, 0)$	$\left[P\left(D_i = d_A\right)\right]^2$	0.857
$(1, 1, 0)$	$2\left[P\left(D_i = d_A\right)\right]\left[P\left(D_i = d_B\right)\right]$	0
$(1, 0, 1)$	$2\left[P\left(D_i = d_A\right)\right]\left[P\left(D_i = d_N\right)\right]$	$\frac{1}{3}$
$(0, 2, 0)$	$\left[P\left(D_i = d_B\right)\right]^2$	-0.857
$(0, 1, 1)$	$2\left[P\left(D_i = d_B\right)\right]\left[P\left(D_i = d_N\right)\right]$	$-\frac{1}{3}$
$(0, 0, 2)$	$\left[P\left(D_i = d_N\right)\right]^2$	0

Let us assume that $EV = 0$ and $v_H - EV = 1$. As $\gamma = \frac{1}{2}$, we have that $P\left(D_i = d_A|S_m = s_O\right) = P\left(D_i = d_B|S_m = s_O\right)$, $P\left(D_i = d_A|S_m = s_H\right) = \frac{1}{2} + P\left(D_i = d_B|S_m = s_H\right)$, and $P\left(D_i = d_A|S_m = s_L\right) = P\left(D_i = d_B|S_m = s_L\right) - \frac{1}{2}$.

Then for the first calendar period

$$E\left(\Delta P_1 | S_m = s_H\right) =$$
$$0.857 \cdot \left[\frac{1}{4} + P\left(D_i = d_B | S_m = s_H\right)\right] + \frac{P\left(D_i = d_N | S_m = s_H\right)}{3} =$$
$$0.405,$$

$$E\left(\Delta P_1 | S_m = s_L\right) =$$
$$0.857 \cdot \left[\frac{1}{4} - P\left(D_i = d_B | S_m = s_L\right)\right] - \frac{P\left(D_i = d_N | S_m = s_L\right)}{3} =$$
$$-0.405,$$

and

$$E\left(\Delta P_1 | S_m = s_O\right) =$$
$$0.857 \cdot \left\{\left[P\left(D_i = d_A | S_m = s_O\right)\right]^2 - \left[P\left(D_i = d_B | S_m = s_O\right)\right]^2\right\} +$$
$$\frac{2P\left(D_i = d_N | S_m = s_O\right)}{3}\left[P\left(D_i = d_A | S_m = s_O\right) - P\left(D_i = d_B | S_m = s_O\right)\right] =$$
$$0.$$

We also have

$$E\left[\left(\Delta P_1\right)^2 | S_m = s_H\right] =$$
$$(0.857)^2 \cdot \left\{\left[P\left(D_i = d_A | S_m = s_H\right)\right]^2 + \left[P\left(D_i = d_B | S_m = s_H\right)\right]^2\right\} +$$
$$\frac{2P\left(D_i = d_N | S_m = s_H\right)}{9}\left[P\left(D_i = d_A | S_m = s_H\right) + P\left(D_i = d_B | S_m = s_H\right)\right] =$$
$$0.34,$$

$$E\left[\left(\Delta P_1\right)^2 | S_m = s_L\right] =$$
$$(0.857)^2 \cdot \left\{\left[P\left(D_i = d_A | S_m = s_L\right)\right]^2 + \left[P\left(D_i = d_B | S_m = s_L\right)\right]^2\right\} +$$
$$\frac{2P\left(D_i = d_N | S_m = s_L\right)}{9}\left[P\left(D_i = d_A | S_m = s_L\right) + P\left(D_i = d_B | S_m = s_L\right)\right] =$$
$$0.34,$$

and

$$E\left[(\Delta P_1)^2 | S_m = s_O\right] =$$
$$(0.857)^2 \cdot \left\{[P(D_i = d_A | S_m = s_O)]^2 + [P(D_i = d_B | S_m = s_O)]^2\right\} +$$
$$\frac{2P(D_i = d_N | S_m = s_O)}{9}[P(D_i = d_A | S_m = s_O) + P(D_i = d_B | S_m = s_O)] =$$
$$0.065,$$

For the second calendar period we note that $\Delta P_2 = P_2 - P_1$ and for each value of P_1 there are six possible values of P_2. To find the expected value of ΔP_2 we first must find $E(\Delta P_2 | P_1 = p_1)$. For example, if the first two trades are at the ask, then $p_1 = 0.857$ and the possible values for P_2 given p_1 are given in Table 3.2.We

Table 3.2: Values and Probabilities with $C_1 = c_A$ and $C_2 = c_A$

$(n, m, \eta - n - m)$	**probability**	$\mathsf{E}(V^{**} \vert \eta_A = n, \eta_B = m)$
$(4, 0, 0)$	$[P(D_i = d_A)]^2$	0.994
$(3, 1, 0)$	$2[P(D_i = d_A)][P(D_i = d_B)]$	0.909
$(3, 0, 1)$	$2[P(D_i = d_A)][P(D_i = d_N)]$	0.939
$(2, 2, 0)$	$[P(D_i = d_B)]^2$	0
$(2, 1, 1)$	$2[P(D_i = d_B)][P(D_i = d_N)]$	0.556
$(2, 0, 2)$	$[P(D_i = d_N)]^2$	0.545

then have

$$E(\Delta P_2 | P_1 = p_1) =$$
$$(0.994 - 0.857) \cdot [P(D_i = d_A)]^2 +$$
$$(0.909 - 0.857) \cdot 2[P(D_i = d_A)][P(D_i = d_B)] +$$
$$(0.939 - 0.857) \cdot 2[P(D_i = d_A)][P(D_i = d_N)] +$$
$$(0.556 - 0.857) \cdot 2[P(D_i = d_B)][P(D_i = d_N)] +$$
$$(0.545 - 0.857) \cdot [P(D_i = d_N)]^2.$$

The signal is incorporated as it was in the case of ΔP_1, in which the probabilities depend on the signal. The unconditional expectation is then

$$E(\Delta P_2 | S_m = s_O) = \sum_{j=1}^{6} E(\Delta P_2 | P_1 = p_j, S_m = s_O).$$

To construct the expected squared price change for the second calendar period, we have

$$
\begin{gathered}
E\left[(\Delta P_2)^2 \,|P_1 = p_1, S_m = s_O\right] = \\
(0.994 - 0.857)^2 \cdot \left[P\left(D_i = d_A|S_m = s_O\right)\right]^2 + \\
(0.909 - 0.857)^2 \cdot 2\left[P\left(D_i = d_A|S_m = s_O\right)\right]\left[P\left(D_i = d_B|S_m = s_O\right)\right] + \\
(0.939 - 0.857)^2 \cdot 2\left[P\left(D_i = d_A|S_m = s_O\right)\right]\left[P\left(D_i = d_N|S_m = s_O\right)\right] + \\
(0.556 - 0.857)^2 \cdot 2\left[P\left(D_i = d_B|S_m = s_O\right)\right]\left[P\left(D_i = d_N|S_m = s_O\right)\right] + \\
(0.545 - 0.857)^2 \cdot \left[P\left(D_i = d_N|S_m = s_O\right)\right]^2.
\end{gathered}
$$

To construct the covariance between squared calendar period price changes, since

$$
Cov\left[(\Delta P_1)^2, (\Delta P_2)^2\right] = E\left[(\Delta P_1)^2 (\Delta P_2)^2\right] - E\left[(\Delta P_1)^2\right] E\left[(\Delta P_2)^2\right],
$$

to find $E\left[(\Delta P_1)^2 (\Delta P_2)^2 \,|S_m = s_O\right]$ we have

$$
\begin{gathered}
E\left[(\Delta P_1)^2 (\Delta P_2)^2\right] = \\
\sum_{j=1}^{6} P\left[(\Delta P_1)^2 = (\Delta p_j)^2\right] (\Delta p_j)^2 \, E\left[(\Delta P_1)^2\right] E\left[(\Delta P_2)^2 \,|P_1 = p_j\right].
\end{gathered}
$$

To provide an idea of the serial correlation patterns in calendar period squared price changes and trades generated by our model, we conduct a set of simulation experiments and simulate sequences of trades and price changes over the course of many information periods. We define a trading day as 32.5 calendar hours, coinciding with a normal NYSE trading week, and assume that a trader arrives to the markets once every five minutes so that there are $\tau = 390$ trader arrivals during a trading day. We measure prices and trades at thirty-minute calendar intervals, corresponding to $\eta = 6$ and $k = 65$, and simulate 195,000 trader arrivals over the course of 500 trading days. For each trading day, m, we randomly determine the signal, S_m, with $\theta = 0.4$ and $\delta = \frac{1}{2}$. The latter specification ensures that asymmetries do not influence the results.

In the set of simulations, we examine the calendar period dynamics of the model over a series of trading days, allowing for random signals and various market conditions. We focus on the results of simulations with $\alpha = 0.2$ and $\epsilon = 0.5$.[10] We

[10] Easley, Kiefer, and O'Hara (1997) estimate $\alpha = 0.17$ and $\epsilon = 0.33$ for an actively traded stock.

focus on the calendar period properties of serial correlation in trades and squared price changes, attempting to reconcile the results of our model with empirical evidence. Throughout, we assume that the frequencies of uninformed trade are equal and that $\delta = \frac{1}{2}$.

We illustrate the basic results of the model in Figures 3.1 and 3.2. Assuming that $\alpha = 0.2$ and $\epsilon = 0.5$, the strike prices of the options are at their respective limits, $\kappa_{C_\mathsf{m}} = v_{L_\mathsf{m}}$ and $\kappa_{P_\mathsf{m}} = v_{H_\mathsf{m}}$, and $\lambda = 1$, we find serial correlation in the total number of trades and serial correlation in an asset's squared price changes, with more persistent serial correlation in total trades.

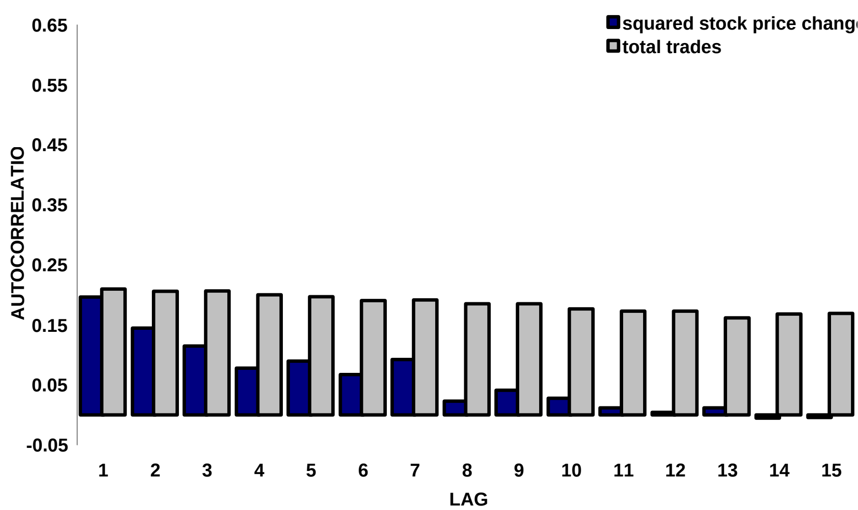

Figure 3.1: Autocorrelation in thirty-minute calendar period squared stock price changes and total trades under random signals, with $k = 65$, $\eta = 6$, $\alpha = 0.2$, $\epsilon = 0.5$, $v_{H_\mathsf{m}} = 100$, $v_{L_\mathsf{m}} = 90$, $\kappa_{C_\mathsf{m}} = 90$, $\kappa_{P_\mathsf{m}} = 100$, and $\lambda = 1$.

We also find evidence of serial correlation in the calendar period trades of an individual asset, indicating that the homogeneous influence of the relatively small proportion of informed traders on trades in an individual asset is not overwhelmed by the noise generated by the uninformed. In accord with empirical features, we find that the serial correlation in calendar period stock trades is more persistent than the serial correlation in squared stock price changes. The serial correlation patterns in calendar period squared price changes for the stock are identical to those for the call and put options, while the serial correlation patterns in calendar period trades for all assets are similar.

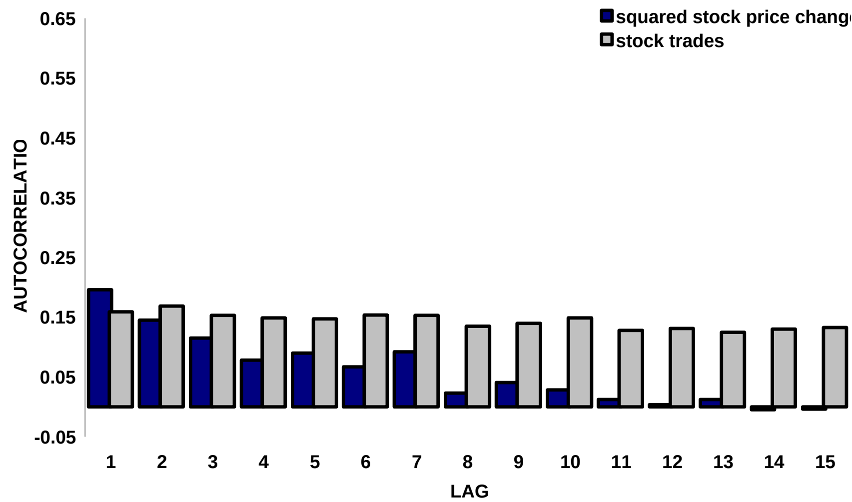

Figure 3.2: Autocorrelation in thirty-minute calendar period squared stock price changes and stock trades under random signals, with $k = 65$, $\eta = 6$, $\alpha = 0.2$, $\epsilon = 0.5$, $v_{H_\mathsf{m}} = 100$, $v_{L_\mathsf{m}} = 90$, $\kappa_{C_\mathsf{m}} = 90$, $\kappa_{P_\mathsf{m}} = 100$, and $\lambda = 1$.

We also analyze the calendar period serial correlation patterns if $\lambda > 1$. In Figures 3.3, 3.4, and 3.5, we illustrate the results for the case in which $\alpha = 0.2$ and $\epsilon = 0.5$, the strike prices of the options are at their respective limits, $\kappa_{C_\mathsf{m}} = v_{L_\mathsf{m}}$ and $\kappa_{P_\mathsf{m}} = v_{H_\mathsf{m}}$, and $\lambda = 1.15$. Comparing Figures 3.1 and 3.3, we find more pronounced serial correlation in calendar period total trades and slightly less persistent serial correlation in calendar period squared stock price changes. Comparing Figures 3.2 and 3.4, we again find evidence of more pronounced serial correlation in calendar period trades for each asset.

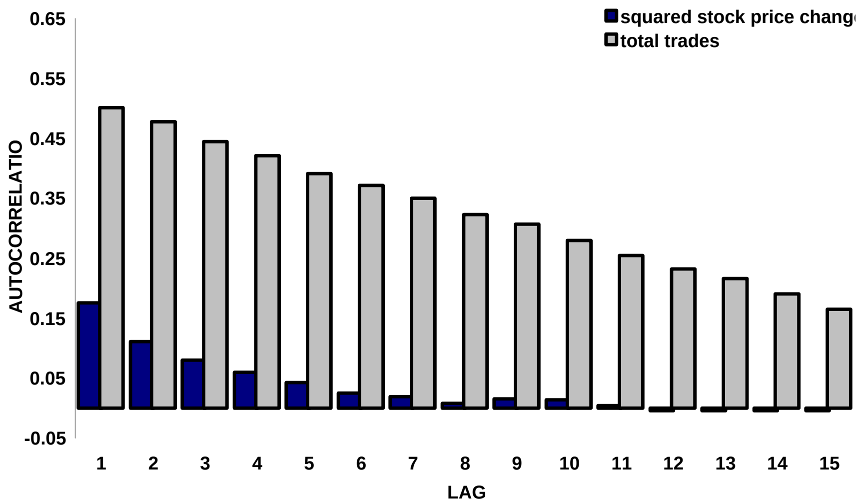

Figure 3.3: Autocorrelation in thirty-minute calendar period squared stock price changes and total trades under random signals, with $k = 65$, $\eta = 6$, $\alpha = 0.2$, $\epsilon = 0.5$, $v_{H_\mathsf{m}} = 100$, $v_{L_\mathsf{m}} = 90$, $\kappa_{C_\mathsf{m}} = 90$, $\kappa_{P_\mathsf{m}} = 100$, and $\lambda = 1.15$.

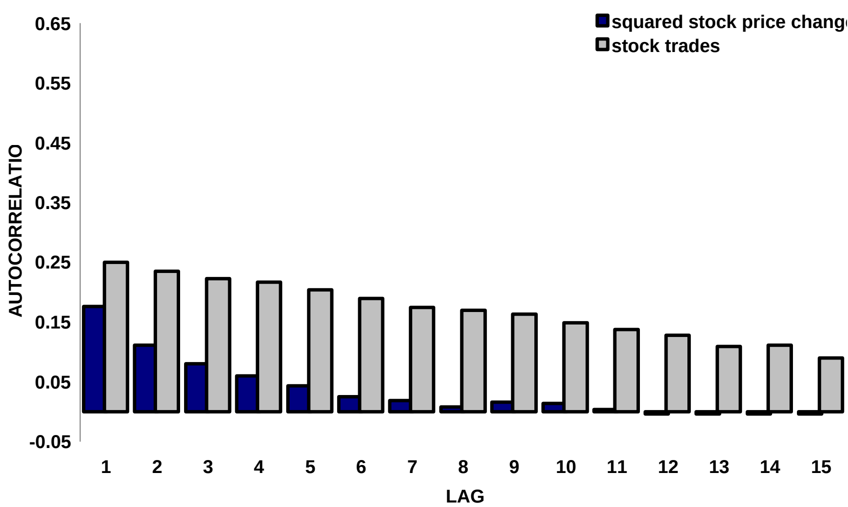

Figure 3.4: Autocorrelation in thirty-minute calendar period squared stock price changes and stock trades under random signals, with $k = 65$, $\eta = 6$, $\alpha = 0.2$, $\epsilon = 0.5$, $v_{H_\mathsf{m}} = 100$, $v_{L_\mathsf{m}} = 90$, $\kappa_{C_\mathsf{m}} = 90$, $\kappa_{P_\mathsf{m}} = 100$, and $\lambda = 1.15$.

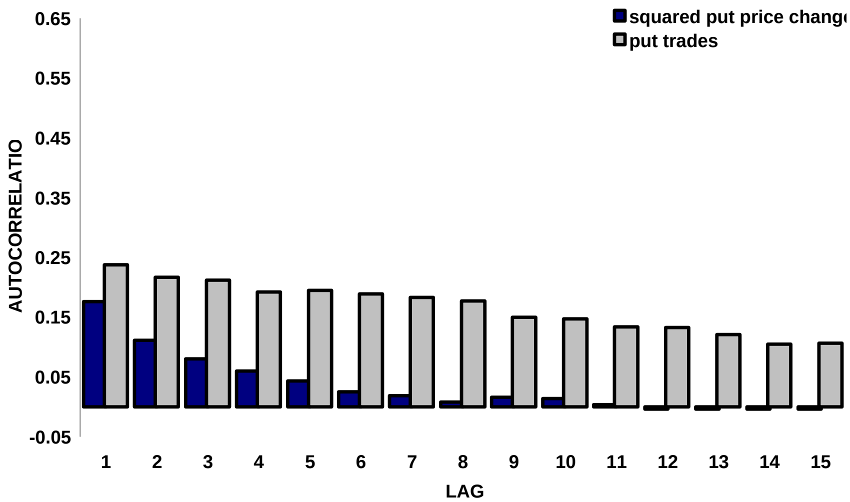

Figure 3.5: Autocorrelation in thirty-minute calendar period squared put price changes and put trades under random signals, with $k = 65$, $\eta = 6$, $\alpha = 0.2$, $\epsilon = 0.5$, $v_{H_{\mathrm{m}}} = 100$, $v_{L_{\mathrm{m}}} = 90$, $\kappa_{C_{\mathrm{m}}} = 90$, $\kappa_{P_{\mathrm{m}}} = 100$, and $\lambda = 1.15$.

4. Conclusions

We focus on the role of private information in the formation of securities prices by developing and analyzing a dynamic, information-based, sequential-trade microstructure model of the markets for a stock and options on that stock. Our model is based on the models of Easley and O'Hara (1992) and Easley, O'Hara, and Srinivas. While such models are gross simplifications, information-based, sequential-trade market microstructure models capture the link between asset prices and informational asymmetries among traders, modeling the bid-ask spread as an adverse selection problem. These models are based on the assertion that trades in an asset are correlated with private information regarding the value of that asset. Such models can offer valuable insight into an asset's price process at a very basic level.

In Easley and O'Hara (1992), as the ask and bid prices converge to the true value of the stock under a given private signal, the bid-ask spread declines over the course of trading as the market maker learns the signal. But in actual markets the arrival and existence of private information is not easily captured, and the theoretical construct of a defined period over which asymmetric information persists is elusive. Moreover, the possibility of multiple, overlapping information events occurring introduces significant complexity. It is not surprising, therefore, that without knowledge of the existence of private information it may be difficult to accurately detect such a pattern in actual data. Further, there is widespread consensus that adverse selection problems faced by market makers are not solely responsible for bid-ask spreads; rather, they are the result of multiple additional factors, including market maker inventory considerations and market power. Nonetheless, our simple economic model provides a theory-based explanation for observed empirical phenomena and, in so doing, establishes an economic foundation for the use of statistical models employed to capture stochastic volatility in asset prices.

References

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

Black, F., 1975. "Fact and Fantasy in the Use of Options," *Financial Analysts Journal* 31, 36-41, 61-72.

[2] Brock, W. and A. Kleidon, 1992. "Periodic Market Closure and Trading Volume: A Model of Intraday Bids and Asks," *Journal of Economic Dynamics and Control* 16, 451-489.

[3] Chan, K., P. Chung, and H. Johnson, 1995. "The Intrday Behavior of Bid-Ask Spreads for NYSE Stocks and CBOE Options," *Journal of Financial and Quantitative Analysis* 30, 329-346.

[4] Cohen, K., S. Maier, R. Schwartz, and D. Whitcomb, 1986. *The Microstructure of Securities Markets*. New Jersey: Prentice-Hall.

[5] Demsetz, H., 1968. "The Cost of Transacting," *Quarterly Journal of Economics* 82, 33-53.

[6] Easley, D., N. Kiefer, and M. O'Hara, 1997. "One Day in the Life of a Very Common Stock," *The Review of Financial Studies* 10, 805-835.

[7] Easley, D. and M. O'Hara, 1987. "Price, Trade Size, and Information in Securities Markets," *Journal of Financial Economics* 19, 69-90.

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

[9] Easley, D., M. O'Hara, and P. Srinivas, 1998. "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance* 53, 431-465.

[10] 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.

[11] Garman, M., 1976. "Market Microstructure," *Journal of Financial Economics* 3, 257-275.

[12] 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.

[13] Harris, L., 1987. "Transaction Data Tests of the Mixture of Distributions Hypothesis," *Journal of Financial and Quantitative Analysis* 22, 127-141.

[14] Harris, M. and A. Raviv, 1993. "Differences of Opinion Make a Horse Race," *The Review of Financial Studies* 6, 473-506.

[15] Jain, P. and G. Joh, 1988. "The Dependence between Hourly Prices and Trading Volume," *Journal of Financial and Quantitative Analysis* 23, 269-284.

[16] Kelly, D. and D. Steigerwald, 2000. "Private Information and High-Frequency Stochastic Volatility," manuscript, University of California, Santa Barbara.

[17] Kyle, A., 1985. "Continuous Auctions and Insider Trading," *Econometrica* 53, 1315-1336.

[18] McInish, T. and R. Wood, 1992. "An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks," *Journal of Finance* 47, 753-764.

[19] O'Hara, M., 1997. *Market Microstructure Theory.* Massachusetts: Blackwell.

[20] Steigerwald, D., 1997. "Mixture Models and Conditional Heteroskedasticity," manuscript, University of California, Santa Barbara.

5. Appendix

Informed Trade Frequencies

We first present the remaining bullish informed trade frequencies. If $S_m = s_H$, then

$$\epsilon_{IA_\mathsf{i}} = \epsilon_{UA} \frac{\left\{ \begin{array}{c} \left[\alpha y_{i-1} + (1-\alpha)\left(\epsilon_{UAC} + \epsilon_{UBP}\right)\right]\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right) - \\ (1-\alpha)\left[\lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UAC} + \lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UBP}\right] \end{array} \right\}}{\alpha y_{i-1}\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UA} + \lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UAC} + \lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UBP}\right]}$$

and

$$\epsilon_{IBP_\mathsf{i}} = \epsilon_{UBP} \frac{\left\{ \begin{array}{c} \left[\alpha y_{i-1} + (1-\alpha)\left(\epsilon_{UA} + \epsilon_{UAC}\right)\right]\lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right) - \\ (1-\alpha)\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UA} + \lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UAC}\right] \end{array} \right\}}{\alpha y_{i-1}\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UA} + \lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UAC} + \lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UBP}\right]}.$$

The bearish informed trade frequencies, which correspond to $S_m = s_L$, are

$$\epsilon_{IB_\mathsf{i}} = \epsilon_{UB} \frac{\left\{ \begin{array}{c} \left[\alpha x_{i-1} + (1-\alpha)\left(\epsilon_{UBC} + \epsilon_{UAP}\right)\right]\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right) - \\ (1-\alpha)\left[\lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UBC} + \lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UAP}\right] \end{array} \right\}}{\alpha x_{i-1}\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UB} + \lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UBC} + \lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UAP}\right]},$$

$$\epsilon_{IBC_\mathsf{i}} = \epsilon_{UBC} \frac{\left\{ \begin{array}{c} \left[\alpha x_{i-1} + (1-\alpha)\left(\epsilon_{UB} + \epsilon_{UAP}\right)\right]\lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right) - \\ (1-\alpha)\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UB} + \lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UAP}\right] \end{array} \right\}}{\alpha x_{i-1}\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UB} + \lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UBC} + \lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UAP}\right]}$$

and

$$\epsilon_{IAP_\mathsf{i}} = \epsilon_{UAP} \frac{\left\{ \begin{array}{c} \left[\alpha x_{i-1} + (1-\alpha)\left(\epsilon_{UB} + \epsilon_{UBC}\right)\right]\lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right) - \\ (1-\alpha)\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UB} + \lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UBC}\right] \end{array} \right\}}{\alpha x_{i-1}\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UB} + \lambda\left(v_{H_\mathsf{m}} - \kappa_{C_\mathsf{m}}\right)\epsilon_{UBC} + \lambda\left(\kappa_{P_\mathsf{m}} - v_{L_\mathsf{m}}\right)\epsilon_{UAP}\right]}.$$

Proof of Theorem 1

We present analysis for $\varepsilon_{IAC_\mathsf{i}}$ and $\varepsilon_{IA_\mathsf{i}}$. Identical logic holds for the remaining informed trade frequencies in the options and stock markets, respectively.

(a) Calculation reveals that $\frac{\partial \varepsilon_{\mathsf{I}\,\mathsf{j}_\mathsf{i}}}{\partial \lambda} > 0$ for j indexing an option trade and $\frac{\partial \varepsilon_{\mathsf{I}\,\mathsf{j}_\mathsf{i}}}{\partial \lambda} > 0$ for j indexing a stock trade. The sign of $\frac{\partial \varepsilon_{\mathsf{IAC}_\mathsf{i}}}{\partial \alpha}$ is the sign of

$$\varepsilon_{UA}\left[\left(v_{H_\mathsf{m}} - v_{L_\mathsf{m}}\right) - \lambda\beta\right], \tag{5.1}$$

which is negative by the greater leverage of options. The sign of $\frac{\partial \varepsilon_{\mathsf{IA_i}}}{\partial \alpha}$ is the sign of

$$\left(\varepsilon_{UAC} + \varepsilon_{UBP}\right)\left[\lambda\beta - \left(v_{H_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right)\right], \tag{5.2}$$

which is positive by the greater leverage of options.

(b) The sign of $\frac{\partial \varepsilon_{\mathsf{IAC_i}}}{\partial y_{\mathsf{i-1}}}$ is the sign of (5.1) while the signs of $\frac{\partial^2 \epsilon_{\mathsf{IAC_i}}}{\partial y^2_{\mathsf{i-1}}}$ and $\frac{\partial^2 \epsilon_{\mathsf{IAC_i}}}{\partial y_{\mathsf{i-1}} \partial \alpha}$ are opposite to the sign of (5.1). The sign of $\frac{\partial \varepsilon_{\mathsf{IA_i}}}{\partial y_{\mathsf{i-1}}}$ is the sign of (5.2) while the signs of $\frac{\partial^2 \epsilon_{\mathsf{IA_i}}}{\partial y^2_{\mathsf{i-1}}}$ and $\frac{\partial^2 \epsilon_{\mathsf{IA_i}}}{\partial y_{\mathsf{i-1}} \partial \alpha}$ are opposite to the sign of (5.2).

(c) Consider $\varepsilon_{IAC_{\mathsf{i}}}$. This informed trade frequency is positive if

$$\varepsilon_{UA}\left[\lambda\left(v_{H_{\mathsf{m}}} - \kappa_{C_{\mathsf{m}}}\right) - \left(v_{H_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right)\right] + \epsilon_{UBP}\left[\lambda\left(v_{H_{\mathsf{m}}} - \kappa_{C_{\mathsf{m}}}\right) - \lambda\left(\kappa_{P_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right)\right] > 0.$$

The first term on the left side is positive because of the greater leverage of options. The second term on the left side is zero because of equal option payoffs. If the uninformed trade each asset with equal frequency, then the remaining inequalities are deduced by inspection of the informed trade frequencies.

(d) For informed trade in the stock market, symmetric option payoffs imply that $\varepsilon_{IA_{\mathsf{i}}}$ is positive if

$$\alpha y_{i-1}\left(v_{H_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right) + (1-\alpha)\left(\epsilon_{UAC} + \epsilon_{UBP}\right)\left[\left(v_{H_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right) - \lambda\beta\right] > 0.$$

Because options offer greater leverage, the second term on the left is negative and the inequality becomes

$$\alpha y_{i-1}\left(v_{H_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right) > (1-\alpha)\left(\epsilon_{UAC} + \epsilon_{UBP}\right)\left[\lambda\beta - \left(v_{H_{\mathsf{m}}} - v_{L_{\mathsf{m}}}\right)\right],$$

from which the bound in the text is easily deduced.

Proof of Theorem 2

For the proof of Theorem 2, let D_j represent $D_i = d_j$. We verify the theorem for U_i, identical logic holds for $U_{C_{\mathsf{i}}}$ and $U_{P_{\mathsf{i}}}$. Proof that $E\left(U_i | Z_{i-1}\right) = 0$ is provided in the text. For general h and i, with $h < i$, the serial correlation in price changes with respect to public information is

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

Thus, price changes are serially uncorrelated with respect to public information.

Recall

$$E\left(U_i^2|Z_{i-1}\right)=\sum_{j=A,B,AC,BC,AP,BP,N} P\left(D_j|Z_{i-1}\right)\left[E\left(V_m|Z_{i-1},D_j\right)-E\left(V_m|Z_{i-1}\right)\right]^2.$$

The upper bound for the conditional variance is

$$\begin{aligned}
E\left(U_i^2|Z_{i-1}\right) &\leq \sum_{j=A,AC,BP} P\left(D_j\right)\left[\hat{A}_i-E\left(V_m|Z_{i-1}\right)\right]^2+\sum_{j=B,BC,AP} P\left(D_j\right)\left[\hat{B}_i-E\left(V_m|Z_{i-1}\right)\right]^2 \\
&\quad +P\left(D_N\right)\left[E\left(V_m|Z_{i-1},D_N\right)-E\left(V_m|Z_{i-1}\right)\right]^2 \\
&\leq \sum_{j=A,AC,BP,N} P\left(D_j\right)\left[\hat{A}_i-E\left(V_m|Z_{i-1}\right)\right]^2+\sum_{j=B,BC,AP,N} P\left(D_j\right)\left[\hat{B}_i-E\left(V_m|Z_{i-1}\right)\right]^2 \\
&\leq \left[\hat{A}_i-E\left(V_m|Z_{i-1}\right)\right]^2+\left[\hat{B}_i-E\left(V_m|Z_{i-1}\right)\right]^2 \\
&\leq \left[\left(\hat{A}_i-E\left(V_m|Z_{i-1}\right)\right)-\left(\hat{B}_i-E\left(V_m|Z_{i-1}\right)\right)\right]^2 \\
&= \left(\hat{A}_i-\hat{B}_i\right)^2,
\end{aligned}$$

where the first inequality follows from the definition of $\hat{A}_i$ and $\hat{B}_i$ and the fourth inequality follows from $B_i \leq E\left(V_m|Z_i\right) \leq A_i$. ■

Proof of Theorem 3

Because the denominator of the learning formulae, conditional on the decision of trader i, is identical for x_i, y_i, and $1-x_i-y_i$, we can construct ratios of x_i and y_i that are recursive linear functions of x_{i-1} and y_{i-1}. 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 $S_m=s_H$, then the relevant ratios are $\frac{x_i}{y_i}$ and $\frac{1-x_i-y_i}{y_i}$; and if $S_m=s_O$, then the relevant ratios are $\frac{x_i}{1-x_i-y_i}$ and $\frac{y_i}{1-x_i-y_i}$.

We present in detail the case $S_m=s_H$, similar logic holds for the other values of S_m. For any $D_i=j$, we have

$$\frac{x_i}{y_i}=\frac{x_{i-1}}{y_{i-1}}\frac{P\left(D_i=j|S_m=s_L\right)}{P\left(D_i=j|S_m=s_H\right)}.$$

If $D_i=d_N$, then

$$\frac{x_i}{y_i}=\frac{x_{i-1}}{y_{i-1}}$$

because $P\left(D_i=d_N|S_m=s_L\right)$ is identical to $P\left(D_i=d_N|S_m=s_H\right)$. Thus

$$\ln\left(\frac{x_i}{y_i}\right)=\ln\left(\frac{x_0}{y_0}\right)+n_A\ln\left[\frac{P\left(D_i=d_A|S_m=s_L\right)}{P\left(D_i=d_A|S_m=s_H\right)}\right]$$

$$+n_B \ln\left[\frac{P\left(D_i = d_B | S_m = s_L\right)}{P\left(D_i = d_B | S_m = s_H\right)}\right] + n_{AC} \ln\left[\frac{P\left(D_i = d_{AC} | S_m = s_L\right)}{P\left(D_i = d_{AC} | S_m = s_H\right)}\right] +$$

$$n_{BC} \ln\left[\frac{P\left(D_i = d_{BC} | S_m = s_L\right)}{P\left(D_i = d_{BC} | S_m = s_H\right)}\right] + n_{AP} \ln\left[\frac{P\left(D_i = d_{AP} | S_m = s_L\right)}{P\left(D_i = d_{AP} | S_m = s_H\right)}\right] + n_{BP} \ln\left[\frac{P\left(D_i = d_{BP} | S_m = s_L\right)}{P\left(D_i = d_{BP} | S_m = s_H\right)}\right]$$

in which n_A is the number of stock trades at the ask in the first i trading opportunities, n_B is the number of stock trades at the bid in the first i trading opportunities, n_{AC} is the number of call option trades at the ask in the first i trading opportunities, n_{BC} is the number of call option trades at the bid in the first i trading opportunities, n_{AP} is the number of put option trades at the ask in the first i trading opportunities, and n_{BP} is the number of put option trades at the bid in the first i trading opportunities.

Because the trader arrival process is i.i.d., as $i \to \infty$ we have

$$\frac{1}{i} \ln\left(\frac{x_i}{y_i}\right) \overset{as}{\longrightarrow} \sum_j P\left(D_i = j | S_m = s_H\right) \ln\left[\frac{P\left(D_i = j | S_m = s_L\right)}{P\left(D_i = j | S_m = s_H\right)}\right].$$

When multiplied by minus one, the right side of this equation is a measure of the distance between the probability measures $P\left(\cdot | S_m = s_H\right)$ and $P\left(\cdot | S_m = s_L\right)$, which is the entropy of $P\left(\cdot | S_m = s_H\right)$ relative to $P\left(\cdot | S_m = s_L\right)$. We denote the entropy distance measure by $\mathcal{I}\left(s_H, s_L\right)$, noting that, by construction, entropy is nonnegative and equals zero only if the probability measures differ solely on a set with measure zero. As $i \to \infty$ we have

$$\frac{1}{i} \ln\left(\frac{x_i}{y_i}\right) \overset{as}{\longrightarrow} -\mathcal{I}\left(s_H, s_L\right).$$

Thus, we find that $\frac{x_i}{y_i}$ behaves as $e^{-i\mathcal{I}(s_H, s_L)}$, so $\frac{x_i}{y_i}$ converges almost surely to zero at the exponential rate $i\mathcal{I}\left(s_H, s_L\right)$. Similarly, we find that $\frac{1-x_i-y_i}{y_i}$ converges almost surely to zero at the exponential rate $i\mathcal{I}\left(s_H, s_O\right)$. Therefore, if $S_m = s_H$ we have

$$\frac{x_i}{y_i} \overset{as}{\longrightarrow} 0$$

and

$$\frac{1 - x_i - y_i}{y_i} \overset{as}{\longrightarrow} 0$$

as $i \to \infty$.

From the convergence properties of these ratios we can deduce the convergence properties of x_i and y_i. Under $S_m = s_H$, $\frac{x_i}{y_i} \xrightarrow{as} 0$ and $\frac{1-x_i-y_i}{y_i} \xrightarrow{as} 0$ imply that $\frac{1}{y_i} - 1 \xrightarrow{as} 0$, which further implies that

$$y_i \xrightarrow{as} 1$$

and, hence,

$$x_i \xrightarrow{as} 0.$$

Then

$$E\left(V_m | Z_{i-1}\right) \xrightarrow{as} v_{Hm},$$

$$E\left(V_{Cm} | Z_{i-1}\right) \xrightarrow{as} v_{Hm} - \kappa_{Cm},$$

and

$$E\left(V_{Pm} | Z_{i-1}\right) \xrightarrow{as} 0.$$

The equal payoff condition implies that

$$\varphi_S\left(\varepsilon_{UAP} + \varepsilon_{UBC}\right) = \lambda\left(\varphi_P \varepsilon_{UAP} + \varphi_C \varepsilon_{UBC}\right).$$

The equal payoff condition, together with convergence of $E\left(V_m | Z_{i-1}\right)$ to v_{Hm}, implies that

$$A_i \xrightarrow{as} v_{Hm} \text{ and } B_i \xrightarrow{as} v_{Hm}.$$

If the equal payoff condition does not hold, then it is most helpful to write the bid as

$$B_i = \frac{\alpha \varepsilon_{IB_i} x_{i-1} v_{L_m} + (1-\alpha)\, \varepsilon_{UB} E\left(V_m | Z_{i-1}\right)}{\alpha \varepsilon_{IB_i} x_{i-1} + (1-\alpha)\, \varepsilon_{UB}}.$$

Because $|\varepsilon_{IB_i}|$ is bounded (between 0 and 1), $B_i \xrightarrow{as} v_{Hm}$.

By similar logic, it follows that

$$A_{C_i} \xrightarrow{as} v_{Hm} - \kappa_{Cm} \text{ and } B_{C_i} \xrightarrow{as} v_{Hm} - \kappa_{Cm},$$

and

$$A_{P_{\mathsf{i}}} \xrightarrow{as} 0 \text{ and } B_{P_{\mathsf{i}}} \xrightarrow{as} 0.$$

■

Proof of Theorem 4

The covariance between $I_{C_{\mathsf{t}}}$ and $I_{C_{\mathsf{t-r}}}$ is

$$Cov\left(I_{C_{\mathsf{t-r}}}, I_{C_{\mathsf{t}}}\right) = E\left(I_{C_{\mathsf{t-r}}}, I_{C_{\mathsf{t}}}\right) - EI_{C_{\mathsf{t-r}}} EI_{C_{\mathsf{t}}}.$$

We first note that $EI_{C_{\mathsf{t-r}}} = EI_{C_{\mathsf{t}}}$. As $EI_{C_{\mathsf{t}}} = \theta\mu_{C1} + (1-\theta)\,\mu_{C0}$ for any t, we have

$$EI_{C_{\mathsf{t-r}}} EI_{C_{\mathsf{t}}} = \left[\theta\mu_{C1} + (1-\theta)\,\mu_{C0}\right]^2.$$

Next we focus on $E\left(I_{C_{\mathsf{t-r}}}, I_{C_{\mathsf{t}}}\right)$. If $r \geq k$, then the independence of the signal process implies that $I_{C_{\mathsf{t-r}}}$ and $I_{C_{\mathsf{t}}}$ are independent. Therefore, we have $E\left(I_{C_{\mathsf{t-r}}}, I_{C_{\mathsf{t}}}\right) = EI_{C_{\mathsf{t-r}}} EI_{C_{\mathsf{t}}}$ and, hence, $Cov\left(I_{C_{\mathsf{t-r}}}, I_{C_{\mathsf{t}}}\right) = 0$. If, however, $r < k$, then there are three possible conditional expectations of $I_{C_{\mathsf{t-r}}} I_{C_{\mathsf{t}}}$ depending on the trading day—or information period—in which $I_{C_{\mathsf{t-r}}}$ and $I_{C_{\mathsf{t}}}$ are measured. With probability $\frac{k-r}{k}$, $I_{C_{\mathsf{t-r}}}$ and $I_{C_{\mathsf{t}}}$ are measured during the same information period, m, and the conditional expectation of $I_{C_{\mathsf{t-r}}} I_{C_{\mathsf{t}}}$ is

$$\theta\mu_{C1}^2 + (1-\theta)\,\mu_{C0}^2.$$

With probability $\frac{r}{k}\theta$, $I_{C_{\mathsf{t-r}}}$ and $I_{C_{\mathsf{t}}}$ are measured over consecutive information periods, m and $m+1$, with $S_{m+1} \neq s_O$, and the conditional expectation of $I_{C_{\mathsf{t-r}}} I_{C_{\mathsf{t}}}$ is

$$\theta\mu_{C1}^2 + (1-\theta)\,\mu_{C0}\mu_{C1}.$$

With probability $\frac{r}{k}(1-\theta)$, $I_{C_{\mathsf{t-r}}}$ and $I_{C_{\mathsf{t}}}$ are measured over consecutive information periods, m and $m+1$, with $S_{m+1} = s_O$, and the conditional expectation of $I_{C_{\mathsf{t-r}}} I_{C_{\mathsf{t}}}$ is

$$\theta\mu_{C0}\mu_{C1} + (1-\theta)\,\mu_{C0}^2.$$

Thus, we have

$$E\left(I_{C_{\mathsf{t-r}}}, I_{C_{\mathsf{t}}}\right) = \left(\frac{k-r}{k}\right)\left[\theta\mu_{C1}^2 + (1-\theta)\,\mu_{C0}^2\right] + \frac{r}{k}\left[\theta\mu_{C1} + (1-\theta)\,\mu_{C0}\right]^2.$$

With the results for $EI_{C_{t-r}}EI_{C_{t-r}}$ and $E\left(I_{C_{t-r}}, I_{C_t}\right)$, we have

$$Cov\left(I_{C_{t-r}}, I_{C_t}\right) = \left(\frac{k-r}{k}\right)\theta\left(1-\theta\right)\left(\mu_{C1}-\mu_{C0}\right)^2 = \left(\frac{k-r}{k}\right)\theta\left(1-\theta\right)\left(\frac{\alpha}{3}\eta\right)^2$$

if $r < k$. Thus,

$$Cor\left(I_{C_{t-r}}, I_{C_t}\right) = \left(\frac{k-r}{k}\right)\frac{\theta\left(1-\theta\right)\left(\frac{\alpha}{3}\eta\right)^2}{\sigma_C^2} \text{ for } r < k$$

and

$$Cor\left(I_{C_{t-r}}, I_{C_t}\right) = 0 \text{ for } r \geq k.$$

As all terms are non-negative, the serial correlation in trades is non-negative. ■

Proof of Proposition 6

We derive $Cov\left(I_{S_{t-r}}, I_{S_t}\right)$ for $k = 2$. Derivation of the general covariance expression follows similar logic. Let $N = 1$ if $t-1$ is the first calendar period in a trading day and $N = 2$ if $t-1$ is the second calendar period. First note that

$$Cov\left(I_{S_{t-r}}, I_{S_t}\right) = E\left(I_{S_{t-r}}I_{S_t}\right) - EI_{S_{t-r}}EI_{S_t},$$

which can be written as

$$Cov\left(I_{S_{t-r}}, I_{S_t}\right) = E\left\{\begin{array}{c}\left[E\left(I_{S_{t-r}}I_{S_t}|N\right) - E\left(I_{S_{t-r}}|N\right)E\left(I_{S_t}|N\right)\right] + \\ \left[EI_{S_{t-r}} - E\left(I_{S_{t-r}}|N\right)\right]\left[EI_{S_t} - E\left(I_{S_t}|N\right)\right]\end{array}\right\},$$

or the sum of the conditional covariance and the covariance of the conditional means. Given that

$$E\left(I_{S_{t-1}}|N=1\right) = \theta\mu_{S1} + \left(1-\theta\right)\mu_{S0} = E\left(I_{S_t}|N=2\right)$$

and

$$E\left(I_{S_{t-1}}|N=2\right) = \theta\mu_{S2} + \left(1-\theta\right)\mu_{S0} = E\left(I_{S_t}|N=1\right),$$

with

$$EI_{S_t} = \frac{\theta}{2}\left(\mu_{S1}+\mu_{S2}\right) + \left(1-\theta\right)\mu_{S0}.$$

Because $P(N=1)=P(N=2)=\frac{1}{2}$, the conditional covariance is

$$
\begin{gathered}
P(N=1)\cdot Cov\left(I_{S_{t-1}}, I_{S_t}|N=1\right)+P(N=2)\cdot Cov\left(I_{S_{t-1}}, I_{S_t}|N=2\right)= \\
\frac{1}{2}\left[E\left(I_{S_{t-1}}I_{S_t}|N=1\right)-E\left(I_{S_{t-1}}|N=1\right)E\left(I_{S_t}|N=1\right)\right]+ \\
\frac{1}{2}\left[E\left(I_{S_{t-1}}I_{S_t}|N=2\right)-E\left(I_{S_{t-1}}|N=2\right)E\left(I_{S_t}|N=2\right)\right]
\end{gathered}
$$

which simplifies to

$$
\frac{1}{2}\theta(1-\theta)(\mu_{S1}-\mu_{S0})(\mu_{S2}-\mu_{S0}).
$$

As $\mu_{S0}<\mu_{S1}<\mu_{S2}$, the conditional covariance is unequivocally positive. The covariance of the conditional means,

$$
E\left\{\left[EI_{S_{t-1}}-E\left(I_{S_{t-1}}|N\right)\right]\left[EI_{S_t}-E\left(I_{S_t}|N\right)\right]\right\},
$$

is

$$
\begin{gathered}
P(N=1)\cdot\left[EI_{S_{t-1}}-E\left(I_{S_{t-1}}|N=1\right)\right]\left[EI_{S_t}-E\left(I_{S_t}|N=1\right)\right]+ \\
P(N=2)\cdot\left[EI_{S_{t-1}}-E\left(I_{S_{t-1}}|N=2\right)\right]\left[EI_{S_t}-E\left(I_{S_t}|N=2\right)\right]
\end{gathered}
$$

and simplifies to

$$
\theta^2\left(\frac{\mu_{S1}-\mu_{S2}}{2}\right)\left(\frac{\mu_{S2}-\mu_{S1}}{2}\right).
$$

As $\mu_{S1}<\mu_{S2}$, the covariance of the conditional means is negative. Ultimately we have

$$
\begin{aligned}
Cov\left(I_{S_{t-1}}, I_{S_t}\right) &= \frac{1}{2}\theta(1-\theta)(\mu_{S1}-\mu_{S0})(\mu_{S2}-\mu_{S0})+ \\
&\quad \theta^2\left(\frac{\mu_{S1}-\mu_{S2}}{2}\right)\left(\frac{\mu_{S2}-\mu_{S1}}{2}\right).
\end{aligned}
$$

We have $Cov\left(I_{S_{t-1}}, I_{S_t}\right)>0$ if $(1-\theta)(\mu_{S1}-\mu_{S0})(\mu_{S2}-\mu_{S0})>\frac{\theta}{2}(\mu_{S2}-\mu_{S1})^2$. By inspection, $\mu_{S2}-\mu_{S0}>\mu_{S2}-\mu_{S1}$, so it is enough to show that

$$
(1-\theta)(\mu_{S1}-\mu_{S0})>\frac{\theta}{2}(\mu_{S2}-\mu_{S1}).
$$

Now, as $\frac{\theta}{2}(\mu_{S2} - \mu_{S1}) = \theta(\overline{\mu_{S2}} - \mu_{S1})$, this is equivalent to showing that

$$(1-\theta)(\mu_{S1} - \mu_{S0}) > \frac{\theta}{2}(\overline{\mu_{S2}} - \mu_{S1}).$$

From the Positive Trade Correlation Condition,

$$(1-\theta)(\mu_{S1} - \mu_{S0}) > \theta(1-\theta)(\overline{\mu_{S2}} - \mu_{S0}).$$

Then

$$\begin{gathered}(1-\theta)(\mu_{S1} - \mu_{S0}) - \theta(\overline{\mu_{S2}} - \mu_{S1}) > \\ \theta(1-\theta)(\overline{\mu_{S2}} - \mu_{S0}) - \theta(\overline{\mu_{S2}} - \mu_{S1}).\end{gathered}$$

The right side of the preceding inequality equals

$$\theta\left[(\mu_{S1} - \mu_{S0}) - \theta(\overline{\mu_{S2}} - \mu_{S0})\right] > 0,$$

and the Postive Trade Correlation Condition implies $(\mu_{S1} - \mu_{S0}) - \theta(\overline{\mu_{S2}} - \mu_{S0}) > 0$.