Identification and Inference for Econometric Models

Essays in Honor of Thomas Rothenberg

Edited by

Donald W. K. Andrews
Yale University

James H. Stock
Harvard University

CAMBRIDGE
UNIVERSITY PRESS

ıool Attendance Affect
)6, 979–1014.
ɔutions of Instrumental

ate on the Dimension,"

ı with a Large Number
Maryland.
ɔf Instruments," *Econo-*

ıe Limited Information

ıriate CLT," *Annals of*

oximations to Tests of

ions of Estimators in a
' *Journal of the Ameri-*

ɔtimators when the De-
e Size," *Econometrica*,

ɔral *k*-Class Estimators
. 27, 575–95.
Limit Theory for Non-

conometric Estimators
ics, Vol. II (ed. by Z.
pp. 881–935.
ıternational Economic

Regression with Weak

s in Linear IV Regres-

s in Linear IV Regres-

Identifying a Source of Financial Volatility

Douglas G. Steigerwald and Richard J. Vagnoni

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. We study identification within the model, paying particular attention to assumptions about the latent trader arrival process. We also derive the stochastic properties of trade-by-trade decisions and prices. Finally, we aggregate trade-by-trade quantities and to show that data generated by the model is consistent with empirical benchmarks from exchange data.

1. INTRODUCTION

Much of Tom Rothenberg's long and insightful career has focused on identification in econometrics. The theme is perhaps most evident in Rothenberg (1973), which has long been the standard for identification in simultaneous equation models. We analyze a market microstructure model, paying particular attention to issues of identification. (The term *micro*structure refers to study of asset markets at the highly disaggregated level corresponding to the arrival of individual traders.) Working from the asymmetric information model in Easley, O'Hara, and Srinivas (1998), we first detail the assumptions needed to identify the parameters. We then derive the stochastic properties of trades and squared price changes for each market and the dynamic pattern of trade across markets. Finally, we use the methods in Kelly and Steigerwald (2004) to construct aggregate trades and squared price changes and compare these to empirical benchmarks. Together, these results provide a theory-based link between asymmetric information, the behavior of market participants, and stochastic volatility.

In Section 2 we first present a model of informed trade in stock and options markets and the resultant likelihood function needed to estimate the parameters. Parameter identification requires specification of the frequency at which traders arrive. We show how misspecification of the arrival frequency imparts bias. In particular, we find that arrival frequency misspecification leads to downward bias of informed trade frequencies. Even with correct specification of the arrival frequency, the likelihood function is sensitive to aggregation and we pinpoint the difficulty. Empirical identification requires a further assumption, by which trades are assigned to a quote. Estimates of the accuracy of the assignment rules typically find an error rate of 15 percent. We determine the bias that arises from such an error rate and again find that informed trade frequencies are biased downward.

In Section 3 we focus on the dynamic pattern of trade within and across markets. We derive (in Theorem 3.1) how frequently, in equilibrium, the informed trade in the options market. Our results nest those of Easley et al. (1998) who implicitly derive conditions under which the informed trade with constant frequency in the options market. We next derive the properties of trade-by-trade price changes. Because informed traders may choose to trade in the options market, option trades can convey information about the stock price (Black 1975; Back 1993; Biais and Hillion 1994). As a result, options are *not* redundant assets as assumed by the Black–Scholes pricing model (Black and Scholes 1973). We detail these linkages and, in Theorem 3.3, we show that the (conditional) variance of price changes in a market is bounded by the squared bid–ask spread for that market. As trade reveals information the bid–ask spread shrinks, thereby reducing the conditional variance. The evolution of the bid–ask spread leads to autocorrelation in the conditional variance, although not specifically of the form modeled in a GARCH process.

In Section 4, we aggregate the trade-by-trade quantities of Section 3 to study the behavior of trades and prices over calendar periods. Three empirical features of stock market data form natural benchmarks for testing the model. 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 larger and to diminish more slowly than serial correlation in squared price changes (Andersen 1996; Harris 1987; Steigerwald 1997). We first show that both trades (or trading volume) and squared price changes are positively correlated. Because the conditional variance of trade-by-trade price changes shrinks as information is revealed through trading, while trade decisions are unaffected, the serial correlation in trades is larger and tends to diminish more slowly than does the serial correlation in squared price changes.

2. IDENTIFICATION IN A MICROSTRUCTURE MODEL WITH 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); Easley et al. (1998). Full details of the model and the derivations that follow are contained in Steigerwald and Vagnoni (2001).

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 $V_m \in \{v_{L_m}, v_{H_m}\}$, with $v_{L_m} < v_{H_m}$. The stock takes the lower value, v_{L_m}, with positive probability δ. Prior to the commencement

ıccuracy of the assignment rules
termine the bias that arises from
:d trade frequencies are biased

ı of trade within and across mar-
ly, in equilibrium, the informed
ıose of Easley et al. (1998) who
ıformed trade with constant fre-
the properties of trade-by-trade
choose to trade in the options
n about the stock price (Black
. a result, options are *not* redun-
ricing model (Black and Scholes
m 3.3, we show that the (condi-
bounded by the squared bid–ask
ıtion the bid–ask spread shrinks,
evolution of the bid–ask spread
ınce, although not specifically of

rade quantities of Section 3 to
ılendar periods. Three empirical
nchmarks for testing the model.
in calendar period squared price
calendar periods, and the serial
be larger and to diminish more
changes (Andersen 1996; Harris
both trades (or trading volume)
related. Because the conditional
inks as information is revealed
ıffected, the serial correlation in
ly than does the serial correlation

TRUCTURE

k and for call and put options on
ıl-trade, asymmetric information
of Easley and O'Hara (1992);
-l and the derivations that follow
)01).
rs over a sequence of trading days,
lizes some per share dollar value,
. with $v_{L_m} < v_{H_m}$. The stock takes
y 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 takes 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 uninformative signal, s_O, provides no information regarding the true value of the stock. Informed traders learn the true value of the stock with probability $\theta > 0$. Proportion α of the traders receives the signal, characterizing the universe of informed traders. The proportion of traders that does not receive the signal characterizes the universe of uninformed 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.[1]

The market makers set an ask and a bid, collectively termed the quotes, for either one share of stock or an option contract that controls $\lambda \geq 1$ shares of the stock. 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. Consider the call option, which provides the owner with the right to buy one share of the stock for a specified strike price, κ_{C_m}, with $\kappa_{C_m} \in \left[v_{L_m}, v_{H_m}\right]$, from the call option writer at the end of the trading day. The value of the call option, V_{C_m}, is $\max\left(V_m - \kappa_{C_m}, 0\right)$.

As all traders are risk neutral, informed traders will trade only if they receive an informative signal. For example, 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 $\epsilon_{IAP} = 1 - \epsilon_{IB} - \epsilon_{IBC}$. Conditional on receiving an informative signal, the informed trader employs the strategy that provides the largest net gain. Uninformed traders are assumed to trade for liquidity reasons and not speculation. The uninformed trade with positive frequency in each market. For example, proportion ϵ_{UB} potentially sells the stock short and proportion ϵ_{UAC} potentially buys λ call options. The sum of the positive frequencies in each market is ϵ, thus $1 - \epsilon$ is the proportion of the uninformed traders that never trade.

Traders randomly arrive to the markets one at a time, so we index them by their order of arrival, i. The ith trader arrives, observes the quotes, and makes a trade decision, D_i. The random variable, D_i, takes one of seven values. For example, if trader i buys the stock at the ask, A_i, then $D_i = d_A$. If trader i writes λ call options at the bid, B_{C_i}, then $D_i = d_{BC}$. If trader i elects not to trade, then $D_i = d_N$. We define the sequence of trading decisions on m as $\{D_k\}_{k=1}^{i}$. Given all publicly available information prior to the commencement of trade on m, Z_0, we specify the publicly available information set prior to the arrival of trader $i+1$ on m 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. The market makers (and uninformed traders) perform Bayesian updating, by which

[1] A trading day captures the interval over which asymmetric information due to a particular signal persists in the markets and is not necessarily coincident with a calendar day.

they learn the signal received by the informed. After witnessing the ith trading decision, the market makers' beliefs regarding the signal that the informed traders received are

$$P(S_m = s_L | Z_i) = x_i \quad \text{and} \quad P(S_m = s_H | Z_i) = y_i.$$

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

Quotes are determined by two equilibrium conditions. The first condition is that a market maker earns zero expected profit from each trade. 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 is that the informed will trade the asset that offers the highest net gain. From the second condition it follows that the quotes are set so that an informed trader earns an equal net gain from each possible trade.

The microstructure model yields the likelihood of each trade decision D_i as a function of the parameters $\Phi = (\alpha, \delta, \bar{\epsilon}, \theta)$, where $\bar{\epsilon}$ is the vector of trade probabilities (for both the informed and uninformed) for each trade decision. As trader arrivals are independent, the likelihood for a sequence of n arrivals is

$$L(\Phi | D_1 = d_1, \ldots, D_n = d_n) = \Pi_{i=1}^{n} P(D_i = d_i | \Phi).$$

From the structure of the model, the probability of each trade decision is straightforward. For example, the probability of a trade at the ask in the stock market is

$$P(D_i = d_A | \Phi) = \theta(1 - \delta)[\alpha\epsilon_{IA} + (1 - \alpha)\epsilon_{UA}] + \theta\delta(1 - \alpha)\epsilon_{UA} + (1 - \theta)(1 - \alpha)\epsilon_{UA}.$$

If $n = (n_A, n_B, \ldots, n_N)$ is the vector of trade counts that correspond to each trade decision, then the corresponding value of the likelihood function is

$$\begin{aligned} L(\Phi) = {} & \theta(1 - \delta)\, p_{1A}^{n_A} \cdot p_{0B}^{n_B} \cdot p_{1AC}^{n_{AC}} \cdot p_{0BC}^{n_{BC}} \cdot p_{1BP}^{n_{BP}} \cdot p_{0AP}^{n_{AP}} [(1 - \alpha)(1 - \epsilon)]^{n_N} \\ & + \theta\delta\, p_{0A}^{n_A} \cdot p_{1B}^{n_B} \cdot p_{0AC}^{n_{AC}} \cdot p_{1BC}^{n_{BC}} \cdot p_{0BP}^{n_{BP}} \cdot p_{1AP}^{n_{AP}} [(1 - \alpha)(1 - \epsilon)]^{n_N} \\ & + (1 - \theta)\, p_{0A}^{n_A} \cdot p_{0B}^{n_B} \cdot p_{0AC}^{n_{AC}} \cdot p_{0BC}^{n_{BC}} \cdot p_{0BP}^{n_{BP}} \cdot p_{0AP}^{n_{AP}} [\alpha + (1 - \alpha)(1 - \epsilon)]^{n_N} \end{aligned}$$

where $p_{1j} = \alpha\epsilon_{Ij} + (1 - \alpha)\epsilon_{Uj}$ and $p_{0j} = (1 - \alpha)\epsilon_{Uj}$ with j indexing trade decisions.

Two assumptions are needed to construct the sequence of trade decisions that identify the parameters. The first assumption identifies the length of time that corresponds to a decision not to trade.[2] As the no-trade decision is designed to isolate periods in which information is not present, the assumption is needed to identify α and ϵ. We first investigate how misspecification of the no-trade interval affects estimation. Let c correspond to the true length of the interval and let $\hat{c}$ correspond to the assumed length of the interval. Because all trades

[2] Specifying the no-trade interval is equivalent to specifying the frequency of trader arrivals.

are observed, only n_N – the number
misspecification. If $c > \hat{c}$, then the nu
if $c < \hat{c}$ the number of no trades is bia
trade decisions are required to record
of the likelihood function, it is not str
bias on individual parameters. To m
and construct estimators under $\hat{c}$. W
later), under which the uninformed
the informed trade frequency in each
use parameter values that correspond
(1997); news arrives on half of the t
more prevelant than good news (δ
($\alpha = .2$), and the overall frequency c
($\epsilon = .8$). The population model ass
a six-hour-trading day, for thirty tra
under each of the alternative assump
four, or five minutes.

As revealed in Panel A of Tabl
interval underestimates the impact
and ϵ is biased upward). The param
θ and δ, are largely invariant to miss
case in which the specified no-trade
no trades declines and days with a
account for the greater relative frequ
To account for the infrequency of
α must decline. If the specified no

Table 7.1. *Impact of misspecificatio*

No-trade interval length	
1 minute	.2
	(.0
2 minutes	.1
	(.0
3 minutes	.1
	[illegible]
4 minutes	.1
	[illegible]
5 minutes	.1
	[illegible]
15% Trade misclassification	.1
	[illegible]

are observed, only n_N – the number of no-trade decisions, is affected by the misspecification. If $c > \hat{c}$, then the number of no trades is biased upward, while if $c < \hat{c}$ the number of no trades is biased downward (as a sequence of actual no-trade decisions are required to record an observed no trade). Given the structure of the likelihood function, it is not straightforward to analytically determine the bias on individual parameters. To measure the bias, we simulate data under c and construct estimators under $\hat{c}$. We use the equal payoff condition (derived later), under which the uninformed trade frequency in each market is $\frac{\epsilon}{6}$ while the informed trade frequency in each market is $\frac{1}{3}$. For the population model we use parameter values that correspond to estimates in Easley, Kiefer, and O'Hara (1997); news arrives on half of the trading days ($\theta = .5$), bad news is slightly more prevelant than good news ($\delta = .6$), 20 percent of traders are informed ($\alpha = .2$), and the overall frequency of trade by uninformed traders is 80 percent ($\epsilon = .8$). The population model assumes a trader arrives every minute during a six-hour-trading day, for thirty trading days. The estimates are constructed under each of the alternative assumptions that a trader arrives every two, three, four, or five minutes.

As revealed in Panel A of Table 7.1, incorrectly specifying the no-trade interval underestimates the impact of informed traders (α is biased downward and ϵ is biased upward). The parameters governing behavior at the daily level, θ and δ, are largely invariant to misspecification of the no-trade interval. For the case in which the specified no-trade interval is too long, the number of recorded no trades declines and days with and without news become more similar. To account for the greater relative frequency of trades on all days, ϵ must increase. To account for the infrequency of no-trade decisions on days without news, α must decline. If the specified no-trade interval is too short, the number of

Table 7.1. *Impact of misspecification on parameter estimates*

Panel A				
No-trade interval length	α	ϵ	θ	δ
1 minute	.2017	.7982	.4667	.5714
	(.0082)	(.0059)	(.0913)	(.1433)
2 minutes	.1633	.8360	.4667	.5714
	(.0095)	(.0064)	(.1173)	(.1409)
3 minutes	.1428	.8583	.4666	.5714
	(.0108)	(.0064)	(.1588)	(.1421)
4 minutes	.1238	.8767	.4667	.5714
	(.0117)	(.0067)	(.2045)	(.1426)
5 minutes	.1114	.8899	.4672	.5712
	(.0104)	(.0063)	(.2011)	(.1405)
Panel B				
15% Trade misclassification	.1862	.7954	.4667	.5714
	(.0093)	(.0063)	(.1021)	(.1435)

recorded no trades increases and, again, days with news become more similar to days without news. Because the relative frequency of trades has declined on all days, ϵ decreases. To account for the infrequency of trade decisions on days with news, α declines. Incorrect specification of the no-trade interval, in either direction, biases the estimator of α downward and makes the presence of informed traders more difficult to detect.

Even if the no-trade interval is correctly specified, empirical identification may be problematic. The analysis of a related likelihood in Easley et al. (1997) is confined to a stock that is not heavily traded. For more heavily traded stocks, numerical difficulties prevent analysis. Rewriting the likelihood makes investigation of the numerical difficulties quite straightforward. Under the equal payoff condition, for which p_{0j} equals $p_0 = (1-\alpha)\frac{\epsilon}{6}$ for all j, the likelihood is

$$p_0^{n-n_N}\,[(1-\alpha)(1-\epsilon)]^{n_N}\cdot\left\{\theta(1-\delta)\left(\frac{\alpha}{3p_0}+1\right)^{n_A+n_{AC}+n_{BP}}\right.$$
$$\left.+\,\theta\delta\left(\frac{\alpha}{3p_0}+1\right)^{n_B+n_{BC}+n_{AP}}+(1-\theta)\left(\frac{\alpha}{c}+1\right)^{n_N}\right\}.$$

The issue concerns the three terms $(\frac{\alpha}{3p_0}+1)^{n_A+n_{AC}+n_{BP}}$, $(\frac{\alpha}{3p_0}+1)^{n_B+n_{BC}+n_{AP}}$, and $(\frac{\alpha}{c}+1)^{n_N}$. For frequently traded stocks, the observed value of trade decisions is quite large. As all three terms are greater than one, these terms dominate the likelihood function when raised to a large power and render the likelihood numerically unstable. (The most common difficulty is simply overflow, the calculated value exceeds the largest number the computer is able to store.) Figures 7.1 and 7.2 reveal the issue. In Figure 7.1, a trader arrives every minute and with 360 trader arrivals in one day no numerical problems are encountered. In Figure 7.2, a trader arrives every twenty seconds, with 1,080 trader arrivals numerical difficulties are prevelant.[3] Because the three terms are increasing functions of α and decreasing functions of ϵ, the likelihood function is correctly computed only for smaller values of α and larger values of ϵ. For the population values $\alpha = .2$ and $\epsilon = .8$ the likelihood function cannot be evaluated with an arrival frequency of twenty seconds.

The second assumption regards the classification of trades. Within the model, all trades occur at a quote. In practice, many trades are recorded at prices between the quotes. To empirically identify the model, all trades must be assigned to a quote. While there are several assignment rules popular in the literature, each of the rules has an estimated error rate of 15 percent. To understand the impact of the misclassification of trades, we randomly misclassify 15 percent of trades. Panel B of Table 7.1 contains the results. Estimation of θ and δ is again largely unaffected. As misclassification of trades does not alter the relative frequency of trades, estimation of ϵ is also unaffected. Yet random misclassification of trades

[3] For ease of viewing, we set the numerically unstable values to an arbitrarily small value, to emphasize that the empirical likelihood is essentially flat.

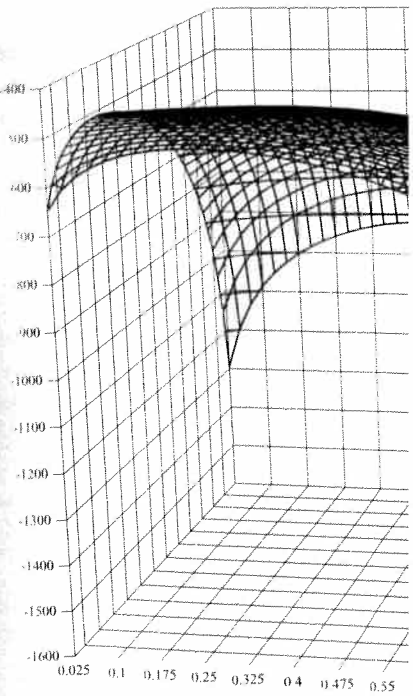

Figure 7.1. Log-likelihood functio

does impact estimation of α. On days wi
is equally likely to affect trades at either
news, for which there are more trades
likely to affect trades at the ask quote
misclassification is more likely to affec
the imbalance of trades (the number of
days is reduced and the presence of info

3. INTRA-TRADING DAY DYN

The evolution of the quotes over the
information revealed through trading. W
day the quotes have bounds that reflect
market makers. We then study the fre
in each market. We show that informed
generally declines over the course of the
the underlying parameters on this frequ
the separating equilibrium derived in Ea
trade only in the options market, will no

ys with news become more similar
: frequency of trades has declined
: infrequency of trade decisions on
lication of the no-trade interval, in
wnward and makes the presence of

specified, empirical identification
d likelihood in Easley et al. (1997)
ed. For more heavily traded stocks,
iting the likelihood makes investi-
ghtforward. Under the equal payoff
$\epsilon)\frac{\epsilon}{6}$ for all j, the likelihood is

$$\delta)\left(\frac{\alpha}{3p_0}+1\right)^{n_A+n_{AC}+n_{BP}}$$

$$1-\theta)\left(\frac{\alpha}{c}+1\right)^{n_N}\bigg\}.$$

$)^{n_A+n_{AC}+n_{BP}}$, $(\frac{\alpha}{3p_0}+1)^{n_B+n_{BC}+n_{AP}}$,
the observed value of trade deci-
ter than one, these terms dominate
e power and render the likelihood
difficulty is simply overflow, the
r the computer is able to store.)
: 7.1, a trader arrives every minute
nerical problems are encountered.
econds, with 1,080 trader arrivals
se the three terms are increasing
ϵ, the likelihood function is cor-
α and larger values of ϵ. For the
hood function cannot be evaluated

cation of trades. Within the model,
des are recorded at prices between
, all trades must be assigned to a
les popular in the literature, each
ercent. To understand the impact
misclassify 15 percent of trades.
mation of θ and δ is again largely
not alter the relative frequency of
andom misclassification of trades

values to an arbitrarily small value, to
al.

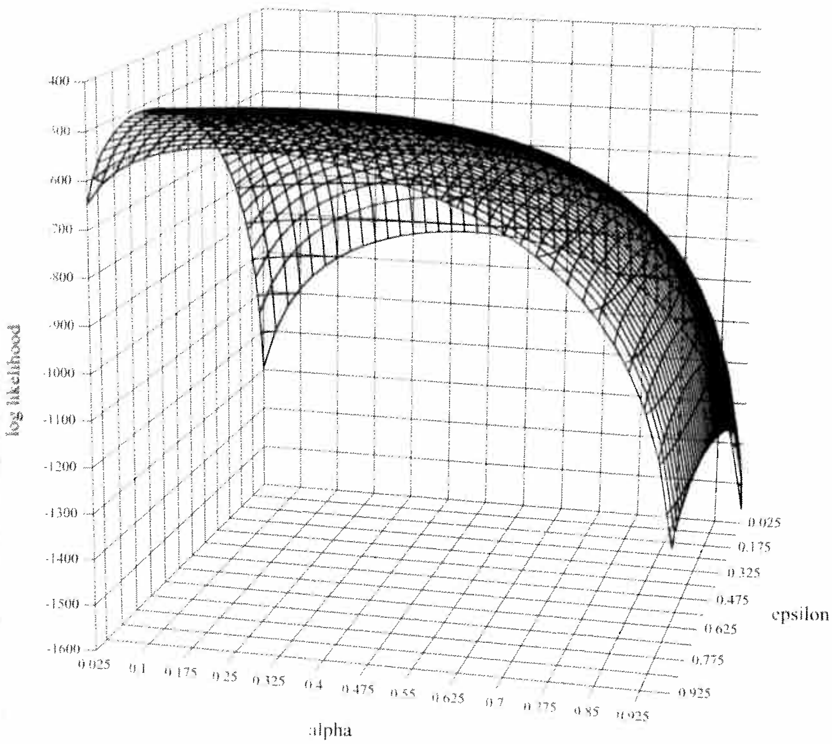

Figure 7.1. Log-likelihood function for 360 arrivals in a trading day.

does impact estimation of α. On days without news, random misclassification is equally likely to affect trades at either set of quotes. Yet on days with good news, for which there are more trades at the ask, misclassification is more likely to affect trades at the ask quotes. Similarly, on days with bad news, misclassification is more likely to affect trades at the bid quotes. As a result, the imbalance of trades (the number of ask trades minus bid trades) on news days is reduced and the presence of informed traders are again hidden.

3. INTRA-TRADING DAY DYNAMICS

The evolution of the quotes over the course of the trading day reflects the information revealed through trading. We show that at each point in the trading day the quotes have bounds that reflect the information asymmetry facing the market makers. We then study the frequency with which the informed trade in each market. We show that informed trade frequency in the options market generally declines over the course of the trading day and we derive the effect of the underlying parameters on this frequency. In doing so, we demonstrate that the separating equilibrium derived in Easley et al. (1998), in which the informed trade only in the options market, will not generally prevail over an entire trading

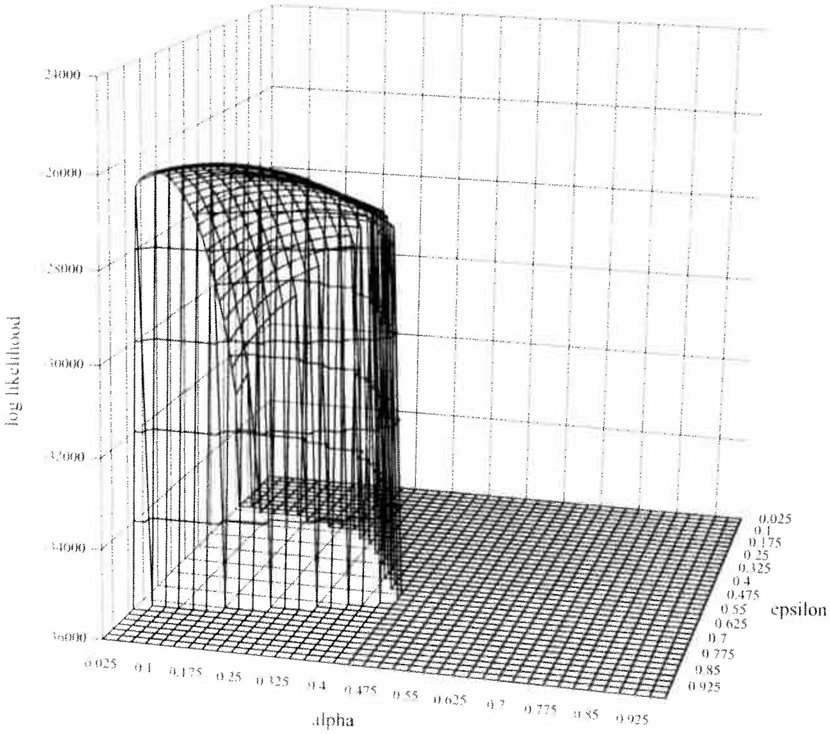

Figure 7.2. log-likelihood function for 1080 trader arrivals.

day. We also show how the bid–ask spread changes as the informed trade frequencies change. The spread in the options market declines more rapidly than does the spread in the stock market, reflecting the flow of informed traders into the stock market. Finally, we derive a condition under which the informed trade with constant frequency in the options market over the course of the trading day; constant informed trade frequencies greatly simplify the analysis of calendar period aggregates in Section 3.

In parallel to the opening quotes for the call, the ith-trade quotes for each asset are obtained as the solution to the zero profit condition with the relevant informed trade frequency. For example, the ith-trade quotes for one share of the stock are

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

and

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

From these equations it is easy to see
the respective limit values of the asset,
maker is certain the informed learn the
We also find that the quotes for the stoc
expected values of the assets, which i
market makers in an effort to offset ex
information.

The quotes process is driven by x_i
beliefs about the signal received by th
according to Bayes' Rule and are determ
formed trade frequencies. In general, inf
out the trading day. The dynamic behavi
is intuitive; as private information is re
gained by the informed through trade ir
the analysis of variable informed trade fr
pirically relevant case in which options
are symmetric, $(\lambda(\kappa_{P_m} - v_{L_m}) = \lambda(v_{H_m}$

Theorem 3.1. *If the options offer greate*
then the informed trade frequencies beh

(a) As λ increases, the informed are
As α increases, the informed
market.
(b) As learning evolves, the inform
stock market. The rate of flow
day. The rate of flow also declin
(c) Informed trade frequencies in t
If the uninformed trade each as
$\epsilon_{IBP_i} > \epsilon_{IA_i}$ and $\epsilon_{IBC_i} = \epsilon_{IAP}$
(d) The ith informed trade frequen
for $j = H, L$,

$$\lambda < \frac{(v_{H_m} - v_{L_m})}{\beta}\left(1 + \frac{\alpha}{1 -}\right.$$

with $\epsilon_H = \epsilon_{UAC} + \epsilon_{UBP}$, $\epsilon_L =$
$b_{L,i-1} = x_{i-1}$.

Proof. See Appendix.

An increase in the proportion of inform
of uninformed traders to informed traders
makes the stock market more attractive t
To understand the dynamic pattern reve

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 market maker 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.

The quotes process is driven by x_i and y_i, which are the market makers' beliefs about the signal received by the informed traders. The beliefs evolve according to Bayes' Rule and are determined in large part by the equilibrium informed trade frequencies. In general, informed trade frequencies vary throughout the trading day. The dynamic behavior of the options market trade frequency is intuitive; as private information is revealed through trading, the advantage gained by the informed through trade in the options market declines. To make the analysis of variable informed trade frequencies concise, we focus on an empirically relevant case in which options offer leverage and the option payoffs are symmetric, $(\lambda(\kappa_{P_m} - v_{L_m}) = \lambda(v_{H_m} - \kappa_{C_m}) \equiv \lambda\beta)$.

Theorem 3.1. *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 $\epsilon_{IAC_i} = \epsilon_{IBP_i} > \epsilon_{IA_i}$ and $\epsilon_{IBC_i} = \epsilon_{IAP_i} > \epsilon_{IB_i}$.*

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

$$\lambda < \frac{(v_{H_m} - v_{L_m})}{\beta}\left(1 + \frac{\alpha}{1-\alpha}\frac{1}{\epsilon_j}b_{j,i-1}\right),$$

with $\epsilon_H = \epsilon_{UAC} + \epsilon_{UBP}$, $\epsilon_L = \epsilon_{UBC} + \epsilon_{UAP}$, $b_{H,i-1} = y_{i-1}$, and $b_{L,i-1} = x_{i-1}$.

Proof. See Appendix. ■

An increase in the proportion of informed traders reduces the depth (the ratio of uninformed traders to informed traders) of the options market, which in turn makes the stock market more attractive to informed traders, as detailed in (a). 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).

Leverage attracts informed traders to the options market. If λ exceeds the separating bound in (d), 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 informed are able to hide more easily in the options market, so the separating bound in (d) decreases and informed trade is more likely to occur only 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.

The bid–ask spread reflects the dynamic pattern of informed trade frequencies. To illustrate the dynamic pattern of the spread, we simulate the arrival of traders over the course of 1,000 trading days on which $S_m = s_H$. We set the information advantage of the informed at 5 percent of the initial value of the asset, so $v_{H_m} = 105$, $v_{L_m} = 95$, and $\delta = .5$. (We ensure that option payoffs are symmetric and set $\kappa_{C_m} = v_{L_m}$ and $\kappa_{P_m} = v_{H_m}$. The greater leverage afforded by options is then captured by $\lambda > 1$.[4]) 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 $\alpha = .2$ and $\epsilon = .75$, noting that the essential features we report hold as α and ϵ vary over $[0, 1]$. In Figure 7.3 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. As the trading day evolves the options spread declines more rapidly than the stock spread, reflecting the movement of informed traders into the stock market.

If the payoff from all three assets is equal, then the informed trade with *constant* frequency throughout the trading day. Because constant informed trade

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

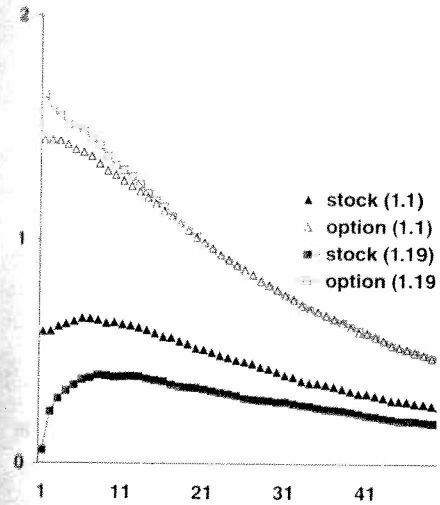

Figure 7.3. Bid–ask spreads with (

frequencies greatly simplify analysis wh gated into calendar periods, we make no

Equal Payoff Condition. The options l

$$v_{H_m} - v_{L_m} = \lambda(v_{H_m} - \kappa_{C_m}) =$$

The constant informed trade frequencie traders in that the informed and uninfor quency in each market

$$\epsilon_{IA_i} = \frac{\epsilon_{UA}}{\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}},$$

and

$$\epsilon_{IBP_i} = \frac{\epsilon_{UBP}}{\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}}.$$

If the informed trade frequencies are recursive. (If the informed trade frequer to obtain a recursive structure.) With cor establish that if there were an infinite n market makers would learn the signal, S_m, converge to the strong-form efficient valu and private information. As transaction p these prices also converge to the respecti assets.

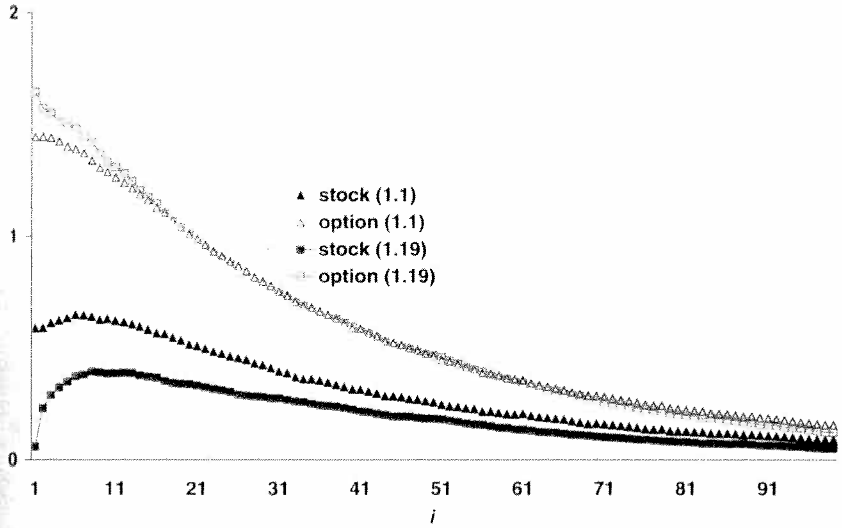

Figure 7.3. Bid–ask spreads with (λ) for $\alpha = .2$ and $\epsilon = .75$.

frequencies greatly simplify analysis when trade-by-trade variables are aggregated into calendar periods, we make note of the condition.

Equal Payoff Condition. The options leverage and strike prices satisfy

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

The constant informed trade frequencies mirror the behavior of uninformed traders in that the informed and uninformed trade with identical relative frequency in each market

$$\epsilon_{IA_i} = \frac{\epsilon_{UA}}{\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}}, \quad \epsilon_{IAC_i} = \frac{\epsilon_{UAC}}{\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}},$$

and

$$\epsilon_{IBP_i} = \frac{\epsilon_{UBP}}{\epsilon_{UA} + \epsilon_{UAC} + \epsilon_{UBP}}.$$

If the informed trade frequencies are constant, then ratios of x_i and y_i are recursive. (If the informed trade frequencies are variable, then it is difficult to obtain a recursive structure.) With constant informed trade frequencies 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 3.2. *If the equal payoff condition is satisfied, then 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 $A_i \xrightarrow{as} v_{L_m}$, $B_i \xrightarrow{as} v_{L_m}$, $A_{C_i} \xrightarrow{as} 0$, $B_{C_i} \xrightarrow{as} 0$, $A_{P_i} \xrightarrow{as} \kappa_{P_m} - v_{L_m}$ and $B_{P_i} \xrightarrow{as} \kappa_{P_m} - v_{L_m}$.
If $S_m = s_H$ then $x_i \xrightarrow{as} 0$, $y_i \xrightarrow{as} 1$, so $A_i \xrightarrow{as} v_{H_m}$, $B_i \xrightarrow{as} v_{H_m}$, $A_{C_i} \xrightarrow{as} v_{H_m} - \kappa_{C_m}$, $B_{C_i} \xrightarrow{as} v_{H_m} - \kappa_{C_m}$, $A_{P_i} \xrightarrow{as} 0$ and $B_{P_i} \xrightarrow{as} 0$.
If $S_m = s_O$ then $x_i \xrightarrow{as} 0$, $y_i \xrightarrow{as} 0$, so $A_i \xrightarrow{as} EV_m$, $B_i \xrightarrow{as} EV_m$, $A_{C_i} \xrightarrow{as} EV_{C_m}$, $B_{C_i} \xrightarrow{as} EV_{C_m}$, $A_{P_i} \xrightarrow{as} EV_{P_m}$ and $B_{P_i} \xrightarrow{as} EV_{P_m}$.

Proof. See 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.

Careful analysis of individual trader price changes reveals three interesting features. First, option trades affect stock prices. Many standard option pricing models assume that the option price is derived from the stock price. Such models are misspecified when informed trade occurs in option markets. Second, price changes are predictable with respect to private information (in contrast to public information). Third, price changes are dependent and heterogenous, and the conditional variance of each price change is bounded by the squared bid–ask spread.

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 $U_i(D_i = d_j) = E(V_m|Z_{i-1}, D_i = d_j) - E(V_m|Z_{i-1})$. Consider a trade at the ask in the stock. Because $E(V_m|Z_i) = x_i v_{L_m} + y_i v_{H_m} + (1 - x_i - y_i)EV_m$, the stock price change is

$$U_i(D_i = d_A) = [v_{H_m} - E(V_m|Z_{i-1})] \frac{\alpha \epsilon_{IA_i} y_{i-1}}{P(D_i = d_A|Z_{i-1})}.$$

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.

Because informed trade occurs in the options market, options are not redundant assets. If trader i elects to buy the call option contract, then

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

Trade in an option affects the price of the stock.

Prices are predictable with respect to private information. Consider the stock price change expected by an informed trader with $S_m = s_H$. The informed trader's expectation differs from that of the market maker because the market

maker is unsure of the signal. The stock p
trader is

$$E(U_i|Z_{i-1}) + \alpha(1 + y_{i-1})[v_H$$
$$+ \alpha x_{i-1}|EV_m - v_{L_m}| > 0.$$

A direct implication is that price change
to private information. If $S_m = s_H$, then
informed trader is

$$E(U_h U_i|Z_{i-1}, S_m = s_H) = U_h E$$

Price changes are conditionally heterc

$$E(U_i^2|Z_{i-1}) = \sum_{j=A,B,AC,BC,AP,BP}$$

As the conditional heteroskedasticity is p
bounds. To do so, we use the effective bi
maximum revision in price resulting from
is simply the bid–ask spread. If, however
and generally made by informed traders
small) then a decision not to trade can yielc
to trade. Hence,

$$\hat{A}_i - \hat{B}_i = \max_{j \in \{AC, BP, N\}} [A_i, E(V$$
$$- \min_{j \in \{BC, AP, N\}} [B_i, E$$

(The effective bid–ask spreads for the call
and $\hat{A}_{P_i} - \hat{B}_{P_i}$, are defined in the same wa

We find that price changes conditional
and not identically distributed, although th
related. An asset's bid–ask spread drives
changes, introducing autoregressive heterc

Theorem 3.3. *Price changes in economic and serially uncorrelated with respect to th*

$$E(U_i^2 \mid Z_{i-1}) \leq (\hat{A}_i - \hat{B}_i)^2, \quad a$$

for $j = C, P$.

Proof. See Appendix.

The fact that the price change variance
spread is an important component of the
and Steigerwald (2004) in the context of a

maker is unsure of the signal. The stock price change expected by an informed trader is

$$E\left(U_i|Z_{i-1}\right) + \alpha\left(1 + y_{i-1}\right)\left[v_{H_m} - E\left(V_m|Z_{i-1}\right)\right] \\ + \alpha x_{i-1}\left[EV_m - v_{L_m}\right] > 0.$$

A direct implication is that price changes are serially correlated with respect to private information. If $S_m = s_H$, then the serial correlation expected by an informed trader is

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

Price changes are conditionally heteroskedastic with

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

As the conditional heteroskedasticity is path dependent, we construct analytic bounds. To do so, we use the effective bid–ask spread, $\hat{A}_i - \hat{B}_i$, which is the maximum revision in price resulting from a trade. In almost all cases, $\hat{A}_i - \hat{B}_i$ is simply the bid–ask spread. If, however, 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. Hence,

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

(The effective bid–ask spreads for the call option and the put option, $\hat{A}_{C_i} - \hat{B}_{C_i}$ and $\hat{A}_{P_i} - \hat{B}_{P_i}$, are defined in the same way.)

We find that price changes conditional on public information are *dependent and not identically distributed*, although they are mean zero and serially uncorrelated. An asset's bid–ask spread drives the conditional variance of its price changes, introducing autoregressive heteroskedasticity.

Theorem 3.3. *Price changes in economic time for each asset are mean zero and serially uncorrelated with respect to the public information set. In addition*

$$E\left(U_i^2 \mid Z_{i-1}\right) \leq \left(\hat{A}_i - \hat{B}_i\right)^2, \quad \textit{and} \quad E\left(U_{j_i}^2 \mid Z_{i-1}\right) \leq \left(\hat{A}_{j_i} - \hat{B}_{j_i}\right)^2$$

for $j = C, P$.

Proof. See Appendix. ∎

The fact that the price change variance is bounded by the effective bid–ask spread is an important component of the model. (This was shown in Kelly and Steigerwald (2004) in the context of a single asset market.) Because the

Table 7.2. $E_H(U_i^2|Z_{i-1}) - E_O(U_i^2|Z_{i-1})$

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

price uncertainty associated with informed trading widens the effective bid–ask spread, Theorem 3.3 suggests that price change behavior is systematically different on days for which the signal is informative.

To show that the price uncertainty is greater on days with an informative signal, we examine the market maker's price uncertainty on a trading day with $S_m = s_H$, $E_H(U_i^2|Z_{i-1})$, relative to the price uncertainty on a trading day with $S_m = s_O$, $E_O(U_i^2|Z_{i-1})$. Straightforward calculations reveal that for the first trader $E_H(U_1^2|Z_{i-1})$ is larger than $E_O(U_1^2|Z_{i-1})$. To determine the sign of $E_H(U_i^2|Z_{i-1}) - E_O(U_i^2|Z_{i-1})$ for $i > 1$, we study the behavior of U_i^2.[5] If α is large, then learning is rapid and largely occurs with the first ten traders. For illustration, in Table 7.2 we calculate $E_H(U_i^2|Z_{i-1}) - E_O(U_i^2|Z_{i-1})$ for $\alpha = .9$, from the exact distributions for U_i^2. We first note that as traders arrive to the market, the market maker learns and the relative price uncertainty decreases. The speed of learning increases as the proportion of uninformed traders who trade, ϵ, decreases. Most importantly, the price uncertainty during a day with an informative signal is always at least as large as the price uncertainty during a day with an uninformative signal.

For smaller values of α, learning is slowed and reduction of an asset's bid–ask spread to zero requires many more trader arrivals. 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. In Figure 7.4 we approximate $E_H(U_i^2|Z_{i-1}) - E_O(U_i^2|Z_{i-1})$ for $\alpha = .2$, with 1,000 simulations. We confirm the results of Table 7.2. Again, learning is more rapid if the uninformed trade with less frequency. Also, we again find that the variance of U_i is higher, uniformly, on a day with an informative signal than it is on a day with an uninformative signal.

4. CALENDAR PERIOD IMPLICATIONS

Aggregation of trader arrivals into calendar periods allows us to compare the model with three empirical benchmarks. For constant informed trade frequencies, we prove that the number of trades has positive serial correlation in each

[5] We assume that the equal payoff condition is satisfied, with $\lambda = 1$.

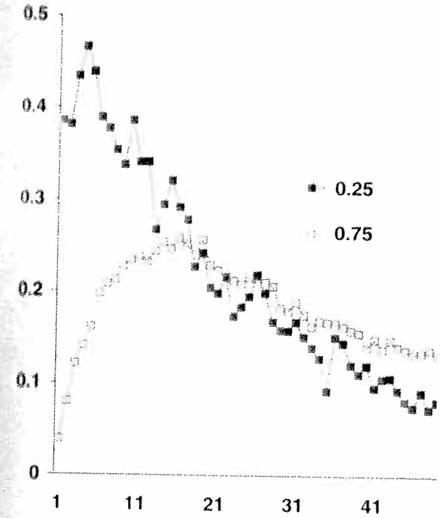

Figure 7.4. $E_H(U_i^2|Z_{i-1}) - E_O(U_i$

market in accord with the first benchma
serial correlation for different levels of ag
rial correlation is higher for data gathere
gathered at hourly intervals. For variable
the formula for trade correlation and a s
to be positive. We then demonstrate that
serially correlated, in accord with the sec
the model is able to satisfy the third benc
in trades that is larger and diminishes mo
squared price changes.

To determine the serial correlation pr
thirty-minute intervals, we divide each tra
let t index, calendar periods. To understa
that $t = 1, \ldots, n$ in which $t = 1$ correspo
trading day. The sample would then consi
drawn from $\frac{n}{k}$ consecutive trading days. E
arrivals (each trader arrival can be though
a given trading day, we have $\tau = k\eta$ trade

We first derive the serial correlation in
the assumption of constant informed tra
are positively correlated. We also find that
market is less than the correlation in total t
markets is not a scale transformation. (Our
the result reported in Kelley and Steigerwa
is analyzed and so informed trade freque
serial correlation directly links the param

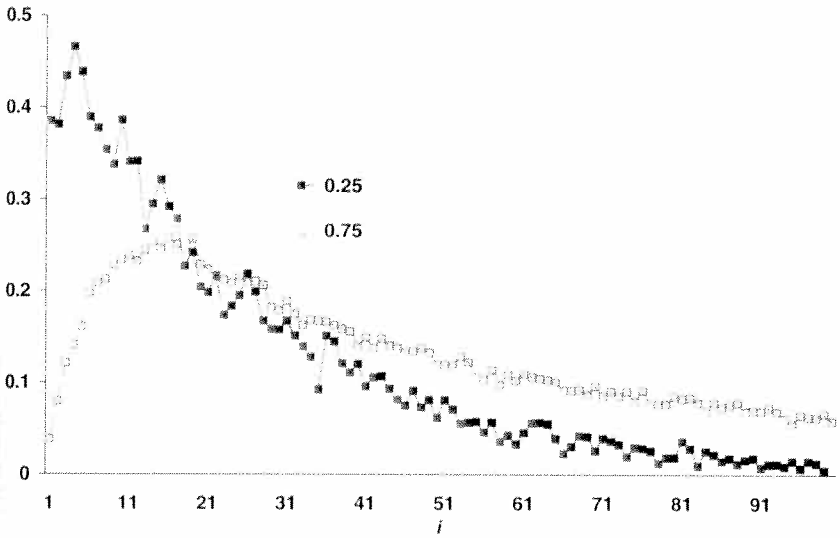

Figure 7.4. $E_H(U_i^2|Z_{i-1}) - E_O(U_i^2|Z_{i-1})$ with $\epsilon = .25$ and $\epsilon = .75$.

market in accord with the first benchmark. We are also able to compare the serial correlation for different levels of aggregation and find that, generally, serial correlation is higher for data gathered at five minute intervals than for data gathered at hourly intervals. For variable informed trade frequencies, we derive the formula for trade correlation and a sufficient condition for the correlation to be positive. We then demonstrate that squared price changes are positively serially correlated, in accord with the second benchmark. Last, we verify that the model is able to satisfy the third benchmark and produce serial correlation in trades that is larger and diminishes more slowly than the serial correlation in squared price changes.

To determine the serial correlation properties for calendar periods, such as thirty-minute intervals, we divide each trading day into k calendar periods. We let t index, calendar periods. To understand how t maps into k and m, suppose that $t = 1, \ldots, n$ in which $t = 1$ corresponds to the first calendar period on a trading day. The sample would then consist of the vectors of k calendar periods drawn from $\frac{n}{k}$ consecutive trading days. Each calendar period contains η trader arrivals (each trader arrival can be thought of as a unit of economic time). For a given trading day, we have $\tau = k\eta$ trader arrivals.

We first derive the serial correlation in trades (per calendar period). Under the assumption of constant informed trade frequencies, we show that trades are positively correlated. We also find that correlation in trades in an individual market is less than the correlation in total trades, as segmenting trades into three markets is not a scale transformation. (Our result for total trades corresponds to the result reported in Kelley and Steigerwald (2004), in which only one market is analyzed and so informed trade frequencies are constant.) The formula for serial correlation directly links the parameters of the market microstructure model to the serial correlation pattern in trades. We analyze the more complex case, in which the informed trade frequency is not constant, in Proposition 4.6. The proposition contains both a formula for trade correlation and a condition that ensures the correlation is positive. The intuition is straightforward: Positive trade correlation arises from the entry and exit of informed traders in response to the arrival of private information.

For correlation in trades in a specific asset, we focus on trades in the call option I_{C_t}. (Analogous results hold for the stock and the put option.) Given η trader arrivals in t, I_{C_t} takes integer values between 0 and η and so is a binomial random variable for which the number of trades in t corresponds to the number of successes in η trials. For each period on trading day m we have

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

and

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

In general, derivation of calendar period trades is quite complicated, as the informed trade frequencies are not constant. To begin, we assume that 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 each informed trade frequency is $\frac{1}{3}$. We then arrive at the following theorem and corollary in which $r > 0$.

Theorem 4.4. *Let the equal payoff condition hold. 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*

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

Proof. Straightforward, but tedious, calculations yield the formula. ■

The correlation in trades for each market is less than the correlation in total trades, which as derived in Kelley and Steigerwald (2004) is

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

The correlations differ because the probability of success for a binomial random variable is not the scale of the random variable, so the variance of trades in a specific asset is not a scale transformation of the variance of total trades.

Corollary 4.5: *Let $r < k$. The p… increasing in α, η, and k. The c… changing the market parameters, … aggregation through $\tau = k\eta$, on th… are ambiguous.*

Proof. The comparative static resu…

As either the frequency of info… η; or the number of calendar per… tion increases through the heighte… traders. In general, increasing the f… correlation, but if ϵ is close to one… impact of the informed trader flow… the probability of an informative s… informative signals are rare. Perha… can compare the serial correlation… lation in five-minute observations… higher in five-minute intervals, but… longer lags, $r \geq \frac{k}{2}$, the serial corre… higher than the serial correlation i…

For the case in which the inform… the day, we focus on trades in the st… we deduce that the frequency of in… trading day evolves. We capture th… (because informed trading frequen…

$$E\left(I_{S_t}|S_m = s_O\right) = \mu_{S0} \text{ f…}$$

and

$$E\left(I_{S_t}|S_m \neq s_O\right) = \mu_{Sj} \text{ f…}$$

with $0 < \mu_{S0} < \mu_{S1} < \mu_{S2} < \cdots$

As any one calendar period is dr… day, the unconditional mean of sto…

$$EI_{S_t} = \theta\overline{\mu_{Sk}} + (1-\theta)\mu…$$

In deriving the serial correlation pr… emerges that ensures the correlatio…

Positive Trade Covariance Condi… *tion is said to hold for period j, wi… for which*

$$\mu_{Sj} > \theta\overline{\mu_{Sk}} + (1-\theta)\mu_{S…}$$

Corollary 4.5: *Let $r < k$. The positive correlation between $I_{C_{t-r}}$ and I_{C_t} is increasing in α, η, and k. The correlation is decreasing in r. The effects of changing the market parameters, ϵ and θ, and of altering calendar period aggregation through $\tau = k\eta$, on the positive correlation between $I_{C_{t-r}}$ and I_{C_t} are ambiguous.*

Proof. The comparative static results follow from differentiation. ■

As either the frequency of informed trade, α, the number of trader arrivals, η; or the number of calendar periods, k, increases, the trade serial correlation increases through the heightened impact of the entry and exit of informed traders. In general, increasing the frequency of uninformed trade reduces serial correlation, but if ϵ is close to one, then further increases in ϵ can amplify the impact of the informed trader flows and increase serial correlation. Increasing the probability of an informative signal, θ, leads to higher serial correlation if informative signals are rare. Perhaps most importantly for empirical work, we can compare the serial correlation in hourly observations with the serial correlation in five-minute observations. We find that serial correlation is generally higher in five-minute intervals, but that the impact is not constant across r. For longer lags, $r \geq \frac{k}{2}$, the serial correlation in five-minute data is unambiguously higher than the serial correlation in hourly data.

For the case in which the informed trade frequencies vary over the course of the day, we focus on trades in the stock market, I_{S_t}. From the results of Section 3, we deduce that the frequency of informed trade in the stock market rises as the trading day evolves. We capture this evolution with a simple structure in which (because informed trading frequencies are zero if $S_m = s_O$)

$$E\left(I_{S_j} | S_m = s_O\right) = \mu_{S0} \text{ for } j = 1, \ldots, k,$$

and

$$E\left(I_{S_j} | S_m \neq s_O\right) = \mu_{Sj} \text{ for } j = 1, \ldots, k,$$

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

As any one calendar period is drawn at random from the periods of a trading day, the unconditional mean of stock trades in a calendar period is

$$EI_{S_t} = \theta\overline{\mu_{Sk}} + (1-\theta)\,\mu_{S0} \text{ with } \overline{\mu_{Sk}} = \frac{1}{k}\sum_{j=1}^{k}\mu_{Sj}.$$

In deriving the serial correlation properties of $\left\{I_{S_t}\right\}_{t\geq 1}$, 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 $E(I_{S_{t-1}}|S_m = s_O) = \mu_{S0}$ and $E(I_{S_t}|S_m = s_O) = \mu_{S0}$. 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 $E(I_{S_{t-1}}|S_m \neq s_O) = \mu_{S1}$ and $E(I_{S_t}|S_m \neq s_O) = \mu_{S2}$. 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 4.6. *Let $r > 0$. The covariance of calendar period stock trades is*

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

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

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

Proof. See Appendix. ■

Serial correlation in squared price changes follows directly from serial correlation in trades if trade-by-trade price changes are i.i.d. As trade-by-trade price changes are not i.i.d, serial correlation in squared price changes is more complex than serial correlation in trades. As a result, formulae linking the parameters of the market microstructure model to the serial correlation are intractable. Positive serial correlation in squared price changes will obtain if squared price changes are higher in periods with higher trading (due to trade by informed traders). We numerically construct the distribution of squared price changes and show that expected squared price changes are higher in periods in which the informed are trading. We then verify that squared price changes are serially correlated, satisfying the second benchmark.

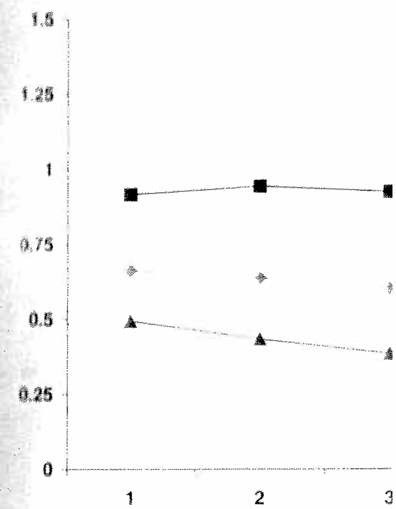

Figure 7.5. Behavior of expec

Serial correlation in an asset's sq
mation content of trades, which in tu
decisions in early economic time c
decisions. We define the (stock) pr
m as

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

A closed-form expression for tl
changes as a function of all of the
tractable. To show that squared pric
we compare price change volatility
change volatility is systematically hi
of information leads to positive seri
illustrate, in Figure 7.5 we present e
days with and without news. A tradi
dar periods, with two trader arrivals
stock market and we set $\alpha = .2$ and
are uniformly higher on news days,
are positively serially correlated.

To show that squared price char
consider a sequence of trading days i

[6] Estimates of $\alpha = .17$ and $\epsilon = .33$ are obtair

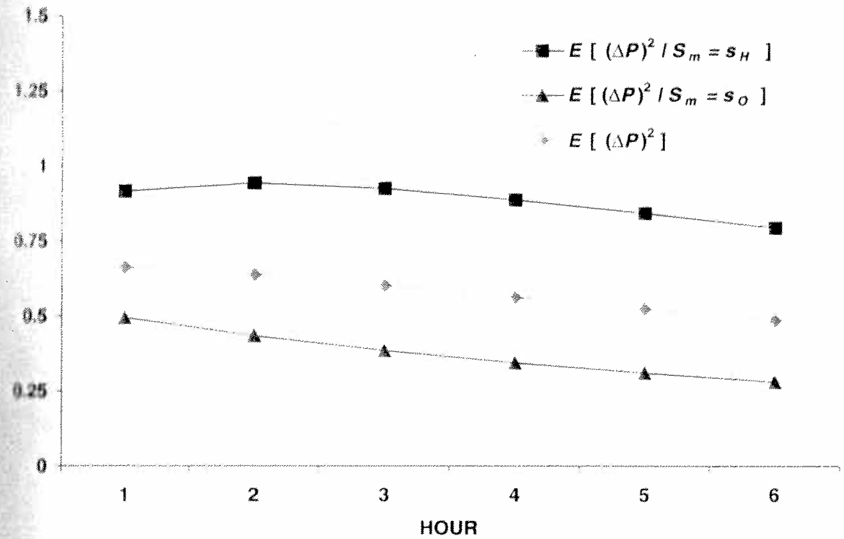

Figure 7.5. Behavior of expected squared price changes.

Serial correlation in an asset's squared price changes stems from the information content of trades, which in turn depends on the history of trades. Trade decisions in early economic time contain more information than later trade decisions. We define the (stock) price change over calendar period t on day m as

$$\Delta P_t = \sum_{i=(t-1)\eta+1}^{t\eta} U_i = E\left(V_m | Z_{t\eta}\right) - E\left(V_m | Z_{(t-1)\eta}\right).$$

A closed-form expression for the population moments of squared price changes as a function of all of the underlying parameters is, in general, intractable. To show that squared price changes have positive serial correlation, we compare price change volatility on days with and without news. If price change volatility is systematically higher on news days, then the random arrival of information leads to positive serial correlation in squared price changes. To illustrate, in Figure 7.5 we present expected squared price changes on trading days with and without news. A trading day is assumed to consists of six calendar periods, with two trader arrivals per period. (In detail, we consider only the stock market and we set $\alpha = .2$ and $\epsilon = .5$.[6]) Expected squared price changes are uniformly higher on news days, which implies that squared price changes are positively serially correlated.

To show that squared price changes are positively serially correlated, we consider a sequence of trading days in which $\theta = .4$ have news (with good and

[6] Estimates of $\alpha = .17$ and $\epsilon = .33$ are obtained in Easley et al. (1997) for an actively traded stock.

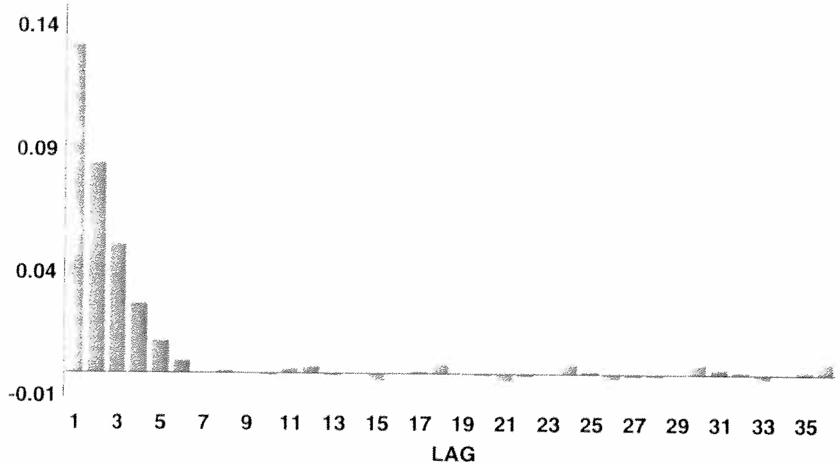

Figure 7.6. Autocorrelation in squared price changes.

bad news equally likely). As is evident in Figure 7.6, the interplay between sequences of squared price changes that lie above the unconditional mean (from news days) with sequences of squared price changes that lie below the unconditional mean (from days without news) leads to positive serial correlation in prices. Further, the serial correlation declines as we move from lag 1 to lag 5, as it is less likely that observations separated by five periods occur on the same trading day. Because the news arrival process is independent across trading days, it would seem that squared price changes are uncorrelated after lag 5. Yet the nonstationarity of the process due to the signal arrival at the start of each trading day leads to correlation in squared price changes at longer lags, which is more pronounced as θ moves away from .5. The first hour of each trading day is noisier than other hours, which leads to serial correlation at lag 6 (and at integer multiples of lag 6) that mirrors the cyclical effects in asset market data.

To verify the third benchmark, we must show that positive serial correlation in trades declines more slowly than does the positive serial correlation in squared price changes. We alter the setting slightly to more closely approximate behavior in a liquid stock traded on the NYSE. We define a trading day to be 32.5 hours, which corresponds to a normal trading week on the NYSE. We measure price and trades at thirty-minute intervals, so there are sixty-five calendar periods in a trading day. A trader arrives every five minutes, so there are six trader arrivals in each calendar period and 390 trader arrivals in a trading day. We simulate 195,000 trader arrivals over the course of 500 trading days.

The strike prices of the options are at their respective limits, $\kappa_{C_m} = v_{L_m} = 95$ and $\kappa_{P_m} = v_{H_m} = 105$, so that $\lambda = 1.15$ captures the greater leverage of an option. In Figure 7.7, we find the positive serial correlation in the total number of trades declines more slowly than does the positive serial correlation in the

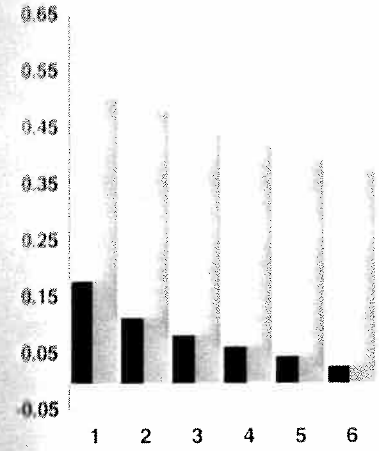

Figure 7.7. Thirty minute au

squared (stock) price changes. A s
trades, rather than total trades, alt
reduced.

5. CONCLUSIONS

We focus on the role of private info
The model captures the link betwe
tries among traders, given a stylize
in actual markets the arrival and e
captured, and the theoretical constr
ric information persists is elusive.
of multiple, overlapping informati
It is not surprising, therefore, that w
information it may be difficult to ac
Further, there is widespread conse
by market makers are not solely r
are the result of multiple additiona
considerations and market power.
provides a theory-based explanatio
so doing, establishes an economic
employed to capture stochastic vol

6. ACKNOWLEDGMENTS

We thank Steve LeRoy and John O

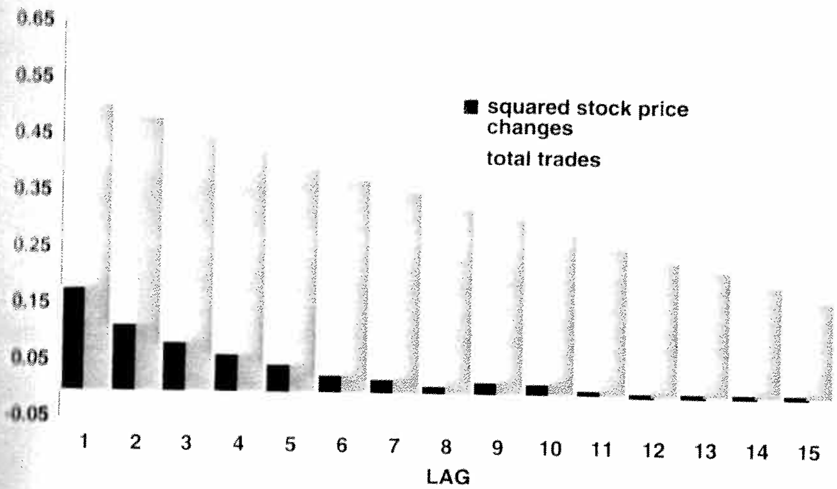

Figure 7.7. Thirty minute autocorrelations ($\lambda = 1.15$).

squared (stock) price changes. A similar picture emerges if we consider stock trades, rather than total trades, although the level of the trade correlation is reduced.

5. CONCLUSIONS

We focus on the role of private information in the formation of securities prices. The model captures the link between asset prices and informational asymmetries among traders, given a stylized arrival process for private information. 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 the occurrence of multiple, overlapping information events 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.

6. ACKNOWLEDGMENTS

We thank Steve LeRoy and John Owens for helpful comments.

APPENDIX

Proof of Theorem 3.1. We present the analysis for ϵ_{IAC_i} and ϵ_{IA_i}. Identical logic holds for the remaining informed trade frequencies in the options and stock markets, respectively.

(a) Calculation reveals that $\frac{\partial \epsilon_{Ij_i}}{\partial \lambda} > 0$ for j indexing an option trade and $\frac{\partial \epsilon_{Ij_i}}{\partial \lambda} > 0$ for j indexing a stock trade. The sign of $\frac{\partial \epsilon_{IAC_i}}{\partial \alpha}$ is the sign of

$$\epsilon_{UA}\left[\left(v_{H_m} - v_{L_m}\right) - \lambda\beta\right],$$

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

$$\left(\epsilon_{UAC} + \epsilon_{UBP}\right)\left[\lambda\beta - \left(v_{H_m} - v_{L_m}\right)\right],$$

which is positive by the greater leverage of options.

(b) The sign of $\frac{\partial \epsilon_{IAC_i}}{\partial y_{i-1}}$ is the sign of the first displayed equation in (a) while the signs of $\frac{\partial^2 \epsilon_{IAC_i}}{\partial y_{i-1}^2}$ and $\frac{\partial^2 \epsilon_{IAC_i}}{\partial y_{i-1}\partial \alpha}$ are opposite to the sign of $\frac{\partial \epsilon_{IAC_i}}{\partial \alpha}$. The sign of $\frac{\partial \epsilon_{IA_i}}{\partial y_{i-1}}$ is the sign of the second displayed equation in (a) while the signs of $\frac{\partial^2 \epsilon_{IA_i}}{\partial y_{i-1}^2}$ and $\frac{\partial^2 \epsilon_{IA_i}}{\partial y_{i-1}\partial \alpha}$ are opposite to the sign of $\frac{\partial \epsilon_{IAC_i}}{\partial \alpha}$.

(c) Consider ϵ_{IAC_i}. This informed trade frequency is positive if

$$\epsilon_{UA}\left[\lambda\left(v_{H_m} - \kappa_{C_m}\right) - \left(v_{H_m} - v_{L_m}\right)\right] + \epsilon_{UBP}\left[\lambda\left(v_{H_m} - \kappa_{C_m}\right) - \lambda\left(\kappa_{P_m} - v_{L_m}\right)\right] > 0.$$

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 ϵ_{IA_i} is positive if

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

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

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

from which the bound in the text is easily deduced. ■

Proof of Theorem 3.2. The proof follows from calculations similar to those in Kelley and Steigerwald (2004) [**Theorem 3.1**].

Proof of Theorem 3.3. For the proof of Theorem 3.3, let D_j represent $D_i = d_j$. We verify the theorem for U_i; identical logic holds for U_{C_i} and U_{P_i}. Proof that $E(U_i|Z_{i-1}) = 0$ and $E(U_h U_i|Z_{i-1}) = 0$ is straightforward. The upper bound

for the conditional variance is

$$E\left(U_i^2|Z_{i-1}\right) \le \sum_{j=A,AC,BP} F$$
$$+ \sum_{j=B,BC,AI}$$
$$+ P(D_N)[I$$
$$\le \sum_{j=A,AC,BP,N}$$
$$+ \sum_{j=B,BC,AP}$$
$$\le \left[\hat{A}_i - E(V_m\right.$$
$$\le \left[(\hat{A}_i - E(V\right.$$
$$= \left(\hat{A}_i - \hat{B}_i\right)^2,$$

where the first inequality follows from inequality follows from $B_i \le E(V_m|Z$

Proof of Proposition 4.6. We derive C general covariance expression follows first calendar period in a trading day a period. First note that

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

or the sum of the conditional covarianc means. Given that

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

and

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

Because $P(N=1) = P(N=2) = \frac{1}{2}$,

$$P(N=1)\cdot \mathrm{Cov}\left(I_{S_{t-1}}, I_{S_t}|N=1\right) + P$$
$$= \tfrac{1}{2}\left[E(I_{S_{t-1}} I_{S_t}|N=1) - E(I_{S_{t-1}}|N\right.$$
$$+ \tfrac{1}{2}\left[E(I_{S_{t-1}} I_{S_t}|N=2) - E(I_{S_{t-1}}|\right.$$

which simplifies to

$$\tfrac{1}{2}\theta(1-\theta)\left(\mu_{S1} - \mu_{S0}\right)\left(\mu_{S2} -\right.$$

for the conditional variance is

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

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

Proof of Proposition 4.6. We derive $\mathrm{Cov}(I_{S_{t-r}}, I_{S_t})$ 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

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

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

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

and

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

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

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

which simplifies to

$$
\tfrac{1}{2}\theta(1-\theta)\left(\mu_{S1} - \mu_{S0}\right)\left(\mu_{S2} - \mu_{S0}\right).
$$

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

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

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

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

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

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

From the positive trade correlation condition,

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

Then

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

The right side of the preceding inequality equals

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

which is positive by the positive trade correlation condition. ■

References

Andersen, T. (1996), "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility," *Journal of Finance*, 51, 169–204.

Back, K. (1993), "Asymmetric Information and Options," *Review of Financial Studies*, 6, 435–72.

Biais, B., and P. Hillion (1994), "Insider and Liquidity Trading in Stock and Options Markets," *Review of Financial Studies*, 7, 743–80.

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

Black, F., and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–54.

Easley, D., N. Kiefer, and M. O'H
Stock," *The Review of Financial*
Easley, D., and M. O'Hara (1992), "
Journal of Finance, 47, 577–605
Easley, D., M. O'Hara, and P. Srin
dence on Where Informed Trade
Harris, L. (1987), "Transaction Dat
Journal of Financial and Quanti
Kelly, D., and D. Steigerwald (2004)
tic Volatility," *Studies in Nonline*
Rothenberg, T. (1973), *Efficient Est*
tion Monograph 23, New Haven:
Steigerwald, D. (1997), "Mixtur
manuscript, University of Califor
Steigerwald, D., and R. Vagnoni (20
Volatility," web manuscript, Univ

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, 805–35.

Easley, D., and M. O'Hara (1992), "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 577–605.

Easley, D., M. O'Hara, and P. Srinivas (1998), "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance*, 53, 431–65.

Harris, L. (1987), "Transaction Data Tests of the Mixture of Distributions Hypothesis," *Journal of Financial and Quantitative Analysis*, 22, 127–41.

Kelly, D., and D. Steigerwald (2004), "Private Information and High-Frequency Stochastic Volatility," *Studies in Nonlinear Dynamics and Econometrics*, 8, 1–30.

Rothenberg, T. (1973), *Efficient Estimation with A Priori Information*. Cowles Foundation Monograph 23, New Haven: Yale University.

Steigerwald, D. (1997), "Mixture Models and Conditional Heteroskedasticity," manuscript, University of California, Santa Barbara.

Steigerwald, D., and R. Vagnoni (2001), "Option Market Microstructure and Stochastic Volatility," web manuscript, University of California, Santa Barbara.