ELSEVIER Journal of Empirical Finance 4 (1997) 159-186

Journal of EMPIRICAL FINANCE

The information content of the trading process [1]

David Easley [a], Nicholas M. Kiefer [a], Maureen O'Hara [b,*]

[a] *Department of Economics, Cornell University, Ithaca, NY 14853-4201, USA*

[b] *Johnson Graduate School of Management, Cornell University, 516 Malott Hall, Ithaca, NY 14853-4201, USA*

Abstract

The trade process is a stochastic process of transactions interspersed with periods of inactivity. The realizations of this process are a source of information to market participants. They cause prices to move as they affect the market maker's beliefs about the value of the stock. We fit a model of the trade process that allows us to ask whether trade size is important, in that large and small trades may have different information content (they do, but this varies across stocks); whether uninformed trade is i.i.d. (it is not); and, whether large buys and large sells are equally informative (they differ only marginally). The model is fitted by maximum likelihood using transactions data on six stocks over 60 days. © 1997 Elsevier Science B.V.

JEL classification: C12; C51; G20; G14

Keywords: Market microstructure; Model estimation

1. Introduction

Market participants draw inferences about the future value of an asset from its trading history. A preponderance of buy orders may signal good news, causing traders to raise their expected value for the asset. A preponderance of sell orders may induce the opposite revision. At a basic level, this learning from the direction

* Corresponding author. E-mail: ohara@johnson.cornell.edu.

[1] We thank Richard Lyons and participants at the High Frequency Data in Finance conference for helpful comments, and Joseph Paperman and Fan Xu for programming assistance. We also thank three anonymous referees and the editor of this journal, Richard Baille, for many helpful suggestions. This research is supported by National Science Foundation Grant SBR93-20889.

PII S0927-5398(97)00005-4

of trades is the fundamental insight of the theoretical microstructure literature, see Kyle (1985) and Glosten and Milgrom (1985). Much more happens in the market, however, than just the arrival of buy and sell orders. Orders may arrive for different quantities of the asset. Does the information content of trades differ depending on their size? During some periods of the trading day, little if any trading may occur. Does the timing of trade provide information to market watchers? Often, a series of buys and sells arrive. Does the sequence of trades provide information beyond that conveyed in the individual orders?

Addressing these issues has taken on increased importance with the advent of high frequency data sets in finance. What is immediately apparent in the data is that patterns in trades abound, and that the natural spacing of trade observations is far from the systematic pattern so desirable for econometric analysis. These patterns in trades, in turn, induce patterns in prices, and so consequently impair our ability to understand price behavior. How to deal with these problems has not been apparent, in part because the information conveyed by the trading process has not been understood.

In this paper we model and estimate the information content of the trading process. Our approach uses the structural framework of microstructure models to formulate the inference problem linking the trading process and beliefs. By estimating this learning problem, we determine the effects on the market maker's beliefs of various types of trade process information. Previous empirical research (see Hasbrouck (1988, 1991); Jones et al. (1994)) has found that *trades* provide substantial insight into the subsequent behavior of prices. What we demonstrate in this research is that there is a wealth of information contained in the *trading process*, and the use of a structural model allows us to extract and interpret this information. Perhaps equally important, our work demonstrates that certain features of the trade process are not particularly informative, giving guidance to researchers regarding what can be excluded in the design of empirical procedures.

Our approach is related to a previous paper (Easley et al. (1993)) in which we demonstrated how the numbers of buys, sells, and no-trading intervals could be used to estimate the parameters of a simple market microstructure model. In this research, we develop an estimation technique to determine the information content of various trade data. Underlying our estimation is the probabilistic structure of trading, and so we first formulate a general model of the trading process. In standard microstructure models, this trading process is quite simple. For example, in both the Kyle (1985) and the Glosten and Milgrom (1985) models, the behavior of the uninformed traders is assumed to be serially independent [2].

[2] In the Kyle (1985) model, for example, the uninformed trades are simply a draw from a mean zero normal distribution. In the multiperiod version of the model these draws are assumed independent across periods. In the Glosten and Milgrom (1985) model, uninformed trades are assumed to be i.i.d. draws from a fixed population. Both models predict, therefore, serially uncorrelated uninformed trades.

Any serial correlation in trades is then interpreted as arising from the trading patterns of informed traders, who either only buy or only sell depending upon their information. If uninformed traders exhibit more complex trading behavior, however, then these microstructure models are mis-specified, and their predictions regarding the information content of specific trade flows may be misleading.

Our model of the trading process avoids this difficulty by incorporating greater complexity in the behavior of traders. We allow the arrival of uninformed traders and their decisions to buy or sell the asset to follow history dependent processes. Alternative specifications, such as the case of serially independent uninformed trade, can then be tested as restrictions of this general model. The model also includes different trade sizes, allowing us to estimate directly the information content of large versus small trades. Further, our model allows testing of differential information content between large buys and large sells, and permits investigation of the informational differences between trade sequences and reversals.

The resulting model has eleven parameters which we estimate using trade data for six stocks. Our analysis reveals three major results. First, we show that trade size generally has information content, but that these trade size effects vary widely. We show that, while for some stocks large trades and small trades have virtually identical effects on beliefs, overall, large trades have approximately twice the information content of small trades. Previous work on trade size using price data has also found significant effects (see Keim and Madhavan, 1994; Seppi, 1992; Barclay and Warner, 1993), but our trade-based estimates provide a means of determining how much of this is directly due to information, and not to other factors such as market maker inventory.

Second, we find that the trading behavior of uninformed traders is highly history dependent. Uninformed traders are more likely to be active when trades have recently occurred, and they are more likely to sell (buy) when the last trade was a sell (buy). This behavior implies that sequences of trades are less informative than is predicted by standard microstructure models, but that reversals in order flow are very informative. We also find autocorrelation in trades, inducing autocorrelation in prices, arising from the history dependent behavior of uninformed traders and the obviously correlated behavior of informed traders. This result complements the finding by Engle and Russell (1995) of complicated serial dependence in time between trades, and provides in addition a possible structural interpretation. This dependence is important because it invalidates standard tests of microstructure models, originally developed by Glosten and Harris (1988) and extended by Berndardt and Hugheson (1993), which rely on the independence of the error term and future trades. Our estimates thus reveal a complexity to trade and trader behavior not recognized in previous research.

Third, we show how informed and uninformed trading differs across stocks. Because we directly estimate the trade process, our results include estimates of informed and uninformed trade propensities, and this allows us to calculate the probability of information-based trading in individual stocks. Our results here provide insight into the 'fine details' underlying the stochastic process of trade, and they suggest the richness of the information that can be gleaned from the trade process.

The paper is organized as follows. In Section 2, we develop the trade model and the price process, and we examine alternative trade process specifications. In Section 3, we formulate the resulting likelihood function, and we discuss the procedures for estimating the stochastic process. In Section 4, we describe the data, and we present estimation results for our general model and for restricted versions of the model. In this section we examine evidence on the information content of trade size, and on the differential behavior of large buy orders and large sell orders. We also investigate the differences in informed and uninformed trade behavior for the stocks in our sample. In Section 5, we examine the robustness of our estimates by investigating several specification tests of our underlying model. In Section 6, we continue this robustness investigation by examining the relation between our trade-based estimates and price behavior. The paper's final section discusses the applications of this approach, the implications of our results, and directions for future research.

2. The trade model

2.1. Trade structure

Individuals trade a single risky asset and money with a market maker over $i = 1, \ldots, I$ trading days. Prior to the beginning of any trading day, nature determines whether an information event relevant to the value of the asset will occur. Information events are independently distributed and occur with probability α. These events are good news (g) with probability $1 - \delta$, or bad news (b) with probability δ. After the end of trading on any day, and before nature moves again, the full information value of the asset is realized.

Let $(V_i)_{i=1}^{I}$ be the random variables giving the value of the asset at the end of trading days $i = 1, \ldots, I$. These values will naturally be correlated. For example, it is likely that the realized value of V_{i-1} would be the prior expected value of the asset at the beginning of day i. We do not make any specific assumptions about the correlations as they are not needed for our analysis. We let the value of the asset conditional on good news on day i be $\overline{V}_i$; similarly it is $\underline{V}_i$ conditional on bad news on day i. The value of the asset if no news occurs on day i is denoted V_i^*. We assume, of course, that $\underline{V}_i < V_i^* < \overline{V}_i$ [3].

[3] In our empirical work we look at several stocks at once. We assume that the random variables giving asset values are uncorrelated across firms. For this reason, the analysis in the text is done for a single firm.

Within any trading day, time is discrete and it is indexed by $t = 1, \ldots, T$. At each time t during the day, at most one trader is given the opportunity to trade, and that trader may enter at most one trade (i.e., he cannot simultaneously enter both a small buy order and a large buy order). Let us first consider a trading day on which an information event has occurred. With probability μ, the trade at time t arises from a trader who has seen the new information. This informed trader will sell the asset if he has observed bad information, and he will buy it if he has observed good information. The quantity that he trades may be influenced by many factors including the price for various quantities, his preferences, and his endowment. Empirical implementation necessitates that we consider only a finite number of trade sizes, and so we consider two trade sizes, small and large. Let γ denote the probability that the informed trader trades the large quantity, with $1 - \gamma$ the probability that he trades the small quantity [4].

There are several aspects of the informed traders' behavior that are important. First, their behavior differs across days according to information events. The market maker does not observe the events directly, nor does he know whether any particular trader is informed. As the distribution of trades differs across days, however, he can use trades to make inferences about the events [5]. Second, within any day, the arrival and behavior of informed traders are independent across time. In previous research (see Easley et al., 1993), we found evidence of dependence in trades within a day. Hasbrouck (1991) also finds that trades are strongly positively autocorrelated. Why informed traders' behavior should depend on the history of trades is not obvious; optimizing informed traders in our framework will ignore this history as they already have the information that makes history useful to the market maker. A more natural belief is that this history dependence arises from the behavior of uninformed traders. Finally, our use of only two trade sizes is clearly restrictive. Our analysis could be performed for any number of trade sizes and for any definition of these sizes. Of course, adding more trade sizes increases the number of parameters to estimate, and this correspondingly reduces our degrees of freedom.

To capture that the number of trades, and thus volume, differs across days, we assume that, when no one is informed, the probability of trade at time t is ε_t. We define how ε_t varies with t shortly. Alternatively, on a day in which an information event has occurred, the probability that a trade arises from an uninformed trader at time t is $(1 - \mu)\varepsilon_t$. This trader can either buy or sell a large or a small quantity of the asset. With probability β, he trades a large quantity, and with probability $1 - \beta$ he trades a small quantity. Similarly, with probability η_t he

[4] A model determining γ endogenously can be given (see Easley and O'Hara (1987)), but does not lead to a useful empirical restriction. Of course, a constant γ is consistent with equilibrium.

[5] In our earlier work, Easley et al. (1993), we found strong evidence that the distribution of trades differs across days in a manner consistent with this structure and inconsistent with simple trinomial model.

sells, and with probability $1 - \eta_t$ he buys. So, for example, the probability of a large sell arising at time t from an uniformed trader on an information event day is $(1 - \mu)\varepsilon_t \beta\eta_t$.

This probabilistic structure allows uninformed trade to differ across trading days, but does not incorporate any within-day dependence [6]. If the uninformed follow trade-based decision rules, or exhibit 'herd' behavior (see Froot et al., 1992), or even follow some more complex dynamic trading strategy, then this specification may be too limited. To allow for dependence in trades in a parametric form, we let the parameters ε_t and η_t depend on what happened in the previous trading interval. This allows the uninformed trading process to be history dependent, and hence permits uninformed trade to follow patterns [7]. Let $o_t \in \{\text{NT}, \text{S}, \text{B}\}$ denote the trading outcome at time t, where NT denotes a no-trade, S is a sell, and B is a buy. Our structural assumption is that $\varepsilon_t = \varepsilon(o_{t-1})$ and $\eta_t = \eta(o_{t-1})$.

Relative to the case where trades are independent, this structure introduces an additional four parameters. This provides a statistical framework for testing the independence of uninformed behavior. In particular, we can estimate our model with and without the restriction that the probability of uninformed trade differs depending upon the history of trade (i.e., $\varepsilon(\text{S}) = \varepsilon(\text{B}) = \varepsilon(\text{NT})$ and $\eta(\text{S}) = \eta(\text{B}) = \eta(\text{NT})$). Rejection of this restriction provides evidence that the behavior of uninformed trade is history dependent. Note that we have not distinguished in indexing these probabilities by trade size, and we have not allowed the trade size parameter β to be history dependent. Doing so is possible, at a cost of degrees of freedom, but we do not believe that the interaction between trade size and history dependence in uninformed behavior is important.

On non-event days, all trades come from the uninformed. In this case, the within day stochastic process of no-trades, and buys, and sells of both sizes, follows a Markov chain. How this chain evolves throughout the day is given by a transition matrix, T_n, which shows how the probability of the trade outcome at time $t + 1$ depends upon the trade outcome at time t. This transition matrix, given in Table 1, shows, for example, that if there were a small buy at time t, then the probability of there being a small buy at time $t + 1$ is $\varepsilon(\text{B})(1-\beta)(1-\eta(\text{B}))$.

[6] The existence of traders who trade for reasons other than speculation is standard in market microstructure models. As is standard in this literature we do not offer an optimizing model of the behavior of these uninformed traders. Each trade an uninformed trader executes occurs at the expected value of the asset given all publicly available information. So although these traders must on average lose money to the informed, they are not being taken advantage of by the market maker. They can gain nothing by observing the pattern of trades as the information is already contained in prices.

[7] This structure is not quite a Markov process as we allow the distribution of uninformed behavior at time t to depend on the outcome of the trade process at time $t - 1$. This process includes both informed and uninformed traders so the behavior of uniformed traders at time t depends on the trade at time $t - 1$ regardless of who generated that trade.

Table 1
The transition matrix for uninformed trades

Time t trade	Time $t+1$ probability				
	NT	LS	SS	LB	SB
No trade (NT)	$1-\varepsilon(N)$	$\varepsilon(N)\beta\eta(N)$	$\varepsilon(N)(1-\beta)\eta(N)$	$\varepsilon(N)\beta(1-\eta(N))$	$\varepsilon(N)(1-\beta)(1-\eta(N))$
Large sell (LS)	$1-\varepsilon(S)$	$\varepsilon(S)\beta\eta(S)$	$\varepsilon(S)(1-\beta)\eta(S)$	$\varepsilon(S)\beta(1-\eta(S))$	$\varepsilon(S)(1-\beta)(1-\eta(S))$
Small sell (SS)	$1-\varepsilon(S)$	$\varepsilon(S)\beta\eta(S)$	$\varepsilon(S)(1-\beta)\eta(S)$	$\varepsilon(S)\beta(1-\eta(S))$	$\varepsilon(S)(1-\beta)(1-\eta(S))$
Large buy (LB)	$1-\varepsilon(B)$	$\varepsilon(B)\beta\eta(B)$	$\varepsilon(B)(1-\beta)\eta(B)$	$\varepsilon(B)\beta(1-\eta(B))$	$\varepsilon(B)(1-\beta)(1-\eta(B))$
Small buy (SB)	$1-\varepsilon(B)$	$\varepsilon(B)\beta\eta(B)$	$\varepsilon(B)(1-\beta)\eta(B)$	$\varepsilon(B)\beta(1-\eta(B))$	$\varepsilon(B)(1-\beta)(1-\eta(B))$

The rows give the probability of an uninformed trader choosing a specific trade at time $t+1$, given the particular trade outcome at time t.

On event days, the process is more complex as informed traders are mixed in. The transition matrix is then formed as a composite of the matrix T_n and the transition matrices that apply when information events occur. In particular, the transition matrix T_n applies when an uninformed trader is drawn. When such a trader is drawn, which happens with probability $1-\mu$, the origin state of the process is the last trade regardless of whether that trade came from an informed or an uninformed trader. Alternatively, when an informed trader is drawn, which happens with probability μ, the probabilities of various trades depend on the event, but not on the last trade. This reflects that informed traders' behavior differs with good and bad news, but is not otherwise history dependent. More precisely, the transition matrix that applies when an informed trader is drawn on a bad event day is I_b, which is given by

$$I_b = \begin{vmatrix} 0 & \gamma & 1-\gamma & 0 & 0 \\ 0 & \gamma & 1-\gamma & 0 & 0 \\ 0 & \gamma & 1-\gamma & 0 & 0 \\ 0 & \gamma & 1-\gamma & 0 & 0 \\ 0 & \gamma & 1-\gamma & 0 & 0 \end{vmatrix} \tag{1}$$

Thus, the informed trader enters a large trade with probability γ, a small trade with probability $1-\gamma$, and otherwise does not trade. Similarly, the transition matrix that applies when an informed trader is drawn on a good event day is I_g which is given by

$$I_g = \begin{vmatrix} 0 & 0 & 0 & \gamma & 1-\gamma \\ 0 & 0 & 0 & \gamma & 1-\gamma \\ 0 & 0 & 0 & \gamma & 1-\gamma \\ 0 & 0 & 0 & \gamma & 1-\gamma \\ 0 & 0 & 0 & \gamma & 1-\gamma \end{vmatrix} \tag{2}$$

Mixing these matrices yields the transition matrix that applies on each day. The transition matrix for no-event days is simply T_{n}. The transition matrix for bad event days is

$$T_{\mathrm{b}} = \mu I_{\mathrm{b}} + (1-\mu) T_{\mathrm{n}} \tag{3}$$

and for good event days it is

$$T_{\mathrm{g}} = \mu I_{\mathrm{g}} + (1-\mu) T_{\mathrm{n}}. \tag{4}$$

The overall process, including the draw of days, is depicted in the tree diagram given in Fig. 1. At the first node of the tree, nature selects the type of day: no-event, bad event or good event. This selection occurs once a day. Then, according to the relevant node of the tree, the $t = 1$ trader is drawn. After this trader acts according to the probabilities on the relevant branch (we arbitrarily set o_0 to NT), we loop back to the node of the tree corresponding to the day drawn by nature. The next trader is selected, and he or she acts according to the probabilities conditioned on the outcome of the first trade. This loop repeats T times. The first day then ends, nature draws a new day, and trading continues anew.

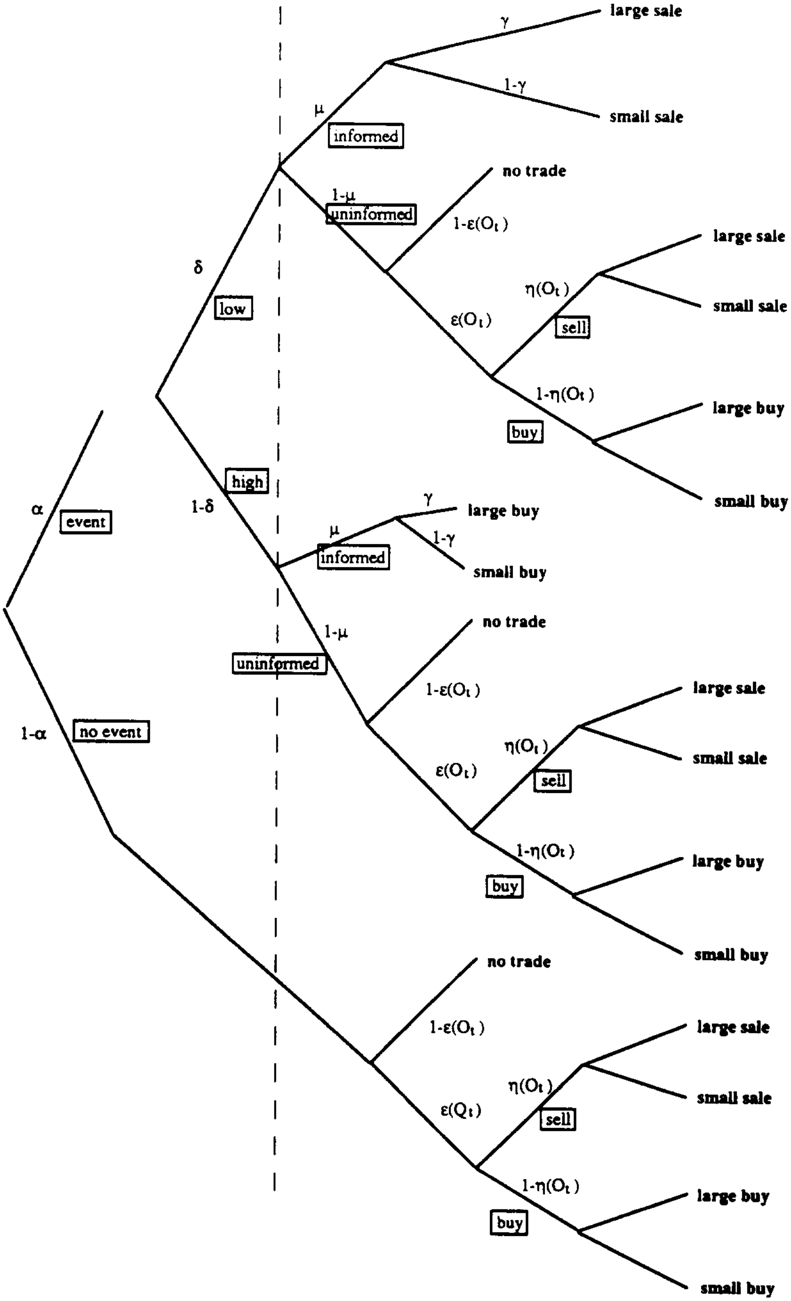
2.2. *Information, beliefs and prices*

At any time, the risk-neutral, competitive market maker's bid and ask prices for various trades are the expected value of the asset conditioned on the market maker's information at that time [8]. This information consists of the information that he has gleaned from the history of the process, and the information contained in the fact that someone wants to execute the proposed trade. For example, at time t in any day, the bid for a large quantity of the asset is the expected value of the asset conditional on the information content of trades and no-trades up through time $t-1$, and on the information contained in the arrival of an order to sell a large quantity of the asset at time t.

Let the trade outcome at time t be $q_t \in \{\mathrm{NT}, \mathrm{LB}, \mathrm{SB}, \mathrm{LS}, \mathrm{SS}\}$, where NT denotes no-trade, LB denotes a large buy, SB denotes a small buy, LS denotes a large sell, and SS denotes a small sell. The posterior probability on good news if, for example, someone wants to sell a large quantity (LS) at time t and if the market maker has observed trade outcomes $(q_1, \ldots, q_{t-1})$ up to time t is denoted $\Pr\{\mathrm{g} \mid q_1, \ldots, q_{t-1}, \mathrm{LS}\}$.

Fig. 1. Information structure and trade. This figure shows the probabilistic structure of trading. Nodes of the tree to the left of the line happen once, prior to the trading day. Nodes to the right apply to each trading interval.

[8] The market maker is also assumed not to be liquidity constrained. The effect of all of these assumptions is that inventory does not affect prices. We do not believe that, in fact, prices are unaffected by inventory. Thus we view the analysis in this section as showing the effects of information in isolation. Any difference in prices between those predicted here and actual prices may be due to inventory effects.



γ
large sale
1-γ
small sale
μ
informed
no trade
1-μ
uninformed
1-ε(Oₜ)
large sale
δ
η(Oₜ)
low
small sale
ε(Oₜ)
sell
1-η(Oₜ)
large buy
buy
small buy
high
γ
large buy
α
event
1-δ
μ
1-γ
informed
small buy
no trade
1-μ
uninformed
1-ε(Oₜ)
large sale
1-α
no event
η(Oₜ)
small sale
ε(Oₜ)
sell
1-η(Oₜ)
large buy
buy
small buy
no trade
1-ε(Oₜ)
large sale
η(Oₜ)
small sale
ε(Qₜ)
sell
1-η(Oₜ)
large buy
buy
small buy

This value can be computed by Bayes rule. Let

$$(\Pr\{g|q_1, \ldots, q_{t-1}, \mathrm{LS}\}, \Pr\{b|q_1, \ldots, q_{t-1}, \mathrm{LS}\}, \Pr\{o|q_1, \ldots, q_{t-1}, \mathrm{LS}\}),$$

where o denotes no news, represent the market makers prior at time t. Then the posterior is

$$(\Pr\{g|q_1, \ldots, q_{t-1}, \mathrm{LS}\} = \frac{\Pr\{g|q_1, \ldots, q_{t-1}\}(1-\mu)\varepsilon_t\eta_t\beta}{\Pr\{\mathrm{LS}|q_1, \ldots, q_{t-1}\}},$$

where the probability of a large sale given $(q_1, \ldots, q_{t-1})$ is

$$\begin{aligned}(\Pr\{\mathrm{LS}|q_1, \ldots, q_{t-1}\} &= \Pr\{o|q_1, \ldots, q_{t-1}\}\varepsilon_t\eta_t\beta + \Pr\{b|q_1, \ldots, q_{t-1}\}(\mu\gamma + (1-\mu)\varepsilon_t\eta_t\beta) \\ &\quad + \Pr\{g|q_1, \ldots, q_{t-1}\}(1-\mu)\varepsilon_t\eta_t\beta\end{aligned}$$

Posterior probabilities on the other states, b and o, and for the other possible trades at time t, are computed similarly. Using the fact that the market maker's prior at time 1 is ($\Pr\{g\} = \alpha(1-\delta)$, $\Pr\{b\} = \alpha\delta$, $\Pr\{o\} = 1-\alpha$) this defines a recursive procedure to compute beliefs for any sequence of trades. Note that if all the trade parameters are strictly between 0 and 1, then the market maker cannot know from any finite sequence of trades the move of nature. Instead, the market maker's beliefs adjust over time in response to the trades he observes.

With these posterior probabilities we can compute the market maker's quotes. The bid at time t for a large quantity of the asset, LS, is

$$\begin{aligned}bL_t = \overline{V}_i(\Pr\{g|q_1, \ldots, q_{t-1}, \mathrm{LS}\} &+ \underline{V}_i\Pr\{b|q_1, \ldots, q_{t-1}, \mathrm{LS}\} \\ &+ V_i^*\Pr\{o|q_1, \ldots, q_{t-1}, \mathrm{LS}\},\end{aligned} \tag{5}$$

The bid for a small sale (bS_t), and the ask prices for large and small buy orders (aL_t, aS_t), are determined by similar expected values conditioned on the relevant trade at time t. The spread for the small trade size is just the difference between the ask and bid prices for a small trade. It is easy to see that these prices are equal if no informed traders use small trades, i.e. $\gamma = 1$, and that the ask is above the bid otherwise. Similarly, the spread for large trades is the difference between ask and bid prices for the large trade size, and it will be non-zero if and only if $\gamma > 0$.

Prices move over time because of trades or, more precisely, because of the information revealed by trades (or the lack of trade). As long as informed trade is possible, but not certain to occur ($1 > \alpha > 0$ and $\mu > 0$), both trade and the lack of trade generally cause prices to move [9]. The absence of trade indicates that no

[9] This feature of quote changes in the absence of trade is common in the data we analyze. If, as our model permits, time between trades carries information any analysis of transaction price data alone, or transaction price and trade data, is mis-specified and give misleading results. For example, the Hasbrouck (1991) analysis of the information content of trades uses a VAR model of prices and trades treats order arrivals as the time index. This is appropriate only if clock time or the level of activity in the stock as measured by the number of orders per unit of clock time carries no information.

potential trader is currently informed, and this causes the market maker to lower his estimate of the likelihood of an information event having occurred. This moves his expected value of the asset towards its unconditional expected value, V_i^*, and thus moves bid and ask prices towards V_i^*. For the same reason, trade which is unlikely to come from informed traders generally causes price revision. For example, suppose that all informed traders trade large quantities, i.e. $\gamma = 1$. Then small trades carry no information about the relative probabilities of good and bad events. Nonetheless, they do carry information: the arrival of a small quantity order at time t in a market in which all informed traders trade large quantities indicates that no information arrived at time t. This lowers the market maker's probability of the occurrence of an information event, which, in turn, causes the bids and asks at time $t+1$ to move toward V_i^*.

2.3. *Information and trade size*

A primary question that we are interested in is the relative information content of these small versus large trades. Information content is defined in our model as the revision induced in the market maker's beliefs. The larger this revision, the greater is a trade's information content. Consider the effect of a large sale versus a small sale at time t on the market maker's beliefs at time $t+1$. If $0 < \gamma < 1$, then both types of orders to sell the asset reduce the market maker's estimated probability of a good event. It is easy to show that if $\gamma > \beta$, then large sales raise the probability of a bad event by more than do small sales as large sales are more likely to come from informed traders than are small sales. Furthermore, following a large sale, the probability of no event is lower than it is following a small sale [10]. These effects dictate that large sales lower beliefs by more than do small sales when $\gamma > \beta$. Of course, the implications are reversed whenever $\beta > \gamma$. Thus, within our model, the implications of trade size for prices are captured by the relation between γ and β.

2.4. *Information and trade patterns*

Patterns of trade will generally have information in our model due to their correlation with the underlying information structure. How to interpret this information, however, is not immediately obvious because the trade process reflects the decisions of both informed and uninformed traders. If uninformed trades are not history dependent, then this inference problem is simplified as only informed trader behavior need be considered. When uninformed trades are history dependent, however, there are patterns in both informed and uninformed behavior, and this can change the information content of the trade process.

[10] In fact, large and small sales can raise the probability of no event, but it is always lower following a large sale than following a small sale.

Our model provides a statistical framework for analyzing these patterns. In particular, a first step is to determine the structure of uninformed trade. As noted earlier, this is captured in the ε_t and η_t variables. If $\varepsilon(\mathrm{S}) = \varepsilon(\mathrm{B}) = \varepsilon(\mathrm{NT})$ and $\eta(\mathrm{S}) = \eta(\mathrm{B}) = \eta(\mathrm{NT})$, then uninformed trade is serially uncorrelated. If this restriction does not hold, then the distribution of the current uninformed trade depends upon the previous trade. Given the structure of uninformed trade, the second step is to characterize the Bayesian updating problem of various trade outcomes. This allows us to determine the information content of various trade patterns.

There are three general patterns of interest, namely, sequences, reversals, and periods of no-trading. We define sequences as consecutive buy orders or sell orders. Because informed trades are perfectly correlated, sequences of orders will always affect beliefs, with buys raising the market makers expectation, or sells lowering it. If uninformed trading is highly positively correlated, however, a sequence of orders is a likely event even on a day with no information. In this case a sequence of buy orders would raise the market maker's expected value of the asset, but the effect will be smaller than if uninformed trade were i.i.d.

Reversals occur when a buy order follows a sell order, or vice versa. To see how these could have information content, suppose that the trade at time $t-1$ was a buy, and the trade at time t is a sell. In a market in which the uninformeds' behavior is highly positively correlated, i.e. $1 - \eta(\mathrm{B})$ is large, such a reversal is unlikely to have come from an uninformed trader at time t. Thus, the reversal of orders (the sell order following a buy order) raises the market maker's probability of an information event, and the current sell order raises the probability that the event was bad. The overall effect, when $1 - \eta(\mathrm{B})$ is large, is to lower bid prices for both small and large orders following a buy, reflecting the market maker's increased belief regarding a bad information event [11].

Finally, no trade intervals affect beliefs because of their correlation with the existence of information. Consecutive no-trading outcomes lower the market maker's belief regarding the existence of new information. This causes the market maker to move bid and ask prices closer to the asset's expected value, thereby reducing spreads. Unlike sequences and reversals, the strength of this effect does not depend on the structure of uninformed trade.

3. Estimating the trade process

3.1. Likelihood function

We now turn to the estimation of the eleven parameters of the trade process. These parameters are $\theta = (\alpha, \delta, \mu, \gamma, \beta, \varepsilon(\mathrm{NT}), \varepsilon(\mathrm{B}), \varepsilon(\mathrm{S}), \eta(\mathrm{NT}), \eta(\mathrm{B}),$

[11] Alternatively, if $\varepsilon(\mathrm{B})$ is large but $\eta(\mathrm{B})$ is small, then a reversal of order flow following a buy is more likely to have come from an uninformed trader. Thus it does not have a large effect on prices.

η(S)). The data set for I days is $D_I=(d_1, \ldots, d_I)$, where d_i is a sequence of T_i trade outcomes $\{q_t\}$ for day i, each of which is an element of (NT, LB, SB, LS, SS). The structure of our model dictates that this estimation is not straightforward. The difficulty arises because we cannot directly observe the arrival of any information events or trades governed by these parameters. Parameters α and δ determine the probabilities of three information events (no news, good news, and bad news), none of which are observable (to us). The remaining parameters refer to probabilities of trade by uninformed or informed traders. We observe trades sorted by size, but we do not observe which traders are uninformed or informed. Estimation of these parameters thus requires a structural model. Our model provides the structure necessary to extract information on the parameters from the observable variables, buys and sells.

Of particular importance is that because we allow trades within a day to be one-step history dependent, we have to keep track of all possible pairs of trade and no-trade observations. There are 25 such pairs of observations $((o_t, o_{t+1})$ for each possible value of observations). Our assumption that the distribution of uninformed trade at date $t+1$ does not depend on whether the trade at date t was large or small allows us to reduce this to 15 relevant pairs. The number of observations of each such pair on any day thus forms a sufficient statistic for the day's data.

To develop our estimating equations, first note that, as days are drawn independently, we can form the likelihood functions for each type of day, and then multiply them together to get the overall likelihood function. To fix ideas, consider the distribution of the sequence of within day trades when a bad event has occurred. On day i, the joint likelihood can be factored as

$$L_i(\theta|q_2, \ldots, q_{T_i}:q_1)=\prod_{t=2}^{T_i}L(\theta|q_t, q_{t-1}), \tag{6}$$

where we condition on q_1, the first trade of the day. Each of the conditional distributions above can be calculated from the tree given in Fig. 1. For example, the conditional distributions following a no-trade are:

$$\begin{aligned}
&L(\theta|NT, NT)=(1-\mu)(1-\varepsilon(NT)),\\
&L(\theta|LS, NT)=\mu\gamma+(1-\mu)\varepsilon(NT)\beta\eta(NT),\\
&L(\theta|SS, NT)=\mu(1-\gamma)+(1-\mu)\varepsilon(NT)(1-\beta)\eta(NT),\\
&L(\theta|LB, NT)=(1-\mu)\varepsilon(NT)\beta(1-\eta(NT)),\\
&L(\theta|SB, NT)=(1-\mu)\varepsilon(NT)(1-\beta)(1-\eta(NT))
\end{aligned} \tag{7}$$

The likelihood for the other origin states and information configurations, good event and no event, can be factored in the same way. The pieces of the likelihood function can be read from Fig. 1, noting that the ε and η appearing in that figure must be indexed by the previous trade. Alternatively, these can be obtained directly from the transition matrices given above.

We form the likelihood function for the day by mixing the likelihoods conditional on bad, good and no events with the probabilities $\alpha\delta$, $\alpha(1-\delta)$ and $(1-\alpha)$. The resulting likelihood function for one day is bi-linear in α and δ. These parameters are probabilities of events that happen at most once per day, and so maximizing the likelihood will put these parameters at 0 or 1, depending on the data. An analogous situation is estimating a binomial probability after observing one trial – the maximum likelihood estimator is either 0 or 1. To estimate α and δ, we need to use multiple days of data.

Combining observations for I days yields the likelihood function (under the assumption of independence across days)

$$L(\theta|d_I) = \prod_{i=1}^{I} L_i(\theta|d_i), \tag{8}$$

where d_i is the data set for day i.

We have only I draws from the first segment of the tree, so we expect our estimates of α and δ to be less precise than our estimates of the other parameters. This, of course, will be reflected in the standard errors of these parameters. All of the parameters of our model are identified, but they are estimated with varying precisions. In particular, precise estimation of γ and β requires a reasonable split of the sample into large and small trades.

Having defined the likelihood function for our model, we can now calculate the parameter values that maximize this function for a given stock. It is easiest to understand how our likelihood function treats the data if we consider uncorrelated uninformed trade. Roughly, what our procedure does is to classify days into buy-led high-volume days, sell-led high-volume days and low volume days (as in Fig. 1). In our structural interpretation, these are good-event days, bad-event days and no-event days, respectively. This ties down the parameters α, δ, ε, η and μ. The difference in the split of trades by size on event versus no event days determines β and γ. Of course, this discussion is heuristic, but it does provide intuition on what data patterns identify the parameters. The likelihood function given in Eq. (8), with independent uninformed trade, is a mixture of multinomials, with specific terms reflecting the numbers of buys, sells, and no-trades and with restrictions across the components of the mixture.

3.2. Estimation and interpretation

We first estimate the eleven parameters of our model without any cross parameter restrictions. To isolate the effect of uninformed behavior on the trade process, we also estimate the parameters with the restriction that uninformed behavior is independent. This restriction is $\eta(\mathrm{B}) = \eta(\mathrm{S}) = \eta(\mathrm{NT})$ and $\varepsilon(\mathrm{B}) = \varepsilon(\mathrm{S}) = \varepsilon(\mathrm{NT})$. In this case, we have a seven dimensional parameter set, and it is easy to see that the total number of no-trades, large buys, large sells, small buys, and small sells per day is a sufficient statistic. This simplifies estimation, and so we use this statistic in the independent case. A likelihood ratio test between the unrestricted and restricted models then allows us to determine whether the independent uninformed trade model or the history dependent uninformed trade model is better supported by the data. The χ^2 test statistic is used to determine statistical significance. Posterior odds ratios (independent/dependent) derived from a neutral prior are reported to aid in interpretation.

To isolate the information effects of trade size, we estimate the general model with the parameter restriction $\gamma = \beta$. This restriction dictates that the information content of large and small trades is identical. We then use the likelihood ratio test to compare this to the unrestricted case. We also consider a variant of the general model in which the informed propensity to buy and sell the large quantity can differ depending upon whether they have learned good or bad news. This corresponds to allowing the informed trades of the large quantity to differ across good and bad news, due perhaps to the effects of short-sale restrictions, or simply due to greater liquidity on one side of the market.

Estimation in both the independent and dependent cases is by maximum likelihood. We first write our parameters, which are restricted to $[0, 1]^{11}$ in the dependent case and to $[0, 1]^7$ in the independent case, as logit transforms of unrestricted parameters. We then maximize over the unrestricted parameters using the quadratic hill-climbing algorithm GRADX from the package GQOPT. The problem is nonconcave, and, for some data configurations and starting values for the parameters, there appear to be directions of increasing likelihood sending parameters toward plus or minus infinity (zero or one in the economic parameterization). To insure that we in fact found global maxima we ran the optimization routine starting from each point in a large grid in the parameter space. In most cases, each of these starting points yielded the same estimate of an interior global maximum. We had no cases with multiple maxima. Numerical derivatives were used throughout. Standard errors for the economic parameters (reported in the tables) were calculated from the asymptotic distribution of the transformed parameters using the delta method.

Except on portions of the boundary of the restricted parameter space all of our parameters are identified. For example, if α, μ or any of the ε's are zero then either there is no informed trade or no uninformed trade and parameters describing the behavior of such traders cannot be identified. We checked for such cases in our search for maxima and found none. For only one parameter, for one stock (γ for DE), does the global maximum occur at a corner ($\gamma = 1$), but this does not create any problems with identification.

A final issue relates to the statistical interpretation of our results. Our model estimations provide likelihood ratios, and statistical comparisons across nested model specifications can be done via χ^2 tests. As our model is based on Bayesian updating, a natural alternative would be to examine posterior odds ratios to determine statistical significance. This test complements the χ^2 and furthermore is applicable in non-nested situations (e.g. $\gamma < \beta$ vs. $\gamma > \beta$). Where applicable, therefore, we report both tests.

4. The data and estimated models

4.1. Data

It is important to remember that our transactions model is a stock specific model. The parameters can vary across stocks and, if they do, we can draw no general conclusions about the informativeness of trade size or trade reversals for 'the market'. This requirement, along with the computational demands of estimating our eleven parameter stochastic process, precludes using a large cross-sectional sample of stocks. We can, however, infer the effects we seek to investigate by estimating the trade process for specific stocks. In this paper we present estimates for six stocks: Ashland Oil (ASH), Bank of New York (BK), Beneficial (BL), Anheiser–Busch (BUD), John Deere (DE), and Goodyear Tire (GT). Our general criteria for sample selection was to include stocks with sufficient trade activity to make estimation feasible. Ashland Oil met this requirement, and was also selected to allow comparison of our estimates here with those we obtained in previous research (Easley et al., 1993). The other stocks in our sample were chosen from stocks trading in the upper two deciles of volume on the New York Stock Exchange.

The transactions data that we use is drawn from the ISSM transactions data base for the period October 1, 1990, to December 22, 1990. There are sixty days of trading in this period. There is obviously an aspect of judgment in selecting the time period: using more days yields lower standard errors, but over more days the stationarity assumptions of our model are less likely to be valid. Experimentation with different trade periods suggested that sixty days is a reasonable sample, and it is these results that we report.

The ISSM database provides a complete listing of quotes, depths, trades, and volumes, together with time stamps for each trade or quote throughout the day. We know when trades occur, but no-trades are obviously not recorded. Our model dictates that clock time and trade time are not identical, and so our first data preparation project is to define the no-trade observations. We do this by measuring the elapsed time since the last trade. If five minutes pass without a trade, we mark a no-trade and initialize the clock. If a trade occurs before five minutes passes, we mark the trade and reset the clock. Thus, a sequence of, say, five no-trades means that 25 min have elapsed since the last trade. We discuss the sensitivity of our results to this specification of a no-trade interval in Section 4 [12].

Our second major data preparation project is the classification of trades. Transactions are recorded in the database, but their direction is not recorded. The

[12] The parameters μ and ε, which are defined as probabilities are affected in the obvious way. The other parameters, mostly defined as fractions of trades of different types, are nearly invariant to small changes in the interval (experiments were done for no-trade intervals of between 2 and 10 min).

Table 2
Trade and price data for stocks in the sample

Stock	Small buys	Large buys	Small sells	Large sells	Price
ASH	13.3 (9.5)	6.0 (4.6)	12.8 (6.1)	5.3 (3.4)	29.14 (1.08)
BK	22.0 (9.4)	20.1 (11.4)	17.0 (6.7)	15.9 (8.0)	17.93 (1.37)
BNL	15.5 (7.9)	7.3 (5.0)	14.7 (6.8)	5.8 (4.1)	40.36 (3.08)
Bud	25.2 (9.1)	36.7 (12.8)	28.4 (9.7)	28.0 (9.8)	39.25 (2.13)
DE	39.3 (14.3)	38.2 (16.5)	32.6 (9.2)	35.1 (16.2)	43.69 (2.64)
GT	34.3 (12.2)	20.6 (14.9)	25.0 (8.7)	18.2 (10.2)	16.00 (0.90)

This table provides a summary of trade and price data for stocks in the sample. The first column gives the stocks in the sample which are Ashland Oil (ASH), Bank of New York (BK), Beneficial (BNL), Anheiser–Busch (BUD), Deere (DE), and Goodyear Tire (GT). Standard errors are given in parentheses.

problem of classifying transactions into buys and sells is ubiquitous in the literature. We use a technique developed by Lee and Ready (1991). Trades at prices above the midpoint of the bid–ask interval are classified as buys; trades below the midpoint are classified as sells. Trades exactly at the midpoint are classified depending on price movements. A midpoint price trade will be a sell if the midpoint is lower than the midpoint of the bid–ask interval at the previous trade. If these midpoints are the same, we look further back until we find a price movement. Correspondingly, if the price had moved up, the midpoint trade is classified as buy.

The final step in data preparation is the classification of trades into large and small. This is somewhat arbitrary, since there are more than two trade sizes. From a market perspective, what is considered large in one stock may not have the same designation in another. Moreover, since trading volume differs across the stocks in our sample, the frequency of 'large' trades differs as well [13]. Inspection of the data suggested that trades of 100–500 shares were most common, with transactions of 1000 shares also not uncommon. We therefore chose 1000 shares as our cut-off, designating trades of amounts less than 1000 shares as small and amounts greater than or equal to this amount as large [14]. We return to this trade size issue later in this section.

The data on numbers of buys and sells of each size, and for each stock in our sample, is summarized in Table 2.

[13] This frequency issue is important because choosing too large a trade split may result in too few large transactions to permit estimation of our parameters. Conversely, too small a split can result in any differential information content being overwhelmed by the uninformed trading volume.

[14] Estimation with larger cut-off levels ran into difficulty for some stocks due to the sample of large trades being too small to estimate our parameters with any precision.

4.2. The structure of uninformed trade

Our general model of the trade process allows the stochastic process of uninformed trade to depend on the previous trade. This structure allows greater latitude in uninformed trading strategies, and in particular allows uninformed trades to be path dependent. An alternative specification is that uninformed trades are independent and identically distributed. This specification can be viewed as a restriction of our model to the case where $\varepsilon(S) = \varepsilon(B) = \varepsilon(NT)$ and $\eta(S) = \eta(B) = \eta(NT)$. Recall that the ε variable is the probability that the uninformed trader chooses to trade, where ε_s, for example, is the probability the uninformed trader chooses to trade given that the last trade was a sell. The η variable captures the probability that he chooses to sell given that he decides to trade.

Table 3 provides the likelihood ratio statistics for both the restricted model and unrestricted model for our sample stocks. Column three of the table gives the relevant χ^2 statistic. As is apparent from the table, the restricted model is strongly rejected. In every case, the unrestricted model outperforms the restricted model in that the likelihood ratio is a larger number. Moreover, the χ^2 statistics are all significant at the 0.005 level, suggesting that independence of uninformed trade is not supported by the data. Finally, the implied posterior odds of independence are all approximately zero.

Our result that uninformed trades are not independent has two important implications. First, sequences of trades are less informative than would be predicted by a standard microstructure model. What is quite informative, however,

Table 3
Testing dependent vs. independent uninformed trade

Stock	Dependent model $\varepsilon(S)$, $\varepsilon(B)$, $\varepsilon(N)$, $\eta(S)$, $\eta(B)$, $\eta(N)$	Independent model $\varepsilon(S) = \varepsilon(B) = \varepsilon(N)$ $\eta(S) = \eta(B) = \eta(N)$	χ^2	Posterior odds independent/ dependent
ASH	−6,415.2	−6,775.8	721.2*	0.0
BK	−10,827.5	−11,119.1	583.3*	0.0
BNL	−7,149.4	−7,577.8	856.9*	0.0
BUD	−14,369.0	−14,751.6	765.3*	0.0
DE	−16,301.5	−16,988.6	1,374.2*	0.0
GT	−12,504.8	−13,016.0	1,022.4*	0.0

This table gives the likelihood ratio statistics for two model specifications. The first column gives the stocks in the sample which are Ashland Oil (ASH), Bank of New York (BK), Beneficial (BNL), Anheiser–Busch (BUD), Deere (DE), and Goodyear Tire (GT). The second column gives the likelihood ratio of the history dependent model. The third column gives the likelihood ratio when uninformed trade is restricted to be independent. The fourth column gives the χ^2 statistic, with * indicating significance at the 0.005 level. The last column gives the posterior odds of the independent to the dependent model with a prior odds ratio of 1. The approximate posterior odds are calculated by using a prior odds ratio of one (the reader can adjust this by multiplying by different prior odds) and approximating the Bayes factor by the likelihood ratio.

is a reversal in order flow. For uninformed traders, reversals are quite rare. Thus a buy following a sale will greatly increase the market maker's probability on good news and raise quotes by more than would a buy following a no-trade interval or another buy. Second, on days when there has been no information event, there are by definition no informed traders, and sequences of no-trades are more likely than if informed traders also existed. Of course, market participants do not know what type of day it is, but as trade progresses through the day, a preponderance of no-trades indicates a greater likelihood that no information exists. Thus, just as reversals were informative, so too are periods of no trading, albeit with the opposite implication. These results indicate the important information conveyed by the trade process, above and beyond that conveyed by the trades themselves [15].

4.3. The information content of trade size

Another dimension of the trade process is order size. If informed traders prefer to trade larger quantities, then the information content of large orders should differ from that of small orders. Our model provides a statistical framework for investigating this effect, and essentially involves a two-stage procedure. First, we test for whether β, the probability that uninformed traders trade the large quantity, equals γ, the probability that the informed traders trade the large amount. This involves comparing the likelihood ratios of the restricted model ($\beta = \gamma$), with the unrestricted model, with significance again given by the χ^2 statistic. Rejection of this restriction implies that large orders have different information content than small orders. We then consider in the second stage how this information content differs. This involves examining the relation between the estimated β and γ. Greater information content for large orders requires finding that $\beta < \gamma$.

Table 4 gives the likelihood ratio statistics for the restricted and unrestricted models for each of the stocks in our sample. In every case, the likelihood ratio is greater in the unrestricted model, indicating the reasonable (and expected) result that a better fit obtains when the uninformed and informed trade probabilities are not required to be equal. The χ^2 statistic shows that the importance of this effect differs across the stocks in our sample. For three of the stocks (BUD, DE, and GT), the differential information content of trade size is strongly supported, with the χ^2 significant at the 0.005 level. For BK and BNL, the effect is less pronounced, but is statistically significant at the 0.05 level. In the case of Ashland Oil, however, there is no significant difference in the restricted and unrestricted

[15] Even if informed trades were history dependent, conditional on an information event, these conclusions about the information content of trade reversals and sequences of no-trades would be unaffected.

Table 4
The information in trade size

Stock	Unrestricted model γ, β	Restricted model $\gamma = \beta$	χ^2	Posterior odds restricted/unrestricted
ASH	−6,415.2	−6,416.2	2.0	0.368
BK	−10,827.46	−10,829.39	3.86 $^+$	0.15
BNL	−7,149.4	−7,151.6	4.4 $^+$	0.11
BUD	−14,369.0	−14,399.1	60.24 *	0.0
DE	−16,301.5	−16,320.3	37.7 *	0.0
GT	−12,504.8	−12,523.6	37.6 *	0.0

This table gives the likelihood ratio statistics for two model specifications. The second column gives the likelihood ratio for the unrestricted model in which trade size can matter. The third column gives the likelihood ratio of the restricted model in which trade size does not matter. The fourth column gives the χ^2 statistic, with * indicating significance at the.005 level and + at the 0.05 level. The last column gives the posterior odds of the independent to the dependent model with a prior odds ratio of 1 (see footnote 16).

models. For this stock, large trades and small trades convey the same information. The posterior odds reported in the first column of Table 4 support this conclusion.

These results provide evidence that order size can convey information to market participants. While not always the case, our results indicate that large and small orders are not traded identically by informed and uninformed traders. How their information content differs can be determined from the estimated parameters in the unrestricted model given in Table 5. These parameter estimates vary widely across the stocks in our sample, indicating that both information-based trading and

Table 5
Testing for trade size effects

Parameters	ASH	BK	BNL	BUD	DE	GT
α	0.546 (0.103)	0.312(0.105)	0.487 (0.099)	0.782 (0.063)	0.227 (0.073)	0.268 (0.074)
δ	0.458 (0.110)	0.121 (0.119)	0.197 (0.105)	0.982 (0.023)	0.711 (0.158)	0.684 (0.142)
μ	0.121 (0.012)	0.160 (0.024)	0.107 (0.013)	0.158 (0.014)	0.114 (0.015)	0.136 (0.017)
ε(N)	0.254 (0.010)	0.480 (0.011)	0.283 (0.010)	0.622 (0.012)	0.708 (0.011)	0.581 (0.010)
ε(B)	0.488 (0.017)	0.643 (0.012)	0.556 (0.015)	0.781 (0.008)	0.849 (0.005)	0.743 (0.008)
ε(S)	0.415 (0.017)	0.652 (0.012)	0.478 (0.016)	0.549 (0.015)	0.824 (0.006)	0.720 (0.009)
η(N)	0.538 (0.025)	0.492 (0.020)	0.573 (0.023)	0.398 (0.020)	0.471 (0.014)	0.434 (0.014)
η(B)	0.182 (0.020)	0.318 (0.017)	0.210 (0.017)	0.220 (0.016)	0.273 (0.008)	0.251 (0.010)
η(S)	0.806 (0.023)	0.645 (0.021)	0.799 (0.022)	0.549 (0.015)	0.670 (0.009)	0.660 (0.012)
β	0.287 (0.014)	0.463 (0.012)	0.283 (0.013)	0.585 (0.008)	0.485 (0.006)	0.364 (0.007)
γ	0.370 (0.048)	0.636 (0.101)	0.444 (0.076)	0.346 (0.035)	1.00 (*)	0.867 (0.082)

This table gives parameter estimates for a model with history dependent uninformed trade and two different trade sizes. Large trades are defined as those of 1,000 shares or greater; small trades are below 1,000 shares. Standard errors are given in parentheses.

market depth vary across stocks. The standard errors show that our parameters are estimated with reasonable precision, as virtually all are statistically significant. Parameter estimates for β and γ, found in the bottom two rows of the table, confirm that in general larger trades have greater information content. For four of the stocks for which $\beta = \gamma$ can be rejected (BK, BNL, DE, GT), γ exceeds β, which dictates greater information content to large trades in those stocks. For these stocks, the probability of informed trading ranges from 1.37 to 2.38 times as great as the probability of uninformed trading in large stocks, with the average across the stocks being approximately 1.8.

One stock, BUD, does not fit this pattern. Here, the paradoxical result emerges that greater information is contained in the smaller quantity, a result at variance with standard models of trader behavior. One possible explanation for this finding is that the trade size cut-off of 1000 shares is not appropriate for this stock. Sufficiently few observations above this level could result in biasing our results in unexpected ways. The precision of the standard errors, however, casts doubt on this explanation. Moreover, experimentation with other cutoffs did not yield different results.

5. Specification tests

The key structural assumption that we use to get information about the parameters of our model from observable data is the information structure. In particular, we assume that information events occur at most once per day and are uncorrelated over time. It is this daily structure that allows us to use differences in the daily distribution of trades to draw inferences about the behavior of informed traders. Given the importance of this structure, we performed two specification tests to examine the validity of these two hypotheses.

We first investigate the hypothesis that information events occur at most once per day. As a simple test of this hypothesis we consider the alternative that events occur at most twice per day; once before the day begins and again at the midpoint (in clock time) of the trading day. The effect of this is to convert our 60 day sample into a sample of 120 half-days. We maximize the likelihood function given in Eq. (8) over this data set for each of our stocks. The results of this estimation are reported in Table 6.

The most interesting result of this alternative specification is that the parameter estimates it produces are remarkably close to those from the 60 day specification. While there are small differences in some of the estimates, these are virtually all within the standard errors of the estimated parameters, suggesting little value to employing the half-day versus full day specification. As the 60 day model is not nested in the 120 half-day model, it is not possible to perform a chi-square test to determine which hypothesis fits the data best. Our conclusion is that while we

Table 6
Parameter estimates and information event horizons

Ticker	ASH		BK		BNL	
	60	120	60	120	60	120
Likelihood	−6415.17	−6374.55	−0.10827.5	−10729.8	−7149.36	−7092.70
μ	0.121 (0.012)	0.122 (0.018)	0.160 (0.024)	0.230 (0.024)	0.107 (0.013)	0.145 (0.017)
ε(N)	0.254 (0.010)	0.253 (0.015)	0.480 (0.011)	0.489 (0.010)	0.283 (0.010)	0.289 (0.010)
α	0.546 (0.103)	0.601 (0.195)	0.312 (0.105)	0.218 (0.050)	0.487 (0.099)	0.372 (0.079)
δ	0.458 (0.110)	0.480 (0.136)	0.121 (0.119)	0.283 (0.115)	0.197 (0.105)	0.109 (0.074
β	0.287 (0.014)	0.280 (0.018)	0.463 (0.012)	0.460 (0.009)	0.283 (0.013)	0.283 (0.013)
η(N)	0.538 (0.025)	0.531 (0.037)	0.492 (0.020)	0.479 (0.015)	0.573 (0.023)	0.587 (0.024)
γ	0.370 (0.048)	0.394 (0.069)	0.636 (0.101)	0.654 (0.049)	0.444 (0.076)	0.441 (0.062)
η(B)	0.182 (0.020)	0.181 (0.028)	0.318 (0.017)	0.306 (0.015)	0.210 (0.017)	0.224 (0.018)
η(S)	0.806 (0.023)	0.808 (0.027)	0.645 (0.021)	0.620 (0.016)	0.799 (0.022)	0.812 (0.023)
ε(B)	0.488 (0.017)	0.486 (0.017)	0.643 (0.012)	0.642 (0.011)	0.556 (0.015)	0.551 (0.015)
ε(S)	0.414 (0.017)	0.412 (0.020)	0.652 (0.012)	0.654 (0.012)	0.478 (0.016)	0.477 (0.016)

This table gives parameter estimates for information event horizons of one day and one-half day. The column labeled 60 corresponds to information events occurring before the beginning of trading in our 60 day sample. The column labeled 120 corresponds to information events occurring before the beginning of trading and again in the middle of the trading day. Standard errors are given in parentheses.

cannot conclude statistically that the daily hypothesis is 'right', at least relative to the half day hypothesis it seems not to matter [16].

To investigate the hypothesis of independence of information events across days, we classify days into no-event days, good-event days and bad-event days and perform runs tests on the classified days. If information events are independent, we would not expect to find runs in event days beyond that due to statistical chance. We classify days by computing the market maker's theoretical posterior beliefs, at the end of the day, using our estimated parameters and the daily trade data. Days are then assigned to be high event, low event or no event based on which of the three 'events' is most likely according to the end of day posterior.

We perform two types of runs tests on this resulting sequence of 60 classified days for each stock. We first look for independence across the occurrence of information events by sorting days into event and no-event days, and then performing a runs test on this sequence. We then consider independence of information across days by looking at runs in the occurrence of good and bad news. To investigate this, we sort the event days into good and bad days, and we

[16] This means that our results regarding sequences and reversals are not artifacts of our time horizon. That is, because of the correlated nature of uninformed trade, trade sequences are not overly informative even if information events are assumed to occur with greater frequency.

Table 7
Testing independence of information events

	ASH	BK	BNL	BUD	DE	GT
# events	32	18	28	46	13	14
# no events	28	42	32	14	47	46
Runs	24	18	28	18	15	19
Mean	30.87	26.20	30.87	22.47	21.37	22.47
Variance	14.61	10.34	14.61	7.45	6.69	7.45
Good events	18	16	23	1	4	4
Bad events	14	2	5	45	9	10
Runs	18	5	9	3	4	5
Mean	16.75	4.55	9.21	2.95	3.85	6.71
Variance	7.49	0.55	2.19	0.04	0.44	2.02

This table presents results of runs tests for independence. Events are days in which new information is predicted to have occurred, with no-events denoting days on which it did not. The event days are divided into good event days and bad event days based on the predictions of the model. The means and variances are those arising from a runs test assuming events are independent.

perform a runs test on that sequence. The results of these tests are given in Table 7.

For ASH, this classification scheme resulted in 18 days classified as high event days, 14 days classified as low event days and 28 days classified as no-event days. According to our estimated parameters, the expected number of days of each type is: 17.75 high event days, 15 low event days and 27.25 no-event days. So our classification scheme seems to work well. In the event versus no-event classification there were 24 runs. Generally, let e be the total number of days in which events occur and n be the total in which no events occur. It can be shown (see, e.g., Moore, 1978) that the total number of runs under the null hypothesis that the series is independent is approximately normally distributed with mean $m = 2en/(e+n)+1$ and variance $s^2 = 2en(2en-e-n)/((e+n)^2(e+n-1))$. Our mean is 30.87 and variance 14.61, so our observed number of runs, 24, is not at all surprising. The hypothesis of independence of information events is thus not brought into question by this test.

Given this result for ASH, we ask whether, when events occur, good and bad events are independent. We do a runs test on the sequence of event days, and we find 18 runs. The approximate normal distribution has mean 16.75 and variance 7.49, so our observation is not at all unusual under the null. Once again, the hypothesis of independence is consistent with our observables.

When we examine the stocks as a group, we see that all of the stocks have fewer runs than expected under the hypothesis of independence and that BK and DE have significantly fewer runs than expected. So there appears to be some positive correlation in events. This correlation is not particularly strong, and with the small number of days that we can reasonably work with it, it would be difficult

Table 8
Price effects of trades

Variable	ASH	BK	BNL	BUD	DE	GT
Constant	−0.600 (0.166)	−0.207 (0.255)	0.468 (0.300)	−0.351 (0.346)	0.110 (0.514)	0.000 (0.206)
LB	0.030 (0.011)	0.044 (0.009)	0.084 (0.021)	0.020 (0.007)	0.027 (0.010)	0.019 (0.004)
SB	0.023 (0.005)	0.005 (0.010)	0.009 (0.012)	0.007 (0.009)	0.018 (0.008)	−0.001 (0.004)
LS	−0.046 (0.013)	−0.034 (0.012)	−0.094 (0.024)	−0.025 (0.008)	−0.054 (0.010)	−0.015 (0.007)
SS	−0.018 (0.008)	−0.013 (0.013)	−0.042 (0.014)	0.006 (0.008)	0.001 (0.013)	−0.004 (0.006)
Adjusted R^2	0.553	0.324	0.400	0.248	0.374	0.259
F	1.49	3.64	3.98	3.05	5.10	4.12

This table gives parameter estimates for the regression of the change in CRSP closing price on a constant and the number of LB, SB, LS, and SS per day. Standard errors are given in parentheses.

to estimate reliably. This suggests that while for some stocks an alternative specification might fit better, the overall assumption of independence is not seriously challenged by the data.

6. Price effects of trades

In our empirical work we have not used prices (except indirectly through the trade classification scheme) in determining our estimates, relying instead only on information derived from trades and no-trade intervals. Our estimated parameters, however, have implications for prices through the pricing equations, and this linkage provides a natural mechanism to examine the validity (and reasonableness) of our results. The most obvious implication of our results is that prices should rise following a buy, either large or small, and fall following a sale, either large or small. Second, we predict that large and small trades have differential effects on prices for all of our stocks except Ashland Oil. These general implication can be explored using daily data [17].

To consider these general implications for prices, we look at the change in price from the close of one day to the close of the next day. For each stock in our sample, we use CRSP closing prices for the 60 trading day period we use in our estimation. We regress the change in the CRSP closing price on a constant and trade data. As we examine daily price changes, it makes sense to look at daily trade data. Our model implies that it is the numbers of large buys, small buys, large sells, and small sells per day that determine the movement of prices. Accordingly, we regress the change in the CRSP closing price from day $i-1$ to day i (ΔCRSPP_i) on a constant, the number of large buys on day i (LB_i), the number of small buys on day i (SB_i), the number of large sells on day i (LS_i), and the number of small sells on day i (SS_i).

The results of these regressions, one for each stock in our sample, are reported in Table 8. We first note that the coefficients generally have the predicted sign. That is, we predict that prices should rise in response to large buys, and an examination of the coefficients reveals that the sign on LB is positive and significant for every stock. Correspondingly, the coefficient on large sells is negative and significant for each stock. The coefficients on small trades always have the predicted sign, or are insignificant. These coefficients on small trades are closer to zero than are the coefficients on large trades.

To examine whether large and small trades differ significantly in information content for price changes, we tested, for each stock separately, the joint restriction

[17] Our model also has implications for the behavior of prices following no-trades and in response to sequences or reversals of trades. To examine these sequence implications in detail would require a greater degree of confidence in the empirical validity of our pricing equations than we feel is justified. There are just too many other things going on, such as inventory effects and public information events, to take the intraday implications of the pricing equations literally.

that the coefficient on LB equals the coefficient on SB (and the corresponding restriction on the sale parameters). The $F(2, 57)$ values for this restriction range from 3.05 to 5.10 for BK, BNL, BUD, DE, and GT, indicating a rejection of this restriction at (at least) the 0.05 level for each of these stocks. Interestingly, we cannot reject this restriction for ASH, the one stock for which our parameter estimates predict no differential information content for small and large trades. The F value for ASH is 1.49, which does not indicate rejection of the no-difference restriction at even the 0.20 level.

These results show that large buys raise prices by more than do small buys, and large sells depress prices by more than do small sells for each stock in our sample except ASH. For ASH there is little value in separating trades by size as there is no difference in their information content. This is exactly what our trade based model predicts. That estimates based on trades alone in conjunction with a simple pricing model could yield insight into price behavior is surprising. We view this as strong support for the validity of our approach.

7. Conclusions

We have developed and empirically tested a model of the trading process. The history dependent structure of our model allows for greater flexibility in the behavior of uninformed traders, and our incorporation of explicit trade sizes allows for testing for information differences in trade quantity. We show that the trade process provides a wealth of information to market watchers. Large trades generally have twice the information content of small ones, while large buy and sell orders only weakly differ in their informativeness. Our estimation reveals that uninformed trades are highly positively correlated. Our results also show how sequences and reversals of trades provide differing information, with the latter being particularly informative.

One conclusion that arises from our work is that the stochastic process of trades and prices is not invariant across time. In particular, because of the differential information conveyed by sequences, reversals, and non-trading periods, observations drawn at different points of the trading day, for example, will not be equivalent with respect to the information they convey. This raises a thorny problem for researchers, as it suggests that sampling transactions data at fixed time intervals is not appropriate. Similarly, our results suggest that removing the patterns in the data to obtain stationarity runs the very real risk of removing the important information contained in those patterns. For some research issues, in particular those involving information-related questions, these difficulties are likely to be severe.

A second contribution of our research is to highlight an enhanced role for trade process data. Our analysis here uses only trade process data, in contrast to the more traditional approach of using price change data to infer the information content of particular orders. For example, the extensive testing of price effects of large trades (see Keim and Madhavan, 1994; Seppi, 1992) uses price data. Barclay and Warner (1993) use price change data for take-over candidates to show that most of the cumulative price change prior to take-over announcements is due to medium-sized trades.

The price-based approach has both advantages and disadvantages relative to our trade-based approach. The most obvious advantage is that, to the extent our interest is in prices and not in trades, knowing how prices respond to trades is clearly important. The most obvious disadvantage with the price-based approach is that prices are influenced by many factors other than the current trade. The most important of these for the questions we consider are inventory and the sequencing of trades [18]. Without precise models of how these factors affect prices, however, it is impossible to determine which price effects are information-based. A second, and related, disadvantage, is that price effects need not relate to discrete events. In particular, if trades (and thus prices) are not Markov, then estimating price effects of individual trades will miss much of the economic effect we seek. Focusing on the indirect evidence in prices rather than the direct evidence in trades thus runs the risk that empirical tests are both mis-specified and mis-interpreted.

Understanding the information implicit in market data is a complex process, and not surprisingly no one approach captures all of its dimensions. The model and approach we develop here provides a mechanism for extracting the information contained in the trade process. For many issues, trade process information provides the best, and most direct, approach for analysis.

References

Barclay, M., Warner, J., 1993, Stealth trading, Journal of Financial Economics.

Berndardt, D., Hugheson, E., 1993, Intraday trade in dealership markets, Working Paper, Queens University.

Easley, D., Kiefer, N., O'Hara, M., 1993, One day in the life of a very common stock, Review of Financial Studies, forthcoming.

Easley, D., O'Hara, M., 1987. Price, trade size, and information in securities markets. Journal of Financial Economics 19, 69–90.

Engle, R., Russell, J., 1995, Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model, Working Paper, University of California, San Diego.

Froot, K., Sharfstein, D.S., Stein, J.C., 1992. Herd on the street: Informational inefficiencies in a market with short-term speculation. Journal of Finance 47, 1461–1484.

Glosten, L., Harris, L., 1988. Estimating the components of the bid/ask spread. Journal of Financial Economics 21, 123–142.

[18] Barclay and Warner (1993) deal with these inventory and sequencing concerns by positing informal models of how these factors affect price changes. At that point, their analysis suffers from limitations similar to those with which we must deal. For example, in arguing that inventory is not the source of their results, they posit a model in which the inventory effect is proportional to the trade size.

Glosten, L., Milgrom, P., 1985. Bid, ask and transaction prices in a specialist market with heterogenously informed traders. Journal of Financial Economics 14, 71–100.

Hasbrouck, J., 1988. Trades, quotes, inventories, and information. Journal of Financial Economics 22, 229–252.

Hasbrouck, J., 1991. Measuring the information content of stock trades. Journal of Finance 46, 179–207.

Jones, C., Kaul, G., Lipson, M., 1994. Transactions, volume and volatility. Review of Financial Studies 7, 631–651.

Keim, D., Madhavan, A., 1994, The upstairs market for large block transactions: analysis and measurement, Working Paper, Wharton School, University of Pennsylvania.

Kyle, A.P., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1336.

Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. Journal of Finance 42, 733–746.

Moore, P.G., 1978. Nonparametric statistics: Runs. International Encyclopedia of Finance, W.J. Kruskal and J. Tanur (eds.), The Free Press, 651–655.

Seppi, D., 1992. Block trading and information revelation around quarterly earnings announcements. Review of Financial Studies 5, 281–305.