THE JOURNAL OF FINANCE • VOL. LI, NO. 3 • July 1996

Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow

DAVID EASLEY, NICHOLAS M. KIEFER, and MAUREEN O'HARA*

ABSTRACT

Purchased order flow refers to the practice of dealers or trading locales paying brokers for retail order flow. It is alleged that such agreements are used to "cream skim" uninformed liquidity trades, leaving the information-based trades to established markets. We develop a test of this hypothesis, using a model of the stochastic process of trades. We then estimate the model for a sample of stocks known to be used in order purchase agreements that trade on the New York Stock Exchange (NYSE) and the Cincinnati Stock Exchange. Our main empirical result is that there is a significant difference in the information content of orders executed in New York and Cincinnati, and that this difference is consistant with cream-skimming.

A STRIKING FEATURE OF global capital markets is the proliferation of trading venues. Whereas established trading centers such as London and New York traditionally controlled the trading volume in securities, new markets have begun to erode that dominance. Frankfurt, Milan, Madrid, and Amsterdam are but a few of the locales now challenging the dominance of the London Stock Exchange for trading in Europe. In the United States, this trend can be seen in the dramatic growth of electronic trading systems such as Instinet and POSIT (the Portfolio System for Institutional Trading), in the increasing volume of the Nasdaq, and in the growth of regional exchanges such as the Cincinnati Stock Exchange (CSE) and the Arizona Stock Exchange.

Whether this fragmentation of trading is desirable is debatable. Increased competition could reduce the monopoly power of price-setting agents and thus result in better execution and prices for traders. But this same competition also reduces the liquidity available in any one setting, thereby potentially limiting any market's ability to provide stable prices. Moreover, liquidity facilitates the crucial price discovery role of markets. As order flow fragments, the ability of prices to aggregate information can be reduced, and with it the efficiency of the market. This problem can be exacerbated if markets choose to compete by focusing on particular components of the order flow. By "cream-skimming" the order flow, new markets could undermine both the viability of old markets, and of the trading process itself.

* Cornell University. We thank Charles Jones, Charles Lee, George Sofianos, and seminar participants at Cornell University, the University of Iowa, the London Business School, Stanford University, the University of Utah, and Washington University for helpful comments. We also thank Joseph Paperman and Fan Xu for programming assistance. This research is supported by National Science Foundation Grant SBR93-20889.

This dichotomy between the beneficial "profit-sharing" role and the destructive "cream-skimming" role of market competition lies at the heart of the debate over purchased order flow. Purchased order flow refers to the practice of dealers or trading locales paying brokers for retail order flow. While perhaps present in some form in the past, this practice increased dramatically in the late 1980s, to the point that the New York Stock Exchange (NYSE) now estimates that 35 percent of small trades are lost to purchased order flow.[1] In the most definitive study of this practice, the National Association of Securities Dealers (NASD (1991)) noted that most payment for order flow agreements provide payments of 1 to 2 cents per share for specific stocks, limit the transaction size they will accept, and require a minimum order flow per month (at least 100,000 shares).[2] Moreover, these agreements generally stipulate that orders will execute at the national best market bid or offer (NBBO) price.[3]

These constraints on the size and composition of trades have led to criticisms that order purchase agreements seek only to divert small, retail trades to their specific locale, leaving the larger, potentially more information-based trades to the NYSE to handle. By taking these profitable "liquidity trades" from the market, order purchasers can afford to match the NYSE bid or offer price, pay rebates to brokers, and still make substantial profits. If this view is correct, prices (and price discovery) on the NYSE may worsen due to the adverse selection in the remaining orders. Moreover, while brokers are enriched, there is no necessary reason why traders are better off. Advocates of the practice contend, however, that it is the large spreads of the NYSE that permit profitable diversions. They argue that for actively traded stocks, the minimum NYSE spread size of 1/8 is too large, and the rebates accompanying diversion of these orders provide one mechanism for reducing this monopoly rent. Under this view, diverting order flow enhances the welfare of traders, and affects the NYSE only by reducing the rents earned by specialists.

Several researchers (Lee (1992; 1993); Blume and Goldstein (1991), McInish and Wood (1992), Huang and Stoll (1994)) have attempted to resolve this conundrum by investigating the price effects of inter-market competition. In general, these authors have found that execution costs on the NYSE are lower, but the reasons for this are not apparent. Blume and Goldstein (1991, p. 14) noted that while the NYSE generally had the best quote, "this result does not

[1] Specifically, Cochrane (1993) found that the NYSE now clears only 65 percent of orders for trades of 100 to 2099 shares.

[2] Lee (1992) argues that the practice began on a large scale in 1988, and grew dramatically in 1989 and 1990. He also notes that order purchase agreements also typically refuse limit orders, program trades, or any institutional orders. An excellent description of the regulatory changes that led to the development of this practice is given in Huang and Stoll (1994).

[3] Blume and Goldstein (1991) note that a criticism of the largest order purchasers (in particular the regional exchanges like Cincinnati and firms such as Madoff Investments) is that they use the quote of the NYSE to set their trading prices. They interpret their empirical finding that the best intermarket quote is generally that of the NYSE to support this criticism. A second issue connected with purchased order flow is whether traders are actually given the best price. Lee (1993) presents evidence that trades on the NASD execute at worse prices than the NYSE. Because our focus in this article is on trades and not on prices, we do not address this issue.

necessarily imply that execution prices on the NYSE are better since the investor should always receive the best intermarket quote regardless of the particular market on which an order is executed."[4] Lee (1992, p. 4), in a thoughtful and careful comparison of eight trading locales, concluded that "in effect, the better performance at the NYSE is due more to greater public participation on that exchange than to lower profit margins on specialists trades." He cautioned, however, that "these results, while shedding light on comparative execution costs, do not address the cream-skimming versus competition issue." This is because even if equal execution prices are found, the cream-skimming problem can still be quite severe. Since order purchase agreements typically promise to match the NYSE bid or offer price, clearing diverted "liquidity" trades at these prices overcharges the trader, while providing large profits to the broker.[5] In effect, all prices are too high, and hence comparisons between them are misleading.[6]

In this research, we develop a test of the cream-skimming versus competition issue that is not based on prices. Our approach uses the more basic information in the trade flow to infer any difference in information content between trading locales. We use the formal structure of a market microstructure model to formulate the learning problem confronting agents watching the trade flow. By estimating this model for a sample of stocks trading on the NYSE and an alternative site, we can determine whether the information content of trades varies by order locale. Estimating the information content of the order flow also allows us to discover any systematic differences between the diverted orders and the *remaining* (non-diverted) order flow, thus shedding light on whether purchases of order flow impose externalities on the market process as a whole.

The particular experiment we undertake is to compare stocks trading on the NYSE and on the Cincinnati Stock Exchange. Lee (1993) found that the price behavior on the Cincinnati Exchange was closest to that of the NYSE, a pattern consistent with the "free riding" on prices allegedly underlying the execution of purchased orders. While the Nasdaq is also used for order diversion, there have been widespread allegations of price fixing, overcharging of customers, and other irregularities.[7] As our goal is to investigate only infor-

[4] Harris, McInish, and Wood (1994) argue that a further difficulty in interpreting price data are hidden limit orders on the NYSE. If specialists do not show all limit orders, then actual liquidity may differ from apparent liquidity, and trading prices may be affected.

[5] This problem is akin to the classic adverse selection problem found in insurance. If a relatively healthy class of insurees is diverted to a particular company, then by charging them the same rate paid by the nondiverted population, the insurance company can make a profit. Examining the premia alone would thus not be informative, but examining the underlying health history or claim experience for each company along with its premia would be.

[6] An additional confounding effect is that price quotes may reflect idiosyncratic factors such as inventory. If regional exchange market makers are more risk averse, for example, then a finding of "worse prices" does not necessarily shed light on the information content of the order flow.

[7] The price-fixing issue is addressed in Christie and Schultz (1994). For a description of alledged customer mistreatment, and the SEC's investigation thereof, see "In-house trades may be costly for small investors", Wall Street Journal, December 20, 1994.

mation-related differences, the comparable behavior found in Cincinnati provides an ideal venue in which to test for differences from the NYSE.

Our main finding is that there is a significant difference in information content between stock trades executed in Cincinnati and those executed on the NYSE, and that this difference is consistent with "cream-skimming." The hypothesis that trades are randomly assigned to New York or Cincinnati, i.e., equal information content of orders on the two exchanges, is strongly rejected when tested against the unrestricted alternative. More important is that the hypothesis of equal information content is also strongly rejected when tested against the alternative hypothesis of cream-skimming by Cincinnati. Conversely, the equal information content hypothesis cannot be rejected (even at the 0.10 level) when tested against the unlikely alternative of reverse cream skimming by Cincinnati (that is, more information content in Cincinnati). For our sample of stocks as a whole, the probability of informed trade in New York is approximately 44 percent higher than that in Cincinnati. These findings give strong evidence that the diversion of orders from the NYSE is not purely competitive. Our trade-based estimates also provide intriguing insights into both the information differences between markets and those between individual stocks. As we argue, these informational differences have important implications both for regulatory policies in the United States, and for market viability in the broader global market.

This article is organized as follows. In section I, we develop a sequential trade market microstructure model that incorporates asymmetrically informed traders. Unlike the standard sequential trade model (see Glosten and Milgrom (1985); Easley and O'Hara (1987)), the model here is in continuous time, a feature that greatly aids in its estimation for the high volume stocks typically found in purchased order flow. We then derive the likelihood function for the model, and discuss its estimation. In Section II, we outline the methodology we use to test for differences in the information content of orders between trading venues. In Section III, we discuss the data used in the analysis, and the potential problems that arise in its interpretation. Our results are presented in Section IV. The article's final section is a discussion of the policy implications of our work, both as it applies to U.S. regulatory issues and to the broader context of competition between market locales.

I. The Model

In this section, we set out a sequential trade microstructure model in which trade takes place in continuous time. The model is most closely related to that developed in Easley, Kiefer, O'Hara, and Paperman (1994) in that it views trading as occurring continuously throughout a sequence of trading days. The model in this article incorporates potentially different arrival rates of informed and uninformed traders, and has the important new feature that it allows for trading in different market locales.

A. Trade Process

In the model, individuals trade a risky asset and money with market makers over a series of trading days. There are two markets in which the security trades, denoted N and C. Within each market, there is a market maker who always stands ready to buy or sell one unit of the asset. We assume that both market makers watch trades in each market, so they have common information.

Within any trading day time is continuous, and it is indexed by $t \in [0, T]$. The asset being traded has an eventual value given by the random variable V. Before the start of any trading day, nature determines whether an information event about V will occur. Information events are independently distributed and occur with probability α. These events are bad news with probability δ, and are good news with probability $1 - \delta$. We denote the expected value of the stock conditional on good news by $\overline{V}$; similarly it is $\underline{V}$ conditional on bad news.[8] The unconditional value of the stock (and its value if there has been no information event) is then $V^* = \delta\underline{V} + (1 - \delta)\overline{V}$.

Orders to trade can arise from either informed traders (those who have seen any signal) or from uninformed traders.[9] On any day, arrivals of uninformed buyers and sellers are determined by independent Poisson processes. Uninformed buyers arrive at rate ε_B, and uninformed sellers arrive at rate ε_s, where these rates are defined per minute of the trading day. On days when there has been an information event, informed traders also arrive. Informed traders are assumed to be risk neutral and competitive, so that a trader who has observed good news will buy if the price is not yet at $\overline{V}$, and one who has observed bad news will sell if the price is not yet at $\underline{V}$. We assume that the arrival of news to one trader at a time, and the arrival of that trader to the market, also follows a Poisson process. The arrival rate for this process is μ. All of the processes are assumed to be independent.

Orders can execute in either market N or in market C. We let β denote the probability that an uninformed order will execute on market N, with $1 - \beta$ the corresponding probability that it executes on market C. Similarly, we let γ denote the probability that an informed order executes in market N, and $1 - \gamma$ the probability that it does so in market C.[10] The Poisson structure of arrivals in our model means that when we split the arrival process by independent probabilities, the resulting arrival processes are exactly equivalent to two independent Poisson processes with rates weighted by these probabilities. Thus, the arrival rate of uninformed buy orders to market N is $\beta\varepsilon_B$, and to

[8] In our empirical work we will look at several stocks at once. Let V_j be the random variable giving the eventual value of asset j. We assume that information events about the V_js are serially and cross sectionally independent. This allows us to consider trade and pricing of each asset separately. In the text we drop the subscript on asset values for this reason.

[9] Traders begin each day with common information. In particular, information from past days is common. The values $\underline{V}$ and $\underline{V}$ are conditional expected values for the day in question. We do not consider prices over multiple days, and so we do not index values by day.

[10] We treat β and γ as exogenous. The notion is that these parameters are the result of choices by brokers who execute traders' orders.

market C it is $(1 - \beta)\varepsilon_B$, while uninformed sell orders arrive at rates $\beta\ \varepsilon_s$ and $(1 - \beta)\varepsilon_s$, to markets N and C, respectively. Similarly, the arrival rate of informed trades to market N is simply $\gamma\mu$, and it is $(1 - \gamma)\mu$ to market C.

This trading process is described by the tree diagram given in Figure 1. The first node of the tree corresponds to nature's decision as to whether an information event occurs. If an event occurs, nature then determines if it is good news or bad news. The three nodes (no event, good event, and bad event) before the dotted line in Figure 1 occur only once per day. Given the node selected for the day, traders arrive at the two markets according to the relevant Poisson processes. So, for example, on good event days, the arrival rates for orders in market N are $\beta\varepsilon_B + \gamma\mu$ for buy orders, and $\beta\varepsilon_s$ for sell orders. On bad event days, the arrival rate in market N is $\beta\varepsilon_B$ for buy orders and $\beta\varepsilon_s + \gamma\mu$ for sell orders. On days with no information event, theses rates are simply $\beta\varepsilon_B$ and $\beta\varepsilon_s$, respectively. The rates for orders in market C are derived similarly.

B. Trades, Beliefs and Quotes

Each day nature selects one of the three branches of the tree. The market makers know the probability attached to each branch, and the order arrival processes for each of the branches, but they do not know which of the three branches nature has selected. Each market maker is a Bayesian who uses the arrival of orders (both those in market N and those in market C) to update beliefs about the occurrence of information events. Since days are independent, we can analyze the evolution of beliefs separately on each day. Let $P_t(G)$, $P_t(B)$, and $P_t(O)$ be each market maker's prior belief about the events "good news" (G), "bad news" (B), and "no news" (O) at time t. Thus, the common prior belief at time 0 is $P_0 = \{\alpha(1 - \delta), \alpha\delta, 1 - \alpha\}$.

If an order arrives at time t, each market maker will update his beliefs using Bayes rule. Let $X(t)$ be the event that an order of type $X \in \{S_N, S_C, B_N, B_C\}$ arrives at time t; where S_N is a sell order in N, S_C is a sell order in C, B_N is a buy order in N, and B_C is a buy order in C. Let $h(t)$ be the history of arrivals of orders prior to time t. So, for example, if an order to sell arrives in N at time t, each market maker's posterior probability on a good event will be

$$P_t(G|S_N(t)) = \frac{P_t(G)\varepsilon_s\beta}{\varepsilon_s\beta + P_t(B)\mu\gamma}. \tag{1}$$

Similarly, the posterior probability on a bad event will be

$$P_t(B|S_N(t)) = \frac{P_t(B)(\varepsilon_s\beta + \mu\gamma)}{\varepsilon_s\beta + P_t(B)\mu\gamma}, \tag{2}$$

and the posterior probability on no event will be

$$P_t(O|S_N(t)) = \frac{P_t(O)\varepsilon_s\beta}{\varepsilon_s\beta + P_t(B)\mu\gamma}. \tag{3}$$

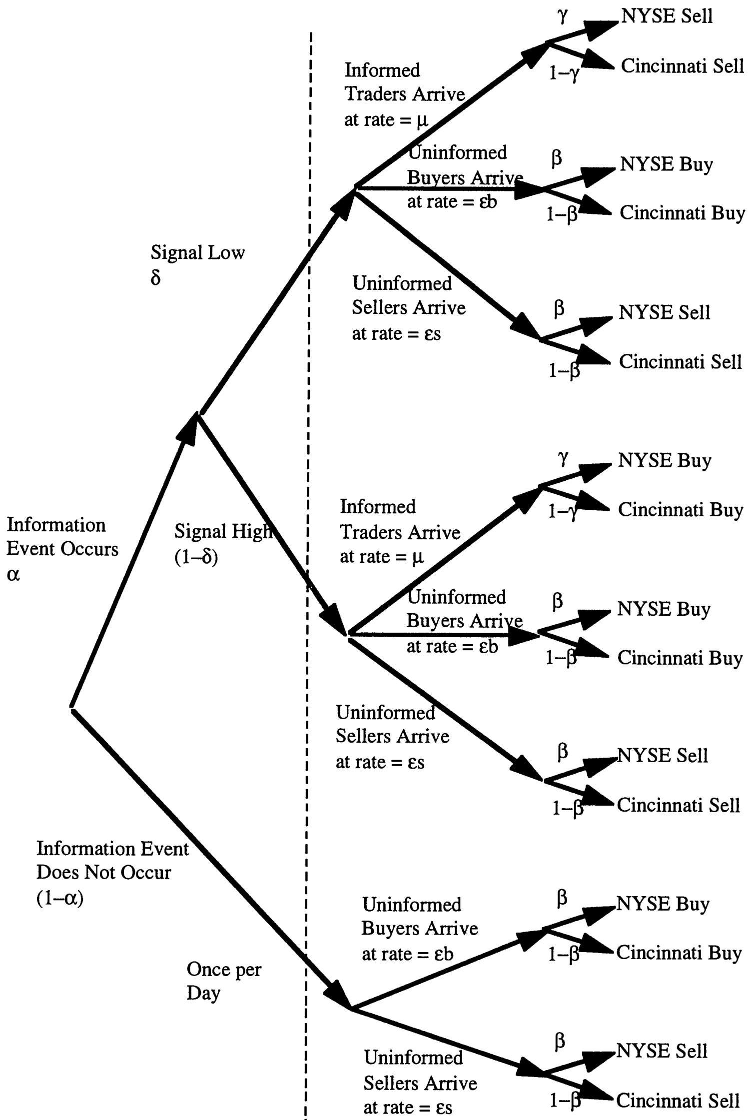

Figure 1. Tree diagram of the trading process. The variables are α, the probability of an information event; δ, the probability that the information is bad news; μ, the arrival rate of traders who know the new information if it exists (i.e., the informed traders); ε_B, the arrival rate of uninformed buyers; ε_s, the arrival rate of uninformed sellers; γ, the probability that an informed trade is executed in New York; and β, the probability that an uninformed trade is executed in Cincinnati. Nodes to the left of the dotted line occur once per day.

Posterior probabilities conditional on the arrival of other orders are calculated similarly.

We compute zero expected profit quotes for the N and C market makers. These are the quotes that should prevail if both market makers are risk neutral and if both markets are competitive. If actual prices diverge from these theoretical zero-expected profit prices, then this divergence may reflect some combination of factors such as risk aversion, inventory costs, transactions costs, price discreteness, and market power. In our empirical work we do not use prices, so no hypothesis on pricing behavior is necessary for our conclusions. Nonetheless, it is instructive to see how the flow of informed and uninformed traders to each market affects competitive prices.

At any time t, the market N zero expected profit bid price, $b_N(t)$, is the expected value of the asset conditional on $h(t)$ and $S_N(t)$. Thus,

$$b_N(t)\mathrm{E}[V \mid h(t),\, S_N(t)]. \tag{4}$$

The market C bid price, $b_C(t)$, at time t is

$$b_c(t) = \mathrm{E}[V \mid h(t),\, S_C(t)]. \tag{5}$$

The ask prices, $a_N(t)$ and $a_C(t)$, are defined similarly.

Using the posterior probabilities in equation (4), we can write the bid price per share, $b_N(t)$, in N as

$$b_N(t) = V^* - (\overline{V} - \underline{V})\left[\frac{(1-\delta)P_t(B)(\varepsilon_s\beta + \mu\gamma) - \delta P_t(G)\varepsilon_s\beta}{\varepsilon_s\beta + P_t(B)\mu\gamma}\right]. \tag{6}$$

Using the posterior probabilities conditional on arrival in market C of a sell order, we have that

$$b_C(t) = V^* - (\overline{V} - \underline{V})\left[\frac{(1-\delta)P_t(B)(\varepsilon_s(1-\beta) + \mu(1-\gamma)) - \delta P_t(G)\varepsilon_s(1-\beta)}{\varepsilon_s(1-\beta) + P_t(B)\mu(1-\gamma)}\right]. \tag{7}$$

Similar equations determine the N and C market makers' ask prices, $a_N(t)$ and $a_C(t)$.

Our focus in this article is on how informed and uninformed trading differs between markets. Given competitive market makers, any differences in the information content of trades would affect quotes in each market. The key parameter determining this in our analysis is β/γ, the ratio of uninformed to informed trading propensities in market N. Calculation shows that in market N, the bid is increasing in β/γ, the ask is decreasing in this parameter, and so the spread is also *decreasing* in β/γ. Conversely, in market C the bid is decreasing in β/γ, the ask is increasing in this parameter, and thus the spread in C is *increasing* in β/γ.

The intuition for these effects can be seen by noting that the probability that a trade at time t is informed, conditional on the occurrence of a trade at time t, is

$$\mathrm{PIN}(t) = \frac{(1 - P_t(O))\mu}{(1 - P_t(O))\mu + (\varepsilon_B + \varepsilon_s)(\beta/\gamma)} \tag{8}$$

in market N, and it is

$$\mathrm{PIC}(t) = \frac{(1 - P_t(O))\mu}{(1 - P_t(O))\mu + (\varepsilon_B + \varepsilon_s)(1 - \beta)/(1 - \gamma)} \tag{9}$$

in market C. Reducing the value of β/γ increases the probability of informed trade in N, and this results in worse prices there. The lower β/γ in market C, however, means that prices should be lower in that market. If $\beta/\gamma = 1$, then prices should be the same in markets C and N. Alternatively, if prices in C are simply set at, or matched to, the level of prices in market N, then the market maker in C would earn expected profits from executing trades if $\beta/\gamma < 1$. This expected profit from trading would allow the purchase of order flow. It would be natural to expect that through the process of entry and exit of markets that fully allocated costs and revenues would balance in each market.

If, like the market makers, we observed the order arrival process and we knew the parameters of the problem, $\theta = (\alpha, \delta, \varepsilon_B, \varepsilon_s, \mu, \beta, \gamma)$, then we could compute the stochastic processes of bids, asks, and spreads. Although we can observe trades and quotes, we do not know the parameters. These parameters can be estimated, however, from the data on order arrivals. This estimation also allows us to determine the information content of orders in our two market settings.

C. The Likelihood Function

Estimating the parameter vector $\theta = (\alpha, \delta, \varepsilon_B, \varepsilon_s, \mu, \beta, \gamma)$ is much more complex than just estimating arrival rates from independent Poisson processes. 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 arrival rates of uninformed or informed traders at various markets. We observe arrivals of trades in each market, but we do not observe which traders are uninformed or informed. Estimation of these parameters thus requires a structural model. We can infer the direction of orders from trades using standard algorithms. Our model then provides the requisite structure necessary to extract information on the parameters from the observable variables: buys and sells in each market.

In our model, buys and sells follow one of three Poisson process on each day. We do not know which process is operating on any day, but we do know that the data reflect the underlying information structure, with more buys expected on days with good events, and more sells on days with bad events.[11] Similarly, on no-event days, there are no informed traders in the market, and so fewer trades arrive. We also know that the buys and sells in each market provide

[11] Note that, unlike in a Kyle framework, trades in our model are not aggregated, so it is the composition and total number of trades that determines beliefs and, thus, prices. See O'Hara (1995) for a discussion of the various theoretical market microstructure models.

information on the split of orders by type of trader between markets. These rates and probabilities are determined by a mixture model in which the weights on the three possible components (i.e., the three branches of the tree reflecting no news, good news, and bad news) reflect their probability of occurrence in the data. The next step in our analysis is to construct this mixture model.

We first consider the likelihood of order arrivals on a day of known type. On a bad event day, sell orders at market N arrive at rate $\mu\gamma + \beta\varepsilon_s$, reflecting that both informed and some uninformed traders will be selling. Similarly, in market C, this arrival rate is given by $(1 - \gamma)\mu + (1 - \beta)\varepsilon_s$. The buy order arrival rates are given by $\beta\varepsilon_B$ in market N and by $(1 - \beta)\varepsilon_B$ in market C. The exact distribution of these statistics in our model is independent Poisson. Thus, the likelihood of observing any sequence of orders that contains B_C buys on market C, B_N buys on market N, S_C sells on market C, and S_N sells on market N on a bad-event day of total time T is given by

$$e^{-\varepsilon_b\beta T}\frac{(\varepsilon_b\beta T)^{B_N}}{B_N!} * e^{-(\mu\gamma+\varepsilon_s\beta)T}\frac{[(\mu\gamma + \varepsilon_s\beta)T]^{S_N}}{S_N!} * e^{-\varepsilon_b(1-\beta)T}\frac{(\varepsilon_b(1-\beta)T)^{B_C}}{B_C!} * e^{-(\mu(1-\gamma)+\varepsilon_s(1-\beta))T}\frac{[(\mu(1-\gamma) + \varepsilon_s(1-\beta))T]^{S_C}}{S_c!} \quad (10)$$

Similarly, on a no-event day, the likelihood is

$$e^{-\varepsilon_b\beta T}\frac{(\varepsilon_b\beta T)^{B_N}}{B_N!} * e^{-\varepsilon_s\beta T}\frac{[\varepsilon_s\beta T]^{S_N}}{S_N!} * e^{-\varepsilon_b(1-\beta)T}\frac{(\varepsilon_b(1-\beta)T)^{B_C}}{B_C!} * e^{-(\varepsilon_s(1-\beta))T}\frac{[(\varepsilon_s(1-\beta))T]^{S_c}}{S_c!} \quad (11)$$

Finally, on a good event day, the likelihood is

$$e^{-(\varepsilon_b\beta+\mu\gamma)T}\frac{[(\varepsilon_b\beta + \mu\gamma)T]^{B_N}}{B_N!} * e^{-(\varepsilon_s\beta)T}\frac{[(\varepsilon_s\beta)T]^{S_N}}{S_N!} * e^{-(\varepsilon_b(1-\beta)+\mu(1-\gamma))T}\frac{[(\varepsilon_b(1-\beta) + \mu(1-\gamma))T]^{B_C}}{B_c!} * e^{-(\varepsilon_s(1-\beta))T}\frac{[(\varepsilon_s(1-\beta))T]^{S_c}}{S_c!} \quad (12)$$

It is evident from equations (10), (11), and (12) that the number of buys and sells in each market (B_N, B_C, S_N, S_C) is a sufficient statistic for the data given T. Thus, to estimate the order arrival rates in each market of the buy and sell processes, we need only consider the total numbers of buys and sells in each market on any day.

The likelihood of observing (B_N, B_C, S_N, S_C) on a day of unknown type is the weighted average of equations (10), (11), and (12) using the probabilities of each type of day occurring. These probabilities of a no-event day, a bad-event day, and a good-event day are given, respectively, by $1 - \alpha$, $\alpha\delta$, and $\alpha(1 - \delta)$. The likelihood for a day of unknown type is thus

$$\begin{aligned} L((B_N, S_N, B_C, S_C)|\theta) = {} & (1-\alpha)e^{-\varepsilon_b\beta T}\frac{(\varepsilon_b\beta T)^{B_N}}{B_N!} * e^{-(\mu\gamma+\varepsilon_s\beta)T}\frac{[(\mu\gamma+\varepsilon_s\beta)T]S_N}{S_N!} \\ & * e^{-\varepsilon_b(1-\beta)T}\frac{(\varepsilon_b(1-\beta)T)^{Bc}}{B_c!} \\ & * e^{-(\mu(1-\gamma)+\varepsilon_s(1-\beta))T}\frac{[(\mu(1-\gamma)+\varepsilon_s(1-\beta))T]^{S_c}}{S_c!} \\ & + \alpha\delta e^{-\varepsilon_b\beta T}\frac{(\varepsilon_b\beta T)^{B_N}}{B_N!} * e^{-\varepsilon_s\beta T}\frac{[\varepsilon_s\beta T]^{S_N}}{S_N!} \\ & * e^{-\varepsilon_b(1-\beta)T}\frac{(\varepsilon_b(1-\beta)T)^{Bc}}{B_c!} * e^{-(\varepsilon_s(1-\beta))T}\frac{[(\varepsilon_s(1-\beta))T]^{S_c}}{S_c!} \\ & + \alpha(1-\delta)e^{-(\varepsilon_b\beta+\mu\gamma)T}\frac{[(\varepsilon_b\beta+\mu\gamma)T]^{B_N}}{B_N!} \\ & * e^{-(\varepsilon_s\beta)T}\frac{[(\varepsilon_s\beta)T]^{S_N}}{S_N!} \\ & * e^{-(\varepsilon_b(1-\beta)+\mu(1-\gamma))T}\frac{[(\varepsilon_b(1-\beta)+\mu(1-\gamma))T]^{B_c}}{B_c!} \\ & * e^{-(\varepsilon_s(1-\beta))T}\frac{[(\varepsilon_s(1-\beta))T]^{S_c}}{S_c!} \end{aligned} \tag{13}$$

For any given day, the maximum likelihood estimator of the information event parameters α and δ will be either 0 or 1, reflecting that information events occur only once a day. Over multiple days, however, these parameters can be estimated from the daily numbers of buys and sells. Because days are independent, the likelihood of observing the data $M = (B_{Ni}, S_{Ni}, B_{Ci}, S_{Ci})_{i=1}^{I}$ over I days is just the product of the daily likelihoods,[12]

$$L(M|\theta) = \prod_{i=1}^{I} L(\theta|B_{Ni}, S_{Ni}, B_{Ci}, S_{Ci}). \tag{14}$$

[12] In Easley, Kiefer, and O'Hara (1993) we tested the hypothesis of independent information events across days, and found that we could not reject the independence assumption. See also Madhavan, Richardson, and Roomans (1994) and Kleidon and Werner (1993) for empirical evidence consistent with the overnight arrival of new information.

To estimate the parameter vector $\theta = (\alpha, \delta, \varepsilon_B, \varepsilon_s, \mu, \beta, \gamma)$ from any data set M, we maximize the likelihood defined in equation (14). This provides direct estimates of the information event structure surrounding a stock, as well as of the rate and location of informed and uninformed trading in that stock. Our interest in this article lies in whether this information-based trading differs in a systematic way between trading venues. Our methodology provides two procedures for testing this, and, in the next section, we derive these tests and determine the appropriate statistical measures of their significance.

II. Hypothesis Testing

If "cream-skimming" is the basis for purchased order flows, then the information content of orders should differ between markets. Alternatively, if purchased orders merely reflect a simple partitioning of the existing order flow, then this difference should not arise. Our model provides two methods for testing this hypothesis. The first involves using restrictions of the general model to test for differences in information content between markets. The second approach uses our parameter estimates to calculate directly the probability of informed trading in each market. We now describe these tests in more detail.

A. Likelihood Ratio Tests and Posterior Odds

The general model derived in Section I allows for differences in informed and uninformed trading between markets. In particular, for any stock the probability that an informed order will execute in market $N(\gamma)$ need not be the same as the probability that it will execute in market $C(1 - \gamma)$, and similarly for the uninformed trade probability, β. There are several reasons why such differences could arise. The first, and most obvious, is simply due to scale. If market N has five times the volume of market C, then it is reasonable to expect that the probability of order execution will, in general, be five times as great. But these trading probabilities may also differ for a less innocuous reason. If there is a systematic diversion of some orders (say, the uninformed orders) to one venue, then both the volume *and* the composition of trade will differ between markets.

We can test for such a difference in the composition of trading by estimating restricted versions of our general model. To understand the intuition for our tests, note that if the mix of uninformed and informed orders in the two markets is the same, then the fraction of informed orders should equal the fraction of uninformed orders in each market. That is, if market C is receiving 5 percent of the informed orders, it should also be receiving 5 percent of the uninformed orders. In our model, this is equivalent to the restriction that $1 - \beta = 1 - \gamma$, or equivalently $\beta = \gamma$. If, instead, market C is receiving more uninformed orders than it is informed orders, then it follows that $\gamma > \beta$.

These parameter restrictions allow direct testing of the cream-skimming hypothesis. If there is no cream-skimming in a stock, then restricting the model to have $\beta = \gamma$ for that stock should have no effect on the model's goodness of fit as measured by the likelihood ratio statistic. If we reject this restriction, there are two possibilities to consider. First, it may be that the information content of market C is less than market N, a scenario consistent with "cream-skimming" by market C. In this case, restricting the model to have $\gamma > \beta$ should have no effect on the goodness of fit, and we will be unable to reject the restriction. Alternatively, if orders in market C have more information content than those in market N, then the model restricted to $\beta > \gamma$ cannot be rejected. This latter case corresponds to diversion of the most informative orders, and it would be difficult to reconcile with either a cream-skimming or profit-sharing hypothesis.

Because we are more interested in differences between the two markets in general, than in the differences for any particular stock, we also compare the log likelihoods for the entire sample of stocks. The maximized log likelihood for the sample of stocks is just the sum of their individual log likelihoods, under our maintained hypothesis that information events and trades are independent across stocks. The appropriate test of the restriction that for each stock the information content is the same in market N and market $C(\beta_i = \gamma_i$ for each stock i) is a chi-square test based on the differences in log likelihoods with and without the restriction. If this restriction is rejected, then it seems appropriate to ask which market generally has more information content. Here it is appropriate to test $\beta_i > \gamma_i$ or $\beta_i < \gamma_i$, for all stocks i, against $\beta_i = \gamma_i$, for each stock i. The appropriate test for these two restrictions is a mixture of chi-squares.

To aid in interpretation of these chi-square tests, and to provide more complete information, we also, where appropriate, report Bayesian statistics. Consider a Bayesian who begins with equal prior probability on the regions $\beta > \gamma$ and $\gamma \geq \beta$. His prior odds ratio, of $\gamma \geq \beta$ (trade in C at least as informative as trade in N) to $\beta > \gamma$ (trade in N more informative than trade in C), would be one. We compute the posterior odds ratio using this prior and the likelihoods under the restrictions $\gamma \geq \beta$ versus $\beta > \gamma$. The reader can easily adjust our reported posterior odds to conform with any prior odds.

B. Probability of Informed Trade

Our second set of tests uses the analysis of beliefs and quotes in section I.B to directly measure the probability of information-based trading for each stock in each market. In particular, equations (8) and (9) provide the market makers beliefs regarding the probability of informed trading in each market. These beliefs are functions of our parameter vector $\theta = (\alpha, \delta, \varepsilon_B, \varepsilon_s, \mu, \beta, \gamma)$, and hence we can use these estimated values for any stock to directly calculate these probabilities for that stock. Again, our interest is in the probability that informed trade in one market is higher than in another. To capture this, we compute for each stock the probability that a trade in market N is informed, conditional on the occurrence of a trade in N. We compute the same statistic for market C. This methodology, as well as that presented in the previous section, allows direct testing of differences in the information content of order flow. In the next section, we detail the data we use to test for these differences, and the potential biases in its interpretation.

III. The Data and Empirical Testing

Our goal is to estimate whether orders traded on the New York Stock Exchange differ in information content from those diverted by purchased order agreements. A natural approach to do so would be to construct a matched sample of stocks that are subject to purchase agreements, and then compare the information in the diverted orders to the orders remaining on the NYSE. Although conceptually reasonable, there are several difficulties in implementing this procedure. First, there are many firms or brokers actively purchasing orders, and obtaining information on all, or even any, of their actual trades is not possible. It is possible, however, to obtain the list of securities which are traded by one of (if not) the largest order purchasers, Madoff Investments. By selecting a sample from these stocks, we can compare the information content for stocks known to be involved in purchased order flow.

A second difficulty is that there are numerous trading venues that could be used for the diverted orders, so that following all of the diverted orders would require detailed trade data for each and every exchange. Because such orders are proprietary to the purchasing firms, it is not possible to obtain such information. In previous work, researchers have attempted to examine differences by trading locale, based on the notion that diverted orders are sent either to the regional exchanges or are traded on the Nasdaq. For the reasons noted in the paper's first section, the Cincinnati Exchange provides the best venue in which to investigate the information-based differences we seek in this study. We assume that the activity in New York and Cincinnati is sufficient to determine our parameters.

There are several potential biases that may emerge from our empirical approach. First, while we know which stocks are used in purchased orders, we do not know that any particular trade is a purchased order. Since the stocks used in purchased orders are those most actively traded, it is likely that the orders we consider in Cincinnati contain both purchased and nonpurchased orders. This difficulty can be viewed as introducing noise into our Cincinnati order data. If the purchased orders actually have lower informational content than normal orders, this would reduce our ability to detect informational differences between New York and Cincinnati. Similarly, purchased orders may be sent to any of the regional exchanges or to Nasdaq, with the choice of location at the discretion of the purchasing firm. Because Cincinnati is the smallest of the regional exchanges, it may be that only a small fraction of the purchased orders are sent there. This, again, would bias our results against finding informational differences.

A. Sample Selection

Purchased order flow is typically concentrated in the most actively traded stocks, reflecting the composition of the retail trade used in this process. To focus on those stocks known to be used in purchased orders, we obtained a list of securities traded by one of the largest purchasers of order flow, Madoff Investments, Inc. These securities include stocks that are normally listed on the NYSE, the American Stock Exchange (AMEX) and Nasdaq, as well as some stocks trading as American Depository Receipts (ADRs). As our focus is on differences in informational content between orders on the NYSE and the CSE, we excluded all stocks not listed on the NYSE, as well as all ADRs. Because Cincinnati is the smallest of the regional exchanges, its trading volume is small in general, and for some stocks, this can lead to serious difficulties with non-synchronous trading. To avoid this difficulty, we selected from this subset the thirty most actively traded stocks on Cincinnati. The resulting sample is listed in Table I.

There are several properties of the sample worth noting. First, as indicated by the 1990 market capitalizations in column 3 of Table I, the sample firms are very large. Using the Center for Research in Security Prices (CRSP) size portfolios as a gauge, nineteen of the thirty stocks are in the largest size portfolio (10), with the average size portfolio for the sample being 9.4. Second, these stocks are also among the most heavily traded, and include such active issues as Philip Morris, IBM, and Citicorp. The 1990 average daily volume for the sample is just over 530,000 shares. During the sample period we consider, approximately 83 percent of trading volume in our securities was executed on the NYSE, with approximately 1.4 percent traded on Cincinnati.

B. Data

Our estimation technique requires both intra-day and inter-day trade data. Trade data for the 30 stocks in our sample was taken from the ISSM database for the period October 1 to December 23, 1990. A sixty day trading interval has a sufficient number of information event nodes to allow reasonably precise estimation of the parameters. The sixty day interval is also short enough so that the stationarity assumptions built into our model are not too unreasonable.

Estimation of the likelihood function given in equation (14) requires knowing the number of buys and sells for each stock in our sample. This can be derived from the Institute for the Study of Security Markets (ISSM) data using two adjustments. First, large orders sometimes have multiple participants on one side of the trade. Reporting conventions may treat such transactions as multiple trades, when in fact only one trade has actually occurred. Hasbrouck (1991) suggests combining trades occurring within five seconds at the same price, and we follow this procedure in the paper. Second, while the ISSM data identifies trades, it does not indicate whether it was initiated by a buyer or a seller. This classification can be done using a technique developed by Lee and Ready (1991). Trades above the midpoint of the spread are classified as buys;

Table I

Sample Firms—Size and Volume

This table presents data from the year 1990 on the stocks used in our empirical work. These are the 30 most actively traded stocks in Cincinnati that are also both traded by Madoff Investments and are listed on the New York Stock Exchange (NYSE).

Name	Stock	1990 Mkt. Value in Thousands	1990 Daily Volume	% Trades NYSE	% Trades CSA
Bankamerica Corp.	BAC	5,650,595.00	840,526	86.53	1.58
Bethlehem Steel Corp.	BS	1,118,817.00	231,029	85.36	1.49
Chrysler Corp.	C	2,845,675.00	644,096	81.62	2.00
Citicorp	CCI	4,195,931.50	1,665,685	80.93	1.84
Champion International Corp.	CHA	2,382,305.00	292,144	90.75	1.30
Centel Corp.	CNT	2,505,515.25	102,255	80.35	1.49
Centerior Energy Corp.	CX	2,491,398.00	224,460	77.22	0.97
Delta Airlines, Inc.	DAL	2,362,127.50	248,774	90.58	1.69
Detroit Edison Co.	DTE	4,150,433.50	292,477	84.43	0.82
Ford Motor Co.	F	12,427,618.00	864,076	84.08	1.20
Federal Express Corp.	FDX	1,798,017.25	218,496	89.91	1.84
FMC Corp.	FMC	1,114,318.13	54,625	84.25	1.27
FPL Group, Inc.	FPL	4,226,112.00	253,207	78.53	1.10
General Mills, Inc.	GIS	8,049,524.00	188,794	84.54	0.91
General Motors Corp.	GM	20,627,440.00	1,162,199	84.50	2.00
Hasbro, Inc.	HAS	884,812.50	118,586	87.93	0.75
Hewlett Packard Co.	HWP	7,726,500.00	581,963	88.30	1.51
International Business Machs.	IBM	64,528,988.00	1,592,895	84.49	2.57
Ingersoll Rand Co.	IR	1,926,756.25	170,164	91.37	0.91
LTV Corp.	LTV	57,036.50	133,092	72.32	2.04
McDonnell Douglas Corp.	MD	1,489,690.00	157,164	85.18	1.11
Philip Morris Cos., Inc.	MO	47,891,984.00	2,073,284	88.73	0.82
Merck & Co., Inc.	MRK	34,815,596.00	793,341	86.33	2.72
NBD Bancorp, Inc.	NBD	2,406,591.00	71,086	88.22	0.75
Public Service Co., NM	PNM	349,857.25	112,968	72.72	0.80
Sara Lee Corp.	SLE	7,320,280.00	530,567	85.27	1.06
Tucson Electric Power Co.	TEP	154,170.00	103,869	62.75	1.82
USAir Group, Inc.	U	711,144.00	233,314	83.90	0.67
Unisys Corp.	UIS	404,645.00	780,983	73.26	1.33
Walmart Stores, Inc.	WMT	34,261,544.00	1,196,714	80.12	1.61
Mean		9,362,514.05	531,095	83.1487	1.3985

those below the midpoint are called sells.[13] Midpoint trades are sorted by referring to the previous different trade price, with trades executed at higher prices called buys, and those at lower prices called sales. This technique

[13] The quotes that we use in this classification scheme are the best bid or offer available in either New York or Cincinnati. The quote is usually the same in New York and Cincinnati, but occasionally the Cincinnati quote is strictly dominated by the New York quote.

undoubtedly misclassifies some trades, but it is standard, and it has been shown to work reasonably well.

IV. Results

We are now ready to estimate the parameters in our model, and to test for informational differences in trading venues. Recall that our model has seven parameters for each stock: α, the probability of an information event; δ, the probability that the information is bad news; μ, the arrival rate of traders who know the new information if it exists (i.e., the informed traders); ε_B, the arrival rate of uninformed buyers; ε_S, the arrival rate of uninformed sellers; γ, the probability that an informed trade is executed in New York; and β, the probability that an uninformed trade is executed in Cincinnati.

A. Tests for Cream-Skimming

Table II gives the (maximized) log likelihood values for four alternative parameter restrictions. The unrestricted log likelihoods correspond to the case where our seven parameters are allowed to vary freely. The three restricted cases correspond to requiring equal information content ($\gamma = \beta$) in the two markets, greater information content in New York ($\gamma > \beta$), and greater information content in Cincinnati ($\gamma < \beta$).[14]

We first test for equality of information content between locales. This test is based on a comparison of the log likelihoods for the unrestricted model and the model where $\gamma = \beta$. For eleven of our thirty stocks the restriction $\gamma = \beta$ is rejected at the 0.05 level. More interesting is asking whether in general New York and Cincinnati trades have equal information content. Because of our assumption of independence across stocks, the appropriate test for this hypothesis is based on a comparison of the sum of log likelihoods for the unrestricted model and the model with the $\gamma = \beta$ restriction. These values are given in the last row of Table II. The log likelihood for the unrestricted case (−30316.32) is larger than that of the restricted case (−30399.55), dictating a better fit for the unrestricted case. The realized test statistic of 166.45 is strongly significant (0.001 cutoff level = 59.7), and the posterior odds ratio is equal to 0.000, meaning that there is a 0 posterior probability that the two markets are the same. The equality of information is thus decisively rejected by the data. It is simply not the case that the information content of orders in Cincinnati and New York are generally equal.

We next test for the direction of information differences, or exactly which market has the greater probability of informed trade? Table II shows that when $\gamma = \beta$ is rejected, it is generally because $\gamma > \beta$. For 20 of our 30 stocks the $\gamma > \beta$ model fits at least as well as the $\gamma < \beta$ model, and for only 3 of the remaining ten stocks is the likelihood with $\gamma < \beta$ significantly better than the likelihood with $\gamma > \beta$. So for the sample as a whole we test the hypothesis $\gamma = \beta$

[14] In this discussion we do not index parameter estimates by stock, but it should be noted that the parameters are never restricted across stocks.

Table II

Testing for Informational Differences in New York and Cincinnati

This table presents maximized log likelihood values for various parameter restrictions related to the information content in order flow. Restricting $\gamma = \beta$ assumes no informational differences. Restricting $\gamma > \beta$ assumes more informed trading in New York, while $\gamma < \beta$ assumes more informed trading in Cincinnati. The reported test statistics are twice the difference of log likelihoods under the indicated restriction and the unrestricted model.

Stock	Log Likelihood Unrestricted	Log Likelihood $\gamma > \beta$	Log Likelihood $\gamma = \beta$	Log Likelihood $\gamma < \beta$	Test Statistic for $\gamma \leq \beta$	Test Statistic for $\gamma = \beta$	Test Statistic for $\gamma \geq \beta$
BAC	−1172.01	−1172.01	−1172.19	−1172.19	0.00	0.37	0.37
BS	−772.15	−774.63	−774.63	−772.15	4.94	4.94	0.00
C	−1574.17	−1574.18	−1574.18	−1574.17	0.01	0.01	0.00
CCI	−1484.40	−1484.48	−1484.48	−1484.40	0.17	0.17	0.00
CHA	−771.84	−772.10	−772.10	−771.84	0.52	0.52	0.00
CNT	−583.34	−586.01	−586.01	−583.34	5.34	5.34	0.00
CX	−518.96	−518.96	−519.13	−519.13	0.00	0.33	0.33
DAL	−819.98	−819.98	−820.02	−820.02	0.00	0.10	0.10
DTE	−713.82	−714.28	−714.28	−713.82	0.92	0.92	0.00
F	−1946.43	−1946.43	−1958.42	−1958.42	0.00	23.97	23.97
FDX	−769.65	−769.65	−776.53	−776.53	0.00	13.76	13.76
FMC	−637.33	−637.33	−643.25	−643.25	0.00	11.84	11.84
FPL	−1092.28	−1092.28	−1103.14	−1103.14	0.00	21.72	21.72
GIS	−766.11	−766.11	−766.74	−766.74	0.00	1.27	1.27
GM	−1418.36	−1418.36	−1427.55	−1427.55	0.00	18.38	18.38
HAS	−643.35	−643.35	−644.00	−644.00	0.00	1.31	1.31
HWP	−1051.91	−1051.91	−1052.00	−1052.00	0.00	0.19	0.19
IBM	−1835.16	−1840.16	−1840.16	−1835.16	10.00	10.00	0.00
IR	−842.64	−842.64	−842.64	−842.64	0.00	0.00	0.00
LTV	−466.28	−466.28	−467.59	−467.59	0.00	2.62	2.62
MD	−886.06	−886.06	−886.08	−886.06	0.04	0.04	0.00
MO	−1495.18	−1495.18	−1498.78	−1498.78	0.00	7.19	7.19
MRK	−1339.46	−1339.46	−1340.69	−1340.69	0.00	2.47	2.47
NBD	−579.79	−579.82	−579.82	−579.79	0.06	0.06	0.00
PNM	−514.86	−514.86	−524.72	−524.72	0.00	19.72	19.72
SLE	−910.81	−910.81	−911.22	−911.22	0.00	0.81	0.81
TEP	−682.00	−682.00	−682.29	−682.29	0.00	0.57	0.57
U	−786.70	−786.70	−788.47	−788.47	0.00	3.54	3.54
UIS	−1442.73	−1442.91	−1442.91	−1442.73	0.35	0.35	0.00
WMT	−1798.56	−1798.56	−1805.54	−1805.54	0.00	13.96	13.96
Overall	−30316.32	−30327.5	−30399.55	−30388.37	22.36	166.45	144.09

against the one-sided alternative $\gamma > \beta$. The appropriate statistical test for this hypothesis is a mixture of $\chi^2(i)$, i = 0, 1, ... 30 with mixing probabilities $\binom{30}{i}2^{-30}$; see Self and Laing (1987) for a discussion of this test. The points corresponding to the upper tail areas 0.10, 0.05, and 0.01 in this mixture of chi-squares are, respectively, 23.18, 26.17, and 32.34. The realized test statistic of 144.09 indicates that the hypothesis of equal information is strongly rejected relative to more information content in New York.

With a large data set, such as the one we use, there is always the concern that any restriction can be rejected. So we also test the hypothesis of equality of information content against a hypothesis of reverse cream-skimming, $\gamma < \beta$. The resulting test statistic is 22.36, which is not significant, even at the 0.10 level. Thus equality of information content cannot be rejected against the alternative of reverse cream-skimming. We view these results as providing strong evidence for the cream-skimming (by Cincinnati) hypothesis.

B. The Probability of Informed Trade

While the hypothesis tests above use goodness-of-fit measures to analyze information content, the parameter estimates of our model provide a direct measure of the probabilities of informed trading in New York and Cincinnati. Recall that for a given stock, equations (8) and (9) give the probabilities of informed trade in each market, conditional on a trade occurring in that locale. These probabilities naturally depend on the trading intensities of informed and uninformed traders, but they are also affected by the probabilities of information events and the type of information in the stock. These composite probabilities can be calculated from the parameter estimates of our unrestricted model.

Table III gives these parameter estimates for the thirty stocks in our sample. As would be expected, the estimates vary across stocks, reflecting the differences that naturally arise between firms. What is more significant are the differences between markets. For some firms, for example, Chrysler and IngersolRand, there are no significant differences in informed trading probabilities between markets. For these firms, the order flow in Cincinnati is smaller, but identical in composition, to the order flow in New York. For the other firms in our sample, however, this is not the case. For 20 of 30 firms, the probability of information-based trades is higher in New York. This difference is very large for some firms, for example, Ford, Florida Power and Light, Federal Express, LTV Corp, and Public Service Co., NM, suggesting that few informed trades in these stocks are executed in Cincinnati.

Of perhaps greater interest are the firms for which this relation does not hold. For 10 firms in our sample, the probability of informed trade is higher in Cincinnati than in New York. Inspection of the data reveals, however, that with two exceptions these differences are quite small, and hence are largely insignificant. For two stocks Bethlehem Steel and Centel Corp. the differences are 0.115 and 0.143, respectively. The greatest difference in trading probabilities of the other eight stocks favoring Cincinnati is 0.06, whereas the corresponding difference favoring New York is more than 0.2 for several stocks. Overall, the probability of informed trade in New York is approximately 44 percent more than that of informed trade in Cincinnati.[15]

[15] The average probability of informed trade in New York is 0.171, whereas it is 0.119 in Cincinnati. These averages should be interpreted with caution. Our model is stock-specific, not portfolio-based, and the aggregation of parameters need not be straightforward.

Table III

Parameter Estimates for Sample Stocks

This table gives parameter estimates for our sample stocks for the period October 1, 1990—December 22, 1990. The variables are α, the probability of an information event; δ, the probability that the information is bad news; μ, the arrival rate of traders who know the new information if it exists (i.e., the informed traders); ε_B, the arrival rate of uninformed buyers; ε_s, the arrival rate of uninformed sellers; γ, the probability that an informed trade is executed in New York; and β, the probability that an uninformed trade is executed in Cincinnati. PI_N and PI_C are the estimated probability of informed trade in that stock in New York and Cincinnati, respectively. Standard errors are given in parentheses.

Stock	α	δ	μ	ε_s	ε_b	γ	β	PI_N	PI_C
BAC	0.2697	0.1236	0.2478	0.1794	0.1954	0.9712	0.9671	0.1519	0.1349
	(0.0578)	(0.0819)	(0.0092)	(0.0033)	(0.0029)	(0.0059)	(0.0020)	(0.0280)	(0.0369)
BS	0.3573	0.1740	0.0861	0.0662	0.0716	0.9353	0.9651	0.1779	0.2927
	(0.0694)	(0.0937)	(0.0057)	(0.0023)	(0.0022)	(0.0123)	(0.0037)	(0.0287)	(0.0623)
C	0.1833	0.4551	0.4061	0.1928	0.2145	0.9610	0.9618	0.1545	0.1572
	(0.0500)	(0.1503)	(0.0121)	(0.0031)	(0.0031)	(0.0057)	(0.0021)	(0.0359)	(0.0425)
CCI	0.5751	0.5508	0.2909	0.5633	0.4506	0.9816	0.9834	0.1414	0.1548
	(0.0002)	(0.0002)	(0.0002)	(0.0004)	(0.0003)	(0.0000)	(0.0000)	(0.0001)	(0.0002)
CHA	0.3689	0.2339	0.0809	0.0851	0.1110	0.9538	0.9634	0.1309	0.1612
	(0.0794)	(0.1020)	(0.0062)	(0.0033)	(0.0026)	(0.0122)	(0.0031)	(0.0235)	(0.0512)
CNT	0.3786	0.2707	0.0435	0.0287	0.0491	0.9185	0.9618	0.1682	0.3112
	(0.0840)	(0.1120)	(0.0039)	(0.0018)	(0.0018)	(0.0180)	(0.0050)	(0.0312)	(0.0716)
CX	0.7434	0.8835	0.0371	0.0350	0.0332	0.9862	0.9793	0.2895	0.2117
	(0.0004)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0005)	(0.0006)
DAL	0.4835	0.0000	0.0970	0.0911	0.1082	0.9736	0.9708	0.1908	0.1754
	(0.0669)	(0.0000)	(0.0050)	(0.0031)	(0.0022)	(0.0075)	(0.0028)	(0.0235)	(0.0529)
DTE	0.4990	0.3976	0.0802	0.0666	0.1114	0.9731	0.9811	0.1823	0.2423
	(0.0700)	(0.0953)	(0.0044)	(0.0024)	(0.0027)	(0.0074)	(0.0024)	(0.0221)	(0.0678)
F	0.4090	0.3877	0.4884	0.7126	0.5751	0.9969	0.9795	0.1363	0.0227
	(0.0665)	(0.1024)	(0.0164)	(0.0064)	(0.0084)	(0.0030)	(0.0009)	(0.0188)	(0.0230)
FDX	0.2938	0.5982	0.0856	0.0969	0.0776	1.0000	0.9782	0.1284	0.0000
	(0.0613)	(0.1194)	(0.0056)	(0.0022)	(0.0022)	(0.0000)	(0.0023)	(0.0241)	(0.0000)
FMC	0.5589	0.2559	0.0585	0.0415	0.0483	1.0000	0.9653	0.2739	0.0000
	(0.0770)	(0.0878)	(0.0036)	(0.0020)	(0.0021)	(0.0000)	(0.0041)	(0.0294)	(0.0000)
FPL	0.5660	0.0884	0.2056	0.1800	0.1300	1.0000	0.9814	0.2766	0.0000
	(0.0660)	(0.0492)	(0.0066)	(0.0051)	(0.0024)	(0.0000)	(0.0016)	(0.0251)	(0.0000)
GIS	0.6130	0.1925	0.0940	0.1184	0.1156	0.9887	0.9801	0.1989	0.1233
	(0.0695)	(0.0688)	(0.0048)	(0.0041)	(0.0025)	(0.0061)	(0.0023)	(0.0209)	(0.0659)
GM	0.1343	0.2480	0.4285	0.5028	0.4110	0.9978	0.9697	0.0609	0.0046
	(0.0443)	(0.1524)	(0.0189)	(0.0050)	(0.0042)	(0.0052)	(0.0012)	(0.0189)	(0.0110)
HAS	0.5015	0.4721	0.0591	0.0467	0.0488	0.9880	0.9764	0.2388	0.1367
	(0.0767)	(0.1034)	(0.0038)	(0.0021)	(0.0022)	(0.0077)	(0.0038)	(0.0292)	(0.0865)
HWP	0.5767	0.5650	0.1517	0.2939	0.2841	0.9555	0.9509	0.1320	0.1205
	(0.0722)	(0.0893)	(0.0070)	(0.0044)	(0.0056)	(0.0095)	(0.0022)	(0.0149)	(0.0288)
IBM	0.4287	0.1944	0.3401	0.6963	0.7388	0.9265	0.9549	0.0897	0.1420
	(0.0656)	(0.0820)	(0.0123)	(0.0086)	(0.0067)	(0.0077)	(0.0013)	(0.0128)	(0.0236)
IR	0.2770	0.2800	0.1197	0.1326	0.1282	0.9906	0.9905	0.1128	0.1118
	(0.0005)	(0.0005)	(0.0003)	(0.0003)	(0.0003)	(0.0000)	(0.0000)	(0.0003)	(0.0005)
LTV	0.2480	0.9327	0.1046	0.0232	0.0417	1.0000	0.9974	0.2860	0.0000
	(0.0563)	(0.0650)	(0.0054)	(0.0010)	(0.0016)	(0.0000)	(0.0013)	(0.0479)	(0.0000)

Table III—*Continued*

Stock	α	δ	μ	ε_s	ε_b	γ	β	PI_N	PI_C
MD	0.4544	0.2912	0.1069	0.1064	0.1196	0.9745	0.9762	0.1766	0.1868
	(0.0684)	(0.0916)	(0.0054)	(0.0028)	(0.0028)	(0.0073)	(0.0024)	(0.0227)	(0.0569)
MO	0.3641	0.5086	0.3107	0.7713	0.6869	1.0000	0.9897	0.0727	0.0000
	(0.0641)	(0.1083)	(0.0130)	(0.0068)	(0.0062)	(0.0000)	(0.0005)	(0.0120)	(0.0000)
MRK	0.4270	0.1992	0.2270	0.4224	0.3959	0.9767	0.9628	0.1073	0.0690
	(0.0679)	(0.0824)	(0.0101)	(0.0060)	(0.0047)	(0.0076)	(0.0016)	(0.0154)	(0.0254)
NBD	0.4667	0.3230	0.0478	0.0517	0.0510	0.9769	0.9800	0.1779	0.2004
	(0.0782)	(0.1067)	(0.0038)	(0.0022)	(0.0020)	(0.0109)	(0.0034)	(0.0259)	(0.0925)
PNM	0.3326	1.0000	0.0445	0.0198	0.0356	1.0000	0.9599	0.2174	0.0000
	(0.0821)	(0.0000)	(0.0045)	(0.0009)	(0.0022)	(0.0000)	(0.0056)	(0.0420)	(0.0000)
SLE	0.5055	0.5943	0.1117	0.2110	0.1738	0.9753	0.9661	0.1290	0.0964
	(0.0698)	(0.1047)	(0.0059)	(0.0046)	(0.0043)	(0.0090)	(0.0022)	(0.0163)	(0.0375)
TEP	0.2107	0.6141	0.1009	0.0494	0.0815	0.9848	0.9770	0.1406	0.0972
	(0.0584)	(0.1434)	(0.0079)	(0.0016)	(0.0025)	(0.0086)	(0.0029)	(0.0327)	(0.0589)
U	0.5105	0.2082	0.0805	0.0747	0.0951	0.9821	0.9621	0.1982	0.1025
	(0.0718)	(0.0788)	(0.0044)	(0.0030)	(0.0023)	(0.0084)	(0.0035)	(0.0243)	(0.0495)
UIS	0.4170	0.1599	0.4396	0.2390	0.3171	0.9863	0.9879	0.2476	0.2718
	(0.0637)	(0.0733)	(0.0092)	(0.0040)	(0.0037)	(0.0023)	(0.0011)	(0.0289)	(0.0508)
WMT	0.3331	0.2496	0.3567	0.4337	0.3721	0.9925	0.9746	0.1306	0.0417
	(0.0609)	(0.0969)	(0.0107)	(0.0049)	(0.0041)	(0.0039)	(0.0013)	(0.0211)	(0.0230)

These data also provide strong evidence in support of the cream-skimming hypothesis. While individual differences arise between stocks, for the sample as a whole there is a clear difference between the information content of trades in New York and Cincinnati. Our data show that the order flow in Cincinnati is less likely to arise from informed traders, suggesting that the selection criteria used to purchase orders is successful in diverting uninformed order flow.

V. Conclusions

Using a sample of stocks typically employed in purchased order flow, we have demonstrated that the information content of orders differs across trading locales. Trades executed on the Cincinnati Stock Exchange are not as likely to be information-based as are trades executed on the New York Stock Exchange. This difference is consistent with the "cream-skimming" of orders to Cincinnati.

This finding has a number of important implications for both the operation and regulation of securities markets. From a competitive perspective, our results suggest that diverting orders to other locales is not uniformly benign. Since the orders diverted are the "least risky," an adverse selection problem arises with respect to the remaining order flow. This, in turn, dictates that prices on the NYSE will worsen to reflect the change in order composition. The fragmentation of orders by type can thus impose costs on existing markets well in excess of the effects arising purely from trade volume.

A second, and perhaps more intriguing, implication is that the more successful the diversion strategy is at segmentation, the more profitable this strategy is for its pursuers. In particular, if order purchasers promise to match the national best market bid or offer (NBBO), the worsening of prices on the NYSE allows even greater profits to be made on diverted trades. This has two important implications. First, while part of this excess may be rebated to brokers or their customers via rebates on the order flow, some (or even all) may be retained by the purchasers themselves.[16] Indeed, since the action of diverting orders worsened prices in New York in the first place, matching the BBO on New York by definition will result in worse prices for customers. With spreads larger, this subtle price effect also means that the profits of brokers or market makers who trade these orders are higher.

Of perhaps more importance are the dynamic effects of this strategy. As more uninformed orders are diverted, the worse prices become on the remaining orders, and the more profit there is to be made on diverted trades. This profit, in turn, allows the payment of greater order rebates, thereby diverting more orders, worsening prices yet again, and so on. Whether such a spiral develops depends in part on the competition among order purchasers, and on the types of orders that remain on the existing exchange.

Yet despite these difficulties, it is not straightforward to determine the welfare effects of competition through order diversion. To the extent that gains are passed on to small retail traders, perhaps through reduced commissions if not through prices, then these traders are made better off. Indeed, precluding order diversion to alternate locales, in effect, shifts the "free rider" problem in the other direction by allowing big traders to benefit from the liquidity provided by the little traders. Which of these effects has the greater impact on welfare depends on how much the market, as a whole, is affected. We conjecture that these market impacts may, in fact, be small. Since purchase agreements do not pursue large orders or institutional (or program) trades, these orders will remain in the original market. Provided these orders are large enough, the effects of purchased orders may be of only secondary importance. This issue requires further study and empirical testing.

Finally, while our results strongly support the existence of "cream skimming," it is not the case that we can therefore reject the simultaneous existence of profit sharing. Our results show that on average the diverted orders are less information-related. For some of the stocks in our sample, however, the order composition between locales is the same. If these stocks are also being profitably diverted, then it must be the case that prices (or spreads) for these stocks are too high relative to competitive levels. This is most likely due to mandated minimum spreads. NYSE rules currently require minimum spreads of 1/8, and for the actively traded stocks in our sample, this may simply be too large. This issue was noted by the Securities and Exchange Commission (SEC) in its

[16] These welfare issues have been the subject of recent regulatory concern by the SEC and the NASD; see "In-house trades may be costly for small investors," the *Wall Street Journal*, December 20, 1994.

Market 2000 study, where the proposal to shift to decimalization was discussed. Our results suggest that this may reduce, but not remove, the incentives for order diversions.

REFERENCES

Blume, M. E., and M. Goldstein, 1991, Differences in execution prices among the NYSE, the regionals, and the NASD, Working paper, University of Pennsylvania.

Christie, W. G., and P. H. Schultz, 1994, Why do Nasdaq market makers avoid odd-eighth quotes? *Journal of Finance* 49, 1813–1840.

Cochrane, J., 1993, US equity market competitiveness, Working paper, New York Stock Exchange.

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

Easley, D., N. Kiefer, and M. O'Hara, 1993, One day in the life of a very common stock, Working paper, Cornell University.

Easley, D., Kiefer, N., O'Hara, M., and J. Paperman, 1994, Liquidity, information, and infrequently traded stocks, *Journal of Finance*, forthcoming.

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

Harris, F. H., McInish, T., and R. Wood, 1994, The competition for order flow in NYSE-listed stocks, Working paper, Wake Forest University.

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

Huang, R., and H. Stoll, 1994, Competitive trading of NYSE listed stocks: Measurement and interpretation of trading costs, Working paper, Owen Graduate School of Management, Vanderbilt University.

In-house trades may be costly for small investors, *Wall Street Journal*, December 20, 1994.

Kleidon, A., and I. Werner, Round the clock trading: Evidence from UK cross-listed securities, NBER Working paper 4410.

Lee, C., 1992, Purchase of order flows and favorable executions: An intermarket comparison, Working paper, University of Michigan.

Lee, C., 1993, Market integration and price execution for NYSE-listed securities, *Journal of Finance* 48, 1009–1038.

Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733–746.

Madhavan, A., M. Richardson, and M. Roomans, 1994, Why do security prices change? A transactions level analysis of NYSE stocks, Working paper, Wharton School, University of Pennsylvania.

McInish, T., and R. Wood, 1992, Price discovery, volume, and regional/third market trading, Working paper, Memphis State University.

New York Stock Exchange, 1992, Auction market trading, Working paper, NYSE.

O'Hara, M., 1995, *Market Microstructure Theory* (Basil Blackwell, Cambridge, Mass.).

Self, S. G., and K.-Y. Laing, 1987, Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions, *Journal of the American Statistical Association*, 82, 605–610.