

# **A Model of Broker's Trading, With Applications to Order Flow Internalization**

**Sugato Chakravarty**

Purdue University

West Lafayette, Indiana, 47906

765-494-6427

Fax: 765-494-0869

EMAIL: SUGATO@PURDUE.EDU

**Asani Sarkar**

Federal Reserve Bank of New York

33 Liberty Street

New York, New York, 10045

212-720-8943

Fax: 212-720-1582

EMAIL: ASANI.SARKAR@NY.FRB.ORG

**Previous version: June 2000**

**Current version: January 2001**

The views expressed in this paper are those of the authors' and do not represent those of the Federal Reserve Bank of New York or the Federal Reserve System.

# **A Model of Broker's Trading, With Applications to Order Flow Internalization**

JEL Classification number: G12

## **Abstract**

Although brokers' trading is endemic in securities markets, the form of this trading differs between markets. Whereas in some securities markets, brokers may trade with their customers in the same transaction (simultaneous dual trading or SDT), in other markets, brokers are only allowed to trade *after* their customers in a separate transaction (consecutive dual trading or CDT). We show theoretically that, informed and noise traders are worse off and brokers are better off while market depth is lower in the SDT market. Thus, given a choice, traders prefer fewer brokers in the SDT market compared to the CDT market. With free entry, however, market depth may be higher in the SDT market provided its entry cost is sufficiently low relative to the CDT market. We study order flow internalization by broker-dealers, and show that, in the free entry equilibrium, internalization hurts retail customers and market quality.

Dual trading occurs when a broker sometimes trades for customers as an agent, and at other times trades for his own account. Personal trading by brokers is pervasive throughout securities and futures markets in the U.S. and the world. There are two types of dual trading: *simultaneous* (where a broker trades for himself and a customer in the same transaction) and *consecutive* (where a broker trades for customers as an agent and for himself at other times, but not in the same transaction).<sup>1</sup>

In futures exchanges, consecutive dual trading is permitted, but not simultaneous dual trading. However, both types of dual trading are allowed in securities markets, currency and interest rate swap markets, and the fixed income market. What are the relative advantages and disadvantages of the two forms of dual trading, and how do they relate to the structure of the securities markets? Our paper highlights the role of brokerage competition in determining the relative costs and benefits of the two types of dual trading. We show that consecutive dual trading markets are associated with many brokers, whereas simultaneous dual trading markets are only viable with few brokers. Therefore, consistent with institutional reality, our model predicts that in futures markets, where typically many dual traders are present in each contract, consecutive dual trading is more likely than simultaneous dual trading. In stock markets, by contrast, the number of dual traders in any stock is few, and simultaneous dual trading is the norm.<sup>2</sup>

In our model, based on Kyle (1985), multiple informed traders and noise traders trade by dividing their orders equally among multiple brokers. The brokers, on receiving the

---

<sup>1</sup> The terminology is due to Grossman (1989), who also surveys dual trading regulations around the world.

<sup>2</sup> In the NYSE, for example, the potential dual traders are--in addition to the specialist--national full-line firms and

orders, may dual trade consecutively or simultaneously. If brokers dual trade simultaneously, each broker submits the net order (his personal trades plus his customers' trades) to the market maker for execution, and the game ends in a single period. If brokers trade consecutively, trading occurs over two periods. In period one, each broker submits the sum of informed and period-one noise trades to the market maker for execution. In period two, each broker submits his personal trade and period two noise trades to the market maker. In all cases, the market maker sets prices to make zero expected profits, conditional on observing the net order flow in the market.

We find that the simultaneous dual trading markets are not viable when there are many dual traders. The reason is that brokers mimic informed trades by trading with insiders in the same direction, causing insiders to trade less and at a higher price (in absolute value), and lowering informed profits.<sup>3</sup> In addition, brokers offset noise trades, thereby increasing noise trader losses. These effects are more adverse, the greater the number of brokers.

When both dual trading markets exist, informed traders have lower expected profits and noise traders have higher losses with simultaneous dual trading. These results appear to imply that, over time, simultaneous dual trading markets should cease to exist. To explain the coexistence of the two types of dual trading, we extend the model by endogenizing the number of brokers and informed traders in the market. In the free entry equilibrium, traders pay a fixed fee to enter the market and, upon entering, choose the number of brokers to trade with. We show that, provided the market entry fee is sufficiently low, enough informed traders enter

---

the investment banks. In 1989, there were six national full-line firms, and ten investment banks (Matthews, 1994).

<sup>3</sup> As a practical example of how brokers' trading may hurt informed traders, large institutional customers in the stock markets have long been concerned that brokers may use knowledge of their orders to trade for their own accounts. See "Money Machine", *Business Week*, June 10, 1991, pages 81-84.

the simultaneous dual trading market to make it viable. Further, traders choose many brokers in the consecutive dual trading market, and only one broker in the simultaneous dual trading market.

To test the applicability of our model, we study order flow internalization by brokers, which primarily affects uninformed retail order flow (Easley, Kiefer and O'Hara (1996), Battalio, Green and Jennings (1997), and the SEC (1997)). We show that, in the free entry equilibrium, internalization reduces market depth and price informativeness, and increases uninformed losses. Further, the number of internalizing brokers is negatively related to the market depth and the number of entering informed traders. Thus, our model predicts that internalization may be more prevalent in thin markets with few informed traders. If thin markets have high spreads, this result supports advocates of purchased order flow who argue that it primarily affects NYSE stocks with large spreads (Easley, Kiefer and O'Hara, 1996). However, contradicting these advocates, we show that market quality is affected adversely.

The existing literature on dual trading does not distinguish between the two types of dual trading. Also, it does not consider multiple informed traders and multiple brokers in the same model, nor does it endogenize the number of traders. Fishman and Longstaff (1992) study consecutive dual trading in a model with a single broker and a fixed order size. Roell (1990) and Chakravarty (1994) have multiple dual traders but a single informed trader. Sarkar (1995) studies simultaneous dual trading with multiple informed traders and a single broker.

Regarding internalization, Battalio and Holden (1996) show that, if brokers can distinguish between informed and uninformed orders (as in our model), they can profit from internalizing uninformed orders. However, unlike our model, they do not focus on the effect of internalization on market quality. Dutta and Madhavan (1997) find that a collusive

equilibrium is easier to sustain with preferencing arrangements. In contrast to the theoretical results, the empirical studies of Battalio (1997), Battalio, Greene and Jennings (1997) and Lightfoot, Martin, Peterson, and Sirri (1997) and the experimental study of Bloomfield and O'Hara (1996) find no adverse effect on market quality.<sup>4</sup>

The remainder of the paper is organized as follows. Section I describes a trading model with multiple customers and many brokers, when the number of informed traders and brokers is fixed. Sections II and III solve the consecutive and simultaneous dual trading models. Section IV endogenizes the number of informed traders and brokers. Section V analyzes order flow internalization. Section VI concludes. All proofs are in the appendix.

## **I. A Model of Trading with Multiple Customers and Multiple Brokers.**

We consider an asset market structured along the lines of Kyle (1985). There is a single risky asset with random value  $v$ , drawn from a normal distribution with mean 0 and variance  $\Sigma_v$ . There are  $n$  informed traders who receive a signal about the true asset value and submits market orders. For an informed trader  $i$ ,  $i=1, \dots, n$ , the signal is  $s^i = v + e^i$ , where  $e^i$  is drawn from a normal distribution with mean 0 and variance  $\Sigma_e$ . A continuum of noise traders also submit aggregate market orders  $u$ , where  $u$  is normally distributed with mean zero and variance  $\Sigma_u$ . All random variables are independent of one another.

---

<sup>4</sup> As Battalio, Greene and Jennings (1997) state, their result may imply that broker-dealers are not systematically skimming uninformed order flows or, alternatively, that internalizing brokers have a cost advantage in executing orders. Macey and O'Hara (1997) survey the literature on preferencing and internalization.

All customers, informed and uninformed, must trade through brokers. There are  $m$  brokers in the market, who submit customer orders to the market maker.  $m$  and  $n$  are common knowledge. We assume that orders are split equally among the  $m$  brokers. By observing informed orders, brokers can infer the informed traders' signals. By observing orders of noise traders, they are aware of the size of uninformed trades. Consequently, brokers have an incentive to trade based on their customers' orders. However, brokers are not allowed to trade ahead of (i.e., front run) their customers.

Brokers may trade in two possible ways. They may execute their customers' orders first, and trade for their own accounts second in a separate transaction---i.e., engage in **consecutive** dual trading. Alternatively, they may trade with their customers in the same transaction--i.e., engage in **simultaneous** dual trading. The sequence of events is as follows: in stage one, informed trader  $i$  (for  $i=1, \dots, n$ ) observes  $s^i$  and chooses a trading quantity  $x^i$ . In stage two, broker  $j$  (for  $j=1, \dots, m$ ) trades consecutively or simultaneously, and places orders of an amount  $z^j$ . In subsequent sections, we describe in more detail how simultaneous and consecutive dual trading differ. Finally, all trades (including brokers' personal trades) are batched and submitted to a market maker, who sets a price that earns him zero expected profits conditional on the history of net order flows realized.

Initially, the number of informed traders and brokers is fixed. Later, we allow informed traders to choose the number of brokers to allocate their orders to, and study the free entry equilibrium where informed traders and brokers decide whether to enter the market, depending on a market entry cost and their expected profits upon entering.

## II. Consecutive Dual Trading.

In this section, we solve for the equilibrium in a market with consecutive dual trading. We assume that brokers do not trade with their customers in the same transaction--i.e., simultaneous dual trading is not allowed, as in futures markets.

### A. The Consecutive Dual Trading Model

Trading occurs in two periods. In period one, brokers receive market orders from  $n$  informed traders and the noise traders, which they then submit to the market maker. In period two, brokers trade for themselves, along with period two noise traders. Each period, a market maker observes the history of net order flow realized so far and sets a price to earn zero expected profits, conditional on the order flow history.

The sequence of events is as follows: in period one, informed trader  $i$ ,  $i=1, \dots, n$  observes  $s^i$  and chooses  $x^{i,d}$ , knowing that his order will be executed in the first period. Accordingly, informed trader  $i$ ,  $i=1, \dots, n$ , chooses  $x^{i,d}$  to maximize conditional expected profits  $E[(v-p_1)x^{i,d} | s^i]$ , where the period one price is  $p_1 = \lambda_1 y_1$ , the period one net order flow is  $y_1 = x_d + u_1$ , the aggregate informed trade is  $x_d = \sum_i x^{i,d}$  and  $u_1$  is the period one noise trade.

In period two, brokers choose their personal trading quantity after observing the  $n$ -vector of informed trades  $\{x^{1,d}, \dots, x^{n,d}\}$ ,  $u_1$  and  $p_1$ . Thus, broker  $j$ ,  $j=1, \dots, m$ , chooses  $z^j$  to maximize conditional expected profits  $E[(v-p_2)z^j | \{x^{1,d}/m, \dots, x^{n,d}/m\}, u_1/m, p_1]$ , where the conditioning is based on each broker observing his portion of the informed and uninformed orders received, plus the period one price.

The  $m$  brokers submit their personal trades to the market maker, who sets  $p_2 = \lambda_2 y_2 + \mu_2 y_1$  where  $y_2 = \sum_j z^j + u_2$ ,  $\sum_j z^j$  is the aggregate trade of all brokers and  $u_2$  is the noise



trade in period two. Finally, the liquidation value  $v$  is publicly observed and both informed traders as well as brokers realize their respective profits (if any).

Define  $t = \Sigma_v / \Sigma_s$ , where  $t$  is the unconditional precision of  $s^i$ ,  $i = 1, \dots, n$ . Note that  $0 \leq t \leq 1$ .

Further, define  $Q = 1 + t(n-1)$ , where  $(Q-1)s^i$  represents informed trader  $i$ 's conjecture (conditional on  $s^i$ ) of the remaining  $(n-1)$  informed traders' signals. Since informed traders have different information realizations, they also have different conjectures about the information of other informed traders.  $t$  also measures the correlation between insider signals. For example, if  $t = 1$  (perfect information), informed signals are perfectly correlated,  $Q = n$ , and informed trader  $i$  conjectures that other informed traders know  $(n-1)s^i$ --i.e., the informed trader believes other informed traders have the same information as he.

Proposition 1 below solves for the unique linear equilibrium in this market.

**Proposition 1: In the consecutive dual trading case, there is a unique linear equilibrium for  $t > 0$ . In period one, informed trader  $i$ ,  $i = 1, \dots, n$ , trades  $x^{i,d} = A_d s^i$ , and the price is  $p_1 = \lambda_1 y_1$ . In period two, each broker  $j$ ,  $j = 1, \dots, m$  trades  $z = B_1 x_d + B_2 u_1$ , the price is  $p_2 = \lambda_2 y_2 + \mu_2 y_1$ , and a broker's expected trading revenue is  $W_a$  where:**

$$(1) \lambda_1 = \frac{\sqrt{nt \Sigma_v}}{(1+Q)\sqrt{\Sigma_u}}$$

$$(2) A_d = \frac{t}{\lambda_1(1+Q)}$$

$$(3) B_1 = \frac{1}{\sqrt{mQ(1+Q)}}$$

$$(4) B_2 = \frac{-\sqrt{Q}}{\sqrt{m(1+Q)}}$$

$$(5) \lambda_2 = \frac{\sqrt{nt \Sigma_v}}{\sqrt{Q(1+Q) \Sigma_u}} \frac{\sqrt{m}}{1+m}$$

$$(6) \mu_2 = \frac{\sqrt{nt \Sigma_v}}{(1+Q)\sqrt{\Sigma_u}}$$

$$(7) W_d = \frac{\sqrt{nt \Sigma_u \Sigma_v}}{\sqrt{Q(1+Q)}} \frac{1}{\sqrt{m(1+m)}}$$

**Discussion:** Period one informed trades are not affected by dual trading in period two since informed traders trade only in period one. The informed trading intensity  $A_d$  is positively related to the market depth ( $1/\lambda_1$ ) and to the precision of the signal. In period two, dual traders piggyback on period one informed trades ( $B_1 > 0$ ) and offset noise trades ( $B_2 < 0$ ). Competition between informed traders leads insiders to use their information less, making piggybacking less valuable. Higher values of  $t$  increases the correlation between insiders' signals, reducing (for  $n > 1$ ) the value of observing multiple informed orders. Consequently, the extent of piggybacking  $B_1$  decreases in the number of informed traders  $n$  and the information precision  $t$ . Broker revenues increase with noise trading, and decrease with the number of brokers.

### III. Simultaneous Dual Trading.

#### A. The Simultaneous Dual Trading Model

Simultaneous dual trading is modeled in a single period Kyle (1985) framework. The notations are the same as in section II. All variables and parameters related to simultaneous dual trading are denoted either with superscript  $s$  or subscript  $s$ .

A group of  $n$  informed traders receive signals  $s^i$  about the unknown value  $v$ , and choose quantities  $x^{i,s}$  knowing that his order will be executed along with the orders of brokers and noise traders in the same transaction. Accordingly, informed trader  $i$ ,  $i = 1, \dots, n$ , chooses  $x^{i,s}$  to

maximize conditional expected profits  $E[(v-p_s)x^{i,s} | s^i]$ , where the price is  $p_s = \lambda_s y_s$ , the net order flow is  $y_s = x_s + mz + u$ , the aggregate informed trade is  $x_s = \sum_i x^{i,s}$ ,  $z$  is the amount each broker trades and  $u$  is the noise trade.

Upon receiving the orders of their informed and uninformed customers, brokers choose their personal trading quantity after observing the  $n$ -vector of informed trades  $\{x^{1,s}, \dots, x^{n,s}\}$  and  $u$ . Thus, broker  $j$ ,  $j = 1, \dots, m$ , chooses  $z^{j,s}$  to maximize expected profits  $E[(v - p_s)z^{j,s} | \{x^{1,s}/m, \dots, x^{n,s}/m\}, u/m]$ , where the conditioning is based on each broker observing his portion of the informed and uninformed orders.

The  $m$  brokers submit their customer trades and personal trades to the market maker, who sets the price that earns him zero expected profits conditional on the net order flow realized. Finally, the liquidation value  $v$  is publicly observed and both informed traders as well as brokers realize their respective profits (if any). Proposition 2 below solves for the unique linear equilibrium in this market.

**Proposition 2: In the simultaneous dual trading case, there is a unique linear equilibrium for  $t > 0$  and  $Q > m$ . Informed trader  $i$ ,  $i = 1, \dots, n$ , trades  $x^{i,s} = A_s s^i$ , broker  $j$ ,  $j = 1, \dots, m$  trades  $z = B_{1,s} x_s + B_{2,s} u$ , the price is  $p_s = \lambda_s y_s$ ,  $W_s$  is the broker's expected revenue, where:**

$$(8) \quad \lambda_s = (1+m) \frac{\sqrt{nt \Sigma_v}}{\sqrt{\Sigma_u}(1+Q)}$$

$$(9) \quad A_s = \frac{(Q-m)t}{\lambda_s Q(1+Q)}$$

$$(10) \quad B_{1,s} = \frac{1}{(Q-m)}$$

$$(11) \quad B_{2,s} = -\frac{1}{(1+m)}$$

$$(12) \quad W_s = \frac{\sqrt{nt \Sigma_u \Sigma_v}}{Q(1+m)}$$

**Discussion.** In contrast to Proposition 1, the informed trading intensity  $A_s$  depends on the number of brokers. The insider trades and equilibrium exists only if  $Q > m$ . Since  $n > Q$ , existence implies  $n > m$ : the number of informed traders must exceed the number of brokers. The intuition behind this result is as follows. Suppose an insider buys. Dual traders also buy in the same transaction, piggybacking on the insider trade (i.e.  $B_I > 0$ ), and increasing the price paid by the insider for his purchase. The order of an individual insider is exploited less as the number of insiders increases, and is exploited more as the number of brokers increases. If the number of brokers is too large relative to the number of insiders, the adverse price effect makes it too costly for insiders to trade.

**B. Comparing Simultaneous And Consecutive Dual Trading**

By combining results from propositions one and two, we can compare the equilibrium outcomes for the two kinds of dual trading, holding the number of informed traders and brokers fixed.

**Corollary 1.** *Suppose  $m$  and  $n$  is fixed. Then:*

- (1) *If  $Q \leq m$ , only the consecutive dual trading equilibrium is viable.*
- (2) *If  $Q > m$ , then both dual trading equilibria are viable. Relative to simultaneous dual trading, uninformed losses and brokers' profits are lower, while informed profits are higher with consecutive dual trading.*

**Discussion.** As discussed in Proposition 2, the simultaneous dual trading equilibrium no longer exists when  $Q \leq m$ , i.e. when there are too many brokers who take advantage of the insider information. Since, with consecutive dual trading, informed trading is independent of the number of brokers, a market with consecutive dual trading continues to exist even with too

many brokers. When the number of brokers is relatively few ( $Q > m$ ), then both types of dual trading exists. Brokers would prefer simultaneous since their profits are higher. Since brokers' profits are at the expense of insiders, informed traders by contrast prefer consecutive dual trading. Aggregate profits of informed traders and dual traders are higher in the simultaneous dual trading market. Since noise trader losses are the negative of these aggregate profits, noise trader losses are also higher with simultaneous dual trading.

#### **IV. Free Entry by Informed Traders and Brokers**

In this section, we study the two forms of dual trading, given that informed and uninformed traders optimally choose the number of brokers to give their orders to, and given that there is free entry into the asset market. The decision-making sequence of agents is as follows: 1) Informed traders and brokers simultaneously decide whether to enter the market; (2) Informed and noise traders choose the number of brokers to give the order to; (3) Informed and noise traders divide their orders equally among the chosen brokers. From then on, the game continues as before.

Informed traders choose the number of brokers to maximize expected profits. Uninformed noise traders choose the number of brokers to minimize their expected losses to informed traders and brokers. The following corollary describes traders' choice of the number of brokers to trade with.

**Corollary 2. (1) When brokers trade consecutively, informed and noise traders choose all the brokers available.**

**(2) Suppose there are at least two informed traders, so that the simultaneous dual trading equilibrium exists. Then, with simultaneous dual trading, informed and noise traders give orders to only one broker.**

**Discussion.** With consecutive dual trading, informed traders' profits are independent of  $m$  but noise trader losses are decreasing in  $m$ . Thus, noise traders choose all available brokers while informed traders are indifferent to the choice of  $m$ . When brokers trade simultaneously, the situation is reversed: informed profits are decreasing in  $m$  whereas uninformed losses are independent of  $m$ . Thus, informed traders choose one broker and noise traders are indifferent to the choice of  $m$ . For concreteness, we assume that informed and noise traders give orders to the same number of brokers.

Next, consider free entry by informed traders and brokers. Let  $k_i$  be the cost of entering market  $i$ ,  $i = d$  (consecutive dual trading),  $s$  (simultaneous dual trading). To obtain analytic solutions, we assume  $t = 1$ . The free entry equilibrium satisfies two conditions. Traders enter a market until their expected profits, net of the entry cost, are zero. And the cost is low enough so that entry is profitable for the minimum number of traders necessary to sustain equilibrium. Proposition three solves for the free entry equilibrium in the two markets.

**Proposition 3. (1) In the market with consecutive dual trading,  $n_d$  informed traders and  $m_d$  brokers enter the market, with  $m_d < n_d$ , and  $m_d$  and  $n_d$  given by:**

$$(13) \quad \sqrt{m_d}(1 + m_d) = \frac{\sqrt{\Sigma_u \Sigma_v}}{k_d} \frac{1}{\sqrt{1 + n_d}}$$

$$(14) \quad \sqrt{n_d}(1 + n_d) = \frac{\sqrt{\Sigma_u \Sigma_v}}{k_d}$$

**At least *two* informed traders and at least *one* broker enter the market if:**

$$(15) \quad k_d \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{3\sqrt{2}}$$

**(2) In the market with simultaneous dual trading,  $n_s$  informed traders and *one* broker enter the market, where  $n_s$  is given by:**

$$(16) \quad \frac{n_s \sqrt{n_s(1+n_s)}}{(n_s - 1)} = \frac{\sqrt{\Sigma_u \Sigma_v}}{2k_s}$$

**At least *two* informed traders and *one* broker enter the market if:**

$$(17) \quad k_s \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{12\sqrt{2}}$$

**Discussion.** Proposition three is intuitive: the number of informed traders and brokers entering a market is inversely related to the cost of entering the market, and positively related to the volatility of the asset value and the size of noise trades (i.e., the depth of the market). In the consecutive dual trading market, a broker expects to make less trading profits than an informed trader since he trades second. Thus, the equilibrium number of brokers is less than the equilibrium number of informed traders in this market.

An important observation is that the participation constraint (17) is more restrictive than the inequality (15), implying that the simultaneous dual trading equilibrium is viable only at lower market entry cost, relative to the consecutive dual trading equilibrium. In the following proposition, we compare the two types of dual trading markets given free entry by brokers and informed traders and the optimal choice of the number of brokers by informed traders.

**Proposition 4.** In the free entry equilibrium of proposition three, suppose  $k_d > 2k_s$ . Then  $n_d < n_s$ . Market depth may be higher with simultaneous dual trading if  $n_s$  is large relative to  $n_d$  and  $m_d$ .

**Discussion.** The proposition says that, if entry costs are sufficiently low, the simultaneous dual trading market may have *more* informed traders and greater market depth in a free entry equilibrium. The reason is that informed traders protect themselves from excessive piggybacking by choosing only one broker, thus maximizing their expected profits in the simultaneous dual trading market. Consequently, if entry costs are low enough, many informed traders enter the simultaneous dual trading market and market depth is higher. Thus, the proposition provides some intuition as to why we see markets with simultaneous dual trading exist in the real world.

## V. Internalization of order flow by broker-dealers

Internalization is the direction of order flow by a broker-dealer to an affiliated specialist or order flow executed by that broker-dealer as market maker. Broker-dealers can internalize order flow in several ways. For example, large broker-dealer firms, particularly NYSE member firms, purchase specialist units on regional exchanges and direct small retail customer orders to them. In off-exchange internalizations, NYSE firms execute orders of their retail customers against their own account, with the transaction taking place in the so-called third market or over-the-counter market. Such transactions, also called 19c-3 trading, have become a major source of profits for broker-dealers.<sup>5</sup>

We use our simultaneous dual trading model to analyze order flow internalization. Since internalizing brokers typically handle order flows of small retail investors, we assume that there are  $m$  internalizing brokers who handle all of the noise trades, and  $m_2$  piggybacking

---

<sup>5</sup> Rule 19c-3 allows NYSE stocks listed after April 26, 1979 to be traded off-exchange. Broker-dealers may have earned over \$500 million in 1994 from 19c-3 trading. See “Report on the practice of preferencing”, the SEC, April 1997, and “In-House trades can be costly for small investors”, the *Wall Street Journal*, December 20, 1994.



brokers who handle all the informed traders, with  $m_1 + m_2 = m$ . As before, the market maker sees the pooled order and, further, cannot distinguish between internalizing brokers and informed-order brokers.

Let  $z_1$  be the trade of an internalizing broker and let  $z_2$  be the trade of informed-order brokers. From (10) and (11):

$$(18) \quad z_1 = -\frac{u}{(I + m_1)}$$

$$(19) \quad z_2 = \frac{x_s}{(Q - m_2)}$$

We compare uninformed losses and the market quality between the internalization model and a model with no brokers' trading. For simplicity, we assume  $t=1$ . The following proposition shows the effect of order flow internalization on noise trader losses and the market, assuming free entry of informed traders and brokers.

**Proposition 5.** **Suppose market entry is free. Then, with internalization of order flow, (i) one piggybacking broker,  $n_0$  informed traders and  $m_0$  internalizing brokers enter the market.  $m_0$  is less than the competitive number of brokers since internalizing brokers make positive expected profits in the free-entry equilibrium. (ii) Relative to a market with no order flow internalization, market depth, price informativeness, expected informed profits and the number of informed traders are lower, while uninformed losses are higher.**

**Discussion.** Since piggybacking hurts informed profits, traders choose just one piggybacking broker. Aggregate profits of internalizing brokers are greater than the aggregate profits of informed traders. Since entry costs must be low enough to allow informed traders to enter the

market at zero expected profits, it follows that expected profits of internalizing brokers must be positive at these costs. The results on market quality and informed profits follow from our earlier result that market quality in the simultaneous dual trading market is worse than in the consecutive dual trading market and in a market without dual trading.

## **VI. Conclusion.**

In this article, we study a wide variety of issues related to brokers' trading. Multiple informed traders and noise traders trade through multiple brokers, who either trade in the same transaction as their customers (simultaneous dual trading) or in a separate transaction (consecutive dual trading).

While the consecutive dual trading equilibrium always exists, the simultaneous dual trading equilibrium fails when the number of brokers is greater than the number of informed traders. The reason is that brokers trade with informed traders in the same direction, thus worsening informed traders' terms of trade. This effect is magnified with many brokers, leading informed traders to stop trading. When both equilibria exist, informed profits and market depth are lower, while uninformed losses and brokers' profits are higher with simultaneous dual trading, relative to consecutive dual trading. Thus, only brokers prefer simultaneous dual trading.

We allow informed and noise traders to choose the number of brokers, and endogenize the number of brokers and informed traders in the market. In the simultaneous dual trading market, informed and noise traders choose only one broker whereas, with consecutive dual trading, informed and noise traders choose all available brokers. If the market entry cost is sufficiently low, more informed traders enter the simultaneous dual trading market, and market

depth may be lower, relative to the consecutive dual trading market. Thus, the adverse effects of simultaneous dual trading on customers and the market are mitigated in the free entry equilibrium. If the market entry cost is high, consecutive dual trading is better.

In the simultaneous dual trading model, we allow some brokers to internalize the uninformed order flow, by selling to noise traders as dealers, out of inventory. In the free entry equilibrium, we find that market depth and price informativeness are lower, and uninformed losses are higher. In addition, the number of internalizing brokers is negatively related to market depth and the number of informed traders.

## References

- Admati, A. and P. Pfleiderer (1988), A theory of intraday patterns: Volume and price variability, *Review of Financial Studies*, 1, 3-40.
- Battalio, R. and C. W. Holden (1996), Why doesn't decimal trading eliminate payment for order flow and internalization? Working Paper, Indiana University.
- Battalio, R. (1997), Third market broker-dealers: Cost competitors or cream skimmers? *Journal of Finance*, 52, 341-352.
- Battalio, R., Greene, J. and R. Jennings (1997), Do competing specialists and preferencing dealers affect market quality? *Review of Financial Studies*, 10, 969-993.
- Bloomfield, R. and M. O'Hara (1996), Does order preferencing matter? Working Paper, Johnson Graduate School of Management, Cornell University.
- Chakravarty, S. (1994), Should actively traded futures contracts come under the dual trading ban?, *Journal of Futures Markets*, 14, 661-684.
- Dutta, P. and A. Madhavan (1997), Competition and collusion in dealer markets. *Journal of Finance*, 52, 245-276.
- Easley, D., Kiefer, N. M. and M. O'Hara (1996), Cream-skimming or profit-sharing? The curious role of purchased order flow. *Journal of Finance*, 51, 3, 811-833.

- Fishman, M. and F.A. Longstaff (1992), Dual trading in futures markets, *Journal of Finance*, 47, 643-672.
- Grossman, S.J. (1989), An economic analysis of dual trading, working paper, University of Pennsylvania.
- Kyle, A.S. (1985), Continuous auctions and insider trading, *Econometrica*, 53, 1315-1335.
- Lightfoot, L.E., Martin, P. G., Peterson, M. A., and E. Sirri (1997), Preferencing and market quality on U.S. equity exchanges, Working Paper, the Securities and Exchange Commission.
- Macey, J. R. and M. O'Hara (1997), The law and economics of best execution, forthcoming, *Journal of Financial Intermediation*.
- John O. Matthews (1994), *Struggle And Survival On Wall Steet*, Oxford University Press, New York.
- Roell, A. (1990), Dual-capacity trading and the quality of the market, *Journal of Financial Intermediation*, 1, 105-124.
- Sarkar, A. (1995), Dual trading: winners, losers and market impact, *Journal of Financial Intermediation*, 4, 77-93.

## ***Appendix***

### **Proof of Proposition 1:**

#### **Consecutive dual trading equilibrium.**

In period one, the insider's problem is identical to the single period model in Kyle (1985), but with multiple informed traders. The solution to this equilibrium is in Lemma 1 of Admati and Pfleiderer (1988). This gives  $A_d$  and  $\lambda_1$ .

#### **The j-th dual trader's problem:**

Let  $x^{m,d}$  denote the n-tuple  $\{x^{1,d}/m, x^{2,d}/m, \dots, x^{n,d}/m\}$ . In period two, dual trader  $j$  observes  $x^{m,d}$ ,  $p_1$  and  $u_1/m$  and trades  $z^j$ . The dual trader's problem is to maximize with respect to  $z^j$

his expected profits  $E[(v-p_2)z^j | x^{m,d}, u_1/m, p_1]$ . Let  $z^{-j} = \sum_{i \neq j} z^i$  for all  $i \neq j$ . Also, define  $s = \sum_{i=1}^n s^i$ .

The first order condition for the dual trader's problem is:

$$(1.1) \quad E\langle v | x^{m,d} \rangle - 2\lambda_2 z^j - \lambda_2 z^{-j} - \mu_2 A_d s - \mu_2 u_1 = 0$$

Using linear projection,  $E[v | x^{m,d}] = ts/Q$ , where  $Q$  is defined in the text. The second order condition is  $-\lambda_1 < 0$ , which is satisfied for  $\lambda_1 > 0$ .

Given that brokers are symmetric,  $z = z^j$  for every  $j$ . Then, solving for  $z$  from (1.1),

$$(1.2) \quad z = \frac{ts}{\lambda_1 Q(1+Q)} \frac{T_\mu}{\lambda_2(1+m)} - \frac{\mu_2}{\lambda_2(1+m)} u_1 \text{ where}$$

$$(1.3) \quad T_\mu = \lambda_1(1+Q) - \mu_2 Q$$

Now,  $y_2 = \sum_j z^j + u_2 = mz + u_2$ . Substituting for  $z$  from (1.2), we have:

$$(1.4) \quad y_2 = \frac{ts}{\lambda_1 Q(1+Q)} T_\mu T_\lambda - T_\lambda \mu_2 u_1 + u_2 \text{ where}$$

$$(1.5) \quad T_\lambda = \frac{m}{\lambda_2(1+m)}$$

Since the period one net order flow  $y_1 = A_d s + u_1$ , we can substitute for  $A_d$  from the text and rewrite as:

$$(1.6) \quad y_1 = \frac{ts}{\lambda_1(1+Q)} + u_1$$

Let  $\Sigma_{01} = \text{Cov}(v, y_1)$ ,  $\Sigma_{02} = \text{Cov}(v, y_2)$ ,  $\Sigma_{12} = \text{Cov}(y_1, y_2)$ ,  $\Sigma_{11} = \text{Var}(y_1)$ , and  $\Sigma_{22} = \text{Var}(y_2)$ . From linear projection,  $\lambda_1$ ,  $\mu_2$  and  $\lambda_2$  are given by the following formulas:

$$(1.7) \quad \lambda_1 = \frac{\Sigma_{01}}{\Sigma_{11}}$$

$$(1.8) \quad \lambda_2 = \frac{\Sigma_{02}\Sigma_{11} - \Sigma_{01}\Sigma_{12}}{D}$$

$$(1.9) \quad \mu_2 = \frac{\Sigma_{01}\Sigma_{22} - \Sigma_{02}\Sigma_{12}}{D}$$

$$(1.10) \quad D = \Sigma_{11}\Sigma_{22} - (\Sigma_{12})^2$$

The results of the Proposition can be obtained after working out the expressions for  $\Sigma_{ij}$ ,  $i, j = 1, 2$  and substituting in (1.7)-(1.10). A less computation-intensive alternative is to assume  $\Sigma_{12} = 0$ , which implies (from 1.7 and 1.9)  $\mu_2 = \Sigma_{01}/\Sigma_{11} = \lambda_1$ . From (1.8),  $\lambda_2 = \Sigma_{02}/\Sigma_{22}$  which gives the expression in the text. To complete the proof, we check that if  $\mu_2$  and  $\lambda_2$  are as given in the proposition,  $\Sigma_{12} = 0$ .

Having obtained  $\mu_2$  and  $\lambda_2$ , we substitute in (1.2). In the resulting expression, the coefficient of  $x_d$  is  $B_1$  and the coefficient of  $u_1$  is  $B_2$ . This gives (3) and (4) in the text.

Finally, the expected profits of the  $j$ -th broker is  $W_d = E[(v - p_2)z]$ . After substituting for  $p_2$ :

$$(1.11) \quad W_d = E[(v - \lambda_2(mz + u_2) - \mu_2(x_d + u_1))z]$$

After appropriate simplifications, we get (7) in the text.

### **Proof of Proposition 2:**

#### **Simultaneous dual trading equilibrium.**

Denote  $x_s = \sum_{i=1}^n x^{i,s}$ . There is only one period in this case, with net order flow:

$$(2.1) \quad y_s = \sum_{j=1}^m z^j + x_s + u$$

Let  $x^{m,s}$  denote the  $n$ -tuple  $\{x^{1,s}/m, x^{2,s}/m, \dots, x^{n,s}/m\}$ . Dual trader  $j$  observes  $x^{m,s}$ , and  $u$  and trades  $z^j$ . The dual trader's problem is to maximize with respect to  $z^j$  his expected profits  $E[(v-p_s)z^j | x^{m,s}, u/m]$ . The first order condition for the dual trader's problem is:

$$(2.2) \quad E\langle v | x^{s,d} \rangle - 2\lambda_s z^j - \lambda_s z^{-j} - \lambda_s x_s - \lambda_s u = 0$$

Using linear projection,  $E[v | x^{s,d}] = (tx_s/QA_s)$ . The second order condition is  $-\lambda_s < 0$ , which is satisfied for  $\lambda_s > 0$ .

Given that brokers are symmetric,  $z = z^j$  for every  $j$ . Then, solving for  $z$  from (2.2),

$$(2.3) \quad z = \frac{x_s}{\lambda_s(1+m)} \left[ \frac{t}{QA_s} - \lambda_s \right] - \frac{u}{(1+m)}$$

Informed trader  $i$  observes  $s^i$  and trades  $x^{i,s}$ . The informed trader's problem is to maximize with respect to  $x^{i,s}$  his expected profits  $E[(v-p) x^{i,s} | s^i]$ . Let  $x^{-i,s} = \sum_k x^{k,s}$  for all  $k \neq i$ . Also, note from linear projection that  $E[v | s^i] = ts^i$  and  $E(x^{-i,s} | s^i) = (n-1)A_s ts^i$ . The first order condition for the informed trader's problem is:

$$(2.4) \quad -2x^{i,s} \left[ \frac{\lambda_s}{1+m} + \frac{m}{1+m} \frac{t}{QA_s} \right] + \frac{ts^i}{Q(1+m)} [Q + m - \lambda_s A_s (n-1)Q] = 0$$

The second order condition is  $\lambda_s + (mt/QA_s) > 0$ . Given  $\lambda_s > 0$ , this requires  $A_s > 0$ .



We solve for  $x^{i,s}$  from (2.4). In the resulting expression,  $A_s$  is the coefficient of  $s^i$  and this gives us (9) in the proposition.

We substitute for  $A_s$  in (2.3). In the resulting expression,  $B_{1,s}$  is the coefficient of  $x_s$  and  $B_{2,s}$  is the coefficient of  $u$ . This gives (10) and (11) in the proposition.  $\lambda_s$  is obtained from the formula  $\lambda_s = \text{Cov}(v, y_s) / \text{Var}(y_s)$ . Finally, the expected profits of the  $j$ -th broker is obtained from the expression  $W_s = E[(v - p_s)z]$ .

**Proof of Corollary 1:**

(1) When  $Q \leq m$ , then  $A_s \leq 0$  and so the simultaneous dual trading equilibrium does not exist.

However, the consecutive dual trading equilibrium is viable so long as  $m \geq 1$ .

(2) Suppose  $Q > m$ . In the consecutive dual trading equilibrium, profits for informed trader  $i$

are  $I_{i,d} = E[(v - p_i)x^{i,d}]$ . Aggregate informed profits are  $I_d = nI_{i,d}$ . Using the results from

Proposition 1 and substituting:

$$(c1.1) \quad I_d = \frac{\sqrt{nt\Sigma_u\Sigma_v}}{(1+Q)}$$

In the simultaneous dual trading equilibrium, profits for insider  $i$  are  $I_{i,s} = E[(v - p_s)x^{i,s}]$ .

Aggregate informed profits are  $I_s = nI_{i,s}$ . Using the results from Proposition 2 and substituting:

$$(c1.2) \quad I_s = \frac{\sqrt{nt\Sigma_u\Sigma_v}}{Q(1+Q)} \frac{Q-m}{1+m}$$

Since  $(Q-m)/Q < 1$  and  $1/(1+m) < 1$ ,  $I_s < I_d$ .

Noise trader losses in the consecutive dual trading equilibrium  $U_d$  is the sum of aggregate informed and dual trading profits (since the market maker makes zero expected profits by assumption). Thus  $U_d = I_d + mW_d$ . Substituting for  $I_d$  from (c1.1) and for  $W_d$  from the text,

$$(c1.3) \quad U_d = \frac{\sqrt{nt\Sigma_u\Sigma_v}}{1+Q} \left[ 1 + \frac{\sqrt{1+Q}}{\sqrt{Q}} \frac{\sqrt{m}}{1+m} \right]$$

Noise trader losses in the simultaneous dual trading equilibrium  $U_s = I_s + mW_s$ . Substituting for  $I_s$  from (c1.2) and for  $W_s$  from the text,

$$(c1.4) \quad U_s = \frac{\sqrt{nt\Sigma_u\Sigma_v}}{1+Q}$$

Clearly,  $U_d > U_s$ . The result for broker profits follows from direct comparison of (7) and (12) in the text.

**Proof of Corollary 2:**

- (1) From (c1.1) in Corollary 1,  $I_d$  is independent of  $m$ . From (c1.3) in Corollary 1,  $U_d$  is decreasing in  $m$ . Thus noise traders choose the highest possible number of brokers and we assume informed traders do the same.
- (2) From (c1.2) in Corollary 1,  $I_s$  is decreasing in  $m$  and, from (c1.4) in Corollary 1,  $U_s$  is independent of  $m$ . Thus, informed traders choose  $m=1$  and we assume noise traders do the same.

**Proof of Proposition 3:**

- 1) Since we assume  $t=1$ ,  $Q=n$ . Since  $I_d$  is decreasing in  $n$ , the equilibrium number of informed traders  $n_d$  satisfies  $I_d = nk_d$ . Using (c1.1) in Corollary 1, we have:

$$(3.1) \quad \sqrt{n_d} (1+n_d) = \frac{\sqrt{\Sigma_u\Sigma_v}}{k_d}$$

Similarly, the equilibrium number of brokers  $m_d$  satisfies  $W_d = k_d$  and, using (7) in the text, we have:

$$(3.2) \quad \sqrt{m_d}(1+m_d) = \frac{\sqrt{\Sigma_u \Sigma_v}}{k_d} \frac{1}{\sqrt{1+n_d}}$$

From (3.1) and (3.2),  $m_d < n_d$ .

If  $m_d = 1$  and  $n_d = 2$  is to be an equilibrium, then  $I_d \geq nk_d$  and  $W_d \geq k_d$  at these values, so that from (3.1) and (3.2) we have:

$$(3.3) \quad k_d \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{3\sqrt{2}}$$

$$(3.4) \quad k_d \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{2\sqrt{3}}$$

Clearly, if (3.3) is satisfied, so is (3.4).

2) In the simultaneous dual trading equilibrium,  $m_s = 1$ . Given  $m_s = 1$ , the condition  $I_s = nk_s$  implies, from (c1.2) in Corollary 1, that:

$$(3.5) \quad \frac{2n_s \sqrt{n_s}(1+n_s)}{n_s - 1} = \frac{\sqrt{\Sigma_u \Sigma_v}}{k_s}$$

If  $m_s = 1$  and  $n_s = 2$  is to be an equilibrium, then  $I_s \geq nk_s$  and  $W_s \geq k_s$  at these values, so that from (3.5) above and (12) in the text:

$$(3.6) \quad k_s \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{12\sqrt{2}}$$

$$(3.7) \quad k_s \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{2\sqrt{2}}$$

Clearly, if (3.6) is satisfied, so is (3.7).

#### **Proof of Proposition 4:**

Assume (3.6) above in Proposition 3 is satisfied. Then, with free entry, both consecutive and

simultaneous dual trading equilibria exist. From (3.1) and (3.5) in Proposition 3:

$$(4.1) \quad \frac{\sqrt{n_s} (1+n_s)}{\sqrt{n_d} (1+n_d)} = \frac{n_s - 1}{2n_s} \frac{k_d}{k_s}$$

If  $k_s \leq k_d/2$ , then (4.1) implies that  $n_s \geq n_d$ .

From (5) and (8) in the text,

$$(4.2) \quad \frac{\lambda_s}{\lambda_2} = 2\sqrt{1+m_d} \frac{\sqrt{1+m_d}}{\sqrt{m_d}} \frac{\sqrt{n_s}}{\sqrt{1+n_s}} \frac{\sqrt{1+n_d}}{\sqrt{1+n_s}}$$

If  $m_d$  and  $n_s$  are both large, then  $(1+m_d)/m_d$  and  $n_s/(1+n_s)$  are approximately 1, and (4.2) is approximately:

$$(4.3) \quad \frac{\lambda_s}{\lambda_2} \approx 2\sqrt{1+m_d} \frac{\sqrt{1+n_d}}{\sqrt{1+n_s}}$$

From (4.3),  $\lambda_s < \lambda_2$  if  $n_s$  is large relative to  $n_d$  and  $m_d$ .

### **Proof of Proposition 5:**

(i) We use the subscript “o” to refer to outcomes in the order flow internalization model. Let  $A_o$  be the informed trading intensity,  $\lambda_o$  be the (inverse of) market depth,  $I_o$  be aggregate expected informed profits,  $m_1 W_o$  be the aggregate expected profits of the internalizing brokers and  $U_o$  be uninformed losses. Analogous to Proposition 2, we can show that:

$$(5.1) \quad A_o = \frac{Q - m_2}{Q(1+m_1)} \frac{\sqrt{t\Sigma_u}}{n\Sigma_v}$$

$$(5.2) \quad \lambda_o = \frac{1+m_1\sqrt{nt\Sigma_v}}{(1+Q)\Sigma_u}$$

$$(5.3) \quad I_o = \frac{Q - m_2}{Q} \frac{\sqrt{nt\Sigma_u\Sigma_v}}{(1+Q)(1+m_1)}$$

$$(5.4) \quad m_1 W_o = m_1 \frac{\sqrt{nt \Sigma_u \Sigma_v}}{(1+Q)(1+m_1)}$$

$$(5.5) \quad U_o = \frac{\sqrt{nt \Sigma_u \Sigma_v}}{(1+Q)}$$

We know from Proposition 3 that the number of piggybacking brokers  $m_2=1$  in equilibrium. Further, equilibrium requires  $n > m_2$ , so that  $n \geq 2$ . The number of internalizing brokers  $m_1$  is determined in equilibrium. Let  $m_o$  be the equilibrium number of internalizing brokers. Let  $k_o$  be the entry cost in the market. Finally, let  $t=1$  so that  $Q=n$ . Thus, the equilibrium number of informed traders  $n_o$  satisfies  $I_o = n k_o$  and (5.3) implies:

$$(5.6) \quad k_o = \frac{\sqrt{\Sigma_u \Sigma_v}}{n_o \sqrt{n_o}} \frac{n_o - 1}{n_o + 1} \frac{1}{1 + m_o}$$

Given  $n_o$  and  $m_2=1$ ,  $m_o$  satisfies  $W_o \geq k_o$  and (5.4) implies:

$$(5.7) \quad k_o \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{n_o \sqrt{n_o}} \frac{\sqrt{n_o}}{n_o + 1} \frac{1}{1 + m_o}$$

Inspection of (5.6) and (5.7) shows that both expressions cannot hold as equalities. So (5.7) must hold as an inequality, implying  $W_o > k_o$ . In other words, the brokerage market is not competitive, since brokers make positive expected profits even with free entry.

(ii) In the market without internalization, and with free entry,  $m_2=1$  and  $m_1=0$ . The equilibrium number of informed traders  $n_w$  satisfies  $I_w = n k_o$  where  $I_w$  is  $I_o$  evaluated at  $m_2=1$  and  $m_1=0$ .

$$(5.8) \quad k_o = \frac{\sqrt{\Sigma_u \Sigma_v}}{n_w \sqrt{n_w}} \frac{n_w - 1}{n_w + 1}$$

Comparing (5.6) and (5.8),  $n_o < n_w$ .

Let  $\lambda_w$  be the (inverse of) market depth in the market without internalization.  $\lambda_w$  is simply  $\lambda_o$  evaluated at  $m_1=0$ . Using (5.2), we have:

$$(5.8) \quad \frac{\lambda_o}{\lambda_w} = n_o - 1 \frac{n_w(1+n_w)}{n_o(1+n_o)} \frac{n_w+1}{n_w-1}$$

$\lambda_o > \lambda_w$  since  $n_w > n_o \geq 2$ .

Price informativeness with internalization,  $PI_o$ , is defined as  $\Sigma_v - \text{Var}(v | p_o)$ , where  $p_o = \lambda_o y_o$  and  $y_o$  is the net order flow with internalization. It can be shown that  $PI = n\Sigma_v/(1+n)$ .

Therefore,  $n_w > n_o$  implies that  $PI_o = n_o\Sigma_v/(1+n_o) < n_w\Sigma_v/(1+n_w) = PI_w$  (price informativeness without internalization).

Expected informed profits without internalization  $I_w$  is  $I_o$  evaluated at  $m_1=0$ . From (5.3),

$I_o < I_w$ .

From (5.5), uninformed losses with internalization are proportional to  $\frac{\sqrt{n_o}}{1+n_o}$  whereas

uninformed losses  $U_w$  without internalization are proportional to  $\frac{\sqrt{n_w}}{1+n_w}$ . Hence,  $U_o > U_w$ .