A model of broker’s trading, with applications to order flow internalization*☆

Sugato Chakravartya,*, Asani Sarkarb

aPurdue University, West Lafayette, IN, 47907-1262, USA
bFederal Reserve Bank of New York, New York, NY, 10045, USA

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. © 2002 Elsevier Science Inc. All rights reserved.

JEL classification: G12

Keywords: Dual trading; Informed traders; Internalization; Piggybacking; Adverse selection

1. Introduction

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 US 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).

In futures exchanges, consecutive dual trading (CDT) is permitted, but not simultaneous dual trading (SDT). 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 CDT markets are associated with many brokers, whereas SDT 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, CDT is more likely than SDT. In stock markets, by contrast, the number of dual traders in any stock is few, and SDT is the norm.\(^1\)

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 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 1, each broker submits the sum of informed and Period 1 noise trades to the market maker for execution. In Period 2, each broker submits his personal trade and Period 2 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 SDT 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.\(^2\) 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 SDT. These results appear to imply that, over time, SDT 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

\(^1\) In the NYSE, for example, the potential dual traders are in addition to the specialist national full-line firms and investment banks. In 1989, there six national full-line firms, and 10 investment banks (Matthews, 1994).

\(^2\) 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, pp. 81–84).
market entry fee is sufficiently low, enough informed traders enter the SDT market to make it viable. Further, traders choose many brokers in the CDT market, and only one broker in the SDT market.

To test the applicability of our model, we study order flow internalization by brokers, which primarily affects uninformed retail order flow (Battalio, Green, & Jennings, 1997; Easley, Kiefer, & O’Hara, 1996; Securities and Exchange Commission [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 et al., 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 CDT in a model with a single broker and a fixed order size. Chakravarty (1994) and Roell (1990) have multiple dual traders but a single informed trader. Sarkar (1995) studies SDT 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 et al. (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.3

The remainder of the paper is organized as follows. Section 2 describes a trading model with multiple customers and multiple brokers, when the number of informed traders and brokers is fixed. Sections 3 and 4 solve the CDT and SDT models. Section 5 endogenizes the number of informed traders and brokers. Section 6 analyzes order flow internalization. Section 7 concludes. All proofs are in Appendix.

2. 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

---

3 As Battalio et al. (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.
There are \( n \) informed traders who receive a signal about the true asset value and submit market orders. For an informed trader \( i, i = 1, \ldots, n \), 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 \( u_i \), where \( u_i \) is normally distributed with mean 0 and variance \( \Sigma_u \). All random variables are independent of one another.

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 CDT. Alternatively, they may trade with their customers in the same transaction — i.e., engage in SDT. The sequence of events is as follows: in Stage 1, informed trader \( i \) (for \( i = 1, \ldots, n \)) observes \( s_i \) and chooses a trading quantity \( x_i \). In Stage 2, broker \( j \) (for \( j = 1, \ldots, m \)) trades consecutively or simultaneously, and places orders of an amount \( z_j \).

In subsequent sections, we describe in more detail how SDT and CDT 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.

### 3. Consecutive dual trading

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

#### 3.1. The CDT model

Trading occurs in two periods. In Period 1, brokers receive market orders from \( n \) informed traders and the noise traders, which they then submit to the market maker. In Period 2, brokers trade for themselves, along with Period 2 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 1, informed trader \( i, i = 1, \ldots, n \) observes \( s_i \) and chooses \( x_i^{\text{sd}} \), knowing that his order will be executed in the first period. Accordingly,
informed trader $i$, $i=1, \ldots, n$, chooses $x^{i,d}$ to maximize conditional expected profits $E[(v - p_1)x^{i,d}|s^i]$, where the Period 1 price is $p_1 = \lambda_1y_1$, the Period 1 net order flow is $y_1 = x_d + u_1$, the aggregate informed trade is $x_d = \Sigma x^{i,d}$ and $u_1$ is the Period 1 noise trade.

In Period 2, brokers choose their personal trading quantity after observing the $n$-vector of informed trades $\{x_1^{i,d}, \ldots, x_n^{i,d}\}$, $u_1$ and $p_1$. Thus, broker $j$, $j=1, \ldots, m$, chooses $z^j$ to maximize conditional expected profits $E[(v - p_2)z^j|x^{1,d}/m, \ldots, x^{n,d}/m, u_1/m, p_1]$, where the conditioning is based on each broker observing his portion of the informed and uninformed orders received, plus the Period 1 price.

The $m$ brokers submit their personal trades to the market maker, who sets $p_2 = \lambda_2y_2 + \mu_2y_1$ where $y_2 = \Sigma z^j + u_2$, $\Sigma z^j$ is the aggregate trade of all brokers and $u_2$ is the noise trade in Period 2. Finally, the liquidation value $v$ is publicly observed and both informed traders as well as brokers realize their respective profits (if any).

Define $t = \frac{\sqrt{m}\Sigma v}{\Sigma u}$, where $t$ is the unconditional precision of $s^i$, $i=1, \ldots, n$. 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 CDT case, there is a unique linear equilibrium for $t > 0$. In Period 1, informed trader $i$, $i=1, \ldots, n$, trades $x^{i,d} = A_ds^i$, and the price is $p_1 = \lambda_1y_1$. In Period 2, each broker $j$, $j=1, \ldots, m$ trades $z = B_1x_d + B_2u_1$, the price is $p_2 = \lambda_2y_2 + \mu_2y_1$, and a broker’s expected trading revenue is $W_d$ where (Eqs. (1)–(7)):

\[
\begin{align*}
\lambda_1 &= \frac{\sqrt{nt\Sigma v}}{(1 + Q)\sqrt{\Sigma u}} \\
A_d &= \frac{t}{\lambda_1(1 + Q)} \\
B_1 &= \frac{1}{\sqrt{mQ(1 + Q)}} \\
B_2 &= -\frac{\sqrt{Q}}{\sqrt{m(1 + Q)}} \\
\lambda_2 &= \frac{\sqrt{m}\Sigma v}{\sqrt{Q(1 + Q)\Sigma u}} \frac{\sqrt{m}}{1 + m}
\end{align*}
\]
\[ \mu_2 = \frac{\sqrt{n} \Sigma_v}{(1 + Q) \sqrt{\Sigma_u}} \]  

(6)

\[ W_d = \frac{\sqrt{n} \Sigma_u \Sigma_v}{\sqrt{Q(1 + Q)} \sqrt{m(1 + m)}} \]  

(7)

3.1.1. Discussion

Period 1 informed trades are not affected by dual trading in Period 2 since informed traders trade only in Period 1. 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 2, dual traders piggyback on Period 1 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.

4. Simultaneous dual trading

4.1. The SDT model

SDT is modeled in a single period Kyle (1985) framework. The notations are the same as in Section 3. All variables and parameters related to SDT 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 \( i, i=1, \ldots, n \), 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 \( x_s = \Sigma x^{i,s} \), \( z \) 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 \( \{x^{1,s}, \ldots, x^{n,s}\} \) and \( u \). Thus, broker \( j, j=1, \ldots, m \), chooses \( x^{j,s} \) to maximize expected profits \( E[(v - p_s)x^{j,s}|\{x^{1,s}/m, \ldots, x^{n,s}/m\}, u/m] \), 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 SDT case, there is a unique linear equilibrium for \( t > 0 \) and \( Q > m \). Informed trader \( i, i = 1, \ldots, n \), trades \( x^{t,s} = A_s s^t \), broker \( j, j = 1, \ldots, m \) trades \( z = B_{1,s} x_s + B_{2,u} u \), the price is \( p_s = \lambda_s y_s \). \( W_s \) is the broker’s expected revenue, where:

\[
\lambda_s = (1 + m) \frac{\sqrt{m \Sigma_v}}{\sqrt{\Sigma_u (1 + Q)}} \\
A_s = \frac{(Q - m)t}{\lambda_s Q(1 + Q)} \\
B_{1,s} = \frac{1}{(Q - m)} \\
B_{2,s} = -\frac{1}{(1 + m)} \\
W_s = \frac{\sqrt{n \Sigma_u \Sigma_v}}{Q(1 + m)}
\]

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., \( B_1 > 0 \)), 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.

### 4.2. Comparing SDT and CDT

By combining results from Propositions 1 and 2, 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 CDT equilibrium is viable. (2) If \( Q > m \), then both dual trading equilibria are viable. Relative to SDT, uninformed losses and brokers’ profits are lower, while informed profits are higher with CDT.

As discussed in Proposition 2, the SDT 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 CDT, informed trading is independent of the number of brokers, a market with CDT 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 SDT since their
profits are higher. Since brokers’ profits are at the expense of insiders, informed traders by contrast prefer CDT. Aggregate profits of informed traders and dual traders are higher in the SDT market. Since noise trader losses are the negative of these aggregate profits, noise trader losses are also higher with SDT.

5. 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 SDT equilibrium exists. Then, with SDT, informed and noise traders give orders to only one broker.

With CDT, 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 \( i \), \( i = d \) (CDT), \( s \) (SDT). 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 3 solves for the free entry equilibrium in the two markets.

Proposition 3: (1) In the market with CDT, \( n_d \) informed traders and \( m_d \) brokers enter the market, with \( m_d < n_d \), and \( m_d \) and \( n_d \) given by (Eqs. (13) and (14)):

\[
\sqrt{m_d(1 + m_d)} = \frac{\sqrt{\sum_a \sum_y(1)}}{k_d \sqrt{1 + n_d}}
\]  

(13)
\[ \sqrt{n_d(1 + n_d)} = \frac{\sqrt{\Sigma_u \Sigma_v}}{k_d} \]  
(14)

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

\[ k_d \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{3\sqrt{2}} \]
(15)

Proposition 3 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 CDT 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 (Eq. (17)) is more restrictive than the inequality (Eq. (15)), implying that the SDT equilibrium is viable only at lower market entry cost, relative to the CDT 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 3, suppose \( k_d > 2k_s \). Then \( n_d < n_s \). Market depth may be higher with SDT if \( n_s \) is large relative to \( n_d \) and \( m_d \).

The proposition says that if entry costs are sufficiently low, the SDT 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 SDT market. Consequently, if entry costs are low enough, many informed traders enter the SDT market and market depth is higher. Thus, the proposition provides some intuition as to why we see markets with SDT exist in the real world.

6. **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.\(^4\)

We use our SDT model to analyze order flow internalization. Since internalizing brokers typically handle order flows of small retail investors, we assume that there are \(m_1\) internalizing brokers who handle all of the noise trades, and \(m_2\) piggybacking 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 (Eqs. (10) and (11)) (Eqs. (18) and (19)):

\[
\begin{align*}
    z_1 &= -\frac{u}{(1 + m_1)} \quad (18) \\
    z_2 &= \frac{x_s}{(Q - m_2)} \quad (19)
\end{align*}
\]

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, \(n_o\), informed traders and \(m_o\) internalizing brokers enter the market. \(m_o\) 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.

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 SDT market is worse than in the CDT market and in a market without dual trading.

\(^4\) Rule 19c-3 allows NYSE stocks listed after April 26, 1979 to be traded off-exchange. Broker-dealers may have earned over US$500 million in 1994 from 19c-3 trading (see SEC, 1997, and “In-house trades can be costly for small investors,” Wall Street Journal, December 20, 1994).
7. 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 (SDT) or in a separate transaction (CDT).

While the CDT equilibrium always exists, the SDT 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 SDT, relative to CDT. Thus, only brokers prefer SDT.

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 SDT market, informed and noise traders choose only one broker whereas, with CDT, informed and noise traders choose all available brokers. If the market entry cost is sufficiently low, more informed traders enter the SDT market, and market depth may be lower, relative to the CDT market. Thus, the adverse effects of SDT on customers and the market are mitigated in the free entry equilibrium. If the market entry cost is high, CDT is better.

In the SDT 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.

Appendix

Proof of Proposition 1: CDT equilibrium.

In Period 1, 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 \( \{x^{1,d}/m, x^{2,d}/m, \ldots, x^{n,d}/m\} \). In Period 2, 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 $E[(v-p_2)\varepsilon'|x^{m,d}, u_1/m, p_1]$. Let $z^{-j}=\Sigma_\varepsilon^d$ 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:

\[
E(v \mid x^{m,d}) - 2\lambda_2 z^j - \lambda_2 z^{-j} - \mu_2 A_d s - \mu_2 \mu_1 = 0
\]

(1.1)

Using linear projection, $E[v \mid x^{m,d}] = ts/Q$, 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 Eq. (1.1):

\[
z = \frac{ts}{\lambda_1 Q (1 + Q)} \frac{T_\mu}{\lambda_2 (1 + m)} - \frac{\mu_2}{\lambda_2 (1 + m)} u_1
\]

(1.2)

where (Eq. (1.3))

\[
T_\mu = \lambda_1 (1 + Q) - \mu_2 Q
\]

(1.3)

Now, \( y_2 = \Sigma z^j + u_2 = mz + u_2 \). Substituting for \( z \) from Eq. (1.2), we have (Eq. (1.4)):

\[
y_2 = \frac{ts}{\lambda_1 Q (1 + Q)} T_\mu T_\lambda - T_\lambda \mu_2 u_1 + u_2
\]

(1.4)

where (Eq. (1.5))

\[
T_\lambda = \frac{m}{\lambda_2 (1 + m)}
\]

(1.5)

Since the Period 1 net order flow \( y_1 = A_d s + u_1 \), we can substitute for \( A_d \) from the text and rewrite as (Eq. (1.6)):

\[
y_1 = \frac{ts}{\lambda_1 (1 + Q)} + u_1
\]

(1.6)

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:

\[
\lambda_1 = \frac{\Sigma_{01}}{\Sigma_{11}}
\]

(1.7)

\[
\lambda_2 = \frac{\Sigma_{02} \Sigma_{11} - \Sigma_{01} \Sigma_{12}}{D}
\]

(1.8)

\[
\mu_2 = \frac{\Sigma_{01} \Sigma_{22} - \Sigma_{02} \Sigma_{12}}{D}
\]

(1.9)

\[
D = \Sigma_{11} \Sigma_{22} - (\Sigma_{12})^2
\]

(1.10)

The results of the proposition can be obtained after working out the expressions for \( \Sigma_{ij} \), \( i, j = 1, 2 \) and substituting in Eqs. (1.7)–(1.10). A less computation-intensive alternative is to assume \( \Sigma_{12} = 0 \), which implies (from Eqs. (1.7) and (1.9)) \( \mu_2 = \Sigma_{01} / \Sigma_{11} = \lambda_1 \). From Eq. (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 Eq. (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 Eqs. (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 $p_2$ (Eq. (1.11)):

$$W_d = E[(v - \lambda_2(mx + u_2) - \mu_2(x_d + u_1))z]$$  \hspace{1cm} (1.11)

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

Proof of Proposition 2: SDT equilibrium. Denote $x_s = \sum_{i=1}^n x^{i,s}$. There is only one period in this case, with net order flow (Eq. (2.1)):

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

Let $x^{m,s}$ denote the $n$-tuple $\{x^{1,s}/m, x^{2,s}/m, \ldots, x^{n,s}/m\}$. 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:

$$E(h v_j x_s; d_i/c_0) = E(z_j x^{m,s}, u/m) = 0$$  \hspace{1cm} (2.2)

Using linear projection, $E[v|x^{s,d}] = (tx_s/Q_A_s)$. 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 Eq. (2.2):

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

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 $x^{-i,s} = \sum_{k \neq i} x^{k,s}$ for all $k \neq i$. Also, note from linear projection that $E[v|s^i] = ts^i$ and and $E(x^{-i,s}|s^i) = (n - 1)A_s t s^i$. The first-order condition for the informed trader’s problem is:

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

The second-order condition is $\lambda_s + (mt/Q_A_s) > 0$. Given $\lambda_s > 0$, this requires $A_s > 0$. We solve for $x^{i,s}$ from Eq. (2.4). In the resulting expression, $A_s$ is the coefficient of $s^i$ and this gives us Eq. (9) in the proposition.

We substitute for $A_s$ in Eq. (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 Eqs. (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 SDT equilibrium does not exist. However, the CDT equilibrium is viable so long as \( m \geq 1 \). (2) Suppose \( Q > m \). In the CDT equilibrium, profits for informed trader \( i \) are \( I_{i,d} = E[(v - p_i) x_{i,d}] \). Aggregate informed profits are \( I_d = n I_{i,d} \). Using the results from Proposition 1 and substituting:

\[
I_d = \frac{\sqrt{m \Sigma u \Sigma v}}{(1 + Q)} \tag{c1.1}
\]

In the SDT equilibrium, profits for insider \( i \) are \( I_{i,s} = E[(v - p_s) x_{i,s}] \). Aggregate informed profits are \( I_s = n I_{i,s} \). Using the results from Proposition 2 and substituting:

\[
I_s = \frac{\sqrt{m \Sigma u \Sigma v} Q - m}{Q(1 + Q)} 1 + m \tag{c1.2}
\]

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

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

\[
U_d = \frac{\sqrt{m \Sigma u \Sigma v}}{(1 + Q)} \left[ 1 + \frac{\sqrt{1 + Q}}{\sqrt{Q}} \frac{\sqrt{m}}{1 + m} \right] \tag{c1.3}
\]

Noise trader losses in the SDT equilibrium \( U_s = I_s + m W_s \). Substituting for \( I_s \) from Eq. (c1.2) and for \( W_s \) from the text:

\[
U_s = \frac{\sqrt{m \Sigma u \Sigma v}}{1 + Q} \tag{c1.4}
\]

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

Proof of Corollary 2: (1) From Eq. (c1.1) in Corollary 1, \( I_d \) is independent of \( m \). From Eq. (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 Eq. (c1.2) in Corollary 1, \( I_s \) is decreasing in \( m \) and, from Eq. (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 \( I_d = n k_d \). Using Eq. (c1.1) in Corollary 1, we have:

\[
\sqrt{n_d(1 + n_d)} = \frac{\sqrt{\Sigma u \Sigma v}}{k_d} \tag{3.1}
\]
Similarly, the equilibrium number of brokers \( m_d \) satisfies \( W_d = k_d \) and, using Eq. (7) in the text, we have:

\[
\sqrt{m_d(1 + m_d)} = \frac{\sqrt{\Sigma_u \Sigma_v}}{k_d} \frac{1}{\sqrt{1 + n_d}}
\]

(3.2)

From Eqs. (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 Eqs. (3.1) and (3.2) we have:

\[
k_d \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{3\sqrt{2}}
\]

(3.3)

\[
k_d \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{2\sqrt{3}}
\]

(3.4)

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

(2) In the SDT equilibrium, \( m_s = 1 \). Given \( m_s = 1 \), the condition \( I_s = nk_s \) implies, from Eq. (c1.2) in Corollary 1, that:

\[
\frac{2n_s \sqrt{n_s(1 + n_s)}}{n_s - 1} = \frac{\sqrt{\Sigma_u \Sigma_v}}{ks}
\]

(3.5)

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 Eq. (3.5) above and Eq. (12) in the text:

\[
k_s \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{12\sqrt{2}}
\]

(3.6)

\[
k_s \leq \frac{\sqrt{\Sigma_u \Sigma_v}}{2\sqrt{2}}
\]

(3.7)

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

Proof of Proposition 4: Assume Eq. (3.6) above in Proposition 3 is satisfied. Then, with free entry, both CDT and SDT equilibria exist. From Eqs. (3.1) and (3.5) in Proposition 3:

\[
\frac{\sqrt{n_s}}{\sqrt{n_d}} \frac{1 + n_s}{1 + n_d} = \frac{n_s - 1}{2n_s} \frac{k_d}{k_s}
\]

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

From Eqs. (5) and (8) in the text:

$$\frac{\lambda_1}{\lambda_2} = 2\sqrt{1 + m_d} \frac{\sqrt{1 + m_d}}{\sqrt{m_d}} \frac{\sqrt{n_s}}{\sqrt{1 + n_s}} \frac{\sqrt{1 + n_d}}{\sqrt{1 + n_d}}$$  \hspace{1cm} (4.2)

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 Eq. (4.2) is approximately:

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

From Eq. (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 $A_o$ (Eq. (5.1)) be the informed trading intensity, $\lambda_o$ be the (inverse of) market depth, $I_o$ be aggregate expected informed profits, $m_1W_o$ be the aggregate expected profits of the internalizing brokers and $U_o$ be uninformed losses. Analogous to Proposition 2, we can show that:

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

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

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

$$m_1W_o = m_1 \frac{\sqrt{nt\Sigma_u\Sigma_v}}{(1 + Q)(1 + m_1)}$$  \hspace{1cm} (5.4)

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

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 \( I_o = nk_o \) and Eq. (5.3) implies:

\[
k_o = \frac{\sqrt{\sum w \Sigma v} n_o - 1}{n_o \sqrt{n_o} n_o + 1 + m_o} \tag{5.6}
\]

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

\[
k_o \leq \sqrt{\sum w \Sigma v} \frac{\sqrt{n_o}}{n_o + 1 + m_o} \tag{5.7}
\]

Inspection of Eqs. (5.6) and (5.7) shows that both expressions cannot hold as equalities. So Eq. (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 \( I_w = nk_o \) where \( I_w \) is \( I_o \) evaluated at \( m_2 = 1 \) and \( m_1 = 0 \).

\[
k_o = \frac{\sqrt{\sum w \Sigma v} n_w - 1}{n_w \sqrt{n_w} n_w + 1} \tag{5.8}
\]

Comparing Eqs. (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 Eq. (5.2), we have:

\[
\frac{\lambda_o}{\lambda_2} = n_o - 1 \frac{n_w (1 + n_w) n_w + 1}{n_o (1 + n_o) n_w - 1} \tag{5.8}
\]

\( \lambda_o > \lambda_w \) 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 Eq. (5.3), \( I_o < I_w \).

From Eq. (5.5), uninformed losses with internalization are proportional to \( \sqrt{\frac{n_o}{1 + n_o}} \) whereas uninformed losses \( U_w \) without internalization are proportional to \( \sqrt{\frac{n_w}{1 + n_w}} \). Hence, \( U_o > U_w \).
References


