Introduction and Policy Summary

This chapter compares trading floor practices and their regulation on futures exchanges and on the New York Stock Exchange (NYSE), analyzes the economic functions carried out on these floors, and considers some public policy issues of interest to the futures exchange industry.¹ Since the regulatory practices of the Commodity Futures Trading Commission (CFTC) and the Securities and Exchange Commission (SEC) are frequently compared and the possibility of combining these agencies has been mentioned, a comparative approach seems particularly appropriate for this chapter. I begin this introduction by considering some similarities and differences between futures and securities that are particularly significant for trading floor practices.

Product Similarities and Differences. Most people are unfamiliar with futures contracts and floor trading on exchanges. They naturally compare them with apparently similar familiar objects. Thus futures contracts are compared with highly levered investments and floor trading to off-floor trading. Although these comparisons have some validity, they are inadequate as a basis for evaluating public policy.

It has been said that securities trades transfer ownership while futures market trades transfer risk. This important idea deserves elaboration. Assets are collections of rights and obligations in which...
the market value of the owner's rights is greater than the market value of his obligations. An asset's market value is the difference between the market values of the rights and obligations. To buy an asset, the new owner pays its market value to the previous owner or to the issuer.

The owner of a fixed-coupon bond, for example, has the right to receive the interest and principal payments specified in the bond indenture. The price of the bond reflects the market value of these and other rights. As conditions change, the market value of the bond changes. Some of the contract conditions of the security may be adjustable. In a floating-rate bond, the interest payments are adjusted from time to time according to a formula specified in the bond indenture. The adjustments may be made monthly, semiannually, or annually. The adjustable contract conditions substantially reduce variations in the price of the bond.

A futures contract is like a floating-rate bond in that it contains an adjustable clause that stabilizes its market value. The adjustment takes place each trading day and produces a contract whose market value is zero.

From the point of view of the buyer, a futures contract is a commitment to accept delivery of a specified good (or a designated substitute) at a designated future time and to pay the delivery price specified. The initial delivery price is determined in the auction market at the time the contract is entered into and is called the futures price. At the end of each trading day, each outstanding contract is adjusted by changing the futures price to that day's settlement price. (Roughly, the settlement price is the futures price at the close of trading.) At the settlement price the market value of the futures contract is again zero. The process of adjustment is called "marking to market" and involves a cash payment from the loser to the winner. Marking to market makes all outstanding futures contracts (for the same commodity and delivery period) identical, even though they may originally have had different futures prices. This system increases the size of the market and facilitates subsequent trading.

The lower the futures price, the more favorable the contract is to the buyer and the less favorable to the seller. If the futures price were very low, the buyer would be willing to pay the seller to enter into the contract. The payment would measure the market value of the contract as an asset to the buyer. If the price were very high, the buyer would enter into the contract only if he received a payment from the seller. In practice, no payments are permitted at the time the contract is entered into. Trades take place only at futures prices.
at which both the buyer and seller consider that the market value of the contract at the trade time is zero.

Of course, as conditions change, if the contract terms were not adjusted, the value of the contract would change, just as the value of a fixed-coupon bond changes. A distinctive feature of futures contracts is that the value of the contract is restored to zero each day by marking to market.

Most assets are also risky. Risk is related to the difficulty of accurately predicting the future value of an asset and the future cash flows that will result from holding it. If we classified assets by amount of risk per unit of value, bank deposits would have a low risk per unit of value, common stocks a higher amount of risk, and call options a very high amount of risk. Futures contracts are unusual in having great risk but zero market value. A major attraction of a futures contract is precisely that a person who becomes a party to it makes no immediate change in the market values of the assets and liabilities in his portfolio but a possibly large change in the riskiness of that portfolio. Although futures contracts are risky, acquiring a futures contract position does not necessarily increase a person’s risk exposure. If unexpected gains in the value of a futures contract are likely to be accompanied by unexpected losses in the value of other assets, then acquiring a risky futures market position may reduce the owner’s overall risk exposure.

The statement that “securities trades transfer ownership while futures market trades transfer risk” means that security market trades transfer the ownership of valuable risky contracts while futures market trades transfer the ownership of valueless risky contracts. Both types of transactions transfer risk, but futures contracts do it without transferring value.

If risk can be transferred without using futures contracts, why are futures contracts used? The answer is that futures contracts are used when they are less costly or faster or in some other respect more convenient than a corresponding asset transaction.

Importance of Transaction Costs. This distinction between futures and other kinds of assets highlights the importance of transaction costs in explaining the very existence of futures markets. There would be little use for futures contracts if it were less expensive and just as fast and convenient to change the risk composition of a portfolio by selling some assets and buying others. This is particularly true if the assets are securities like common stock, in which the transaction costs are relatively low.
TRADING FLOOR PRACTICES

On all exchanges, low transaction costs require high trading volumes because trading markets are subject to economies of scale; that is, other things being equal, trading costs are lower in more active markets. Many of the differences between the trading practices of securities and of futures exchanges can be explained by the need for low transaction costs on futures markets. Futures exchanges have many fewer trading vehicles than stock exchanges but many more transactions per vehicle. The volume of transactions in most futures contracts is large enough for several exchange members to derive their sole livelihood from trading in that market. The more active contracts attract several hundred members. By contrast, there are few common stocks in which the trading volume is large enough so that a stock exchange member can make a living by trading only one issue.

When a large number of persons trade in a single active market, a considerable division of labor is possible. The individual trader can select a trading role suited to his talents and circumstances. In each possible role he is likely to face competing traders. A detailed description of the activities of members who trade on the floor of a futures exchange is presented later in this chapter; in broad terms they may be brokers or dealers or both. The brokerage function seems to be well understood. The dealer’s function on exchange floors is frequently misunderstood.

It is commonly assumed that dealers on the floor of an exchange do what traders off the floor do but have an intrinsic advantage because they are at the center of the action. This is a plausible assumption, but it is not accurate for most floor dealers. The used-car dealer is a better model for the role of an on-the-floor dealer. Do car dealers buy used cars for the same reason that we do? Of course not. We buy used cars to drive. They buy used cars for resale. We look at a car and ask, “How will this drive?” The dealer looks at a car and asks, “How will it sell?” There may be a connection between driving well and selling well. But the two criteria are not identical. The dealer may not think a particular car drives well, but if he knows that his customers think it does, he may buy it. Or he may think it is a fine driving car, but if it is unpopular with his customers, he is reluctant to buy it. Floor dealers are similar to used car dealers in that both aim to profit by earning a dealer’s spread on a rapidly turning inventory. There are important differences, including the lack of standardization for used cars, which creates trading opportunities that do not exist when trading standardized futures contracts.

The consequences of one or the other of these views of the floor dealers’ function are far reaching. The concept of the on-the-floor
dealer as a speculator whose access to more timely information gives him an unfair advantage over "the public" has formed the basis for much of the regulation of floor trading on securities exchanges. A consequence of this regulatory pattern for securities markets is that trading is more expensive than it would otherwise be. The same approach in futures markets might well eliminate some of these markets. The fundamental difference is that there is no good substitute for many securities markets, while the existence of futures markets depends on their providing less costly transactions than the underlying assets. A futures contract that does not attract enough volume is likely to disappear, but low-volume common stocks will continue to be traded somewhere.

Futures markets are efficient means of transferring risk. In the United States they are characterized by relative freedom from insider trader problems, a high volume of transactions per trading vehicle, and settlement by offset. All three characteristics are conducive to low transaction costs. Futures exchanges are expert at creating and operating exchange marketplaces with low transaction costs. The stock exchanges are expert at operating marketplaces that require more nontransaction services, especially surveillance and regulation. The high cost of the additional services is one reason stock markets are likely to have higher transaction costs. Both types of exchange serve substantial numbers of individuals and small businesses.

The next two sections of this chapter present detailed descriptions of "Trading on Futures Exchange Floors" and "Trading on the New York Stock Exchange Floor." They devote particular attention to the specialized roles performed by exchange members and to the similarities and differences between the degree of specialization in the two types of exchange as it is influenced by the volume of trading, the physical arrangement of trading, trading techniques, and, particularly for the New York Stock Exchange, the regulatory environment. These descriptive sections are followed by "Economic Analysis of Trading on Exchange Floors," which considers some features common to both types of exchange and relevant to public policy issues, with special attention to market making and other kinds of floor trading. The last section discusses economic criteria for evaluating the regulation of exchange markets and compares the regulation of stock exchanges and futures exchanges with respect to floor trading, trading procedures, and trading vehicles. Some of the policy issues considered are mentioned briefly here.

Futures exchanges usually permit any qualified members to trade for their own accounts from the floor. This has not been a matter of controversy, although it is a major point of difference between the
futures and securities industries. The issue of dual trading—whether persons who act as floor brokers should also be allowed to trade for their own accounts—has generated controversy within the futures industry. Exchanges regulated by the SEC generally exclude from the trading floor persons trading primarily for their own accounts unless it can be shown that they perform a market-making function. If they do, their floor trading is usually restricted to this market-making function. The rationale for the SEC’s approach to this area of regulation is described and critiqued.

Recently the CFTC proposed that exchanges be given the alternative of eliminating dual trading or recording trades to the nearest minute. It is difficult to evaluate this proposal since the CFTC gave no indication of the extent of the dual trading problem or of how the data it proposes to require might be used. At a minimum, this information should be provided and an effort made to estimate the costs and benefits of the proposed new regulations.

Futures exchanges tend to be more restrictive than stock exchanges in regulating trading procedures. This has made experimentation and innovation in this area difficult. Two possible areas of innovation, block trading and computer-assisted trading, are discussed, and some suggestions for experimenting with new trading techniques on futures exchanges are advanced.

The present regulatory framework prevents futures trading in individual stocks. A limited number of stocks would be suitable for futures trading. Allowing such trading would tend to increase the operating efficiency of both futures and stock exchanges. These advantages should be considered when the regulatory framework is reconsidered.

Trading on Futures Exchange Floors

Trading Floors. The trading floor of an active futures exchange is a large room, sometimes nearly as large as a football field. Trading takes place in roughly circular areas called pits, which are flat in the center and have wide steps rising in concentric stages to the edge. Traders stand in the center facing out or on the steps facing in. The steps help them to see one another. A trading floor usually has several trading pits, near which are booths with telephones or other
electronic equipment, which are assigned to member brokerage firms. Orders from customers are sent to a booth used by the brokerage firm and transmitted by messenger to a broker in the appropriate trading pit. Trading floors also have a variety of devices to display information. The kind of information varies, but reports about trading (on that exchange or others) are always prominent.

For an actively trading futures contract, the pit may contain over 200 members. Trading is by open outcry. Members shout and wave their arms to attract attention to their bids and offers; they use hand signals to communicate the details of the bids and offers and to acknowledge their acceptance. Because the members trading with one another may be twenty or thirty feet apart, both buyer and seller record each trade separately. The CFTC "requires every contract market to adopt rules requiring that every trade by a floor broker or trader be confirmed promptly with the opposite floor broker or floor trader." After being confirmed, each side is reported through the trader's clearing member to the clearinghouse, which attempts to match them. "Unmatched trades, also known as out-trades, generally result from misunderstandings between traders on the exchange floor or errors by traders, clerks, or key-punch operators in reporting trades." Out-trades are ordinarily resolved before the next day's trading begins. They are an expensive but not an insurmountable problem.

When a clearing house accepts a trade for clearance it interposes itself as a principal between the buyer and seller, or between the clearing members handling their accounts if the buyer and seller are not themselves members of the clearing house. Until the clearing house accepts a trade for clearance, the individual exchange members who effected the trade are obligated to one another for the performance of the contract they have made. At such time as the clearing house accepts the trade for clearance, the clearing house assumes the rights and obligations of each exchange member with respect to the other . . . and the exchange members are no longer obligated to one another.

Until the trade has been accepted, the member has only a claim against another member that is not substantiated by any jointly acknowledged written document. Thus members prefer to trade with other members whose reputations for competence and reliability are known to them.

The volume of trading in an active futures markets regularly exceeds the volume of trading in all but the most active stocks. The numbers of units (one contract or one round lot) traded in the ten
most active exchange-traded products on January 11, 1984, are shown in table 2-1. With a six-hour trading day, 45,000 contracts are equivalent to an average rate of over two contracts per second. (On the previous day, T-bond trading approached 82,000 contracts.) A convenient record of the number of separate transactions would show an increased difference between stocks and futures, since a substantial part of the stock volume consists of block trades (transactions exceeding 10,000 shares), which are prearranged off the stock exchange floor. Prearranged trades are generally prohibited on futures exchanges.\(^13\)

Corresponding to the differences in trading volume are different trade reporting systems. The stock market attempts to record the price and volume of every trade. In futures markets price changes are reported but not every trade and not volume.\(^14\) Volume estimates are prepared during the trading sessions at half-hour intervals, but accurate data become available from clearinghouse reports only after the market has closed.

The activity in a trading pit is not steady. It is high near the open and close of trading and when there is an important change in the environment. If a runner with an order from a brokerage firm approaches the pit, traders may begin shouting to attract attention even before the order has been handed to the broker. Only traders on the top step can see the runner. When they become active, others lower down guess the reason and respond in a similar way. Still others who cannot guess the reason know that the chance of a trade

<table>
<thead>
<tr>
<th>Market</th>
<th>Volume</th>
<th>Minimum Dollar Change</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-bonds</td>
<td>45,000</td>
<td>31.25</td>
<td>futures</td>
</tr>
<tr>
<td>Soybeans</td>
<td>44,868</td>
<td>12.50</td>
<td>futures</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>40,166</td>
<td>25.00</td>
<td>futures</td>
</tr>
<tr>
<td>AT&amp;T (wi)</td>
<td>37,523</td>
<td>12.50</td>
<td>stock</td>
</tr>
<tr>
<td>Corn</td>
<td>36,300</td>
<td>12.50</td>
<td>futures</td>
</tr>
<tr>
<td>Gold (Comex)</td>
<td>36,000</td>
<td>10.00</td>
<td>futures</td>
</tr>
<tr>
<td>AT&amp;T (old)</td>
<td>17,184</td>
<td>12.50</td>
<td>stock</td>
</tr>
<tr>
<td>Texaco</td>
<td>17,103</td>
<td>12.50</td>
<td>stock</td>
</tr>
<tr>
<td>Baxter Trav</td>
<td>16,966</td>
<td>12.50</td>
<td>stock</td>
</tr>
<tr>
<td>Nat'l Semi</td>
<td>11,374</td>
<td>12.50</td>
<td>stock</td>
</tr>
</tbody>
</table>

\(^a\) When issued.  

in the next few moments is greater and join in the chorus if they want to trade. Whether or not a trade takes place, the commotion may subside as quickly as it started. Increased activity may originate in the center of the pit as well. Traders there may be the first to notice a change in one of the information displays surrounding the floor, making them more anxious to trade, perhaps at a price that would not have been acceptable before. Trading in a stock index futures pit may respond to information about a price change for one of the heavily weighted stocks (especially IBM); trading in a grain pit may respond to price changes in one of the other grains.

Interpreting what is happening in a large, active trading pit is impossible for an inexperienced observer. Even experienced traders differ in their ability to absorb and make sense of the activity that is occurring; no one can absorb it all. Therefore economic theories that assume fully informed, rational participants cannot accurately describe behavior on futures market trading floors.¹⁵

One strategy for coping with the information overload is to focus on certain kinds of information and ignore the rest. One could focus on a particular trader whose actions are considered important or whose behavior is typical of a larger group or on certain kinds of orders. One brokerage firm keeps track of all large orders in the contracts its customers trade and summarizes them every half-hour. The amounts sold by large traders in the past half-hour and the price response may be useful information about the state of the market.

The trading floor of a futures exchange is characterized by a complex division of labor that has not been given adequate attention or economic analysis. General equilibrium analyses tend to assume undifferentiated economic actors or to make only gross distinctions based on broadly defined roles (producer versus consumer, buyer versus seller) or on the amount of information available to different individuals performing the same role. Microanalyses of futures markets are more likely to distinguish among different roles (hedgers, speculators) but give at best a sketchy indication of the connections among them. In this section I attempt to give adequate recognition to the fundamentally different behavior of different groups of participants in the futures market and to the relations among them.¹⁶

The several floor trading roles are relatively well defined as ideal types. An individual may combine more than one role at one time or may switch from one role to another. One important way of describing a market is by the number of participants and the proportion who carry out each major role. Traders may migrate from one trading pit to another or may leave the trading floor. The freedom to switch or to combine roles is limited by the rules of some exchanges.
On futures exchanges role-limiting rules are relatively rare. On stock exchanges subject to SEC regulations, they are more significant.\(^{17}\)

Trading Roles.

_Floor broker._ A floor broker transacts business on behalf of a client in accordance with the client’s orders. Retail customers ordinarily deal directly with their account executives and have no direct contact with the broker who handles their orders on the floor of the exchange. Commercial customers and very active individual customers are more likely to have some direct contacts with a floor broker. The floor broker may be a salaried employee of the brokerage firm whose orders he handles or an independent who works on commission for one or more brokerage firms.

Floor brokers act under clients’ orders and may have less decision-making freedom than others in the pit who are responsible only to themselves. In practice, some clients delegate considerable responsibility to their floor brokers to carry out their trading objectives, since the floor broker may have a better feel for the market.

In principle a broker should always put the client’s interests first. Some of the most sensitive and difficult issues concerning the trading floor pertain to situations in which a floor broker may be tempted to deviate from this rule.\(^{18}\) An obvious example is a situation in which a broker may be tempted to trade for himself before executing a client’s order. Another is a situation in which a broker holding orders from two customers is tempted to favor one of them at the expense of the other. The similarity between the two situations is increased if the broker has an economic incentive to favor one of the customers. These issues are discussed in more detail later.

_Scalper._ Scalpers are independent traders who seek to identify temporary imbalances in the flow of orders to the pit from which a profit can be earned by buying and selling for their own account. Scalpers avoid exposing themselves to large losses by limiting the size of their positions. One recent study found that a scalper in an active market held positions for an average of just under two minutes. Positions that were liquidated quickly were usually profitable; positions held for more than three minutes were usually unprofitable.\(^{19}\)

The scalper provides a service to other traders who are anxious to trade by making it possible for them to complete their orders quickly at a price close to the last price. Scalpers provide liquidity or “predictable immediacy” to the extent that they buy at the bid and sell at the ask price. They may also provide “depth.” When they close out losing positions, however, by selling at the bid price or
buying at the ask price, they "consume" liquidity and may cause temporary instability.

*Position trader.* Position traders take larger positions than pure scalpers and hold them for a longer time. Having a lower turnover and more risk, they must aim for a higher profit margin.

Why should any position traders be members of a futures exchange or trade on the floor if they are members? Certainly most position traders are not floor traders or even exchange members. Most nonhedging transactions that originate off the floor of an exchange are from nonmembers who would be classified as position traders. It may be surprising to find that any position traders are also floor traders.

Three considerations might influence a trader in choosing whether to do position trading from the floor of an exchange. Cost savings are one possible consideration. Joining an exchange might be cost-effective if the trader's volume were large enough that the savings from trading at the member's rate covered the added fixed costs of membership. Similarly, a member might chose to trade from the floor if the savings from not paying floor brokerage to other members were great enough to cover the added opportunity costs of spending time on the floor. Futures exchange rules would not prevent a traders from joining an exchange if their main motive was to save money by executing trades from the floor, but current SEC rules would not permit them to do so on a stock exchange.

A second consideration is the extent to which relevant information is more easily available on or off the floor. This is best discussed in connection with the various kinds of position trading.

A third consideration is the trader's primary occupation. If the main source of income is an activity that requires being on the floor, any position trading is likely to originate from the floor, even if being there is a disadvantage to this trading. If the main source of income is an activity best conducted off the floor, position trading will also originate off the floor, even if being on it would be an advantage. A position trader may be off the floor most of the time, even though preferring to be there under some conditions.

The ability to combine position trading with other activities that require a presence on the floor lowers the cost of providing the other competitively priced services with which it is combined. Thus I expect that prohibiting floor brokers from trading for their own accounts would increase the cost of floor brokerage services. The extra cost needs to be considered in evaluating the net benefits of proposed trading restrictions.
Position traders can be classified, on the basis of the information that they use, as fundamental traders or technical traders. A fundamental trader makes decisions based on information about the demand and supply for the commodity or security underlying the futures contract—corn, wheat, Treasury bills, common stock prices. Can fundamental traders obtain an information advantage by trading from the floor of an exchange? For most fundamental traders at most times, the answer appears to be no. There may be exceptions. No one would deny that a trader on the floor may sometimes acquire relevant fundamental information that is not so readily available elsewhere. In deciding whether to trade on or off the floor, the fundamental trader has to compare the amount of relevant information available on the floor with the amount off it. It seems clear that most (perhaps all) persons who are mainly fundamental traders spend most of their time off the floor, although some may spend some time on it when the particular information they are seeking may be more readily available there.

The distinction between fundamental information and technical information is not always clear. In analyzing the price of corn, a decrease in the supply of other feed grains would certainly be considered fundamental information, since it would tend to increase the demand for corn as a livestock feed. But suppose corn oil processors increase their demand because they anticipate a future decrease in the supply of other feed grains. Is this fundamental information or technical information? Whichever it is, it will be of interest to many who consider themselves fundamental traders, and it is the kind of information more readily available around the Chicago Board of Trade than in Washington, D.C., or Miami. It is also likely that many fundamental position traders are "impure" in that they also engage in other types of trading, particularly scalping.

A technical trader studies the past behavior of price and volume, often with the aid of charts, in an effort to anticipate the future course of the market. The existence of profitable technical trading is a matter of controversy among academic economists. Profitable technical trading is ordinarily interpreted as a violation of the efficient market hypothesis, because it suggests that future prices are predictable from publicly available information. Considerable academic literature attempts to replicate technical trading rules and work out profitable strategies from them. Most such attempts have proved unsuccessful, and most academic economists are extremely skeptical about the possibility of successful technical trading strategies.

The strongest evidence to the contrary is indirect. Many futures market traders claim to be technical traders. The amateurs can be dismissed: most amateur futures traders lose money, and there is
no evidence that amateur technical traders do better than other amateurs. But the existence of professional traders who claim to be technical traders is harder to dismiss. If they make a living from trading, it must be profitable. It is sometimes suggested that their trading is profitable despite, rather than because of, their efforts at technical trading or that they are earning a profit but less than the opportunity cost of their capital or labor. These explanations are hard to accept. It is easier to accept the tentative hypothesis that technical traders operating from the floor of an exchange have access to nonfundamental information about the state of the market that has not been captured successfully in the published academic studies, which rely largely on closing prices.

If some professional traders are successful as technical traders, it is probably because they have succeeded in detecting some regular patterns of behavior of other market participants. Conversations with knowledgeable traders have suggested plausible ways in which knowledge about "the state of the market" rather than fundamental information could be the basis of profitable trading. The required knowledge identifies the existence of an unstable rather than a stable equilibrium, where a price decrease is likely to stimulate additional selling or a price increase to stimulate additional buying. Usually the market will be unstable with respect to price increases or to price decreases but not to both. These may be temporary conditions, but while they exist, the expected change in the market price is not zero for those able to identify the required conditions.

*Spreader.* Floor traders may also be spreaders. Spreading is taking offsetting or opposing positions in two or more related commodities. A specific kind of spreading in which one of the goods can be transformed into the other is called arbitrage. An example would be buying a T-bill that could be delivered to satisfy a short position in a T-bill future. Spreading takes place between many different markets: between maturities of the same future; between futures and the cash markets underlying the futures; between different futures markets; between futures and options on the same cash commodity; and between futures and options on the futures.

Spreading is a strategy that can be used by scalpers as well as position traders, but some traders specialize in it. I use the term "spreader" to refer to this group. Most spreaders specialize in spreading between a particular set (usually a pair) of markets. For any pair of markets, the spreading role can be further classified as scalping or position trading. Thus spreading is not a single trading role but a collection of related roles.

Spreading is important in a number of respects. First, it links
prices in related markets so that news or trading pressures that manifest themselves first in one market rather quickly influence prices in related markets as well. Because transaction costs are nearly always lower in futures markets than in the underlying cash markets, spreads involving futures on at least one side are likely to respond to smaller influences than other markets. (Some markets, most notably the money market, are rather closely related to all markets.) Second, spreading adds immensely to the complexity of the trading influences that affect any particular market and therefore reduces the extent to which any single trader can obtain a comprehensive picture of all the influences that might affect any market or market pairs. Third, spreading between an active and an inactive market adds liquidity to the inactive markets.

Hedger. Finally, floor traders who represent commercial firms are hedgers. Hedging is usually defined as taking a position in one market (in this case a futures market) that is the opposite of a position in another market (usually the cash market) with the objective of reducing risk. I much prefer an alternative definition. A hedge is "the purchase or sale of a futures contract as a temporary substitute for a transaction to be made at a later date." Hedging is related to spreading. The difference is one of emphasis. The hedger's primary business is in the cash market; futures transactions are made to facilitate the cash business. Spreading does not imply any primacy of one market or the other. In practice, most spreading between cash and futures is done by hedgers; spreaders operate in other market pairs. The hedger incurs transaction costs in the cash market as part of a regular business; the extra transaction costs from hedging are only those of the futures market part of the spread.

Scalping. The scalper's objective is to buy at the bid and sell at the offer and thus earn a gross profit equal to the spread as often as possible while minimizing exposure to inventory losses. Sometimes it will be necessary to close out a position at a loss to avoid the risk of an even larger loss.

The expected net gains to scalping are equal to the expected number of round trips per day times the expected net profit per round trip. The number of round trips depends on the total volume of trading in the market, the number of competing scalpers, and the ability of the scalper. The expected net profit per round trip depends on the proportions of profitable or unprofitable round trips and the expected profit or loss from each kind. These in turn depend on the scalper's skill in relation to those of other scalpers and on such factors
as the variable costs scalpers must pay per trade, the size of the minimum price movement, and the characteristics of the flow of incoming orders.

Rules for scalping. Let us consider a simple set of rules that give a crude approximation of how a scalper might operate. In these rules $K$ is a positive integer representing the largest number of contracts the scaler is willing to hold; $K$ could be as small as one. For expository convenience, the trader is assumed to be male.

1. If the scaler has an inventory (long or short) of less than $K$ contracts, he publishes competitive quotations on both sides of the market until he completes a transaction. (A competitive quotation is one that at least matches the best quotation on that side of the market.)

2. If the scaler has a long or short position equal to $K$ contracts, he publishes a competitive quotation on the side of the market that would move his inventory toward zero if it were accepted and withdraws his quotation from the other side of the market. Thus with a long position he publishes a competitive offer, with a short position a competitive bid. He maintains these quotations until he completes a transaction or the quotations move against him. If he completes an inventory-reducing transaction before the quotations change against him, he returns to rule 1. He has completed a profitable round turn. If the quotations change against him, he follows rule 3. (A quotation change is against the scaler if the competitive offer declines when he is trying to sell or the competitive bid rises when he is trying to buy.)

3. If the quotations change against the scaler when he is holding a position, he closes out his position by accepting the best available bid (if he is long) or offer (if he is short). He has then completed an unprofitable round turn and returns to rule 1.

One important determinant of a scaler's profits is the number of round turns he is able to complete. Consider a scaler who is willing to hold a position of no more than one contract ($K$ equals 1). Suppose that the flow of orders and the competition from other scalpers are such that a scaler can complete a transaction every four minutes on average if he publishes competitive quotations on one side of a market. If he publishes competitive quotations on both sides, his average waiting time to complete the first transaction will be two minutes. The turnover frequency under these assumptions can be calculated from the rules described above. If he starts out with no inventory, rule 1 applies, and he publishes two-sided quota-
tions. His average waiting time to the first transaction will be two minutes. Then, following rule 2, he offers quotations on only one side of the market. His average waiting time to the next transaction will be four minutes. Thus a round turn will require an average of six minutes. His expected turnover will be ten round turns per hour.

When these conditions prevail, a scalper can increase the volume of his transactions by taking larger inventories. Suppose that all transactions are for one contract and that \( K \) equals two contracts. Two-sided quotations are offered if the inventory is zero or one contract, one-sided quotations only if the inventory is two contracts, long or short. The average waiting time for the next transaction will be only two minutes when the scalper's inventory is zero or one and four minutes when it is two contracts. The average waiting time per round turn will be five minutes, and the expected inventory turnover will be twelve transactions per hour.  

**Rules versus judgment.** The scalping rules described are completely mechanical. The behavior of a scalper who followed them literally would depend only on his inventory position and the market quotations. Although it would be possible to scalp by following these rules exactly, it is not clear how profitable it would be. Scalpers generally follow similar rules but depart from them to fit the particular circumstances. That is, most of the time they do what the rules say, but sometimes, relying on market feel or judgment, they do something different. Although rule 1 says to publish competitive quotations on both sides of the market when you have no inventory, a scalper may on occasion choose to publish only an offer. He will do so when his judgment tells him that the market is much more likely to decline than to rise. Some of the information that might be used to arrive at such judgments is described later. Scalpers differ in their ability to use judgment to modify standard scalping rules effectively.

**Effects of order types.** Scalpers have an opportunity to trade with anyone willing to buy or to sell at the market. Limit orders held by floor brokers constitute competition for scalpers. A limit order is an order for a specific quantity to be executed at a specific price or better. A limit to buy ten contracts at $3.10 or less would be a limit buy order.

Limit orders at prices away from the current market price have ambiguous effects. Consider a limit order to buy at 0.5 cent below the current bid. If the bid drops to that level, the floor broker holding the limit order competes with the scalper for the right to deal with anyone selling at the market. In that sense limit orders offer potential competition, but they also provide a safety net for a scalper who
may wish to dispose of his position at a loss (under rule 3). A scalper feels safer holding a long inventory if he knows there are a large number of limit buy orders in the market at prices just below the current bid. To the extent that limit orders at prices away from the market increase the willingness of scalpers to hold long positions, such orders may reduce the probability that the price will drop to a level at which they can be executed. A market with many limit orders close to the current price would be a relatively stable market.

A stop order is “an order which becomes a market order if and when the market reaches a designated price. A ‘buy stop’ is placed above the market, and if the market price rises to the level designated on the order (the ‘Stop Price’), the order becomes a market order.”

I have never seen statistics on what kinds of orders are submitted by brokerage customers, but it appears that market orders are more common than limit orders in futures markets and that stop orders are frequently used in an attempt to limit losses. This has some important implications.

It is a serious violation of exchange rules to disclose another person’s order. It appears, however, that traders on the floor often have a general idea of the existence of unfilled orders. This does not necessarily imply any violation of exchange rules. A floor broker who also trades for his own account will know about the orders in his hand and may assume that other brokers have similar decks of orders. It is a violation of the rules for a broker to trade for his own account when holding unfilled customers’ orders at the same price and on the same side of the market. It is not a violation (unless trading by floor brokers is prohibited entirely) to buy or sell when holding orders that are away from the market. A main objective of charting and other technical analysis systems is to discover where the limits and stops are. Moreover, the behavior of the market when it enters certain price levels may provide information about the existence of limit orders.

Order flows. The ideal market for a scalper is one with many at-the-market orders exactly balanced between buy and sell orders. Under those conditions scalpers can have high inventory turnover and low inventories. Such conditions are not likely to obtain for long in any market. Even if market orders are generated at random with equal probabilities of buy and sell orders, the number of buy orders and the number of sell orders will eventually differ by a large amount. More important, during some periods the probabilities of buy and sell orders will not be equal.

When equal numbers of buy and sell orders are arriving, some
scalpers may have long positions and others short ones, but the consolidated inventory position of all the scalpers will be close to zero and will remain there. As long as the order flow remains balanced and there is no important news, there will be no reason for the bid and ask quotations to change.

If the order flow shows a predominance of sell orders, the scalpers quoting two-sided markets will find that their bids are more likely to be accepted than their offers, and their inventories will increase. As a scalper's long inventory position reaches its upper limit, he will withdraw his bid. Selling pressure will then be more concentrated on the remaining scalpers, whose inventory positions will tend to increase faster. When nearly all scalpers have large inventories on the same side of the market, their opportunities for profitable transactions are reduced, because all the scalpers are competing to take the other side of any market order that would reduce their inventory, and their opportunities for losses are increased, because scalpers are less willing to take the other side of inventory-increasing market orders. This situation increases the chance that such orders can be completed only through an unfavorable price movement.

If a scalper becomes aware of such a market condition, he may decide to reduce his inventory by selling at the bid rather than taking the chance of incurring a loss if the price declines. Of course, his action may cause such a decline. He may nevertheless prefer a small loss to the possibility of an even larger loss. If the price declines, either because scalpers begin to recognize the weak position of the market and seek to reduce their inventories or because the previous order imbalance continues, other scalpers are likely to dispose of their inventories aggressively, causing further price declines.

Scalper inventory liquidation. The term "scalper inventory liquidation" refers to a situation in which scalpers as a group seek to reduce their inventories aggressively. Such events have important implications for market prices and scalpers' profits.

When the consolidated inventory position of scalpers is near zero, scalpers add stability to the market because of their willingness to take the other side of either buy or sell market orders with no more than the minimum price change associated with moving from one side of the bid-ask spread to the other. Under those conditions, small imbalances, if they are not associated with obvious news developments, are unlikely to cause a change in the market quotations. As the consolidated inventory position moves away from zero, however, to long or short positions near the upper limits that scalpers
as a group are comfortable in holding, the market becomes potentially unstable. Relatively small imbalances in order flows or minor news may lead to a scalper inventory liquidation, which will cause price changes that are likely to be large in relation to the imbalance that served as the immediate trigger.

**Price and profit consequences.** The price consequences of scalping can be summed up briefly as follows. If the flow of market orders is balanced or there is an imbalance that is not too large, the presence of scalpers causes market orders to be executed at prices more favorable to the persons initiating them than if there were no scalpers. These conditions are also favorable to the scalper, who tends to make a profit equal to the spread on each round turn and does not experience inventory losses.\(^{31}\) If the flow of market orders is sufficiently unbalanced, the presence of scalpers may briefly delay a price change but not prevent it. The price change may be more concentrated in time than it would have been in a market without scalpers. Insofar as price changes are caused by imbalances in the flow of orders, they tend to occur when scalpers have relatively large inventory positions and to cause inventory losses to scalpers.

The combined inventory of scalpers reflects the cumulative imbalances in limit orders over some period of time. When this imbalance reaches the maximum amount scalpers are willing to hold, they seek to limit their risk by aggressively reducing their inventories.

**Number of scalpers.** A large market that can support many scalpers has two possible adjustment mechanisms. If scalping is profitable, existing scalpers try to get additional business and may lower the spread between their bid-ask quotations if the spread is large enough to cover their variable costs. In all conventional organized markets, the minimum price variation is set by exchange rules. A spread equal to the minimum price variation is the minimum spread possible without changing the rules. If the spread is at the minimum allowed and scalping is still profitable, additional scalpers will be tempted to enter the market. If each scalper can attract an approximately equal amount of business, the profit per scalper will decline as more scalpers enter a market whose total volume is fixed.

**Economic Relationships.** A futures contract is a zero-sum relation. The gross profits of the winners are equal in absolute amount and opposite in algebraic sign to the gross profits of the losers. With this in mind I next consider the factors that determine the number of professional traders who adopt each role or style of trading.
Since futures markets have few rules that limit the ability of a member to move from one style of trading to another, members can change roles whenever they believe it is profitable to do so. They can also move from one trading pit to another. The price of memberships is determined by the market and should reflect the marginal value of a membership, given the number available. Treating the opportunity cost of the membership as a cost, in equilibrium members should distribute themselves among futures contracts and among trading styles within a pit so that no member should expect to be able to increase expected profits by transferring from one pit or from one style of trading to another.

Certain conditions are favorable to a particular style. The use of market orders creates favorable opportunities for scalpers. As the conditions favorable to a particular style increase, the number of members who use that style should increase. Similarly, certain conditions are unfavorable to a particular style of trading. Changes in the equilibrium price level not caused by sudden publicly announced news events are unfavorable for scalping. As such changes increase in frequency and size, the number of scalpers should decrease. Although the factors that favor or discourage a particular style of trading are difficult to observe directly and often difficult to quantify, the numbers of traders following particular styles may be easier to observe and may indicate some features of futures markets that would otherwise be unobservable. Comparisons of the numbers for different futures or for the same contract at different times may help identify important changes in the markets and give us a much richer understanding of how they work. There is some reason to believe, for example, that the proportion of traders who are scalpers is lower now than in previous decades. It would be interesting to know why.

Three groups of traders cannot necessarily be assumed to earn profits (in the sense defined above) from their participation in any particular futures market. (This does not mean that they do not benefit from their participation.) These groups are hedgers, spreaders, and amateur traders (all trading styles). For each group the reason is that the traders do not earn a living only from trading in one futures market. These three groups as a whole must have net losses equal to the net members’ profits plus operating costs. It would help our understanding of futures markets if we knew more about the flow of profits and losses among groups of traders. Some evidence suggests that, at least for American futures markets in grains, amateur traders have usually been losers but hedgers have not. There is also some indication of a high turnover among amateur traders, with losers dropping out. This turnover is inconsistent with the idea of
markets as a gambling casino in which the customers pay to play but consistent with the idea of markets as a school to which traders pay tuition to discover whether they have talent.

In addition to information about the profits or losses of each group of traders, it would be interesting to know what specific circumstances or events are profitable or unprofitable for each. There are frequently great differences in the size, liquidity, and volume and cost of trading among the markets in which hedgers and spreaders operate. If events that directly affect a limited market (for example, an increasing popularity of Italian food, which increases the demand for durum wheat) and events that directly affect many markets (for example, a change in monetary policy) are distinguished, it is reasonable to assume that events of general importance affect active, liquid markets more quickly than inactive, less liquid markets. Unexpected general news may create profitable opportunities for professional spreaders since they can take advantage of lags in price changes in the less active markets. Hedgers and some amateur traders may be in a position to identify specific events before they become generally recognized. One might therefore expect changes in specific circumstances to create profitable trading opportunities for these groups.

Conclusions. The market for an actively traded futures contract is an extremely complex environment subject to rapid change. No one trader can absorb and process all the information "publicly available" in such a market. Models assuming that all participants play essentially the same role with at most differences in the kinds of information they have available miss important features of the marketplace environment. Traditional economic models assume that competitive markets are efficient in the sense that prices reflect all publicly available information. In futures markets professional technical traders are apparently able to trade profitably on the basis of widely available information about the state of the market.

I have chosen an alternative approach, which corresponds more closely to how traders think about themselves. It is also similar to the approach used in many sociological and biological models. Traders accommodate themselves to the environmental complexity by adopting a style of trading incorporating one or more specialized roles. In each role they seek to profit by adapting their trading and information-acquiring behavior so as to take advantage of certain aspects of the market environment. Each role carries the risk of losing money to other traders when changes take place in aspects of the environment to which it is less well adapted.
Each role can be thought of as performing a function in a complex and continually changing environment, analogous to the function of specialized cells in an organic structure or species in an ecological context. Other things being equal, the more traders performing similar roles, the lower the average profits of each. The profits traders can obtain from performing a particular role depend on their relative abilities. The number of traders performing the role depends on the aggregate demand for the functions it performs.

Although there are useful similarities between the biological models and the economic model sketched here, there are also important differences. Successful species reproduce until a balance is reached between their numbers and what the environment can support. Traders can switch from one role to another as long as the expected gross profits from the new role exceed those from the old. They will remain members as long as the expected gross profits from their most desirable role exceed the opportunity cost of owning a seat. In an economic model adjustments to changed circumstances may be faster than in the biological model. The possibility of overadjusting in some circumstances exists in both models.

Trading on the New York Stock Exchange Floor

The purpose of this section is to compare the trading process on a stock exchange with the process on futures exchanges and thus to set the stage for the description and analysis of the regulation of floor trading by the CFTC and the SEC. The New York Stock Exchange (NYSE) provides a marketplace for stocks and bonds issued by certain corporations. Most of the transactions involve common stock. The shares of about 1,500 companies are listed and traded on the NYSE, which is the primary market for most of them.33

The stock exchange trading post corresponds to a futures exchange trading pit, but there are marked differences. First, the rate of trading is relatively slow. In 1983, when the volume was very high by historical standards, the average stock had one transaction every 9.19 minutes. Since the average of common stocks per specialist is less than four, an average specialist may observe about twenty-five transactions per hour at his trading location.34 Second, not many members are at any one post. At the end of 1981, approximately 406 members of the NYSE were specialists, and 839 were floor brokers—about two floor brokers for each specialist. Although specialists tend to be at a trading post a large part of the time they are on the floor, floor brokers spend a substantial part of their floor time traveling to and
from trading posts and telephone booths. Sizable crowds gather at a few posts where the most active stocks are traded, but most of the time at most trading posts there are only the stock exchange specialist assigned to that post and from zero to three members waiting to trade.\textsuperscript{35}

As on all American exchanges, trading on the floor of the NYSE can take place only between members. At the end of 1982 the NYSE had 1,462 members.\textsuperscript{36} Members who work on the floor can be divided into three groups: floor brokers, specialists, and floor traders.

Floor brokers form the largest group. At the end of 1981 commission house brokers—floor brokers employed by and working exclusively for particular brokerage firms—numbered 639. There were 200 independent floor brokers, who work on a commission basis for any member firm that has orders to be filled.\textsuperscript{37}

Specialists constitute the next largest group of members. In October 1983 there were 412 specialists on the NYSE, grouped into fifty-nine specialist units with a median of twenty-one common stocks assigned per unit.\textsuperscript{38} A specialist unit is a partnership or similar organization. It may perform several other functions, including making markets over the counter or providing investment advisory services; but some units rely mainly on the specialist activity for their income. Each specialist unit has certain responsibilities and privileges in connection with the markets for the stocks assigned to it. These are described in detail later.

The smallest category of members on the floor consists of floor traders, officially known as registered competitive market makers (RCMMs). RCMMs and specialists are permitted to buy or sell stocks for their own accounts on the floor of the exchange.

Each security is assigned to a particular trading post, at which all trading on the exchange in that security must take place. A broker who has an order for a particular stock can go to the location at which it is traded and be reasonably sure that any member with an interest in buying or selling that stock will be represented there.

Securities are assigned to locations in such a way that the expected volume of trading is similar at all locations.\textsuperscript{39} Typically, five to eight stocks are assigned to one trading location; if a stock is very active, however, it may have a trading location to itself. The trading volume of some active stocks is higher than the combined volume of any post with a collection of stocks.

Four methods of trading are used on the NYSE: a mechanical limit order system (the cabinet system, used exclusively for bonds); a call market (used for openings); prearranged negotiated trades (used for large blocks); and the continuous two-sided auction (the
normal and still dominant form). In 1979 it was estimated that 21 percent of the share volume of transactions was arranged through block trades, 7 percent took place at the opening, and the remaining 72 percent took place through the auction market. In 1983 the share of block transactions had increased to 46 percent.

Specialists. Historical background. Initially the specialist was a floor broker’s broker. In time he acquired the role of dealer, and he now performs both functions, which are about equally important. Understanding how this evolution took place provides some insights into the trading process.

Orders from nonmember customers must be executed by members acting as brokers. Most floor broker members handle orders in any stock. They must therefore frequently move around the trading floor. The specialist role was created when some floor brokers decided to remain at specific trading locations. Although they could handle orders only for the stocks traded at the specific locations, they could execute orders in those stocks whenever conditions allowed.

Suppose that a floor broker handling orders for many stocks had a limit order for XYZ stock that could not be executed immediately because the limit price was away from the market. If he left the XYZ trading location, the price might move through the limit price when he was not there to execute the order. If he waited at the XYZ trading location, he would forgo the opportunity to earn commissions by executing orders in other stocks. In these circumstances he might give the order to a specialist broker. He would do so when sharing his commission with the specialist was less expensive than waiting to execute the order himself and forgoing other business.

Active trading posts often had several competing specialists. A specialist could attract business from competing specialists by offering immediate execution on orders close to the market by taking the other side himself. This gave the floor broker’s customer faster execution at a good price and made the broker look good to his customer. Presumably the broker accommodated in this way would reward the specialist by giving him a larger share of his future brokerage business. Thus the specialist, who began as a broker’s broker, added the dealer’s function as a sideline to attract additional brokerage business.

In the 1920s some specialists operated pools. The typical stock pool acquired a large supply of stock and attempted to distribute it to the public at a higher price. The pool might obtain new stock or
options to buy new stock from the company, from an estate or other large holder, or in the open market. Specialists did not necessarily form the pool, which may have required considerable capital; but being on the floor at the stock’s trading location all the time, they were logical persons to run it from day to day.

To dispose of stock profitably, the pool had to increase the demand for the stock. One way to do so was to spread favorable information about the company. Of course, information is valuable only if it is not generally known. To help make the information believable, a pool might increase activity in the stock, using wash sales and similar techniques to make it appear that insiders were aware of the information and were already buying the stock.

Disseminating information that is known to be false as a part of a pool amounts to fraud, but proving fraud is difficult and expensive. A person who had bought stock while a pool was in operation would be very unlikely to try to recover his losses by suing the pool members. It would be easier to accept the losses as an educational experience. The existence of pools was not a well-kept secret; undoubtedly many persons bought stock knowing that a pool was in operation. They might be hoping to be free riders, or they might believe that the pool operators were truly insiders with access to special information, which was not always wrong. The creation of the SEC was in part a reaction to the allegations of fraud and manipulation by pools.

In the past seventy-five years, three major trends have affected the specialist business. First, the dealer function has increased in relative importance. In recent years NYSE specialists have earned approximately equal amounts from their brokerage and dealership activities.42 Second, the number of competing specialists has declined. The 228 stocks with competing specialist units in 1957 decreased to thirty-seven by 1963 and to zero shortly thereafter.43 Efforts to revive direct competition among specialists in the 1970s were not successful. Third, the regulation of specialists has increased.

Regulation of specialists. Regulation of specialists increased significantly in the 1930s through passage of the Securities Exchange Act of 1934, which directed the new Securities and Exchange Commission “to regulate or prevent floor trading by members of national securities exchanges” and to study whether the brokerage and dealer functions of the specialist should be segregated.44 In his history of the NYSE, Sobel writes that

the two central issues in 1935 were a proposal to segregate broker and dealer functions and another to curb specialists

© 1985 by the American Enterprise Institute for Public Policy Research, Washington, D.C. All rights reserved.
and floor traders. The former was an "inside" matter in that it was not recognized as a problem by most Americans, while the latter was discussed not only by financial columnists but on the front pages of the New York press.45

After studying the segregation issue, the SEC concluded that the specialist in acting both as broker and dealer has an inherent conflict of interest. From the present evidence it is not possible to conclude whether that conflict can be eliminated in such a way as to enable him to function more adequately in the public interest. . . . pending acquisition of further knowledge, emphasis should be placed on:

(1) Insistence upon the observance of rules against unjustified trading by the specialist for his own account. Trading for his own account should meet an affirmative proof of justification. . . .

(2) The development of appropriate restrictions governing the conditions under which the specialist may trade with his book.46

A commission appointed by Governor Hughes of New York, as a result of the panic of 1907, to investigate speculation on the New York Stock Exchange considered essentially the same issue and reached essentially the same conclusions. The Hughes commission report states:

We have received complaints that specialists on the floor of the Exchange, dealing in inactive securities, sometimes buy or sell for their own account while acting as brokers. Such acts, without the principals' consent are illegal. . . . Notwithstanding that the system of dealing in specialties is subject to abuses, we are not convinced that the English method of distinguishing between brokers and jobbers serves any better purpose.47

After the Hughes report, the New York Stock Exchange adopted rules prohibiting members acting as brokers from buying or selling stock as principals to complete the brokerage orders they had accepted from their customers. It did not prohibit a member from dealing as a principal with one party while holding a brokerage order from another.48

As a result of its comprehensive special study, the SEC became concerned that the specialists might not be dealing enough. An attempt was therefore made to require them to undertake the dealings necessary to contribute to a fair and orderly market. In the Institutional Investor Study Report of 1971, it was again claimed that some specialists
SEYMOUR SMIDT

were not dealing enough to perform their market-stabilizing function. The NYSE subsequently tried to get more comprehensive information about specialists and to use the stock allocation process to reward specialists who performed well. Some unsuccessful efforts were made to encourage direct competition among specialists. The flow of orders at the typical trading location on the NYSE floor is not high enough to support two dealers. Competition would require consolidating trading locations or providing effective access to the flow of orders to persons not at the trading location.

The tick test is used to determine if the specialist's principal transaction is justified. The "tick" is a price change, positive for increases and negative for decreases. If a transaction is at the same price as the previous transaction (a zero tick), the price change is measured from the last different price. Specialists are allowed to buy on negative ticks (or zero negative) and to sell on plus ticks (or zero plus). Negative ticks occur on sell orders; if the specialist is buying, he is preventing the price from going lower. Hence he is stabilizing, as he is when selling on upticks. The specialist is allowed to violate the tick test when liquidating positions to limit losses.

Several comments can be made about the tick test. First, it is operational. Records are kept, and the test can be applied to every transaction to see whether it complies. It is the only specialist performance criterion with this clearly operational characteristic. Second, a scalper who followed the simple rules described in the previous section would also comply with the tick test; the test is thus consistent with at least some profitable trading strategies. Third, the test prohibits the specialist from taking the initiative to cause price changes in situations in which there would otherwise be no transactions. Fourth, the test does not, as is commonly believed, require that the specialist's inventory changes be "stabilizing." As long as there are both buy and sell orders, the specialist can increase his inventory while the price is rising and decrease it or increase a short position while the price is falling and still satisfy the tick test on every transaction in which he participates. The trick is to be selective. Suppose you wish to increase your inventory: be the most competitive buyer so that you take the other side on all sell orders, but be a noncompetitive seller, so that the book or other brokers take the other side of all buy orders. By only buying you are assured of increasing your inventory. If there are enough buy orders, the prices will rise.

The specialist is one kind of market maker. The economics of market making are discussed in detail in the section "Economic Analysis of Trading on Exchange Floors."
**Floor Trading.** Floor trading refers to purchases and sales initiated on an exchange floor by members for their own account or the account of a firm with which they are affiliated. Transactions by stock exchange specialists for a market-making account are excluded. Floor trading may never have been as important on the NYSE as it is on futures exchanges. Nevertheless, floor trading on stock exchanges is worth considering for at least two reasons. First, it has a long history of controversy, giving rise to a substantial number of studies. Second, the SEC's regulation of floor traders and its rationale will become very important in any move to unify the regulation of exchange trading practices on futures exchanges and stock exchanges.

**Historical background.** In 1917 a knowledgeable observer estimated that an average of 500 to 600 members of the stock exchange were in attendance on the floor, of whom 50 to 100 were floor traders, buying and selling for their own account. He wrote:

> There are between 50 and 100 of these Room Traders who enjoy the privilege of being all the time on the floor of the Board-Room, and thus able to take advantage at once of every opening. They know the prices even before they are recorded on the tape, and they are able to join in every upward movement the moment it begins, and to abandon it the moment it shows any signs of wavering. They are in and out of the market perhaps a dozen times a day. They constitute an important element in it.

Between 1935 and 1961 the number of members of the NYSE engaged primarily in floor trading ranged from thirty to forty-eight. Some member firms also had active floor trading accounts. The special study staff estimated that as many as 300 members may have participated in floor trading to some extent. Floor trading decreased from 6.8 percent of total volume in 1937 to 2.1 percent in 1961. In 1981 registered competitive market makers, the modern successors of the old floor traders, participated in 1.57 percent of total volume.

The SEC found that floor trading on the NYSE is concentrated in actively traded stocks. One reason for the limited floor trading may be the small number of stocks that have sufficient volume to accommodate that style of trading. A second reason, although the evidence is not as clear, may be that institutional trading, which is a growing proportion of the total, is not as conducive to floor trading as trading by individuals. A third reason for the decline, especially since 1961, may be the regulations imposed by the SEC. I discuss these after I consider the studies that formed their basis and justification.
SEC studies. In the special study in the early 1960s the SEC undertook an extensive look at floor trading on the NYSE and to a lesser extent on the American Stock Exchange (AMEX). It also reviewed previous studies it had undertaken.  

The special study considered whether floor traders contribute to price stability. It measured the net purchases or sales of floor traders during a day in relation to the change in prices during that day and found that

floor traders tend to have a destabilizing influence on prices. On at least 15 separate occasions since 1934, studies conducted by the Commission and the Division of Trading and Exchanges, confirmed by studies made by the Special Study, have shown that floor traders are generally buyers in rising markets and sellers in declining markets, with respect to both the market as a whole and to individual stocks. Their trading, as a result, is inimical to the orderly functioning of the market, tending to accentuate rather than to stabilize price movements.

Regulation. The attitude underlying the regulation of floor trading is best illustrated by quoting from the special study's summary and conclusions on this topic.

Of all classes of exchange members on the floor, the floor trader stands alone in having no fiduciary status, no duty to execute transactions, and no market responsibilities or obligations in relation to the operation of the market as a public institution. . . . That any individual who purchases a seat thereby becomes entitled to do his personal trading on the floor of an exchange . . . raises . . . a fundamental question . . . as to the extent to which a public market may be permitted to maintain this vestigial "private club" aspect, even apart from very serious questions as to the net impact of floor trading on the orderly functioning of the market.

The report continues by describing the perceived unfair advantages of the floor trader over nonmember investors. These advantages are not documented, but they offer the study's explanation of why floor trading is successful.

The privilege of access to the floor provides trading advantages of a substantial nature; the commission cost of trading . . . is appreciably lower than for off-floor trading, trading activity may be observed minutes before it appears on the tape, and bids or offers may be entered or withdrawn in a
matter of seconds. . . . Presence on the floor carries with it the benefit of . . . a "feel of the market"—a heightened sense of market tenor and trend. In part . . . attributable to the constant exchange of observations among floor members. . . . More subtle factors also add significantly to the floor members' awareness; familiarity with the trading techniques of specialists or floor brokers . . . combined with knowledge that a large block of stock is being accumulated or distributed . . . facilitates the trading activities of the floor trader.56

The special study made two major recommendations regarding floor trading.

1. Floor trading in its present form is a vestige of the former "private club" character of stock exchanges and should not be permitted to continue on the NYSE or Amex. . . .
2. . . . the financial capacity of some specialists or of the specialist system generally is in need of strengthening, and it is possible that some present floor traders could perform a highly useful function as "auxiliary specialists."57

New York Stock Exchange Rule 107, which was made permanent by the SEC in February 1981, essentially implements these recommendations for the NYSE.

Several comments may be made concerning the SEC studies. First, the inventory behavior of stock exchange floor traders in relation to price trends is distinctly different from that of specialists or over-the-counter market makers and from what one would expect to find for scalpers.58 This supports the idea that stock exchange floor traders perform a different role in the price discovery process from that of specialists. The behavior of stock exchange floor traders is similar to the behavior attributed to position traders on the floor of the futures exchanges.

Second, the special study's use of the term "destabilizing" to describe the price effects of stock exchange floor trading implies that in the absence of this trading the price might not have moved and that with the trading a temporary movement occurs that is subsequently reversed. Nothing in the systematic data in the various SEC studies supports either of these implications of the word "destabilizing." The evidence is also consistent with the hypothesis that floor trading accelerates price movements that would otherwise have taken place more slowly.

Third, the SEC statement about the trading advantages of nonspecialist members in initiating transactions from the floor appears to apply equally to futures markets.

Fourth, neither the SEC nor Pratt appears to give any serious
thought to explaining the number of nonspecialist floor traders. All the SEC studies were done at a time when a member of the NYSE was free to choose floor trading instead of floor brokerage as a vocation. (He was not free, during most of this period, to choose to be a specialist.) The advantages of being a floor broker seem to have outweighed the attractions of being a floor trader for most members. Apparently, floor trading has some disadvantages and difficulties not adequately explained by the SEC studies.

Fifth, the SEC studies shed little light on the behavioral patterns of other traders that make profitable floor trading possible. Except that floor traders trade in active stocks and appear to trade with the trend, we know very little about how they operate and even less about the trading behavior of those with whom they interact. If floor trading is profitable, it must rely on some regular patterns of behavior. An economic analysis of the price effects of trading that considers the state of the market is presented in the section “Economic Analysis of Trading on Exchange Floors.”

Block Trading. A block trade is a single transaction consisting of 10,000 shares or more. The NYSE began to keep records on block trading in 1965, when an average of nine block trades took place per trading day, amounting to 3 percent of the share volume. By 1983 more than 1,400 block trades per day accounted for 46 percent of the share volume.59

It is not just size that differentiates block trades from other transactions. Block trades are arranged in advance, usually off the floor of the exchange, in a complex process presided over by a broker who specializes in handling them. After the buyers and sellers have agreed on the essential terms of the transaction, the process is completed and announced to the world on the floor of the exchange. If the marriage ceremony takes place on the exchange floor, the courtship that preceded it took place upstairs, in a telephone market.

For an auction market to work well, the orders flowing to it must be such that transactions are not too infrequent, not too easily predictable from previous transactions, and not too different from one another in size. These conditions are likely to occur if a sufficient number of stockholders each hold a relatively small proportion of the stock and make independent decisions. Exchange listing requirements are designed in part to identify stocks in which an auction market is likely to work well. Because stockholders who plan to buy or sell only a small proportion of the normal daily volume in a stock can reasonably believe that their decisions will have little or no influence on the market price, they will not be afraid to expose their
interest in buying or selling by transmitting an order to the exchange floor.

By contrast, an institutional investor may wish to buy or sell an amount of stock equal to or many times larger than the normal daily auction market volume. Such large position changes can be accomplished through the auction market only by conducting many small transactions spread over several days or weeks. Even then, making the position change through the auction market is likely to cause at least a temporary change in the market price, because it changes the balance of buy and sell orders. If other traders recognize that a large position change is under way, they may suspect that the desire to trade is motivated by information and may search for that information. Whether they discover it or merely change their estimate of the value of the stock on the assumption that the information must exist, the attempt to trade is likely to lead to a change in the equilibrium price of the stock.

In summary, significant disadvantages are incurred by institutions in using the auction market to arrange large position changes. For this purpose the auction market is time consuming, entails high transaction costs in processing many small transactions and in monitoring the floor traders carrying them out, and is likely to lead to unfavorable temporary price changes because the balance of orders in the auction market is changed. If the position change was motivated by special information, the information may be discovered by others, and the market price may adjust before the position change can be completed. In spite of these disadvantages the auction market was frequently the best game in town until block trading developed.

Over the long term the proportion of stock held by institutional investors has increased, and in the mid-1960s the turnover rate of institutional holdings increased substantially. Sometimes while one institution was surreptitiously disposing of a position in a stock, another was secretly acquiring a position in the same stock. For every institution changing positions, another one or two might be considering making such changes if a favorable trading opportunity developed. Block trading becomes possible when an institution wishing to initiate a large position change has a reasonable chance of being able to find institutions willing to take the other side of the trade.

The key element in arranging a block trade is the broker. Brokers who make a specialty of this activity are known as block traders. An institution that wishes to initiate a large position change discloses its intentions to a block trader. The trader searches for one or more institutions to take the other side of the transaction and acts as an intermediary in getting the sides to agree on the terms of the trade.
(mainly price and quantity). If an agreement is achieved, the trade is executed on the floor of the exchange. Usually one broker represents both the buyer and the seller, and the trade is considered a cross. Special procedures are involved in executing crosses. In principle, a third party could break up a cross if the transaction were conducted at an unreasonable price.

The main practical difficulty in arranging block trades is that institutions are reluctant to disclose their trading plans even to a broker, let alone to permit the broker to advertise them to other institutions. Part of the skill of the block trader is in overcoming this hesitation. Institutions are more likely to disclose their trading plans to a broker who has a reputation as a successful arranger of block trades. Block traders work hard to enhance their reputations, by such means as arranging very large blocks that are sure to be well publicized. One technique a trader can use is to guarantee the institution a minimum price by offering to buy the block himself, if necessary. If the block trader can arrange a block at a better price, the institution gets the benefit (and the block trader gets commissions on both sides of the transaction). A block trader who cannot find institutions to take the other side and who buys the block from the institution and resells it in the auction market can expect to lose money on stock acquired in this way. Successful block trading firms make enough in commissions to offset their trading losses.60

The block trader relies on the auction market in three very important respects: pricing, position disposal, and legitimization. First, it would be much harder to get agreement on the price of a block trade without relatively frequent auction market transactions to provide current prices that can be used as a basis for negotiation. Second, the block trader would be unable to guarantee prices or to take up the difference in quantities between the amount offered and the amount sold without an auction market in which the excess quantities could be resold.61

Third, executing a block transaction on the floor of the NYSE serves to legitimize the transaction. Why is there a need for legitimization, and how is that need satisfied by bringing a prearranged transaction to the floor? Legitimacy would not be a problem if block transactions were typically arranged between individuals using their own money or at prices very close to the last price in the auction market, but neither of these conditions is typically met in block trades. The participants are acting as agents for institutions or as employees of their firms, and most blocks are executed at a price different from the last price in the auction market. What assurance can be provided to a skeptical outsider that the price is reasonable?62
A degree of assurance is provided by taking the trade to the floor of the exchange and exposing it so that others who think the price is attractive have an opportunity to buy. If, as is usual, the block price is below the last auction market price (minus tick blocks), some limit orders may be on the specialist's book at prices between the last auction market price and the block price. The specialist may also want to acquire some stock at the block price for his own account. Usually the specialist and limit orders are accommodated. This process helps ensure that the pricing processes on the auction market and the block market are integrated and helps provide legitimation. Although the system is certainly not foolproof, it is useful. Transactions conducted on the third market or on regional exchanges do not convey the same degree of legitimacy, although there is no evidence that those markets have been used for improper transactions.

Price changes accompanying block trades. The price changes associated with block trades fall into two categories: changes that reflect new information and changes that reflect costs of providing liquidity. The empirical evidence indicates that both types are associated with block trades. On the average prices change in the expected direction (increase for blocks initiated by buyers and decrease for blocks initiated by sellers) for several days before as well as immediately before the block transaction. For minus tick blocks only, a partial rebound occurs in the few minutes following the block. The rebound suggests that persons who buy minus tick blocks are given some slight compensation for supplying liquidity to the seller. Within at most a few minutes after the block, prices adjust to a level at which they are on the average unbiased estimates of the market equilibrium price.

Economic Analysis of Trading and Exchanges

In this section I move from description to economic analysis. A definition of market makers is followed by a description of theories about market making, the two main theories being the bid-ask spread theory and the inventory management theory. I then summarize some important empirical studies of market making, focusing on two questions: Is market making a natural monopoly? and Are market makers better informed than other traders? Finally I summarize what is known about the price effects of floor trading and speculation, the economics of exchanges, and describe differences between cash markets and futures markets.
Market Makers Defined. A dealer makes a business of buying and selling. Investors buy securities with the intention of holding them; dealers buy securities with the intention of selling them.

A market maker is a special kind of dealer who helps customers complete transactions by being willing to take the other side. Customers usually take the initiative in transactions with dealers. Block traders do not usually call up mutual funds and say, "I'd like to buy 20,000 shares of your Atlas stock." They ask, "What do you want to sell today?" If the customer wants to sell, the dealer buys; if the customer wants to buy, the dealer sells.

Market makers almost never initiate transactions that increase their long or short positions. Sometimes they initiate transactions to reduce their existing positions. Such transactions are usually loss reducing or loss limiting.

The role of the market maker depends on the market. The primary interest in this report is on futures markets. American futures markets are continuous two-sided auction markets. Sellers compete to get the highest selling price; buyers compete to pay the lowest buying price. The stock market has essentially the same structure. I consider other market structures only as they are directly relevant to the operations of a futures market.

Theoretical Analyses. Having defined market makers as a special kind of dealer, I now present a more comprehensive picture of what they do and why they do it. Three distinct theories about market makers are the legal theory, the bid-ask spread theory, and the inventory management theory. The legal theory visualizes market makers in the context of a set of legal rights and obligations. The bid-ask spread theory sees the business of the market maker as providing liquidity to traders and the major problem as what price to charge for the service. The inventory management theory also sees market making as a business, one whose major problem is inventory management. The theories are not mutually exclusive. Each could be true, although they emphasize different factors.

Legal theory. The legal theory sees the market maker as a participant in a complex social structure with reciprocal rights and obligations vis-à-vis other participants. The issue to which the theory is addressed is why there are or should be differences in those rights and obligations. The theory sees the market maker as a volunteer who has agreed to accept some onerous obligations to stabilize the market; and as a reward or incentive for accepting those obligations, he is given some special privileges. Some persons would prefer to
reserve the term "market maker" for those who have a legal obligation to provide market-making services, but this seems unnecessarily restrictive; evidence from more than one market shows that dealers regularly provide market-making services even when they have no legal obligation to do so.

Bid-ask theory. Harold Demsetz, in a path-breaking article, said that "the ask-bid spread is the markup that is paid for predictable immediacy of exchange in organized markets." A "prospective purchaser either submits a market order, in which case he concludes an exchange at the lowest limit ask price, or, if he is prepared to stand ready and waiting, he submits a limit bid order stipulating an upper limit to the purchase price he is willing to pay." Thus the more anxious a purchaser is to buy, the higher is the bid. The market bid reflects the highest bid that has not yet been satisfied. Demsetz hypothesized that the cost of waiting was closely and inversely related to the rate of transactions and that it would therefore exhibit economies of scale. The empirical evidence on this point is discussed below. The bid-ask spread in this view reflects both the value to the trader of not waiting and the cost to the dealer of taking a position and waiting until it can be turned over.

One of Demsetz's major contributions was to identify a service, "immediacy," that is bought and sold along with financial instruments and to define the price for that service (the bid-ask spread) and the persons who demand and supply it. (The demanders are traders who submit market orders; the suppliers are market makers and other traders who submit limit orders.) By using the standard apparatus of economists—prices and quantities, demand curves and supply curves—Demsetz introduced market making into the mainstream of economic thinking. He also recognized there were costs of supplying immediacy both to investors who submitted limit orders and to the market maker. He believed that competition among alternative suppliers of immediacy on the NYSE would reduce the price of immediacy to its cost of production; he left the detailed explanation of the nature of that cost to later writers. Another important contribution was providing support for his ideas through statistical analyses relating the size of the spread to the frequency of transactions and other factors. Demsetz's ideas are presented in a static framework, with nothing to indicate that the behavior of any participant might reflect some transient state of the market or of the participant's knowledge compared with that of other market participants.

Jack Treynor, writing under the pseudonym Walter Bagehot, identified information asymmetries as one of the important cost factors
in determining the bid-ask spread. According to Bagehot, informed traders would use market orders to buy stock before its price rose or sell stock before its price declined. As a result, the market maker would regularly suffer losses in the value of his inventory, losses that must be recovered if he was to stay in business. The bid-ask spread, according to Bagehot, must be large enough to recover those expected inventory losses.

Although inventory losses can be considered a cost of producing immediacy, they differ from other costs, such as telephone bills, in that the revenue is received by the market maker’s customers. To measure the cost of market making to the customers as a group, we must subtract the average rebate (inventory loss) from the gross spread, to arrive at a net spread.

Bagehot emphasized that the market maker could not recover these costs solely from the well-informed traders who caused them, because he could not identify which traders were well informed. The costs were charged to everyone, and the net costs were borne only by the ill-informed. Bagehot also argued that ill-informed traders could escape paying the costs by using limit orders in place of market orders. Thus well-informed traders imposed inventory losses on market makers that were recouped from poorly informed traders who mistakenly believed that they were well informed.

The importance of depth as part of the service provided by market makers was emphasized in an article by Smidt. A spread could be small and good for only 100 shares, or it could be small and good for several thousand shares. If the depth is not specified when a market maker gives quotations, it is frequently requested by the potential customer.

Although individual investors buy in smaller quantities than institutions, they are more susceptible to harm from a lack of depth if their orders bunch on one side of the market, because an individual investor is less aware of what other investors are doing. By contrast, an institution dissatisfied with the price it could obtain from the market maker in a large transaction might choose not to transact or try to arrange for a block trade or spread the transaction over a longer period. Although Smidt was primarily concerned with the stock market, with its unitary specialist system, depth is also a concern in markets with multiple market makers. Ho and Stoll in particular have pointed out that comparisons between markets with a single market maker and markets with multiple market makers may not be valid if different degrees of depth are involved. This could apply, of course, even to markets with a single market maker. In principle, the quotations published by one market maker could be narrower.
than those of another but valid for a greater depth. There is no objective standard for determining which market is better. (In practice, this may not be a problem: there is no evidence that specialists who offer less depth compensate by offering narrower quotations.)

The work described above is representative of the main static analyses that have been developed from Demsetz’s path-breaking study. Related work includes empirical studies of bid-ask spreads and theoretical studies that present the behavior of market makers and their customers in a dynamic context and as the outcome of some maximizing behavior.

The bid-ask spread theories characterize the market maker as a financial intermediary producing a service: predictable immediacy of exchange. The spread is the cost of this service. These theories have stimulated empirical studies that have added substantially to our understanding of market making.

Nevertheless, some cautions should be kept in mind when interpreting a spread as the price of immediacy. First, the bid-ask spread may contain an element of monopoly rent, in which case the magnitude of the spread may overestimate the cost of producing the service of immediacy. Second, the amount of depth included in the immediacy service may vary from one market to another or from one time to another. This is particularly a problem in markets with multiple dealers, since the combined depth of all dealers may be the appropriate measure. Third, the relevant spread, from the point of view of nondealer traders in such markets, is the market spread, which may be smaller than the spread of any individual market maker. Finally, the quoted spread may be quite different from the realized spread because of inventory losses incurred by the market maker. Since these losses are matched by trading gains by at least some of the market maker’s customers, the quoted spread is not an accurate measure of the average price paid by the purchasers of immediacy.

Inventory management theory. In a continuous market transactions take place one at a time. The static models previously described abstract from this lumpiness by using averages: the average price or the average bid-ask spread or the average rate of transactions over some period of time. Some important features of real marketplaces, however, cannot be understood without considering the discrete character of trading.

To illustrate the erratic price changes that would occur in the absence of market makers, Stigler simulated a continuous market in which orders arrived randomly and transactions took place only between investors. He recognized that specialists reduce the extent
of these apparently random fluctuations and that their task would be substantially easier if the equilibrium price was not subject to unforeseen changes. Stigler also warned that even if the equilibrium price was fixed and known, long runs of buy or sell orders might on occasion exhaust the resources of the specialist.

Garman demonstrated that even if prices are set so that expected demand is equal to expected supply, the random arrival of buy and sell orders is sufficient to guarantee that the market maker will fail in a fairly brief time because he has reached the limits of his capital. This result "makes clear that the specialists must pursue a policy of relating their prices to their inventories in order to avoid failure. It cannot be the case that they simply respond to temporary fluctuations in demand and supply as some stock exchange propaganda would suggest." The conclusion that a specialist who kept prices at the equilibrium level might fail had been anticipated by Stigler, but before Garman no one seems to have realized how frequent such failures would be even under the most favorable assumptions of a known and unchanging equilibrium. The absence of frequent failures among specialists or other market makers reinforces Garman's conclusion that market makers must move their quotations away from the equilibrium level to control their inventory positions.

Garman's paper did not present a model of how the specialist would respond to inventory fluctuations, but it challenged other economists to create such models. A number of papers have been written since Garman explaining how specialists would adjust their bid and ask prices as their inventories changed.

Two concepts are essential to understanding the relation between the bid and ask prices that a monopolistic market maker quotes and his inventory: target inventory and equilibrium price. The equilibrium price is the price at which the long-run demand and supply for the item are equal. (In any period the actual demand will differ from the actual supply; but at the equilibrium price the ratio of actual demand to actual supply will almost certainly be very close to one if the period is long enough.)

If the market maker holds his target inventory, his only objective in setting the bid and ask prices is to maximize the expected profit from the jobber's turn, buying at the bid and selling at the ask. If both the bid and the ask prices equaled the equilibrium price, the market maker would earn no jobber's turn. He sets the bid price a little lower than the equilibrium price and the ask price a little higher so that the expected profit is as high as it can be while the expected inventory change is zero (that is, inventory increases and decreases are equally likely). If a market maker's actual inventory is above or

87
below the target inventory by a sufficiently large amount, he is
willing to forgo some profit to increase the chance that the inventory
will move closer to the target. This is because some costs or disutilities
are associated with holding an inventory that deviates from the target.
These costs may be due to a variety of factors, including the risk of
loss from an unfavorable change in the equilibrium price, increased
carrying costs for the inventory, and limits on the capital available
to the market maker.

It is worth considering in detail how a market maker might react
to an inventory that deviates from the target level. If the market
maker's inventory is above the target, he is willing to give up some
expected profits from his jobber's turn to reduce it. He does so by
reducing the bid and ask prices so that expected demand is greater
than expected supply and the expected inventory change is therefore
negative. If his inventory is only slightly above the target, the benefit
gained by moving closer to the target is relatively small. He is there­
fore not willing to give up much of his expected jobber's profit to
move the inventory closer to the target. In this case the bid and ask
prices will be only slightly below the equilibrium level; expected
demand will be only slightly below expected supply, and the expected
rate of inventory change will be only slightly negative. If the market
maker's inventory is very much greater than the target, the benefits
gained by moving closer to the target are correspondingly larger.
The dealer will then be willing to give up more of his expected
jobber's profit to move the inventory closer to the target. In this case
the bid and ask prices will be substantially below the equilibrium
level, and the expected rate of decline in inventory will be corre­
spondingly larger.

If the market maker's inventory is below the desired level, he
will attempt to increase his quotations above the equilibrium level
so that the inventory is more likely to increase than decrease. Again,
the larger the inventory deviation below the target (including short
positions), the higher the quotations above the equilibrium level.

The general shape of the market maker's inventory demand
function is illustrated in figure 2–1. The bid and ask quotations are
denoted by \( bb \) and \( aa \) respectively. When inventory is at its target
level, the quotations bracket the equilibrium price.

*Price dynamics.* To understand how market making affects price
levels and price changes, consider a simple equation. Let \( P_t \) represent
the price at which a transaction takes place. Three influences can
affect that price. One is the equilibrium price, \( P_e \). In a perfect market
\( P_t = P_e \) for every transaction. If a market maker's inventory is not
at its target level, there may be a deviation between the midpoint of the bid and ask quotations and the equilibrium price. Let $D$ represent the amount of that deviation. $D$ is not a constant but depends on the market maker's inventory and the shape of the inventory demand functions. If transactions took place at the midpoint of the bid and ask quotations, the transaction price would be $P_t = P_e + D$. Most transactions, however, take place at either the bid or the ask. Let $S$ represent either plus or minus one-half of the bid-ask spread; $S$ will be positive if the transaction is at the ask or negative if the transaction is at the bid. Thus

$$P_t = P_e + D + S \quad (1)$$

For example, if $P_e = $55 per share, $D = -$2 per share, and $S = \pm $0.125, $P_t$ would equal $53.125$ for purchases by customers and $52.875$ for sales. The expected value of $P_t$ for the next transaction would be a little more than $53$ (because $D$ is negative, the quoted prices are centered below the equilibrium price, and purchases are more likely than sales).

Under reasonable assumptions the expected value of $D$ and $S$
will be zero or very close to it for almost any imaginable market, so that the transaction price will be an unbiased estimate of the equilibrium price. This is true from the point of view of someone who does not know the state of the market maker's inventory. Someone who did know the value of the inventory would consider the price a biased estimate of the true equilibrium, the direction of the bias depending on whether the market maker's inventory was above or below its target level. Furthermore, a trader knowing the market maker's inventory would expect the transaction prices to exhibit a trend in relation to the equilibrium price that would reduce the size of the bias. In equation 1 the transaction price, $P_t$, is the sum of three elements. The $S$ component produces oscillations between the bid and the ask; the $D$ component produces short-term trends, sometimes positive and sometimes negative; the $P_e$ term changes from time to time in relatively unpredictable ways because of unexpected events affecting fundamental values.

To evaluate the accuracy of $P_t$ as an estimate of $P_e$, we can look at the variance of the difference between the transaction price and the true equilibrium price:

$$\text{Var}(P_t - P_e) = \text{Var}(D) + \text{Var}(S) + 2\text{Cov}(D,S)$$

where $\text{Var}(.)$ and $\text{Cov}(.)$ are the variance and covariance functions respectively. Consider briefly the main influences on the accuracy of the transaction price as an estimate of the equilibrium price. The variance of $S$ depends mainly on the size of the bid-ask spread (it should be very close to $S$-squared). The variance of $D$ depends on the characteristics of trading in the stock and on the shape of the inventory demand function. Other things being equal, the flatter (more elastic) the inventory demand function, the smaller the variance of $D$. The covariance between $D$ and $S$ should be negative, since public purchases ($S > 0$) are more likely to occur when $D$ is negative.

Observations can be weighted by volume or time. Since the volume of transactions is likely to be greater when there are large deviations between the average transaction price and the equilibrium price, one would expect the volume-weighted variances to be larger than the time-weighted variances. As an example of these differences, consider a block trade. Since block trades are large by definition and are associated with larger deviations from the midpoint of the bid-ask spread, $S$ is large. On a volume-weighted basis, a block trade contributes to a large deviation between transaction prices and equilibrium prices. If prices quickly return to the equilibrium
level after a block trade, they may not contribute very much to a
time-weighted deviation between market prices and equilibrium prices.
The alternative to a block trade may be a large number of small
transactions in the auction market over a period of hours, days, or
even weeks. If these transactions move the market maker’s inventory
away from its target, they will increase the value of D. Although the
increase may be smaller than would have occurred with a block trade,
it is likely to last for a longer time. Therefore, on a time-weighted
calculation, block trades may cause a smaller deviation from equilib­
rium prices than a position change carried out by small auction
market transactions.

The weighting used could affect averages as well as variance.
The calculation of the spread is an important example. Let $B_t$ and $A_t$
represent the time-weighted average bid and ask quotations. Let $B_v$
and $A_v$ represent the volume-weighted bid and ask prices at which
the market maker conducts transactions. Then

$$S_t = A_t - B_t$$

represents the time-weighted average spread. This is the variable
used in most empirical studies to determine the effect of market
volume, competition, risk, and other characteristics on the cost of
immediacy. The correct estimate of the cost of immediacy to traders,
however, and of the gross profits of market makers is

$$S_v = A_v - B_v$$

$S_t$ is the average quoted spread, $S_v$ the average realized spread. If
market makers had inventory losses on average, the realized spread
would be less than the quoted spread. The evidence on this point
is reviewed later.

**Empirical Studies of Market Makers.** This section summarizes the
results of some empirical studies of market making that have contrib­
uted to our understanding of many questions. I focus on two ques­
tions that are of concern to public policy issues connected with floor
trading on futures exchanges:

- Is market making a natural monopoly?
- Do market makers have a trading advantage because they are
  better informed than other traders?

Is market making a natural monopoly? A firm is a natural monopoly
if expanding its rate of output will lower its average costs. The usual
arguments for competition between firms do not apply in the natural
monopoly case since if several competing firms share a market each will have a lower rate of output and therefore higher costs than if all the output were produced by one larger firm.

Two types of evidence are relevant to this question. One is provided by studies of the bid-ask spread: under some conditions studies of the spread can be used to estimate the shape of a market maker’s cost curve. The second is based on observations of the number of competing market makers in various situations.

Numerous studies have been made of the size of the quoted bid-ask spread of equity market makers. All of them find that the size of the quoted spread decreases as the size of the market (measured by volume of trading or number of stockholders) increases. This is the relation one would expect if market making were a natural monopoly. The issue is not so simple, however, since other factors could produce a downward-sloping relation that would not indicate a natural monopoly. The following points are relevant:

- The rates of return that NYSE specialists earn indicate that their quoted spreads include an element of monopoly profit. Therefore, the lower spreads associated with higher-volume issues may be due in part to the effects of more competition.
- Controlling for other factors, it has been found that the greater the number of competing dealers in a stock, the lower the spread for that stock. This has been interpreted as indicating that the economies of scale are external to the market-making firm. (If they were internal to the firm, dividing up a market of a given size among several firms should increase the costs of each and result in larger spreads.) External economies do not lead to natural monopoly. One possible source of external economies is the ability of dealers to trade with one another to adjust their inventories; with more dealers in a market it may be easier for any one dealer to keep inventories at the desired level.
- A finding “that spreads fall faster with market size on the OTC (where there are several competing market makers) than on the NYSE (where there is only one specialist per stock)” has been interpreted as “evidence that the market making function is not a natural monopoly.”

The authors of a recent review of the literature on this topic conclude that “the empirical evidence thus quite clearly suggests that market making is not a natural monopoly.”

The second type of evidence concerning the natural monopoly argument is the extent of competition observed in situations in which no severe barriers to entry exist. If production of a product or service
is a natural monopoly, firms have an incentive to merge, or one firm will be able to drive its competitors out of business. The persistence of competition suggests the absence of the kinds of economies of scale that lead to natural monopoly.

In the over-the-counter market, stocks with very low trading volumes have one market maker, but at volumes much less than those reached by most listed stocks multiple market makers appear.\textsuperscript{82} In contrast, on exchanges that use the specialist system, competing market makers have disappeared even though the volume of trading on most listed stocks is greater than the volume at which multiple dealers would have appeared in the over-the-counter market.

One explanation is that the relevant economies of scale depend on the trading environment and that a higher volume of trading is needed to support two market makers on an exchange than in the over-the-counter market. A key element of the environment is where stocks are traded on the exchange floor. On U.S. stock exchanges a stock can be traded only at its assigned location. Locations seem to be assigned in such a way as to keep the trading evenly spread over the entire trading floor. Very few locations regularly have enough volume to support more than one market maker, and the locations at which the volume is high enough are not likely to be adjacent to one another. Therefore, a firm that wished to compete with existing specialists either would have its activity spread over the trading floor, making communication and coordination more difficult, or would have to compete at many locations in which the volume of trading was inadequate. The physical arrangement of trading on stock exchange floors may create an important barrier to competition. Fewer trading locations and a higher volume of trading at each might make competing dealers economically feasible. In futures markets the volume in one trading pit is usually sufficient to support multiple traders, and the number of scalpers seems to be closely related to the volume.\textsuperscript{83}

The available evidence indicates that market making is not subject to such substantial economies of scale as to preclude multiple market making in the financial assets that are commercially important. The trading volume at which competition appears, however, depends very much on the structure of the market.

The evidence also indicates that competition, whether from other market makers or in other forms, reduces the prices charged to users for the market-making service. It does not necessarily follow that moving an asset from a trading environment with competing market makers to one with only one is beneficial to the public customers. Other costs may also vary. If the brokerage costs of searching for the best price were higher in the over-the-counter market than on
an exchange, the benefits of less expensive market-making services might be offset by higher-cost brokerage services. Customers are interested in the total costs of trading.

_Are market makers better informed than other traders?_ A better-informed trader will buy before prices rise and sell before they decline. Since market makers typically sell at the bid and buy at the ask, they should earn the quoted spread if there is no change in the average level of prices between the time they buy and the time they sell. If market makers are better informed, their gross profits should be greater than the spread; if they are poorly informed, their gross profits should be less than the spread.

Three types of evidence could be brought to bear on this question. One is the relation between changes in market makers’ inventories and subsequent changes in prices. This directly addresses the issue whether the market makers are well informed on average for the time period in question. A second type of evidence is a comparison of quoted and realized spreads. This addresses the question whether market makers are better informed on average than their customers over all time periods. The available evidence of these two types is discussed below.

A third type of evidence is the cost incurred by traders when they deal with a market maker. The gross spread paid by all traders should equal the gross spread earned by the market maker. Comparing the gross realized spreads of a particular group of traders with the quoted and realized spreads of the market maker should tell us something about the information advantage or disadvantage of that trader compared with the market maker. Unfortunately, but not surprisingly, practically no studies of this kind are available.

If market makers could anticipate price changes, they would increase their inventories before price increases and decrease them before price decreases. The institutional investor study collected data on overnight inventories of NYSE specialists for ninety-three stocks for fifteen months beginning July 1, 1968. Similar data were collected for dealers who were not members of the stock exchange but made markets for any of these stocks in the over-the-counter market. This over-the-counter market in listed stock is known as the third market. The relevant data are shown in table 2–2.

Consider the first row in the table. If NYSE specialists and third market dealers were able to forecast the prices of the stocks in which they made inventories correctly, they would increase their inventories on days when large price changes were imminent. On the day before the price of the stock increased by 5 percent or more, NYSE
specialists decreased their inventories in the stock by approximately $24,000, third market dealers by an average of $3,500. The average inventory change was negative before stock price increases and positive before stock price decreases. Similar behavior has been observed for over-the-counter market makers in stocks not listed on exchanges.84

These data indicate that market makers, both on exchanges and in the over-the-counter market, lose money on their inventories because of their inability to anticipate day-to-day price changes. The evidence covers only day-to-day price changes in stock markets. Market makers might be better at predicting price changes over shorter intervals or in other markets.

In a very interesting study, Hans Stoll estimated the quoted and realized spreads for NYSE specialists as a group.85 The average quoted spread was 0.65 percent of the stock price (average 1975–1979). The realized spread was 0.327 percent of the stock price (average 1975–1980), or approximately one-half the quoted spread, indicating that specialists on balance suffer losses because of price changes in their inventories. This is the opposite of what one would expect if specialists had better information than other traders.

Holbrook Working analyzed the trading record of “the leading floor trader” on the New York Cotton Exchange in March and April 1952. The average realized profit was 0.93 points per pound. A point equals one one-hundredth of one cent ($0.0001).86 He does not say

### TABLE 2–2
AVERAGE VALUE PER STOCK OF NET INVENTORY CHANGES OF NYSE SPECIALISTS AND THIRD MARKET DEALERS, BY NEXT DAY'S CHANGE IN STOCK PRICE, 1968–1969

<table>
<thead>
<tr>
<th>Percentage Change in Stock Price in Next Day's Market</th>
<th>Change in Market Maker's Inventory Today (dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NYSE specialists</td>
</tr>
<tr>
<td>5.0 or more</td>
<td>−23,974</td>
</tr>
<tr>
<td>3.0 to 4.9</td>
<td>−12,083</td>
</tr>
<tr>
<td>1.0 to 2.9</td>
<td>−9,585</td>
</tr>
<tr>
<td>−0.9 to 0.9</td>
<td>798</td>
</tr>
<tr>
<td>−2.9 to 1.0</td>
<td>7,913</td>
</tr>
<tr>
<td>−4.9 to −3.0</td>
<td>10,362</td>
</tr>
<tr>
<td>−5.0 or less</td>
<td>21,553</td>
</tr>
</tbody>
</table>

what the quoted spread was, but since the minimum fluctuation in
the cotton futures contract was one point, the realized profit was
just about equal to the minimum possible spread. If this person was
the "leading floor trader" on the cotton exchange, other floor traders
probably had lower profit margins.

Silber analyzed the trading records of two scalpers on the New
York Futures Exchange.87 One had an average profit of $13.59 per
contract, the other an average profit of $12.12 per contract. These
data are for the same four-day period for the two traders, who had
367 and 823 round turns respectively during this period. For one of
the traders the average gross profit for thirty-one trading days, during
which he participated in 2,178 round turns, was $10.56 per round
turn. The market spread during this period was at least $25 per
contract. (The difference between the spread and the average profit
is statistically significant at the 0.01 percent level.)

The evidence on realized spreads from the stock market covers
only a short time but covers all NYSE specialists. The evidence from
the futures markets is based on self reports of only three traders in
two different markets, but it is consistent with the evidence from
the stock market in rejecting the idea that floor traders profit by
anticipating future price changes. For the stock market the evidence
points strongly in the opposite direction. Stock exchange specialists
systematically lose money because of their inability to forecast price
changes as well as the average traders with whom they deal.

Price Effects of Speculation and Floor Trading. The price conse­
quences of one style of trading cannot be determined without know­
ing how other market participants behave. In this section I characterize
certain basic trading styles and indicate the potential price conse­
quences of markets with various proportions of each. I also consider
the trading strategies that might be profitable for floor traders in
each of these circumstances.

Since the main interest in this section is in factors determining
price changes, I do not explicitly consider the bid-ask spread and
the supply and demand for liquidity. Instead, I consider each spec­
ulator as buying or selling at the midpoint of the bid-ask spread and
determine gross speculative profits or losses on this basis. Some
expenses (brokerage fees and payments for liquidity) or supplemen­
tary earnings (from supplying liquidity) are not considered here.

I distinguish three types of demand: demand based on funda­
mentals; demand from speculating on fundamentals; and demand
from speculating on prices. For convenience I refer to fundamental
traders, fundamental speculators, and price speculators. Of course, the same person may follow more than one strategy sequentially or in combination. The important point is that the price consequences of each strategy are different.

The demand in question is for an asset rather than a service (a stock as distinct from a flow). The total inventory of assets available to the market is fixed. The total inventory of common stock is the amount issued by the corporation, and the price must adjust so that the public is willing to hold that amount. The total supply of a futures contract is zero, and the price must adjust so that the demand for long and short positions is equal.

Hirshleifer and Riley define speculation as "purchase with the intention of re-sale, or sale with the intention of repurchase, where the uncertainty of the future spot price is a source of both risk and gain. The probabilistic variability of price is in turn due to anticipated emergence of information." Analysis of price consequences of speculative trading must consider both acquisition and disposition of a position.

Fundamental demand. Figure 2–2 illustrates the familiar form of a fundamental demand function for an individual trader. Such a trader takes the market price as given. The quantity demanded is determined by two factors: the market price and "fundamentals." A change from one point to another point on the same curve illustrates the effect of a change in the market price when there is no change in the fundamentals. The shift from one curve to another illustrates the change caused by a change in fundamental factors.

The "market" demand function is derived by holding fundamental factors constant and, for each possible price, adding up the quantities that would be demanded by each trader at that price. Although individual traders are free to adjust the quantities they demand and consider the price given, at the market level the relations are reversed.

The important characteristic of the fundamental demand function is the inverse relation between price and quantity; that is, as the price decreases, the quantity demanded increases. If all the individual demand functions have this characteristic, the market demand function must also have it. Therefore, if all the individual demand functions are of the fundamental type, the market demand function must show an inverse relation between price and quantity. This characteristic is important because it produces a unique and stable equilibrium price. The equilibrium price is stable in the sense that when a temporary shock causes a small shift away from it, the price
Demand Function for a Fundamental Trader

will tend to return to that equilibrium. If fundamental factors change frequently, the equilibrium price and the market price may exhibit great variability and still be stable in this sense.

It is interesting to consider how prices would change in an otherwise frictionless market dominated by fundamental demand. If a change in fundamentals was instantly known to all traders, the price would adjust immediately to the new equilibrium price. If the information became available to traders one at a time, the price would adjust gradually. Copeland has analyzed the price and quantity changes that would occur in such a situation. 89

Speculating on fundamentals. Speculators buy or sell because they believe they can anticipate the emergence of new information that will change the equilibrium price. I call the emergence of such information an information event, bullish or bearish according to whether the information will increase or decrease the equilibrium market price. The fundamental speculator who anticipates a bullish (bearish) information event will take a long (short) position until the event
occurs. If the event turns out as anticipated, the speculator will be able to liquidate the position with a gain. If the speculator is wrong about the price effect of an information event, the position will be liquidated with a loss.

A critical relationship for my analysis is the role of the current market price in the demand function of the fundamental speculator. The fundamental speculator makes decisions based on anticipated information events. In the interval between the opening of a speculative position and its close after the event occurs, I assume that the fundamental speculator either ignores the current price or responds to it like a fundamental trader. If the speculator is bullish and the price declines without new fundamental information, the speculator concludes that the position is even more attractive than it was. (For a bearish trader, a price decline would be less attractive.) Thus if the speculator makes any response to price changes, it will be qualitatively the same as the response of a fundamental trader—buying if the price declines and selling if it rises. The fundamental speculator does not interpret changes in the current price as revealing information about the occurrence or likelihood of information events.

Assuming that the market would otherwise be dominated by fundamental traders, the price effects of fundamental speculating depend on the accuracy of the speculators’ anticipations. The fundamental speculator who is correct causes prices to change in the correct direction sooner than they otherwise would. Resource allocations influenced by current prices also take place sooner. When the correctly anticipated information event occurs and the demand functions of other traders are shifting in one direction, the demand functions of fundamental speculators are shifting in the opposite direction. They liquidate their positions to realize profits by selling on correctly anticipated good news and buying on correctly anticipated bad news.

The consequences of incorrect anticipations are just the opposite. They cause prices to move in an inappropriate direction before the information event. When the event occurs, it is at least as much of a surprise to the fundamental speculators (who were expecting the opposite) as it is to the fundamental traders (who may not have been expecting any news). Both groups tend to buy on good news and sell on bad news. Since the trading of the fundamental speculators when they are wrong is on the same side of the market as the trading of the fundamental traders, it increases the size of the price adjustments that must be made when the information event occurs.

Between opening and closing a speculative position, fundamental speculators ignore the market price or use it to adjust the size of their positions. In this interval a market demand function that is
constructed from the demand functions of fundamental traders and fundamental speculators is certain to have an inverse relation between prices and quantities.

In summary, accurate speculation on fundamentals has desirable price consequences and is profitable to the speculators; inaccurate speculation has undesirable price consequences and is unprofitable to the speculators. A market demand function including only the demands of fundamental traders and fundamental speculators will have an inverse relation between price and quantity and thus result in a stable equilibrium price. Those who defend speculation apparently have this kind of speculation in mind.

Speculating on prices. Speculating on prices is similar to speculating on fundamentals in many ways. One important difference is in the way price speculators respond to price changes in the interval between opening and closing a position. Both kinds of speculators initiate positions on the basis of their anticipations about future information events. But fundamental speculators do not use intermediate price changes to update their opinions about whether the information event will occur (or has occurred).

In contrast, price speculators are willing to change their opinions about whether the information event will occur (or has occurred) on the basis of the price changes they observe before they receive definitive information about the occurrence of the event. The rationale for this behavior is simple and appealing. If some traders learn about fundamental information events sooner than others, the trading of those who learn early causes price movements that tend to reveal the outcome to those who have not yet learned and who are willing to learn from prices. Therefore bullish price speculators will interpret a rise in price as evidence that their forecasts are more likely to be correct and a fall in price as evidence that their forecasts are less likely to be correct. They can respond to an intermediate price increase in one of two ways. If the new price is still substantially below the level they expect it to reach, they may consider the ratio of risk to reward more favorable and add to their positions. As the price approaches the level they anticipated, they may become nearly certain that the information event has occurred and begin profit taking. If the price moves in the opposite direction from that expected, they will take it as evidence that their anticipations are not correct and reduce their positions. The demand curve of a price speculator after taking a position has the general shape illustrated in figure 2–3.

Suppose, for example, that the actual price is at \( P_a \). A price speculator believes that the price will either rise to \( P_u \) or fall to \( P_d \) in
the near future. After studying the situation, the speculator further decides that the odds are favorable for a bullish price move and acquires a long position of $Q_a$ contracts.

Curve $DD$ represents the demand of this speculator, assuming the only relevant information that is received comes from subsequent price changes. Any price rise above $P_a$ will be interpreted as indicating a higher probability that the price will eventually rise to $P_u$. Any price decline below $P_a$ will be interpreted as meaning that there is a higher probability that the price will eventually fall to $P_d$.

Every time the price changes, the price speculator reconsiders the appropriate size of the long position by reevaluating both the gains and losses that could be suffered and their probabilities; both their size and their probabilities will change as a result of the price change. Assume that the speculator is satisfied that a long position of $Q_a$ units is the appropriate response to a price of $P_a$. Each change in price changes the risk-reward ratio.

A price increase has two effects: it increases the probability that
the price will eventually reach \( P_{0} \); and it reduces the profit potential from a bullish move and increases the loss potential from a bearish move. The speculator balances these two effects. The price \( P_{0} \) represents the largest speculative position that the speculator will take. At prices above \( P_{0} \) the long position is gradually reduced as the price increases because the decrease in the gain is more important than the increase in its probability. At prices below \( P_{0} \) the long position is gradually reduced as the price declines because the decrease in the probability of a gain is more important than the increase in its size.

Price speculation is especially appealing to speculators who believe that others will learn about information events sooner than they. Fundamental speculation is appealing to speculators who believe that they will learn about information events at least as soon as other market participants.

Assuming a market dominated by fundamental traders and fundamental speculators, the price effects of price speculation at the time information events occur are the same as the price effects of fundamental speculation and depend on the accuracy of the anticipations. If information events are accurately anticipated, price speculation is profitable to the speculator, causes prices to adjust partially to the new information sooner than without the speculation, and reduces the magnitude of the price adjustment at the time of the information event. If information events are inaccurately anticipated, price speculation is unprofitable to the speculator, causes prices to adjust inappropriately before the information event, and increases the magnitude of the price adjustment at the time of the event.

Price speculation and fundamental speculation may differ significantly from the time speculative positions are opened until the information event occurs. To the extent that using intermediate price information to update anticipations increases the price speculator's accuracy, it may increase the frequency of good consequences and decrease the frequency of bad ones. During this intermediate period the price speculator's demand function tends to reduce the downward slope of the market demand function in a limited neighborhood. See figure 2-4 for an illustration.

If price speculators are sufficiently important, the market demand function may contain a section that slopes upward to the right, as illustrated in figure 2-5. The market "equilibrium" price will exhibit limited instability in the neighborhood of this bulge. That is, if the price is in this neighborhood, a temporary shock causing a small shift from the equilibrium price may cause a large price movement—up or down depending on the direction of the temporary shock—
FIGURE 2-4

Market Demand Function Illustrating the Effect of a Small Number of Price Speculators

until a new point of equilibrium has been reached. Since small shocks are almost certain to occur, such bulges are likely to add to the price variability. Since they can cause large price movements not related to changes in fundamental factors, they are likely to reduce the gain in accuracy that the price speculator hopes to obtain by updating expectations through the use of prices as well as direct fundamental information.

Trend trading. Trend trading is similar to speculation in that the positions taken are temporary. It is different in that the positions are not taken in anticipation of an information event. The trend trader believes that there are patterns in price changes that can be recognized soon enough to profit by trading on them.

Two examples will illustrate what may be involved. First, consider a market dominated by fundamental traders and fundamental spec-
Markets Demand Function Illustrating the Effect of a Large Number of Price Speculators

Price

Quantity

Speculators who do not all learn about the occurrence of information events at the same time. Although both respond appropriately as soon as they learn about the event, prices will tend to take time to adjust. The path of price changes from the time the first trader learns of the information event to the time the last person learns of it can be called a simple trend. The existence of trends is one possible source of profit opportunities for price speculators and for trend traders.

Second, consider a market in which there are market makers and the price is initially at equilibrium. As Garman has demonstrated, if buy and sell orders arrive at random, imbalances are certain to accumulate from time to time. Such imbalances will lead to changes in the market makers’ inventories. With an imbalance of sell orders, market makers accumulate a larger long inventory than they want and, as the inventory accumulates, lower the market price to induce an offsetting excess of buy orders. The resulting price pattern might be called a bulge, a movement in one direction followed by one in the opposite direction. A continuous market with market makers
must have bulges. A trend trader capable of recognizing the formation of bulges with sufficient accuracy and speed might profit by trading accordingly.

Profitable trading involving simple trends and bulges would increase the speed with which simple trends ran their course and reduce the magnitude of bulges. Both effects are desirable and make markets more nearly perfect. Whether trend trading on this sort of pattern was profitable would depend on the frequency of simple trends and bulges, their size, the skill of the trader in recognizing them, the size of the positions a trader could take, transaction costs, and so on.  

A kind of profitable trend trading might also be associated with the existence of points of unstable equilibrium. In this case the price consequences might not be considered desirable. Whether one would fault the trend trader who exploited the situation or the overabundance of price speculators that caused it is certainly arguable.

Economics of Exchanges. An exchange provides specialized facilities for trading. To enjoy the use of those facilities, a person must buy or rent "a seat," that is, a membership in the exchange. The value of a seat at any given time reflects the present value of the future economic rents that the potential member can hope to receive as a result of owning the seat.

If the exchange industry were characterized by perfect competition and freedom of entry, the value of the seat would also reflect the value of the capital resources employed in providing the specialized facilities. Exchanges clearly compete with one another in a variety of ways, but their competition has not reduced to zero the economic rents associated with successful exchanges. The market value of exchange seats is substantially greater than the market value of the assets employed in the exchange. The difference between the market value of the assets and the market value of the right to use the facilities can be thought of as the capitalized value of some "market imperfection." It would be helpful to know the nature of the imperfection. In part, it appears to be a result of some economies of scale.

Economies of scale. One view is that seat values reflect the economies of scale associated with operating an exchange. The average cost per unit of output of providing exchange services is lower in large exchanges than in small ones, but the costs of operating an exchange are a relatively small part of the total costs of transacting. It is doubtful if the economies achieved in operating large exchanges are the main source of their monopoly power. If they were a major
consideration, small exchanges in related activities would be under great pressure to merge. But the several futures exchanges in New York City maintain their independence even though they now share some common facilities, and the American Stock Exchange has successfully resisted efforts to merge with the New York Stock Exchange. Many brokerage firms who use the facilities of each of the potential merger partners would benefit from any economies that might result from a merger.

Another explanation is that the significant economies of scale that create seat values are the economies associated with a high rate of transactions in securities or futures contracts traded. These economies appear to be external to the market-making firms. That is, insofar as volume influences the unit cost of providing market-making services, it is mainly the total volume of trading in the market for the particular security or futures contract that is important, rather than the rate of trading of the particular market maker. If the economies of scale were mainly a result of the volume of business done by a single dealer, one would expect to see the largest dealer in a market drive all the others out of business by providing services at a lower cost. But competing market makers persist indefinitely in over-the-counter stock markets, in which the volume per issue is relatively low, and in futures exchange pits, in which the volume of trading per contract is relatively high.

If the external economies associated with a high volume of trading in a particular security or futures contract create the monopoly position of an exchange, exchanges do not need to be large to survive. What is critical is that they be able to attract a high proportion of the volume in the items that they trade. This explains why the same (or even very similar) contracts seldom trade successfully on two exchanges at the same time.

Seat values. To maintain or increase seat values, the officers of exchanges seek to generate more volume for the exchange, to restrict the number of seats, to reduce costs, and to lobby for favorable legislation and government policies. Although exchanges have a monopoly on their existing products, they compete vigorously with other exchanges or nonexchange marketplaces for new trading vehicles. One might say that they compete for the right to establish monopolies. The NYSE seeks to encourage companies that meet its listing standards to list their common stock on the exchange. In this respect it competes mainly with the over-the-counter market. Futures exchanges devote substantial resources to developing new contracts. They have been so successful that they have had to expand their
memberships, which they have done in part by creating restricted forms of membership so that existing members' interests are not diluted.

Some kinds of cost reduction are desirable from the point of view of members of an exchange, but not all. In a purely competitive industry, firms that do not use the most economic technology suffer substandard profits. The economies of scale, however, may allow an exchange to use less than the best technology without losing business to competing exchanges.

Suppose, for example, that a certain technological improvement would lower the cost of supplying ticker tape services and that persons with easy access to a ticker tape do more trading than those without such access. An exchange might be able to lower the price of the ticker tape service if the extra profits from increased sales of the service plus the extra profits earned by members from the increased volume of trading were greater than zero. Selling the service below cost might be desirable if the losses were more than made up by extra economic rents earned by members through the increased volume.

Suppose, however, that a technological improvement enabled each member to handle twice as much volume as before at no extra cost. If such an innovation were adopted, competition between members would force the price they charge for their services to decline until either (1) the price was equal to the marginal costs of providing the services and economic rents were zero or (2) volume had doubled before condition (1) occurred. If, at half the price, the quantity of members' service demanded was more than twice as much as at the old price, the value of seats would increase. If it were less than twice as much, the value of seats would decline. Since the costs of members' services are only part of the costs of trading, cutting the price of those services in half would not reduce the cost of trading by one-half. Therefore, unless the demand for trading services was extremely elastic, it is doubtful whether such an innovation would increase seat prices.

Members should be expected to oppose cost-saving innovations that might lower the value of their seats. Innovations that reduced the need for members' services in producing transactions would be dangerous to their economic interests in most circumstances. An exception would be a situation in which an exchange might lose business to a competing exchange if it did not introduce the innovation. The choice might then be to innovate and suffer a decline in profits or not to innovate and suffer an even larger decline. A second exception would be a situation in which the exchange was operating at or near capacity and a further increase in volume was likely. This
exception apparently explains why certain cost-saving innovations have been introduced into the trading process on the NYSE in recent years. (In the late 1960s the NYSE faced public criticism and loss of business and risked the loss of its independence because it was unable to handle the volume of business available.)

Since the form in which exchanges are organized makes certain kinds of cost-saving innovations undesirable, one might wonder why exchanges do not change from membership organizations to ordinary for-profit corporations or alternatively why new firms with this form of organization have not entered the industry. A completely satisfactory answer is not available. Some attempts at entry have been made, at least in the stock market. A firm known as Instinet provides services similar to those of an exchange. None of the new entrants have been outstandingly successful. One advantage of a membership organization may be that it is more effective at controlling certain kinds of antisocial activity.

Some futures exchanges permit the owner of a seat to lease it to another. The lessee has the right to trade, but the owner continues to benefit from any increases in the value of the seat. This appears to be a step in the direction of separating ownership from participation.

Cash versus Futures Markets. This section considers some similarities and differences between a cash market and a futures market. Imagine a futures contract that called for the delivery of stock certificates. The common stock would be called a cash (or spot) commodity. Most of the distinctions I draw here are general and would apply to any cash item and the corresponding futures contract, although the details would differ. Three points of comparison are especially relevant. These are trading, settlement, and arrangements for the interim from execution to settlement. If the cash commodity was not a security, an additional basis of comparison would be the degree of heterogeneity between the items traded in the cash and futures markets.

Trading. There is no necessity for different trading methods for cash and futures. Common stock and common stock futures contracts are traded in the same way in Paris. The "bargains" arrived at on the London Stock Exchange are something between futures contracts and cash contracts. In practice, different methods are often established, either because of the nature of the contract or because of the volume of trading. Because each batch of grains, for example, is different, physical examination may be desirable, but it is not necessary in a grain futures contract.
Settlement. A trade may be settled in two ways after it has been executed: by delivery or by offset. In cash markets nearly all trades are settled by delivery; in futures markets nearly all are settled by offset. Both types of settlement occur, however, in both types of markets. Settlement costs can be an important component of the overall cost of transacting. Whether settlement is by offset or by delivery depends on the procedures adopted by the market as well as on the characteristics of the traders. One market can compete with another by adopting a lower-cost settlement method.

Delivery may be of the item contracted for, an acceptable substitute, or the cash equivalent. For common stock the item contracted for might be a properly endorsed stock certificate; an acceptable substitute might be the liability of a recognized security depository or trust company. Settlement by offset means that a purchase is settled by a sale or a sale by a purchase. Every exchange that has a clearinghouse allows settlement by offset. Clearinghouses are the exception in cash markets. When they do not exist, settlement by offset requires an offsetting trade between the parties to the original trade or their successors.

All trades that take place within the same settlement period are settled at a common future time. On the NYSE, for example, the settlement period is one trading day, and the common future time is five business days later. If a trader makes a purchase and a sale in the same settlement period, the two transactions are offset, and no delivery is required. If a purchase and a sale take place in different settlement periods, two deliveries are required to settle the trades: the trader must accept delivery on the purchase and make delivery on the sale.

Markets vary greatly in the length of their settlement periods. In most cash markets each transaction creates its own settlement period, and settlement by offset is usually impractical. This is no hardship if buyers almost never resell and sellers almost never repurchase. The common practice in U.S. stock markets is for the settlement period to equal one trading day. On the London Stock Exchange the settlement periods are two or three weeks long.

On U.S. futures markets, the settlement period should properly be described as the life of the contract. Only open positions that have not been offset when the contract expires need to be settled by delivery.

Interim arrangements. During the interim from execution to settlement the unfulfilled contract must be monitored. This entails risks and costs.
The practice on both futures exchanges and stock exchanges is to rely on oral contracts at the time a trade is executed on the floor. No written document is created that is acknowledged by both parties and constitutes proof that the trade occurred. This practice saves time during the trading session. Each party reports the trade to a clearing firm, which attempts to match the traders' reports. When two reports are matched, the buyer and seller have acknowledged claims on each other. Until then each side bears a noncomparison risk. In futures exchanges, an effort is made to complete the comparison process before the next day's trading begins.

After trades have been compared, each side has an unfulfilled contractual obligation to the other side or its successors until the obligation is eliminated by settlement or offset. As with any unfulfilled contract, monitoring the contract entails default risks and, more important, costs; each side must make sure it fulfills its obligations to the other and receives what is due from the other. In the stock market the period from execution to settlement is relatively short, and no special steps are taken. In the futures markets a position may remain open for months. Therefore, futures exchanges have developed special procedures to reduce the cost and risk of monitoring the performance of unfulfilled contracts. Positioning the clearinghouse as one side of every trade is one part of the process. Closely related is marking-to-market each day. That is, any gains or losses that occur to an open position are transferred in cash, through the clearinghouse, from the loser to the gainer each night, so that at the start of trading each day there are no unrealized gains or losses on any position.

Economic relationships. If market institutions do not evolve in ways that minimize the cost of transacting, they may be replaced by alternative institutions with lower costs. Since monitoring unfulfilled contracts is expensive, short settlement periods make sense. But delivery is usually more expensive than offset, and the longer the settlement period, the more likely offsetting will occur; therefore, long settlement periods make sense. In other words, increasing the length of the settlement period increases some costs and reduces others. There may be an optimal settlement period whose length varies with the characteristics of the object being traded and of the trader. If the optimal period is very short, there is a cash market. If the optimal period is very long, there is a futures market. Some markets are in between. Legally, it may be important to have a sharp dividing line between cash and futures; economically, the difference is a matter of degree.
Futures exchanges are marketing institutions that have specialized in reducing transaction costs by having long settlement periods and specialized procedures to minimize the cost of monitoring the unfulfilled contracts. They ought to be used if they can reduce the cost of the existing transactions and if they make possible transactions that would be too costly under existing procedures.

Regulation of Trading on Futures and Stock Exchange Floors

Until recently the futures industry and the securities industry were relatively separate. Each had its own exchanges and a separate regulatory agency in Washington. Many brokerage firms handled only securities or only futures contracts; firms that handled both might have separate departments for the two. Many participants in one of these industries had only a slight knowledge of the other, especially of trading procedures on the floors of the exchanges. In recent years, however, the separation between the two industries has been breached. Futures exchanges now trade products closely related to traditional securities industry products, and stock exchanges have created futures exchange subsidiaries.

The contact between the two industries raised questions about which regulatory standards would be applied to new products. Three broad answers could be given to these questions. One would be to unify the regulatory framework. A single regulatory agency, the Super Commission (SC), would be assigned responsibility for all the functions now performed by the SEC and the CFTC and would set up uniform standards. A second answer would be to divide the regulatory responsibilities between two or more agencies. The boundary lines could be drawn in several ways. A third answer would be to permit the regulated groups to have some choice about which agency would regulate them.

The regulatory regimes administered by the SEC and the CFTC have substantial differences. If a unified regulatory framework is chosen, some groups will face considerable adjustments. This is especially true of the trading practices of exchanges. Among brokerage firms some movement has already begun to create a unified regulatory framework. Brokerage firms that have not maintained a strict separation between futures business and security business prefer not to provide two separate rule books for their account executives, one applicable to securities transactions and the other to futures transactions. They have sought to create a set of rules that satisfy both regulatory standards. To the extent that they succeed, they will set the pattern for more closely integrated regulation. Among exchange
floors no such process of integration has yet taken place. At most some pairs of futures and securities exchanges have taken steps to facilitate the flow of orders from the members of one such exchange to the members of the other.

If two or more regulatory agencies are to exist, each with a clear field of responsibility, some decisions must be made about how to draw boundaries. One way is to assign some products to one commission and some to the other. A major step in this direction has already been taken; whether it will or should be the pattern for future decisions is one of the questions addressed in this section. Under this approach one regulatory agency would be responsible for all or nearly all the regulatory functions pertaining to a particular product.

Another way would be to create boundaries along functional lines and assign each agency responsibility for a function or a group of related functions for all products. One functional area might be the regulation of security issuers; the agency assigned this function would be responsible for disclosure by security issuers and for regulating insider trading and conflicts of interest between the issuers (managers) and the investors (stockholders). A second function might be the regulation of brokerage firms that do business with the general public. The agency assigned this function might be concerned with deceptive sales practices, asset protection, margins, asset segregation, and similar issues. A third function might be the regulation of exchanges and other secondary markets. One agency might be created for each function, or one agency might take two of the functions and a second the third. In the short run, dividing the territory by product was easier to accomplish. In the long run, however, it may or may not prove satisfactory.

A third answer is to permit competition among regulators by allowing the regulated groups to choose which jurisdiction would apply to them. This solution would produce a system somewhat like that for regulating banks.102

The following subsections describe some economic criteria that can be used to evaluate regulations; rules about members trading for their own account on the exchange floor (including dual trading on futures exchanges); rules about trading procedures (including block trading and computer-assisted trading); and rules about trading vehicles (including the possibility of futures trading in individual stocks). The regulatory approaches of the CFTC and the SEC are compared and, when appropriate, evaluated by the economic criteria.

Economic Criteria for Evaluating Regulation. In evaluating industries, economists commonly use two groups of criteria: structure and
performance. The performance criteria are primary but not always easy to measure and are more likely to be affected by short-term conditions. Some industrial structures have been shown to be more conducive to effective economic performance, and structural characteristics are relatively stable. Competitive industries, for example, generally perform better than monopolies. The structure of an industry, though an indirect indicator of performance, is always a helpful supplement to direct measures and is frequently more reliable.

*Performance criteria.* Two measures of performance are of major concern: pricing efficiency and operating efficiency. Operating efficiency refers to the costs of arranging transactions. Brokerage commissions and realized spreads are costs paid by the ultimate traders (hedgers, speculators, investors) for the services of others. Ideally transaction costs should also include the costs the traders themselves incur in deciding what to trade and in monitoring the services of others, insofar as those costs are affected by some characteristic of the trading system. In practice, the costs paid to others are more susceptible to measurement, and they usually receive the main attention.

Even when operating costs can be easily measured, the problem of evaluation remains. Are they as low as possible? Structural characteristics are particularly important here. The opportunities for new firms to enter the industry and gain customers by offering more satisfactory services (lower costs or higher quality for the same price or new and desirable features) is one basis for judging whether operating costs are as low as they can be. Secondary evidence, such as the development of new products or services or the introduction of new and lower-cost production techniques, may be an additional basis for making judgments. But many opportunities for cost reduction are external to the firm (because they require industry-wide coordination, for example).

Pricing efficiency refers to the price information produced as a result of transactions and quotations. Ideally, the prices should adjust promptly to reflect fully all the information available in the economy. Prices are a byproduct of the trading system. But transaction prices and bid-ask quotations are useful to others who are not directly parties to the transaction. Bankers use the prices of commodity futures and common stocks to value collateral; common stock prices influence decisions about capital equipment expenditures and financing and about dividend policy; farmers and processors, who do not themselves buy or sell futures contracts, may make production and processing decisions based on the prices of futures contracts. The
pricing efficiency of marketplaces has a special element of public interest because the prices are useful to persons not directly parties to the transactions.

In practice, these two goals are closely related and complementary. High operating efficiency (that is, low transaction costs) increases the number of persons who willingly participate in a market and the frequency of transactions and thus tends to keep prices up to date and responsive to new information. In other words, ease of access and low transaction costs encourage speculators to acquire information about supply and demand. Exchanges and futures commission merchants (FCMs) collect and distribute this information because it encourages trading. It also tends to increase pricing efficiency.

Operating and pricing efficiency are the two performance measures that are conventionally mentioned in connection with exchange markets. A third service produced by exchange markets, particularly futures exchanges, may be called "education." Many persons are attracted to futures trading (and to some extent to trading on other types of exchange) because it offers them the possibility of getting rich quickly. This applies to member traders as well as to nonmembers. Some discover that the possibility is a reality for them; many more discover that, for them, losses are easier to come by than gains. Those who are not successful tend to leave the activity, perhaps poorer and almost certainly with a better understanding of their abilities and limitations. Tuition at postgraduate educational institutions that offer opportunities to learn by experience is often expensive, but enrollment is voluntary. I leave to others the question whether higher "admission" standards are an appropriate form of consumer protection. My concern here is how the performance of this educational function affects the pricing efficiency of the markets.

One concern is that pricing efficiency may depend on the proportion of experienced and inexperienced traders. In a market dominated by experienced professional and nonprofessional traders, the presence of a moderate number of neophytes should not adversely affect pricing efficiency. Markets dominated by inexperienced traders, however, may not achieve the pricing efficiency of which they are capable. There is evidence that some new types of futures or options markets improve pricing efficiency as they become older, in part because of learning by market participants. It is reasonable to expect that even old markets may decline in pricing efficiency if they become overwhelmed for a time by a sudden large influx of inexperienced traders.

Another concern is the educational behavior of the experienced traders. If inexperienced traders behave in ways that contribute to
market inefficiency, experienced traders are likely to teach them the error of their ways. Shall we punish the teachers for disciplining their students or the students for their inexperience?

A critical issue in trading of all kinds is whether profitable trading and pricing efficiency are connected. If profitable trading contributes to pricing efficiency, there appears to be no economic reason to regulate trading. This issue has received considerable attention. Our present understanding of the admittedly complex issues involved leads to the following conclusions. One cannot say unconditionally that profitable trading increases pricing efficiency. Whether it does depends on the state of the market before the new entrants began to trade and on their interactions with other groups of traders. Two market states are important for this analysis.

First, suppose that before the new entrants begin to trade, the market is already characterized by pricing efficiency in the sense that prices respond appropriately to new information and to random shocks unrelated to new information (such as a temporary imbalance of demand or supply due to random factors). Suppose that the new entrants begin to trade and that their trading is profitable. The following conclusions seem warranted. The new traders are supplying liquidity at a lower cost than existing suppliers, or increasing the pricing efficiency of the market, or both. In brief, in an efficient market profitable trading contributes to increased pricing or operating efficiency.

Now consider a situation in which the market does not necessarily respond appropriately to new information and random shocks. Again assume that some new entrants begin to trade and that their trading is profitable. In this situation there is no necessary connection between profitable trading and pricing efficiency. Under the conditions assumed, it may be possible for traders to profit from a course of trading that does not contribute to operating efficiency or to short-run pricing efficiency.

These conclusions may appear somewhat paradoxical. If markets are efficient, profitable trading makes them more efficient. If markets are not efficient, some trading may be profitable even if it does not contribute to efficiency. The solution to the paradox is as follows. If markets are not efficient, it is because a sufficiently large group in the market is behaving in an unprofitable way. That can certainly happen from time to time, but it is an unstable and self-correcting situation. While it obtains, a trader can earn profits in three ways: by supplying lower-cost liquidity; by contributing to pricing efficiency through superior anticipation of future information events; and by exploiting the weaknesses of the unprofitable traders and increasing
their losses. The third way does not necessarily contribute to short-run pricing efficiency, but it does increase the speed with which unprofitable traders change their trading styles or leave the market and by so doing contributes to long-run efficiency. Regulatory schemes to prevent so-called destabilizing speculation are likely, if they are successful, to increase the length of time during which the market is inefficient.

Structural criteria. Structural questions of relevance here can be examined at two levels: the level of the exchange and the level of the trader on the exchange. Market structures that expose firms to challenges from other firms are likely to lead to better economic performance than structures that provide firms with a more or less assured market.

Regulation of Floor Trading. Exchanges regulated by the SEC usually exclude from the trading floor persons trading primarily for their own account unless it can be shown that they perform a market-making function. If they do, their trading from the floor is usually restricted to the market-making function. In contrast, futures exchanges usually permit any qualified member to trade on the floor of the exchange without imposing any so-called public purpose test.

The arguments for regulating or eliminating floor trading by members (other than market making) fall into two categories. One is that such trading has undesirable economic consequences, the other that it is unfair to the public in its use of special knowledge available on the floor. 106

Undesirable economic consequences. The SEC has developed tests to judge trading initiated by members from the floor of an exchange. Such trading is considered desirable if it is stabilizing and undesirable if it accentuates price trends or is speculative. Stabilization has been measured transaction by transaction by the tick test and on a daily basis by comparing net purchases or sales of a trader during the day with the change in prices during the day. 107 In a falling market net purchases are considered stabilizing and desirable; net sales are considered to accentuate the price trend and to be undesirable. The hidden premise, never explicitly stated or defended, is that prices should not change, at least not as a result of transactions initiated on the floor by members of the exchange. 108

One example is the finding of the SEC's Special Study that floor traders on the NYSE trade in a "destabilizing" manner. 109 The basis for this conclusion was evidence that floor traders bought during price increases and sold during price declines. This kind of trading
pattern would be expected if the floor trading was speeding up the rate at which prices adjusted to new information. It might also occur if the floor trading was exploiting price instabilities created by the trading habits of other traders. To distinguish between the two (given that floor trading was profitable), it would be necessary to determine whether the price changes to which floor traders contributed tended to persist (which would be expected from information trading) or to be offset (which would be expected from trading that exploited price instabilities). Information on these points is not provided by the SEC studies.

Economically stabilization is not an appropriate criterion. Trading is not necessarily good when it limits price changes or bad when it accentuates them. Whether it is good or bad depends on whether the changes are appropriate in light of new information about the value of the asset being traded.

*Unfairly exploiting special knowledge.* Two kinds of knowledge might be of interest to a trader. Fundamental information is knowledge that might help estimate the intrinsic value of the asset traded; for example, the earnings of a company or the quantity of wheat in various parts of the world. A second kind of knowledge is information about the portfolio position or trading plans of other market participants. I use the terms "market conditions" and "state of the market" interchangeably to refer to this kind of information. Unfair exploitation of fundamental information (inside information, fraud, manipulation) is not particularly related to trading on the floor of an exchange and is not discussed here.

The idea that trading initiated on the floor might make inappropriate use of market knowledge depends on any of three arguments, two of which have been used by the SEC. One is that members on the floor are better informed because of their position on the floor. It is not using the information but using it before others have it that is wrong. A second argument is that traders on the floor have access to information about market conditions that should not be used as a basis of trading because the use violates a contractual relationship. A third argument is that the use of certain kinds of knowledge of market conditions should be discouraged as a matter of public policy because it harms the market.  

The idea that members derive an unfair competitive advantage from their position on the floor is clearly expressed in the following quotation about floor traders from a 1936 SEC study:

Members who trade for their own account while on the floor enjoy certain competitive advantages over the general public,
in that (a) the cost of effecting purchases and sales is materially lower for them than for nonmembers; (b) they can trade extensively with less capital than nonmembers; and (c) their presence on the floor enables them to make and revise their market appraisals more promptly than nonmembers upon the first manifestation of any trend or shift in trend.\textsuperscript{111}

The same ideas are repeated, more than forty years later, in the SEC’s options study.\textsuperscript{112}

The economic logic of this argument is an embarrassment. If the competitive advantages attributed to floor traders exist, it is because these traders have invested time and money to become members and to learn their trade and because they engage continuously in it. Being present on the floor during most of the day, they cannot practice law or medicine, sell furniture, or repair automobiles. By the logic of this quotation, lawyers should not be allowed to practice law, because they can look up cases and write briefs in less time than their customers and their subscriptions to information on the latest court decisions gives them a better chance of winning their cases. If the members on the floor have chosen to specialize in trading and are better at it than others because of their talent and investments of time and money, should they not be encouraged to concentrate their time and energy on that activity and let the rest of us devote ours to activities in which we have comparative advantages?

The second objection to floor trading is that it may use information about the state of the market acquired as a direct or indirect consequence of a confidential professional relationship to the person contemplating the trade in a way that is harmful to that person. The main issue on futures exchanges is whether floor brokers should be allowed to engage in dual trading, that is, to trade for their own account on the floor. The objection to dual trading is that the broker may trade ahead of a customer, that is, may trade for his own account while holding an unfilled order for a customer at the same price and on the same side of the market. No one defends trading ahead, and exchange rules prohibit it, but trading procedures make it extremely difficult to detect abuses, especially occasional small-scale violations. Given the existence of the temptation and the uncertainty of detection, it is reasonable to assume that such violations occur.

The CFTC addressed this issue in its recent \textit{Insider Trading Report}:

While Commission and exchange rules require . . . that floor brokers execute customers’ orders before trading for their own accounts, many exchanges have deficiencies in their trade reconstruction systems that have created difficulties
in enforcing such rules in a timely and effective manner. The Commission, therefore, intends to propose a rule requiring each exchange to implement a system of trade sequencing for rapid and accurate trade reconstruction. . . . the Commission will consider as an alternative the merits of prohibiting dual trading.¹¹³

Nowhere in its report does the CFTC provide evidence on the possible extent of trading ahead.¹¹⁴ Traders are now required to record the half-hour time bracket in which a trade occurred. From these data it should have been possible to determine how frequently a floor broker trades for his own account and executes a trade for a customer on the same side of the same futures contract in the same time bracket. Not all such trades would constitute instances of trading ahead, but practically all instances of trading ahead would be included.¹¹⁵ These data would have set an upper limit to the possible fraction of trades that present a trading-ahead concern.

The commission gives no details on how trade-sequencing information would be used to enforce the existing trading-ahead rules. It would have been helpful if it had reported how exchanges that now require trade-sequencing records use them in policing trading-ahead violations. The report seems to take the position that detecting such violations is the major problem. The information the commission proposes to collect is not sufficient to establish a violation; the information might show that a floor trader bought ahead of his customer but not whether he did so after receiving the customer’s order. To detect violations mechanically would require a file that contains not only timed trades but timed orders. Orders are timed when they reach the floor. The costs of constructing and analyzing the necessary files might be very large, or they might be modest. It would be helpful to know.

The CFTC may have intended that the sequenced trading records not be routinely examined but be consulted in response to allegations of trading-ahead violations. This would substantially change both the costs and the benefits to be expected from trade-sequencing data.

Private remedies exist that reduce or eliminate the trading-ahead problem. The CFTC report mentions that private enforcement may be more appropriate than federal regulation.¹¹⁶ A customer seriously concerned about the trading-ahead problem could attempt to use a floor broker who did not trade for his own account. This would be relatively easy to enforce. Some floor brokers do not trade for their own account, in order to attract customers who prefer such an arrangement. If some futures commission merchants adopt and publicize a policy of using only floor brokers who do not trade for
themselves, they might attract retail customers who prefer that policy. The CFTC could facilitate this form of self-enforcement by requiring such merchants to disclose their policies in this regard to their customers.

Retail customers would almost never be in a position to detect trading-ahead violations even if they occurred regularly. An experienced retail customer might be able to detect consistently poor executions, but the appropriate response to poor executions, whatever the cause, is to select another futures commission merchant. Commercial customers might be in a better position to identify trading ahead as a possible cause of poor executions, but their normal response would also be to get a new floor broker.

The effect of the CFTC's recommendations depends on how the data collected are used. If the records were consulted mainly in response to allegations, they would be of little or no help to a customer who is not in a position to identify trading-ahead violations on his own. They would help provide legal proof in a lawsuit, if it came to that. If the CFTC recommendations resulted in the routine reconstruction of the sequence of trades and orders for all floor brokers on every market every day and surveillance by the exchange to identify potential violations, trading ahead might be too dangerous to attempt because of the high probability of detection.

Even if the CFTC's recommendations are implemented in such a way as essentially to eliminate trading-ahead violations, it does not follow that the welfare of the customers would be improved. Whether it would depends on the economic effects of trading-ahead violations.

The short-run effect is to reduce the benefits the customer derives from a trade and to increase the benefits the floor broker derives by about the same amount. If every transaction for a customer were equally likely to be subject to a trading-ahead violation, what would be the effect of an increase in the frequency of trading-ahead violations?

The effect might be similar to an increase in floor brokerage commission rates. The expected cost per transaction to the customer and the expected revenue to the floor brokers would increase by comparable amounts. It is unrealistic to assume, however, that the net effect would be a simple income transfer from customers to floor brokers, for the same reasons that a rise in commission rates is not a simple transfer from one group to the other. Rather customers would do less trading, because it would be more costly and less profitable, and more members would become floor brokers rather than spreaders or floor traders. The price of seats on the exchange
might change, until the extra benefits of being a member were equal at the margin to the extra costs of seats.

The analysis in the preceding paragraph assumes that trading ahead effectively increases the commission rates paid by customers. It may not. The brokerage rates paid by customers are competitively determined. If floor brokers earned significant amounts by trading ahead, competition among them for the right to handle brokerage business might lead to a lower commission rate, so that the net revenue of the broker (the stated rate and the expected amount earned by trading ahead) and the net cost to the customer would be the same as if no trading ahead occurred.

Although, as a first approximation, changes in the frequency of trading-ahead violations are similar to changes in brokerage commission rates, there are some important differences. I mention a few of the most important.

First, the effect of trading-ahead violations is not equal for all customers. Experienced customers who are expected to continue trading are likely to bear less than an average share of the costs, because they are better able to protect themselves. They are more likely to be able to identify poor executions and to respond by switching their business to another floor broker. If they do a large volume of trading, the possibility of losing their business will be given more weight. Inexperienced customers are likely to bear more than an average share of the costs. They are less likely to be able to identify poor executions; and if their trading is unprofitable, the possible loss from losing their business a little sooner may be less than the extra gains that can be obtained from trading ahead.

Second, dual trading may well offer some real benefits to brokers or their customers. It may provide some flexibility in the use of the members’ time; it may lower the cost of learning a different trading role; it may help to conceal the identity of a principal. Most important, the competence of brokers who engage in dual trading may be different from the competence of brokers who do not. The case for eliminating dual trading would be strong if we could be certain that the only reason for the practice is that it gives opportunities to “trade ahead.” If there are other legitimate reasons for dual trading, however, eliminating the practice may not help public customers.

Third, the existence of dual trading creates a suspicion on the part of some persons that they may be treated unfairly in futures marketplaces. Whether or not the suspicions are correct, their existence is a damaging fact, which undoubtedly serves to discourage some persons from participating in futures markets. The image of
the markets would undoubtedly improve if dual trading were abolished.

These paragraphs are certainly not a definitive economic analysis of the effects of trading ahead. They will serve their main purpose if they indicate that the consequences are not obvious and not necessarily as injurious to customers as might initially be thought.

A difficult question is whether further restrictions on dual trading, even its outright elimination, are appropriate or whether the problem should be left to private enforcement. Although a case for federal regulation could be made, it has not been effectively made by the CFTC. Given the importance of efficiency to futures markets, we should be very reluctant to impose cost-increasing regulations on them without a strong case. There is certainly no evidence in the CFTC report that the problem is common or is growing worse.

I believe the CFTC should encourage the exchanges to make effective use of the data they already collect. Exchanges that use the bracket system could estimate the extent of the problem with the data already being collected. Exchanges that now require more precise data could be asked how they are actually used; an effort could be made to estimate the cost of analyzing the data and to determine whether they have made a difference in trading patterns or in the extent of dual trading. Future public discussion of the issue could then be conducted on the basis of more facts and fewer assumptions.

**Regulation of Trading Procedures.** Exchanges possess certain economies of scale in providing trading facilities that can lead to a degree of monopoly once an exchange is able to establish itself as the dominant market in a particular trading vehicle. These economies of scale make it difficult for new entrants (other exchanges or other firms providing competing services) to compete effectively with an established exchange. Since the established exchanges recognize that successful competitive threats are unlikely, their incentive to reduce costs is diminished. It should be no surprise, then, that trading procedures have been notably lacking in innovation and technological change. Many possible innovations in this area are actively resisted by exchanges because they economize on members' services.

These conditions create a classic dilemma for regulators, if they are concerned with the problem at all. (The mandates of the SEC and the CFTC are not concerned mainly with monopoly profits or lower rates of productivity increase in the exchange industry.) In the short run, costs for consumers may be lower if production processes take advantage of the lower costs achievable with a larger market. In the long run, costs may be higher if productivity increases are
forgone because of the lack of competition. The appropriate choice depends in part on the extent of the economies of scale and in part on the potential for increased productivity.

The SEC has a slightly better record than the CFTC of permitting alternative trading locations and procedures to develop. It has restrained NYSE efforts to restrict trading on other exchanges and in the over-the-counter market in NYSE-listed securities and has encouraged or required the development of common clearing facilities for stocks and listed options. It has generally encouraged alternative trading mechanisms when they have appeared. In contrast the open outcry requirements and restrictions on prearranged trading have restricted the location at which futures contracts can be traded and the process of trading.

Two major innovations in trading procedures have been introduced in recent decades. One is block trading, which now accounts for over 40 percent of the volume on the NYSE. The second is the National Association of Security Dealers automated quotation system (NASDAQ) in the over-the-counter market. Both of these affect how trades are arranged, not how they are completed. Both took place on stock markets. Futures markets have been much more innovative in developing new products; stock markets have been more open to innovative procedures for arranging trades. Regulatory obstacles probably do not entirely explain these differences, but obstacles to innovation in trading procedures appear to be somewhat greater in the futures industry.

I now consider the potential advantages of two innovations relevant to futures markets and the regulatory obstacles to them. One is block trading and the other a computer-assisted trade completion system.

**Block trading.** In a face-to-face auction market there is no way to transact business except by publicly announcing a willingness to buy or sell. In some circumstances, however, the expected cost of publicly announcing a willingness to buy or sell is greater than the expected gain from trading. In that case, if there is no way to trade without publicly announcing a willingness to trade, the person may refrain from trading entirely or use an alternative trading mechanism if one is available. Trading might be made possible in these circumstances by allowing persons to arrange trades in private. (The New York Stock Exchange, for example, allows member firms to trade bonds in private without bringing them to the floor of the exchange even though the bonds are listed on the exchange. It allows members to arrange stock trades in private although they must be brought to
some exchange to be consummated.) Auction markets are effective in arranging trades if each participant is small in relation to the market or if a large participant is seeking to trade with a large number of small participants but not as effective in arranging transactions between large participants. In the stock market prearranged block trades complement the auction market. They are initiated by a large participant who believes that one or more large participants can be found to take all or most of the other side of the transaction.

In traditional commodity futures contracts, situations in which large participants are on both sides of the market are relatively rare. A substitute for block trades, the exchange of “futures for physicals,” is available in most markets.

In the financial futures markets, especially the stock index and foreign currency markets, the situation is different. First, exchange of futures for physicals is not available (for example, cash settlement contracts) or not appropriate. Second, the potential for large participants to be on both sides of the market must be great, because large participants are on both sides of the underlying cash markets. Third, existing cash markets provide a well-developed means of arranging cash trades between large principals at acceptable costs. Futures markets will be used only if they cost less. Conventional futures market transactions may sometimes be more expensive and less convenient than cash market transactions because of the cost and inconvenience of concealing the desire to trade. To be successful, futures market block trades would need to combine the privacy and low cost of cash market transactions with the flexibility given by access to a well-developed auction market.

Apart from any regulatory considerations, the procedures used to execute block trades in the stock market cannot be used without change in a futures market because of differences between the specialist system and a competitive market maker system. In the stock market the block trader can have an associate on the floor of the exchange ask the specialist how many shares he would take at a possible block price. This facilitates arranging the block since the block trader can inform all parties to the transaction in advance, with a high degree of certainty, of the number of shares they will be able to buy or sell at the proposed block price. Checking with the floor in advance works in the stock market because the specialists have learned to treat the information they receive about forthcoming blocks in confidence and with discretion. This procedure would not work in a futures market since the information needed could not be obtained in advance without disclosing it to everyone on the floor, which would defeat the purpose of the block trade.
Consider two methods that might be used to execute block trades on the floor of a futures exchange. Both would require regulatory changes. Other methods might be devised that would be superior to those described.

Suppose that the block trader has a customer who wants to sell 10,000 futures contracts and is willing to accept a discount of at most 2 cents per unit below the price of the last transaction. The block trader has found customers willing to buy 8,200 contracts at a discount of at least 1.5 cents below the price of the last transaction. Suppose that each contract represents 5,000 units and that the minimum price variation is one-fourth of a cent per unit ($12.50 per contract). At any discount between 1.5 cents and 2 cents below the last sale, there is an excess supply of 1,800 contracts among the block trader’s customers.

The first method is a cross transaction. The block trader decides to cross 10,000 contracts at 1.75 cents below the last sale. This is acceptable to the seller, since the discount is less than 2 cents, and to the buyer, since the discount is more than 1.5 cents. To execute the block on the floor of the exchange, the block trader would offer 10,000 contracts at 1.75 cents below the last sale. Suppose the rules of the exchange require that he give priority to any bids received on the floor at that price up to the amount offered. Ideally, he would like to receive bids for exactly 1,800 contracts. If he actually received bids for 3,200 contracts at that price, he would bid for 6,800 contracts for the account of his customers. Then he would cross 1,400 contracts, going short from his own (or his firm’s account) by that amount to satisfy the needs of his customers. The total block would amount to 11,400 contracts at 1.75 cents below the last sale. (If the block trader had known the demand would be that large, he would have offered the block at a higher price.)

The second method is a dutch auction. Suppose that the block trader were permitted to interrupt the auction market briefly to conduct a dutch auction for the block. In a dutch auction the total quantity is announced in advance, and the price is decreased one increment at a time until the quantity offered is sold. Again, assume that the rules require the block trader to give preference, at any price, to bids from the floor over bids from his own customers. The possible results of a dutch auction for 10,000 contracts are summarized in table 2–3. The block trader would bid for the 7,100 contracts remaining from the dutch auction. To satisfy the requirements of his customers, who wanted at least 8,200 contracts, he would sell short for the account of his firm 1,100 contracts at 1.75 off the last auction market price through a cross transaction.
TABLE 2–3
RESULTS OF HYPOTHETICAL DUTCH AUCTION

<table>
<thead>
<tr>
<th>Discount below Last Sale</th>
<th>Quantity Bought</th>
<th>Cumulative Purchases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>1.25</td>
<td>300</td>
<td>350</td>
</tr>
<tr>
<td>1.5</td>
<td>1,050</td>
<td>1,400</td>
</tr>
<tr>
<td>1.75</td>
<td>1,500</td>
<td>2,900</td>
</tr>
</tbody>
</table>

Block trades are viewed unfavorably by many futures market participants, perhaps partly because they are seen as competing with and possibly undermining the auction market. In circumstances in which block trading might be feasible in a futures market, the alternative is not likely to be the same volume in the regular pit market. Most of the potential transactions would not take place at all, or substitute transactions would take place in other markets.

For block trading to be successful, the block trader's customers must be assured that they will be able to buy or sell at an agreed quantity and price. Otherwise the customer risks disclosing his desire to trade before the trade is complete. This is likely to be expensive. To remove price and quantity risks for his customers, the block trader must be willing to assume them himself. He may have good reason to do so, because of the brokerage commission earnings possible with block trades. Consider the foregoing example. Suppose the block trader negotiates a commission of $5 per contract for each part of the transaction (not a round turn). In the cross transaction, the block trader earns commissions on 10,000 contracts on the sell side and on 8,200 contracts on the buy side, for a total of $91,000. This also explains why the block trader is willing to take the risk of a speculative short position of 1,400 contracts. He can afford to lose some money on the short position and still profit on the block trade. His cushion on a marginal cost basis is 1.3 cents per unit ($0.013 per unit $\times$ 5,000 units per contract $\times$ 1,400 contracts = $91,000).

Although block trading in futures markets faces no legislative barriers, exchange and CFTC regulations dealing with prearranged trades, cross trades, and taking the other side of a customer's order would need to be changed to permit it. Computer-assisted trading. The processing required for a custom-

126
er's order can be divided into three stages: (1) from the customer to the brokerage firm (usually by telephone to an account representative); (2) from the brokerage firm to the trading floor; and (3) from one side of the trade to the other on the floor of an exchange. Additional operations are required to complete the trade (clearing, billing), check margins, and report to the customer. Each manual step involves someone's time and the possibility of error; additional time is required to detect and correct errors.

In a computer-assisted trading system (CATS), basic information is entered only once. Thereafter all processing is electronic. All the manual processing that would be eliminated is essentially mechanical; no judgment is involved. A CATS should be distinguished from a computer-aided decision system (CADS). A computer system for making trading decisions could be coupled with a manual system for carrying out the trades; a computer system for carrying out trades could be coupled with a manual system for deciding what trades to undertake. There is no necessary connection between the two and probably no great advantage to coupling them.

In high-volume operations, such as an active futures market, the potential savings from a CATS might be significant. Furthermore, since the software and hardware for one futures contract would be very similar to that for another, the overhead could be spread over a large number of products.

In its simplest form an electronic trading system for a futures contract would consist of a file of buy orders, a file of sell orders, and programs for keeping the files up to date, for matching buy and sell orders to produce trades, and for reporting prices, volumes, quotations, and other state-of-the-market information. The orders in each file would be arranged by price and within each price level by some priority system, such as size or arrival time. Users would be able to receive information from the system and transmit orders to it from remote locations through computer terminals, personal computers, or similar devices.

In the electronic trading system described here, the computer is substituting for pencils and telephones, not making decisions about what should be traded or when. Of course, the kind of information available to traders would be different in this system.

Decisions about stages 1 and 2 of the order-handling process are to a considerable extent in the hands of individual brokerage firms, which are relatively free to introduce cost-saving innovations when they seem desirable. Since brokerage firms compete, they have an incentive to keep costs down.

Decisions about stage 3 may be dealt with in at least two ways.
If a CATS were introduced by an exchange, brokerage firms and trading floor members would be required to conform, even if conformance changed their way of doing business substantially. Such changes are difficult to arrange under the best of circumstances. Alternatively, individual firms could be allowed to introduce in-house trading systems, which would, in effect, compete with the exchange floor. One firm might make its systems available to others. Arbitrage would be necessary to keep prices more or less in line and to transmit trading imbalances between the computer systems and the floor.

Allowing brokerage firms to develop in-house systems would provide an evolutionary path for trading innovations. If individual firms could experiment with computer-assisted trading systems, experience about various methods of operating them would be gained without committing the whole exchange community to a radical change with uncertain results.

One advantage of a floor trading system—that much activity takes place in full view of the other traders—can also be a disadvantage, as I indicated in connection with block trades. The designers of a computer-assisted trading system have more control over what information is released. They can provide more of some kinds of information than would be available on the floor, less of other kinds, and some rather new kinds. Real-time volume is an example. More interesting kinds of information could be obtained by summarizing volume for classes of accounts. The Chicago Board of Trade has a system that provides, each evening, a summary of the purchases and sales each hour by class of account (speculators, hedgers, locals). The information comes from the clearinghouse records and is therefore available only after the close of trading. Some brokers try to estimate similar information by observing trading on the floor. A CATS offers the opportunity to provide such information during the trading day. The chances of designing the perfect system the first time are not great. But experimentation is almost impossible in an exchange-operated system to which everyone must conform.

Although many new regulatory problems are created by the new technology, some old ones are eliminated. Monitoring the behavior of a member who acts as broker and dealer is easy when the timing of every transaction can be recorded to the nearest hundredth of a second.

The main purpose of this section is neither to defend nor to criticize a possible CATS. Such a system is an excellent example of an innovation that is extraordinarily difficult to introduce in an exchange environment, where regulators can stifle innovation by insisting that
all possible problems be considered in detail in advance. Procedures must be developed to permit experimentation at reasonable costs.

Congress is the ultimate regulator. One useful step might be to require the regulatory commissions to report to Congress periodically on the innovations that have been attempted in their industry. With communication and data processing technology changing rapidly, it is unlikely that opportunities for innovation will not appear. If successful innovations do occur, progressive regulation deserves part of the credit; if they do not, regulators and Congress should consider whether regulation is partly to blame.

**Regulation of Trading Vehicles.** The most difficult issue in recent years has been how to demarcate the jurisdictions of the SEC and the CFTC with respect to certain equity-related products. The *modus vivendi* in effect, based on the Shad-Johnson accord, had the desirable consequence of permitting some new products to come into existence without interminable court battles.

The accord temporarily eliminated the possibility of trading futures contracts on individual stocks. I believe that allowing futures exchanges to trade such contracts, with actual certificate deliveries, would tend to increase the operating efficiency of both futures and stock exchanges. Since most persons would presumably agree that improving operating efficiency would be desirable, this is one of the consequences that should be weighed in the balance when the accord is reconsidered, as it is likely to be.

*Possible benefits of futures contracts for individual stocks.* The similarities and differences between futures trading for individual stocks and trading for cash delivery as now carried out on U.S. stock exchanges are discussed in the section "Economic Analysis of Trading on Exchange Floors." Futures trading tends to occur when it results in lower transaction costs and sufficient volume. Transaction costs might be reduced in trading futures on individual stocks because of lower trading costs, if pit trading with competitive market makers turned out to be less costly than the specialist system, or lower settlement costs, because some trades would be settled by offset rather than by delivery.

For pit trading to be viable, high transaction volumes are necessary. To substitute offset for delivery, high turnover is necessary. Therefore, the stocks that might make attractive candidates for futures trading would be stocks with high volume and high turnover.

The trading volume of at least some NYSE stocks is in the range
associated with successful futures trading. Although the common stock of about 1,500 companies is listed on the NYSE, the top 50 companies (about 3 percent of the total) accounted for 23 percent of the share volume in 1983. The volumes of some companies in the top group appear to be in the range needed to support successful futures trading. Volume among the top 50 companies ranged from about 64 million to 420 million shares per year. Some of these are large companies with average turnovers; some are moderate-sized companies with high turnovers.

If futures markets attempted to offer trading facilities for some of the stocks in this group, the futures exchanges might maintain their traditional trading techniques (pit trading, relatively unrestricted access to the floor by qualified members, and reliance on competition rather than regulation to control behavior) and adopt settlement techniques and interim arrangements that correspond more closely to those used by the stock exchanges. This approach does not seem to be prohibited by the accord. A futures exchange (or its subsidiary) could apply for designation as a stock exchange and trade stocks listed on other exchanges. Securities legislation does not require stock exchanges to use the specialist system. This approach would probably be unwise, however. The SEC has a nearly half-century tradition of willingness to sacrifice competition among members trading for themselves on the floor of an exchange for unenforceable legal responsibilities of unproven worth. A futures exchange would want to use all its competitive advantages, including the cost reductions from the special form of a futures contract.

The possibility that a futures exchange might capture a substantial part of the transactions now taking place on the NYSE should improve efficiency on the futures exchanges and the NYSE, even if the threatened competition never succeeded. As long as the NYSE is not completely confident that it is invulnerable to competition, it has an incentive to try harder to improve efficiency. As long as the futures markets believed that it was worth trying to get the additional business (if not now, then at some future time when conditions might be more propitious), they would have an incentive to improve the efficiency of the relevant parts of their operations. Thus the efficiency-improving consequences of this proposal depend not on one competitor's successfully entering the other's business but on the existence of a believable threat.

Some objections. Several objections are commonly raised to the possibility of futures trading in individual stocks. One is that it would
increase the possibility of insider trading violations and manipulations, but monitoring trading for these problems need not be harder on a futures exchange than on a stock exchange. Most stock that might be traded on futures exchanges is already traded on regional exchanges, and options on stock are traded on options exchanges. The SEC has considerable experience in handling these problems. Its special study of the options exchanges stated, "These [options] exchanges have developed techniques to detect related option and stock trading that might be manipulative, prearranged and fictitious option trading, misuse of material market and inside information, and violations of certain other exchange rules. . . . improvements must be made." If stock futures are traded, the statutes and regulations should be amended if necessary to treat owning a long or short position in a future or option as the equivalent to owning a long or short position in the actual security with respect to insider trading, fraud, and manipulation. Responsibility for this area should continue to be in the hands of one agency, but that agency need not be responsible for all exchange functions to do an adequate job.

A second concern is that prices in the underlying cash market might be distorted if owners of a long or short open position in futures held for delivery. This problem is of vital concern with any futures contract, not only stock futures, and also with stock options. The futures exchanges and the CFTC have had considerable experience in dealing with it. With appropriate monitoring of positions, the relations between futures prices and options prices should be less subject to manipulation than those between options prices and stock prices. I suggest that in the twenty trading days before a stock futures contract expires, the required margin be raised by 5 percent per day until it equals 100 percent. (On short futures positions, unmargined stock could serve as margin for the futures contracts.) Before a futures contract expired, the cash required for the futures position would then be the same as the cash required to own the underlying stock in an unmargined account. Under these conditions anyone attempting to manipulate the price of a stock could do so as well in the stock market as in the futures market. The rule would impose no hardship on a person who intended to accept delivery on a long position. If such a person had Treasury bills equal to the amount required to buy the stock, they could be used to satisfy the margin requirements. Persons who did not intend to deliver would have an extra incentive to switch their positions to a nonmaturing future.
Notes

1. Time limits did not permit a systematic study of trading floor practices on options exchanges.

2. Liabilities are collections in which the market value of the obligations is greater than the market value of the rights.


4. Evidence on this point is discussed in the section "Economic Analysis of Trading on Exchange Floors."

5. A comparison of floor brokerage on futures and stock exchanges illustrates the greater division of labor that exists on futures exchanges. There are floor brokers on both stock and futures exchanges. But on a stock exchange few, if any, floor brokers concentrate on a single stock or even on all the stocks in a single trading location. The typical floor broker moves from one trading location to another to execute his orders. As a consequence, at most trading locations there is no floor broker who is sure to be there. If a stock brokerage firm receives a limit order that cannot be filled quickly because the price limit is too far from the current market, it has no practical choice but to leave the order with the only member who is always at the trading location, the specialist. Specialists can "specialize" in particular stocks because they do not specialize in particular functions; they both trade for their own account and provide brokerage services for others.

6. Trading regulations that increase the costs of trading common stock may reduce their trading volume but rarely eliminate the market completely. Disclosure regulations, as distinct from trading regulations, may sometimes completely eliminate the market for a particular common stock. Some foreign securities may not trade in the United States because of disclosure problems. And the cost savings and flexibility that stem from not having to meet SEC disclosure requirements may sometimes be a significant consideration for a small firm in deciding whether or not "to go private."

7. The floor of the Chicago Mercantile Exchange is 40,000 square feet. See Wall Street Journal, January 3, 1984. An earlier version of this section was presented at a conference sponsored jointly by Tel Aviv University and the Tel Aviv Stock Exchange in January 1984 and at finance workshops at Dartmouth and Cornell. The author acknowledges helpful comments from Richard Bower, Menachem Brenner, Haim Mendelson, George Oldfield, and Richard West.

8. Thomas A. Russo, Regulation of the Commodities Futures and Options Markets (Colorado Springs, Colo.: Shepherd's-McGraw-Hill, 1983), sec. 1.8. A member trading for himself would record for each trade (1) quantity, (2) delivery month, (3) price, (4) the opposite member's name (or number), (5) clearing firm, and (6) time on a card or order containing (7) member's name and (8) clearing firm, (9) date, (10) commodity being traded, and (11) whether the member was the buyer or seller. The time is recorded in half-hour intervals called brackets. The complete record of a trade includes these eleven items.
Requiring additional information would take time, which could be a burden during active trading. This is a sensitive subject since the absence of a precise time record makes it difficult to reconstruct an exact trading sequence. See the later discussion of dual trading.

9. Ibid., sec. 2.02.

10. Ibid.


12. There is a question how to compare volume in securities or futures contracts. For studies of trading behavior it seems reasonable to base comparisons on the amount of the gain or loss that can occur on the minimum price fluctuation of the smallest amount traded. This minimum dollar change is indicated after each of the futures and stocks.

13. See section "Regulation of Trading on Futures and Stock Exchange Floors." See also Russo, *Regulation of Futures and Options Markets*, secs. 1.07, 1.09, 12.50. An exception to the general prohibition of prearranged trades is an exchange of futures for the underlying commodity, which may be permitted in some circumstances.

14. Rule 528 of the Chicago Mercantile Exchange, for example, states in part, "The buyer and seller in a pit transaction must report immediately to the pit observer any change in the last sales price or last quotation."


17. Rules on some futures exchanges limit the ability of a trader to combine acting as broker and trading for his own account. Some exchanges have more than one class of members, and members not in the most privileged class may be restricted in which futures contracts they may trade. Stock and options exchanges, by contrast, all have rules that limit the members' freedom of choice of roles and usually what securities they may trade.

18. In legal parlance such situations are called conflicts of interest; in economic jargon they are referred to as agency problems.


20. With scalpers the issue does not arise. A scalper must constantly pay
attention to the market in order to respond instantly to the flow of orders. With very large trading volumes, the total costs to a scalper of paying even a very small floor brokerage fee per transaction would be significant. With some electronic trading systems scalping might be done even by nonmembers.

21. In connection with this study, Mitchel Abolafia and I made an unsuccessful attempt to collect questionnaire data on where professional traders spend their time.


24. Examples of the kinds of information referred to are (1) information that scalpers' inventories are near their maximum acceptable levels, so that scalper inventory liquidations are easily possible; (2) information about the levels at which there may be concentrations of stop orders; and (3) information about the price levels at which there may be concentrations of limit orders.


26. Models of the inventory behavior of a single dealer are discussed in the section "Economic Analysis of Trading on Exchange Floors." Ho and Stoll present a model of competing dealers in an over-the-counter stock market (Tom S. Y. Ho and Hans R. Stoll, "The Dynamics of Dealer Markets under Competition," *Journal of Finance*, vol. 38 [September 1983], pp. 1053–74). Some aspects of each of these models are applicable to scalpers competing on a futures exchange. In particular, Ho and Stoll's model recognizes that over-the-counter stock dealers' quotations are typically competitive on only one side of the market whereas my model assumes that scalpers' quotations are typically competitive on both sides. The over-the-counter dealer can increase his expected rate of transactions by increasing the number of stocks in which he is active. The upper limit on his rate of transactions is determined by the time it takes him to respond to an inquiry. By publishing competitive quotations on only one side of the market, as in the Ho and Stoll model, the dealer avoids spending time on less desirable inquiries. Once a scalper has selected a pit location, however, he suffers no opportunity cost by publishing competitive quotations on both sides of the market.

27. This assumes that traders are aware that the orders exist. See below for a discussion of how a floor trader could be aware of limit orders held by floor brokers. For an analysis of the effect of limit orders on the trading behavior of a stock exchange specialist, see James Bradfield, "A Formal Dynamic Model of Market Making," *Journal of Financial and Quantitative Analysis*, vol. 14 (June 1979), pp. 279–91.
29. See, for example, Chicago Mercantile Exchange Rule 532.
31. Of course, persons initiating market orders could have obtained executions at more favorable prices if they had taken the risk of using limit orders. The main risk associated with a limit order is that the desired execution may never take place, because the market price may never reach the limit price. If the price moves in the opposite direction, the trader loses a potential profit by trying to lower his transaction costs.
35. Most corporate bonds are owned by financial institutions, and trading in institutional-sized quantities takes place in the over-the-counter (OTC) market. Members of the NYSE are permitted to participate in OTC trading in bonds even if the bonds are listed on the exchange.
39. To my knowledge, this has never been an explicit policy of the NYSE. It does, however, describe the results of whatever explicit policy or tradition has been consciously followed.
42. Stoll, "Specialist System."
48. Ibid., p. 395.
49. The description of floor trading in this section relies heavily on the descriptions and analyses in SEC, *Special Study*, pt. 2, chap. 6, sec. F.
51. SEC, *Special Study*, p. 204. The years 1941 to 1949 are excluded since no data are available.
53. SEC, *Special Study*, app. 6H, lists fifteen SEC studies or reports dealing with floor trading between 1936 and 1962, of which eleven were nonpublic. Some data from the nonpublic reports are disclosed in the special study.
55. Ibid., pp. 238–39.
56. Ibid., p. 239.
57. Ibid., pp. 241–42.
62. A standard way in which two agents could steal from their principals in these circumstances would be to arrange trades at "unfair" prices with a third party who would resell the shares at a fair market price and earn a nearly riskless profit. The third party could be a fictitious entity that the two agents had set up or a confederate who would rebate part of the profits. An example was reported in the press in 1984 regarding the London operations of an American bank. The transactions were taking place in the Eurobond market.
64. Ibid., pp. 1625–29.
65. In principle, in a frictionless market mechanism, if new information were available simultaneously to all market participants, the price of an asset could change instantaneously without any transactions occurring. (The price
at which transactions could occur would be reflected in the current bids and offers.) In practice, because not all market participants learn of new information at the same time and because of frictions of various sorts, price changes reflecting new information are usually associated with some transactions; persons who learn of the news first attempt to buy or sell at the old prices to profit by the change in price that they think will occur as the news is disseminated. These information-based transactions help alert other market participants to the existence of new relevant information. To say that the transactions cause the price change is inaccurate. If the news would eventually be released to the market, the possibility of transacting may cause it to be publicized earlier. It is only with respect to information that requires effort to be discovered that one could say that the transaction caused the price change.


70. Some trading procedures have been developed that are designed to assure market makers that the traders are not well informed. Presumably the market maker can afford to charge a smaller fee for serving traders who have no special information, since he is less likely to suffer inventory losses from dealing with such traders. See Larry J. Cuneo and Wayne H. Wagner, "Reducing the Cost of Stock Trading," *Financial Analysts Journal* (November–December 1975), pp. 35–44.


75. Ibid., p. 267. Garman also demonstrates that the expected time to failure is very sensitive to the amount of capital employed by the specialist.


78. For a review of this literature, see Cohen et al., "Market Makers and the Market Spread."

79. Smidt, "Which Road."


81. Ibid.


84. Stoll, "Dealer Inventory Behavior."


90. In the theoretical literature, this is described by saying that prices are
"revealing." That is, price changes provide information about the true state of nature. In the extreme case, there may be a one-to-one relation between future price changes and the true state of nature, so that prices are "fully revealing."

91. The demand functions of speculators depend on the current market price and the anticipated information event (anticipated fundamentals). There is always an inverse relation between price and quantity when the effect of anticipated fundamentals is held constant. Since for fundamental speculators a change in market price has no effect on anticipated fundamentals, we can be assured of an inverse relation between price and quantity in their demand functions. For price speculators a change in market price affects anticipated fundamentals. For them a change in the market price changes the quantity demanded directly through the price variable and indirectly through changes in anticipated fundamentals. The two effects may not be in the same direction. If they are in opposite directions and the shift effect is larger than the direct price effect, the effective relation between price and quantity may be positive rather than inverse.

92. I am indebted to Patti Bailey and Rob Goodridge of the National Futures Association for suggesting the use of this example.

93. Consider the following statement describing the relationship between U.S. commercial banks and the Chicago Mercantile Exchange's International Monetary Market (IMM):

Moreover, closer contact with the IMM provides important information. Many participants in the futures markets rely on the same charting techniques to predict exchange rate movements. The widespread use of this approach can generate similar behavior which in turn sometimes causes sharp, short-term price movements that spill over into the interbank market.


98. They are not always successful. For a study of the effect of government


100. See Abolafia, "Self-Regulation as Market Maintenance."


106. A third group of arguments claimed that restrictions on access to the floor were necessary to preserve the fixed minimum commission rate structure. These arguments are anachronistic and are not considered here.

107. See earlier section "Trading on the New York Stock Exchange Floor."

108. The SEC has never, to my knowledge, proposed that stabilization tests be applied to trading initiated off the floor by members or nonmembers of the exchange. If the tick test were universally required, only transactions that did not change the price of the item traded would be allowed, and perfect price stability would be ensured. The fact that a universal stabilization test has not been proposed is consistent with the idea that it is only price changes initiated by members from the floor of an exchange that are inherently "unfair." On this interpretation, price stability is not really the goal but merely the means used to limit the ability of on-floor traders to use their special knowledge.

109. See "Trading on the New York Stock Exchange Floor."

110. This kind of argument does not appear to have been used by the SEC as a reason for controlling trading from the floor of an exchange. But
the reverse of the argument, requiring certain types of knowledge to be disclosed, appears to be at the basis of some CFTC rules that limit prearranged trades. This is discussed in the section on block trading.

113. Commodity Futures Trading Commission, A Study of the Nature, Extent, and Effects of Futures Trading by Persons Possessing Material, Nonpublic Information (Submitted to Committee on Agriculture, House of Representatives, and Committee on Agriculture, Nutrition, and Forestry, Senate, pursuant to Section 23(b) of the Commodity Exchange Act as amended [September 1984], p. 10).
114. The commission did collect data indicating the possible magnitude of other potentially important "insider trading" problems, particularly that of large traders with potential access to nonpublic information. Partially as a result of examining this data, the CFTC concluded that no action was warranted.
115. It is technically possible for a trade-ahead violation to result in a trade for the trader’s own account in one time bracket and the customer’s trade in the subsequent time bracket. The length of the time bracket should ensure that this is a rare occurrence.
117. The following quotation illustrates the advantages of concealment.

A large scale speculator may find it advantageous, however, to do part of his trading in person on the trading floor. When he wishes to buy heavily, for example, he may place appropriate buying orders with brokers and, by doing some selling in person, frustrate the efforts of those who seek to profit by following his lead.

119. The value of the monopoly position is reflected in the price of seats. Exchanges would like to innovate in ways that increase the value of seats and to prevent changes that would reduce their value.
120. Not surprisingly, exchanges have done what they can to add to the natural difficulties of their competitors. For example, the clearing association is of critical importance in controlling competition in a futures market. (For a description and analysis of clearing associations, see Franklin R. Edwards, “The Clearing Association in Futures Markets: Guarantor and Regulator,” in Ronald W. A. Anderson, ed., The Industrial Organization of Futures Markets (Lexington: Mass.: D.C. Heath, Lexington Books, 1984), pp. 225–54.) Since the clearinghouse takes the other side of every open position, a futures exchange can control competition by controlling access to its clearinghouse.
121. See Sobel, N.Y.S.E., pp. 89–98. The Instinet system is an interesting example. When it first applied to the SEC, the question arose whether it should be regulated as an exchange or as a broker-dealer. Given its proposed
line of business, either position would have been legally defensible. Practically, regulation as an exchange would have been a death sentence. The SEC chose to consider Instinet a broker-dealer.

122. The development of a link between the clearinghouses of futures exchanges in Chicago and Singapore is important, but it affects clearing procedures, not trading procedures.

123. This is likely to be the case if trading by the person whose willingness to trade is being announced is on a large enough scale to influence the market price significantly. This will be the case for small transactions in thinly traded markets, for large transactions even in large markets, and for relatively small transactions when the size of the issue is small. One important characteristic of an over-the-counter market is that there is less publicity about willingness to trade. Such markets occur in the common stock of small companies, in bonds, and in block trades. It is interesting that the London Stock Exchange, which lists many small companies, has announced that it is adopting a computerized equity trading system similar to NASDAQ. See Wall Street Journal, July 20, 1984, p. 25.

124. The specialist would buy or sell for his own account and for the account of the brokers whose limit orders he was holding. He does not disclose how many shares go for each purpose, and he cannot speak for brokers in the crowd, if any. But for most stocks most of the time, there is no "crowd." Practically all the stock bought or sold by the specialist for himself and his customers receives the benefit of the block price.

125. In one way or another specialists are compensated for their cooperation in arranging blocks. One way is through the opportunity to buy or sell stock at the favorable block price. Another is by receiving floor brokerage commissions on at least part of the block, even though there is no need for their brokerage services.

126. See Russo, Regulation of Futures and Options Markets, Sec. 1.07.


128. That the ban may be temporary is clearly emphasized in the following passage from a joint SEC-CFTC release:

Pending further review of appropriate regulatory systems for trading in such futures, the legislative proposal would foreclose the trading of futures (and options on futures) on individual corporate or municipal securities. . . . Nevertheless, the two agencies intend to study further the issues raised by such trading with a view toward a future recommendation to lift this restriction.


130. SEC, Options Markets, p. 243.

131. Ibid., pp. 173–82.