Third Markets and the Second Best

Craig Pirrong
Oklahoma State University
Stillwater, OK 74078
405-744-1243
pirrong@okstate.edu

October 15, 2001

Abstract. The costs and benefits of "third markets" for financial instruments are widely debated. This article shows that a third market actually improves welfare when entry into the third market is unrestricted and it can free ride on the exchange's price discovery. A third market reduces welfare if its membership is limited unless the exchange expands membership to compete with it. The microstructure of financial trading allows exchanges to restrict inefficiently the supply of liquidity. Third markets can mitigate this source of inefficiency, and usually do so most effectively when they free ride. Thus, although free entry to the exchange would maximize welfare, encouragement of a free entry third market may be a second-best response to exchange market power.

JEL Classification: G10, G20, L11, L12, L31 Key Words: Financial exchanges, third markets.

1 Introduction

Although trading on exchanges is typically the focus of academic and popular attention, virtually all organized exchanges face competition from other trading venues. For example, "third market" dealers account for approximately 8 percent of the volume and 10 percent of the transactions in NYSE listed equities. Moreover, a large fraction of listed shares are transacted off-exchange in the block market. Derivatives exchanges face competition from the OTC market. Nor is this a new phenomenon. In the early days of stock trading, the NYSE faced competition from the New York Open Board. Similarly, the Chicago Open Board and the notorious "bucket shops" once competed with the Chicago Board of Trade in the grain futures business. Thus, the third market phenomenon is ubiquitous.[1]

Regardless of the era or instrument traded, third markets typically adopt rules and practices that are intended to preclude, or at least sharply curb, informed trading. For example, third market stock dealers in the United States often accept only small orders that are most likely submitted by uninformed retail traders, and sometimes buy such orders from brokers who have a largely retail clientele. Similarly, the bucket shops of old accepted only small trades. At the other end of the trade-size spectrum, block traders also attempt to screen out the informed (Seppi, 1990).

The efforts of off-exchange dealers to avoid adverse selection are frequently and pejoratively referred to as "cream skimming." Moreover, it is sometimes claimed that the existence of a third market that "free rides" on the price dis-

[1]Hereafter I will utilize the term third market to refer to any off-exchange trading mechanism.

covery performed on primary exchanges reduces market quality and impairs the informational efficiency of prices.[2] It is claimed that a cream skimming third market can initiate a "death spiral" in the primary market, whereby the movement of uninformed traders from the exchange to the third market exacerbates the adverse selection problem on the exchange, spurring more traders to leave the exchange. This supposedly reduces exchange liquidity, price efficiency, and exchange member profits, and in the extreme, the hemorraging of uninformed traders makes the exchange unviable.

These views have not gone unchallenged. Stoll (1994), Oesterle (1994), Hagerty and McDonald (1996), McInish and Wood (1996), Macey and Haddock (1985), and Seligman (1985) argue that satellite markets are beneficial because they increase competition faced by exchanges that exercise some market power. Such beliefs were the basis for the creation of the National Market System (NMS) and the Intermarket Trading System (ITS), and motivate the SEC's longstanding (and eventually successful) efforts to induce the NYSE to repeal its Rule 390 that precludes member firms from trading some listed stocks in the OTC market.

The costs and benefits of third markets therefore depend crucially on the interaction between externalities and competition. Information externalities would indeed lead to inefficiencies *if* but for the existence of a third market the exchange would supply the first best level of risk bearing and liquidity. If for some reason the exchange provides a suboptimally small level of output, however, it does not follow that free riding reduces welfare. In these circum-

[2]This argument is set out in Telser (1984), Mulherin, Netter and Overdahl (1991), Bronfman and Overdahl (1993), Bronfman, Lehn, and Schwartz (1994), Coffee (1996), Mulherin (1996), Miller (1996), and Easley *et al* (1996).

stances, we are in the world of the second best, in which case the existence of another "market failure" such as an externality can actually improve welfare. In particular, if the primary exchange is able to exercise market power and thereby restrict the supply of liquidity, an externality that enhances competition may improve welfare.

There are strong reasons to believe that exchanges may well supply less than the first best quantity of output. It has long been believed that exchanges possess strong natural monopoly tendencies (Stoll, 1992). Pirrong (1999) (1) shows that exchanges may indeed be natural monopolies that include too few members and inefficiently restrict the supply of risk bearing and liquidity and (2) presents empirical evidence consistent with these predictions. These results weaken the case against third markets.

Pirrong (1999) does not include informed trading. Any analysis of third markets must. This article does so in the context of a batch auction model with risk averse market makers, and is able to analyze the positive and normative implications of third markets. Moreover, unlike previous research, this article endogenizes financial market structure and third market policies. Rather than assuming that the exchange is a monopoly or competitive, it solves for market "macrostructure"–the number of exchanges, exchange size (the number of liquidity suppliers on the exchange) and the strategies of the third market dealers–from fundamental microstructural considerations.

The results of the analysis are striking. First, when third markets are precluded, with informed trading all transactions occur on a single exchange. Second, this exchange limits its membership (and thereby its supply of risk bearing services) in order to increase the profits of its members. Thus, the basic results of Pirrong (1999) carry over when informed trading is allowed. Third, a third market always exists if the monopoly exchange maintains the same membership when the third market enters as when it is precluded regardless of whether the third market "free rides" on exchange price discovery or not. Fourth, if the monopoly exchange maintains the same membership when the third market enters, and entry to the third market is unrestricted, total execution costs paid by uninformed traders decline regardless of whether information free riding occurs, although the costs of some uninformed traders rise. Fifth, total welfare is higher when a free riding open entry third market exists, and welfare is unchanged if there exists an open entry third market that cannot free ride. Total surplus rises because risk is borne more efficiently when the free riding third market exists. Sixth, if the third market dealers limit entry, the existence of a third market reduces welfare unless its competitive threat induces the exchange to expand. Seventh, prices are actually more informative when the third market exists regardless whether or not the third market limits entry. Eighth, if the share of trading accounted for by the third market is large enough, the exchange may increase membership to eliminate this competition. If this occurs, the execution costs paid by the uninformed fall and total welfare rises due to more efficient risk bearing.

All of these results imply that competition from a third market improves welfare when there is free entry onto the third market, and may do so even if the third market limits the number of dealers. Moreover, the increase in welfare is typically largest when there is free flow of price information from the exchange to the third market. Therefore, the theory implies that policies intended to facilitate the flow of information from markets where informed trading occurs increase rather than reduce welfare.

The remainder of this article is organized as follows. Section 2 extends the model of Pirrong (1999) to allow for informed trading to derive the equilibrium number and sizes of exchanges when there is no third market. Section 3 analyzes the effects of open entry third market trading under varying assumptions about what information from the exchange is available to third market dealers. Section 4 analyzes the effects of a third market that limits the number of dealers that trade on it. Section 5 summarizes the work.

2 Market Structure With No Third Market

To understand the role of third markets and their efficiency implications, it is first necessary to understand the nature of equilibrium when there is no third market. The welfare effects of the introduction of the third market depend on the efficiency of the no-third market equilibrium. If the no-third market equilibrium is first best, introduction of the third market cannot improve welfare, and if the creation of a third market generates externalities its creation will actually reduce welfare. Conversely, if the no-third market equilibrium is not first best, the introduction of a third market can improve welfare and externalities may lead to further improvements.

To understand these points, consider trading in a security or financial contract. The unconditional distribution of the value of the traded instrument is normal with a mean of 0 and a variance equal to σ^2.

Two types of agents desire to trade the instrument. First, there is a large (but finite) number of noise traders. Net noise trader demand for the asset is perfectly inelastic, and is a normal random variable with mean 0 and variance

S. Individual noise trader demands are uncorrelated, so the variance of the sum of several noise trader's demands is equal to the sum of the variances of their individual demands. Noise trader demand and the value of the asset are orthogonal. There are also K risk neutral informed traders who know the true value of the asset.

I first consider the case where the instrument must be traded on an exchange. For example, the Commodity Exchange Act requires that all futures contracts be traded on a government-approved exchange; this has effectively eliminated bucket shops. Following Pirrong (1999), I treat exchanges as coalitions of intermediaries who supply liquidity to a securities or derivatives market. There is a set of potential liquidity suppliers $\mathbf{L} = \{1, 2, \ldots, N\}$. Each liquidity supplier $j \leq N$ is risk averse, with a constant absolute risk aversion coefficient α_j. Equivalently, the risk tolerance of intermediary j is $t_j = 1/\alpha_j$. Moreover, wlog $t_j > t_k$ for $j < k$. That is, intermediaries are ordered by decreasing risk tolerance. The total supply of risk bearing capacity (i.e., aggregate risk tolerance) is $T^A = \sum_{i=1}^{N} t_i$.

The assumption of risk averse market makers is realistic and important.[3] Limits on the capital of market makers constrain their ability to bear inventory risk and induce them to act as if they are risk averse. It is well documented that market makers in securities are compensated for bearing risk, which would not occur if they were risk neutral. Moreover, the existence of limits to market makers' risk bearing capacity implies that the size of exchanges has efficiency implications; risk is borne inefficiently if exchanges

[3] See Diamond-Verrecchia (1991), Admati-Pfleiderer (1991), Subrahmanyam (1991), and Brown-Zhang (1997) for examples of models involving market maker risk aversion.

restrict the number of members to be suboptimally small.

The trading process is as follows. First, the intermediaries form coalitions–exchanges. Once exchanges are formed, the noise traders choose which one to trade on. Noise traders choose to trade on the exchange that minimizes their expected execution costs.[4] Trade then takes place on the exchanges in batch auctions. In the auctions, the the noise traders submit market orders to the exchange they have chosen, whereas the informed traders can submit market orders to all exchanges. Market makers on an exchange observe total order flow on that exchange and condition their trades on this information. The real value of the asset is revealed after exchange trading ends.

Consider the trading process when two exchanges form; the analysis can be extended readily to incorporate an arbitrary number of exchanges. The total risk tolerance (i.e., the sum of the risk tolerances) of the members of exchange 1 is T_1, and the total risk tolerance of exchange 2 is $T_2 < T_1$. Assume initially that fraction q_1 of the noise traders have chosen to trade on exchange 1, and $q_2 = 1 - q_1$. Due to the independence of noise trader demands, the variance of noise trader order flow on exchange 1 is $S_1 = q_1 S$, and the variance of noise trader order flow on exchange 2 is $S_2 = q_2 S$.

Upon learning v the informed traders conjecture that the price on ex-

[4]This is a simplified version of the Chowdhry-Nanda (1991) framework. In some versions of their model they preclude some noise traders from choosing where to trade. In contrast, all noise traders in the present model are "discretionary" in their terminology. Chowdhry-Nanda also include a large noise trader who can split orders between exchanges. When all noise traders in their model can choose where to trade, the large noise trader ends up trading on a single market. Since this result would obtain in the present model, I simplify the approach by considering only small discretionary noise traders. Unlike Chowdhry-Nanda, I assume that market makers are risk averse.

change i, $i = 1, 2$ is a linear function of order flow:

$$P_i = \lambda_i(\sum_{k=1}^{K} w_{ik} + z_i) \tag{1}$$

where w_{ik} is the order that the informed trader k submits to exchange i, z_i is net noise trader demand on exchange i, and λ_i is a constant. Given this conjecture, the informed trader l chooses w_{il}, $i = 1, 2$ to maximize:

$$V_i = w_{il} E[v - \lambda_i(w_{il} + z_i + \sum_{k \neq l} w_{ik})] \tag{2}$$

where the expectation is taken over z_i. Given that v and z_i are orthogonal, the symmetric solution of the informed traders' maximization problems implies that:

$$w_{il} = \beta_i v = \frac{v}{(K+1)\lambda_i} \quad \forall \quad l \leq K \tag{3}$$

Conditional on order flow liquidity supplier j on exchange i chooses his trade y_j to maximize his risk-adjusted profit. This implies:

$$E\Pi_j = \max_{y_j}\{y_j E[v - P | K\beta_i v + z_i] - \frac{.5\hat{\sigma}^2 y_j^2}{t_j}\} \tag{4}$$

where $\hat{\sigma}^2$ is the variance of v conditional on $K\beta_i v + z_i$, and where P is given by (1). Note that due to the normality of v and z_i, $E[v|K\beta_i v + z_i]$ is given by the regression of v on $K\beta_i v + z_i$. Thus,

$$E[v|K\beta_i v + z_i] = \frac{K\beta_i\sigma^2}{K^2\beta_i^2\sigma^2 + S_i}(K\beta_i v + z_i) \tag{5}$$

Moreover, by (1), $E[P|K\beta_i v + z_i] = \lambda_i(K\beta_i v + z_i)$, and

$$\hat{\sigma}^2 = \frac{S_i\sigma^2}{K^2\beta_i^2\sigma^2 + S_i} \tag{6}$$

Therefore,

$$y_j = \frac{t_j[\frac{K\beta_i\sigma^2}{K^2\beta_i^2\sigma^2 + S_i} - \lambda_i](K\beta_i v + z_i)}{\hat{\sigma}^2} \tag{7}$$

Call $\mathbf{L_i}$ the set of intermediaries on exchange i. Market clearing implies:

$$z_i + \sum_{j \in \mathbf{L_i}} y_j + K\beta_i v = 0. \tag{8}$$

Thus,

$$\frac{T_i[\frac{K\beta_i\sigma^2}{K^2\beta_i^2\sigma^2+S_i} - \lambda_i](K\beta_i v + z_i)}{\hat{\sigma}^2} + K\beta_i v + z_i = 0 \tag{9}$$

where $T_i = \sum_{j \in \mathbf{L_i}} t_j$. This, in turn, implies:

$$\lambda_i = \frac{\hat{\sigma}^2}{T_i} + \frac{K\beta_i\hat{\sigma}^2}{S_i} \tag{10}$$

This expression shows that the sensitivity of price to order flow in exchange i consists of two parts. The first part is the cost that intermediaries incur to absorb order flow imbalances. The second term is the adverse selection cost incurred when trading with informed traders.[5]

Substituting from (6) for $\hat{\sigma}^2$ and from (3) for λ_i produces a quadratic equation in β. The positive root of this equation gives the equilibrium β:

$$\beta_i = -\frac{(K+1)S_i}{2KT_i} + \frac{1}{\sigma}\sqrt{\frac{S_i^2\sigma^2(K+1)^2}{4K^2T_i^2} + \frac{S_i}{K}} \tag{11}$$

Substituting this expression for β_i into (10) gives an expression for λ_i.

Taking the derivative of expression (10) after substituting $\lambda_i = 1/(K+1)\beta_i$ implies that $d\beta_i/dS_i > 0$, that is, the intensity of informed trading on exchange i is increasing in the variance of noise trader order flow. Therefore, $d\lambda_i/dS_i < 0$. Moreover, since $dS_i/dq_i = S > 0$, $d\lambda_i/dq_i < 0$. This means that the sensitivity of price on exchange i to order flow is smaller, the larger the fraction of noise traders select to trade on exchange i. It is also straightforward to show that $d\lambda_i/dT_i < 0$. That is, price on exchange i is less sensitive

[5] This is similar to the result in Brown and Zhang (1997).

to order flow, the larger the total risk tolerance of the members of exchange i. Finally, $d\hat{\sigma}^2/dS_i > 0$; conditional variance is increasing in the variance of noise trader order flow.[6]

These results determine where noise traders choose to transact. Each noise trader takes the expected cost of execution on each exchange as a given and chooses to trade where the per-noise trader cost of execution is smallest. The per-noise trader expected execution cost on exchange i is given by:

$$x_i(q_i, T_i) \equiv \frac{\lambda_i S_i}{q_i} = \frac{\lambda_i q_i S}{q_i} = \lambda_i S \tag{12}$$

Since λ_i is decreasing in q_i, exchanges are subject to increasing returns to scale; per uninformed trader expected execution costs are smaller, the larger the number of noise traders that choose to trade on that exchange.

This analysis implies that there are three possible equilibria in this market when noise traders choose where to trade simultaneously. Figure 1 illustrates these equilibria. The horizontal axis in the figure is q_1, the fraction of noise traders that choose to trade on exchange 1. The downward sloping curve is $\lambda_1 S$, the average noise trader execution cost on exchange 1; the downward slope indicates the economies of scale. The upward sloping curve is $\lambda_2 S$, the average noise trader execution cost on exchange 2. The upward slope also indicates economies to scale, as an increase in q_1 implies a decrease in q_2, and thus a rise in execution costs on that exchange.

Figure 1 indicates that there are three potential equilibria if all noise traders choose where to trade simultaneously. The first equilibrium, which is

[6] This last result is proved in the appendix. With risk neutral market makers the conditional variance does not depend on order flow variance. With risk averse market makers, in contrast, prices are less informative, the greater the noise trading volume.

unstable, occurs at the intersection of the two curves. The second equilibrium occurs at $q_1 = 1$, i.e., all noise traders congregate at exchange 1. The third equilibrium is $q_1 = 0$, i.e., all noise traders choose to trade on exchange 2.

This analysis indicates that exchange markets with informed trading are "tippy." That is, all traders choose one exchange or the other. The intermediate equilibrium with $1 > q_1 > 0$ is not stable; any perturbation of q_1 away from this point tends to "tip" the noise traders towards one exchange or the other. Thus, stable equilibria in this market are monopoly equilibria.[7]

With M exchanges there may be M stable monopoly equilibria. How is one to choose among them? There are two approaches that imply that the only relevant equilibrium results in all noise traders choosing the exchange with the highest aggregate risk tolerance T_i.

In the first approach, the noise traders choose where to trade sequentially rather than simultaneously. Moreover, each noise trader knows all relevant market characteristics including σ^2, S, K and the T_i. This implies that payoffs are common knowledge. Each also can observe the choices of those who select an exchange prior to him. In this case, traders would like to coordinate their move to the highest T_i exchange due to the fact that execution costs for all noise traders are minimized when they choose to trade there. This implies that the corollary to Proposition 1 and Proposition 3 in Farrell and Saloner

[7] Pagano (1989), Admati-Pfleiderer (1988) and Chowdhry and Nanda (1991) present models in which trading consolidates. In Pagano, risk sharing provides the motive, whereas in Admati-Pfleiderer and Chowdhry-Nanda mitigation of adverse selection drives consolidation. Both forces are at work in the present model. Neither the Pagano, Chowdhry-Nanda nor Admati-Pfleiderer models include risk averse market makers which are the source of the key results that follow. Their "intermediary free" models do not permit analysis of intermediary coalitions such as exchanges or of the welfare consequences of exchange membership limits.

(1985) both hold. These in turn imply that the unique perfect equilibrium if noise traders move sequentially is to select the exchange with the highest aggregate risk tolerance. This result holds if the sequence in which noise traders move is specified exogenously or determined endogenously.

In the second approach, there is some mechanism to coordinate the movement of noise traders to the exchange that minimizes their trading costs.[8] One way to rationalize this result is most readily seen in the market for a corporation's stock. The corporation desires to minimize liquidity costs. It can influence the cost of liquidity by choosing where to list its stock. The firm can coordinate the flow of noise traders to the low cost exchange by listing its stock on that exchange.[9]

If $T_1 > T_2$, the fact that execution costs are decreasing in an exchange's total risk tolerance implies that $x_1(1, T_1) < x_2(1, T_2)$. Therefore, in this case, the lowest cost equilibrium involves all noise traders choosing to trade on the exchange with the greatest risk bearing capacity–exchange 1.

This fact influences the equilibrium allocation of intermediaries among ex-

[8]This assumption is often used in multi-stage games when a multi-equilibrium coordination game is played at one of the stages (after the first stage). A prominent example is Katz and Shapiro (1986). Fudenberg and Tirole (1999) claim that this is the "standard equilibrium selection in static network models."

[9]The multiplicity of equilibria in network coordination games leads some, notably David (1985), Arthur (1994), and Krugman (1994), to speculate that markets may become trapped in inefficient equilibria. In contrast, Leibowitz and Margolis (1999) and Katz and Shapiro (1986) argue that efficient producers who "own" their network have an advantage in coordinating movement of consumers to their network. In this view, the high T_i-low cost equilibrium is focal. There are recent examples of exchanges coordinating a "tipping" of a financial market. For example, in 1997-1998, through extensive marketing efforts and introductory price discounts, the Swiss-German exchange Eurex induced all of the users of the previously dominant London International Financial Futures Exchange (LIFFE) to trade German government bond futures on Eurex instead.

changes. This allocation must satisfy several equilibrium conditions.[10] First, in equilibrium no additional exchanges must be able to enter profitably. That is, no coalition of intermediaries outside the equilibrium exchange(s) can earn a profit for each of its members by forming an exchange. Second, the members of an equilibrium exchange cannot increase their profits by altering the size of their exchange's membership. Third, if a total of $\hat{L}$ intermediaries belong to exchanges, then the equilibrium allocation requires intemediaries $\{1, \ldots, \hat{L}\}$ to belong to exchanges. This condition reflects the fact that exchange memberships are transferrable. If intermediary j is a member of an exchange, and intermediary $i < j$ is not, there is a price at which i could buy the membership from j that makes both parties better off.[11]

The only coalition of intermediaries that satisfies these conditions is $\mathbf{L}^* = \{1, 2, \ldots, L^*\}$, where $\sum_{j=1}^{L^*} t_j > .5T^A$, and $\sum_{j=1}^{L^*-1} t_j < .5T^A$. The intermediaries in $\mathbf{L}^*$ account for just over half of the total risk tolerance; if intermediary L^* were excluded from the coalition, the exchange would offer less than one-half of total risk tolerance. This exchange can attract all noise traders because no other exchange can offer lower execution costs (since any other exchange has lower total risk tolerance). Therefore, this exchange is immune from entry. Moreover, an exchange consisting of some strict subset of the intermediaries in $\mathbf{L}^*$ would attract no noise traders because another exchange with greater total risk tolerance would enter, capture all of the order flow, and earn a profit; thus, such a subset cannot be an equilibrium exchange.

Furthermore, it is possible to show that the members of $\mathbf{L}^*$ are harmed

[10] See Pirrong (1999) for a more formal statement of these conditions.

[11] Expression (16) below shows that a member's profit is increasing in t_j, which implies the stated result.

by the addition of more members. To see why, first note that by (7) and (9), the position of trader $j \in \mathbf{L}^*$ is equal to

$$y_j = -\frac{t_j}{T_1}(K\beta v + z) \tag{13}$$

where subscripts are suppressed because there is only a single exchange.

The expected risk-adjusted profit of any member $j \in \mathbf{L}^*$ is given by:

$$E(\Pi_j) = E[y_j(v - \lambda_1(K\beta v + z)) - \frac{.5\hat{\sigma}^2 y_j^2}{t_j}] \tag{14}$$

where this expectation is taken over the unconditional joint distribution of v and z. Therefore,

$$E(\Pi_j) = -\frac{t_j K\beta\sigma^2}{T_1} + \frac{t_j}{T_1}[\lambda_1 - \frac{.5\hat{\sigma}^2}{T_1}](S + K^2\beta^2\sigma^2) \tag{15}$$

After some additional substitution, this reduces to:

$$E(\Pi_j) = \frac{.5 t_j \sigma^2 S}{T_1^2} \tag{16}$$

Note that the expected profit of the exchange member does not depend on informed trading in any way. Noise traders bear informed trading costs, and exchange members earn profits by supplying liquidity to the noise traders.

Expression (16) implies that $dE(\Pi_j)/dT_1 < 0$. Since this holds for $T_1 = T_1^*$, the profitability of an exchange member $j \in \mathbf{L}^*$ declines if additional members are added; increasing membership beyond $\mathbf{L}^*$ increases the competition faced by those in $\mathbf{L}^*$, and thereby reduces their profits.

Together, these results imply that in equilibrium, the exchange consists of the intermediaries $j \in \mathbf{L}^*$. This implies that total equilibrium risk tolerance is $T_1^* = \sum_{j \in \mathbf{L}^*} t_j \approx .5T^A$. Given the formation of such a coalition, no other exchange can enter profitably. Moreover, both increases and decreases in the membership of this coalition reduce the profits of its members. Due to the restriction in membership, exchange members earn rents.

There is empirical evidence that is consistent with these predictions. First, historically exchanges have limited membership, and exchange members have resisted expansions in membership because of its effect on their profits. Second, as documented in Pirrong (1999), exchange members earn economic rents. Indeed, these rents are large even in comparison to those earned by manufacturing firms that plausibly exercise market power.

Thus, the equilibrium exchange is a monopoly that limits the number of intermediaries it admits to increase the profits of its members. Optimal risk bearing requires the exchange to admit all intermediaries $\{1, 2, \ldots, N\}$. The appendix shows that total cost with the monopoly exchange is $.5\sigma^2 S/T_1$. Total costs equal the expectation of execution costs minus member profits minus informed trading profits. The cost of operating the market is minimized, and surplus is maximized, when $T_1 = T_A$. The exchange has no incentive to grow this large, however. By limiting membership to $\mathbf{L}^*$, it is immune from competitive entry by another exchange and does not dissipate profits as would be the case if more intermediaries were admitted. Therefore, limits on risk bearing cause deadweight losses.

3 The Third Market: The Free Entry Case

The foregoing analysis implies that intermediaries $\{L^* + 1, \ldots, N\}$ are excluded from membership of the exchange. If allowed, however, these intermediaries can supply liquidity off exchange by serving as third market dealers.

It is evident that these third market dealers cannot survive if they offer to trade with anyone and everyone, including the informed traders. If they do not exclude the informed trader, the analysis of the prior section implies that the execution costs noise traders pay when trading with third market dealers are higher than the execution costs incurred when trading on the exchange because when an exchange with membership $\mathbf{L}^*$ forms, individually and collectively the third market dealers have lower risk tolerance than the members of the exchange. Therefore:

Result 1 *To survive, the third market dealers must be able to exclude informed traders.*[12]

Formally, assume that fraction q^* of the noise traders can prove that they are not informed, and fraction $1-q^*$ cannot prove to the third market dealers that they are uninformed. To survive, the third market dealers must limit their dealings to the noise traders who can prove they are not informed.[13] That is, third market dealers have developed a technology that allows them to identify some, but not all, of the uninformed. For example, small noise traders may be able to represent credibly that they are not informed, whereas large traders may not be able to so represent. By dealing only with small noise traders, third market dealers avoid being "picked off" by the informed. Reputation and trading constraints (such as the "no bagging" constraint

[12] This analysis only considers execution costs, and as a result traders defect to the third market only if it offers them more liquidity than the exchange. The fixing of supercompetitive commissions by exchanges prior to 1975 provided another incentive for trading on third markets, especially by institutions. Since the elimination of fixed commissions, according to Stoll (1994) the primary motive for trading on third markets is to avoid adverse selection costs as modeled herein.

[13] Admati-Pfleiderer (1991) also assume the existence of an exogenous number of noise traders who can credibly disclose that they are uninformed.

discussed in Seppi, 1990) are other means by which some (but not all) *large* uninformed traders can identify themselves as such.

Note that the objective of this article is not to analyze how screening mechanisms work. It takes their existence as given. The literature on cream skimming third markets assumes such mechanism exist, and to join this debate, I make the same assumption.[14] Moreover, Result 1 implies that without some screening mechanism, the third market cannot exist. Finally, there is abundant empirical evidence that off-exchange venues sharply limit informed trading. Easley *et al* (1996) present empirical evidence showing that orders executed on one third market (Cincinnati) are substantially less informative than orders submitted to the NYSE; Hasbrouck (1995) estimates that NYSE trades account for 93 percent of the information revealed by trading; Huang and Stoll (1994) and Bessembinder and Kaufman (1997) show that NYSE trades have larger persistent price impacts than off-NYSE trades in listed stocks; Madhavan and Cheng (1997) present evidence consistent with they hypothesis that block trades intermediated "upstairs" are preferred by traders who can signal credibly that they are uninformed; and Smith *et al* (2001) demonstrate that "upstairs" trades of listed shares on the Toronto Stock Exchange have virtually no information content, whereas trades on the exchange trading mechanism proper do. All of these results are all consistent with cream skimming. Therefore, the screening assumption is descriptively accurate and strongly empirically supported. Consequently, when analyzing the effects of cream skimming alternative trading venues it is appropriate to

[14]There is widespread agreement that off-exchange venues attempt to exclude the informed. See O'Hara (1997) for a discussion of the "cream skimming" efforts of third market participants such as Madoff Securities and crossing networks such as Posit.

assume the existence of a screening device without analyzing its mechanics in detail.[15]

The addition of the third market requires modification of the trading timeline. First, a coalition of intermediaries form an exchange; those excluded from the exchange become third market dealers. The analysis of section 2 implies that only one exchange forms. Once the exchange forms, the noise traders who can use the third market choose whether to trade on exchange or the third market. Noise traders choose to trade in the venue that minimizes their expected execution costs. The noise traders who choose to transact on the exchange and the informed trader submit market orders to the exchange. Exchange members observe total order flow on that exchange. Trade then takes place on the exchange in a batch auction. Trading then takes place on the third market. Noise traders who choose to transact on the third market submit market orders there. These market orders are executed in a batch auction. The asset's true value is revealed after the third market clears.

Since third market dealers exclude informed trading, they bear no adverse selection costs. The cost that they incur to supply liquidity to those who can use the third market therefore depends on third market dealers' estimate of the variance of the value of the instrument. This depends on the information available to the third market dealers. I consider two information regimes. In the first regime, third market dealers observe the price in the exchange market auction. In the second regime, they cannot observe the exchange price. Therefore, in the first regime, the price variance estimated by third market dealers is $\hat{\sigma}^2$, whereas in the second regime it is $\sigma^2 > \hat{\sigma}^2$.

[15]Seppi (1990) presents a model of a screening process.

Since there is no informed trading in the third market, an analysis like that used to derive (10) implies that the λ of the third market is $\lambda_{31} = \hat{\sigma}^2/T_3$ in the first regime, and $\lambda_{32} = \sigma^2/T_3$ in the second, where T_3 is the total risk tolerance of third market dealers. Therefore, in the first regime, the expected execution cost of each trader who chooses to trade in the third market is:

$$x_{31}(T_3) = \frac{\hat{\sigma}^2 S}{T_3} \tag{17}$$

whereas in the second regime it is:

$$x_{32}(T_3) = \frac{\sigma^2 S}{T_3} \tag{18}$$

The relative costs of trading on exchange and on the third market depend on the sizes of the exchange and the third market. Assume initially that that exchange membership is given by the coalition $\mathbf{L}^*$, where as before this coalition offers just more than half of the total risk tolerance. Also assume that there is free entry onto the third market. That is, unlike the exchange, the third market does not restrict the number of dealers who trade on it.[16] This assumption accurately reflects the history of many third markets, including the New York Open Board (the "Open" in the title signifying that the Board was open to all, unlike the NYSE), the Chicago Open Board, the curb market, the bucket shops, and modern third markets.

With free entry to the third market and an exchange that maintains membership at $\mathbf{L}^*$, $T_3 = T_A - T_1^* \approx T_1^*$. A comparison of (17) to (10)-(12) shows immediately that average execution costs on the exchange assuming all noise traders trade there is higher than average execution cost on the third

[16]It is an equilibrium for the exchange to maintain membership at L^* for some q^*. Section 4 discusses the implications of restrictions on the size of the third market.

market under the first information regime. That is, $x_1(1, T_1^*) > x_{31}(T_3)$. Moreover, since $x_1(q_1, T_1^*)$ is decreasing in q_1, $x_1(1 - q^*, T_1^*) > x_1(1, T_1^*) > x_{31}(T_3)$. Average execution costs are lower on the third market than on exchange in the first information regime because those who trade in the third market bear no adverse selection costs.

This analysis implies that in the first information regime all noise traders who can use the third market will do so if the membership of the exchange remains unchanged. When exchange membership is $\mathbf{L}^*$ and third market dealers can observe the outcome of exchange trading, switching to the third market reduces noise trader execution costs. Thus:

Result 2 *If (a) third market dealers observe the price determined in the exchange auction and (b) exchange membership is* $\mathbf{L}^*$*, fraction* q^* *of noise trading takes place on the third market.*

The appendix proves that the same outcome occurs in the second regime:

Result 3 *If (a) third market dealers cannot observe the price determined in the exchange auction and (b) exchange membership is* $\mathbf{L}^*$*, fraction* q^* *of noise trading takes place on the third market.*

The foregoing implies that if an exchange chooses the same membership in the presence of the third market as it would in its absence, the third market attracts fraction $q^* > 0$ of noise trading and the exchange attracts fraction $1 - q^*$. Results 2 and 3 and the fact that λ_1 is increasing in q^* together imply that the creation of a third market reduces execution costs for the noise traders who can switch to the third market, but raises the execution costs of those who cannot. The effect of the entry of a third market on total noise trader execution costs therefore depends on which effect dominates.

Total noise trader execution costs on exchange and third market are:

$$x_i^*(T_1^*) = S[(1-q^*)\lambda_1(T_1^*, 1-q^*) + q^*\lambda_{31}(T_3)] \tag{19}$$

where λ_1 is given by (10) and (11) with $S_1 = (1-q^*)S$, and λ_{3i} is given above with $i = 1, 2$; the notation is expanded to recognize the dependence of the λ's on q^* and T_1 and T_3. When third market dealers can observe the outcome of exchange trading,

$$x_1^*(T_1^*) = S\frac{\hat{\sigma}^2(1-q^*)}{T_1^*} + K\beta_1(1-q^*)\hat{\sigma}^2(1-q^*) \tag{20}$$

where $\hat{\sigma}^2$ and β_1 are now written as functions to recognize explicitly their dependence on q^*. Note that as shown in section 2, $\hat{\sigma}^2(1-q^*) < \hat{\sigma}^2(1)$ and $\beta_1(1-q^*) < \beta_1(1)$. Therefore, $x_1^*(T_1^*) < x_1(1, T_1^*)$. This proves:

Result 4 *If exchange membership is* $\mathbf{L}^*$ *and third market dealers can observe the outcome of exchange trading, introduction of a third market unambiguously reduces total noise trader execution costs.*

Thus, although some noise traders are harmed by the introduction of a third market, in aggregate noise traders are better off when a third market is introduced and third market dealers can condition their trades on the price determined in exchange trading. Therefore, contrary to the critics of cream skimming, the creation of a free riding third market makes liquidity traders better off in aggregate.[17]

[17] It is possible to show that noise trader execution costs are lower in the second regime

Indeed, the introduction of the third market increases total surplus if the third market free rides, but has virtually no welfare effect when free riding is precluded. Given the inelasticity of noise trader demands, total surplus is maximized by minimizing the cost of operating the market. Cost equals noise trader execution costs minus informed trader profits minus risk-adjusted market maker profits. The appendix shows that with free riding total cost equals:

$$TC_{31} = \frac{.5\sigma^2(1-q^*)S}{T_1^*} + \frac{.5\hat{\sigma}^2 q^* S}{T_3}$$

Since $\hat{\sigma}^2 < \sigma^2$ and $T_3 \approx T_1^*$, TC_{31} is smaller than the total cost incurred when there is no third market, $.5\sigma^2 S/T_1^*$. Thus, the free riding open entry third market unambiguously improves welfare.[18] This improvement is attributable to the fact that the third market improves the efficiency of risk bearing. The third market dealers supply additional risk bearing capacity to the market. Although this reduces the profits of the exchange members, their loss is more than offset by the gains realized by noise traders and third market dealers.

When there is no free riding, total cost equals:

$$TC_{32} = \frac{.5\sigma^2(1-q^*)S}{T_1^*} + \frac{.5\sigma^2 q^* S}{T_3} \approx \frac{.5\sigma^2 S}{T_1^*}$$

than when a third market is precluded. Some empirical evidence suggests that execution costs are actually smaller on exchange than on the third market. This is a puzzling result; why do those who can trade on a cheaper market trade on a more expensive one? This result is likely due to an incomplete comparison of executions on exchange and the third market. Using a more comprehensive comparison of execution costs and attributes, Battalio, Hatch, and Jennings (2000) show that execution costs for trades made by one third market dealer (Trimark) are smaller (by a statistically significant amount) than execution costs on exchange if payment for order flow is netted out (as it should be). This finding is consistent with the results of the analysis herein.

[18] This is similar to the Rothschild-Stiglitz-Wilson result that separating equilibria can sometimes improve welfare. Note that $dTC_{31}/dq^* < 0$, i.e., cost is decreasing in the size of the third market.

Therefore, the creation of an open entry, *non-free riding* third market does not improve welfare. (In fact, since T_3 is slightly smaller than T_1^*, welfare declines slightly). Although the creation of a third market increases risk bearing capacity, these new market makers are relatively inefficient due to their information disadvantage. The cost disadvantage almost exactly offsets the beneficial effect of the entry of new risk bearing capacity on surplus.

Given K, the exchange's price is actually more informative when the third market exists because (1) the conditional price variance decreases as S_1 declines, and (2) the creation of the third market reduces S_1.[19] Holding T_1^* fixed, reducing S_1 reduces the magnitude of noise trading-driven price fluctuations. Thus, again contrary to the arguments against third markets, in this model they actually improve the informational efficiency of prices.

Moreover, if information is costly and the number of informed traders is therefore endogenous, the existence of a third market improves welfare for another reason; it reduces excessive expenditures on information. The expression for total cost derived in the appendix implies that informed trading

[19] Holding exchange membership fixed, prices become completely uninformative if $q^* = 1$; total cost is the same under this outcome as when the third market is precluded. However, as I show below, if $q^* = 1$ the exchange will always expand membership rather than suffer a complete loss of business to the third market. Thus, a complete breakdown of the exchange is not possible in this model. It should also be noted that the conclusions regarding price informativeness depend on the assumptions that information is exact and costlessly obtained, and the informed trader is risk neutral. Different results may obtain under different assumptions. The analysis is sufficient to show, however, that cream skimming does not necessarily imply a reduction in price informativeness. Note, moreover, that since in this model price informativeness provides no social value in the absence of a third market, prices are necessarily *too* precise when private information is costly. Also note that if informed traders are risk averse, reducing noise trader order flow to the exchange reduces execution risk, induces the informed to trade more intensively, and thereby increases price informativeness. Thus, assuming of risk neutral informed traders causes understatement of the impact of the third market.

does not improve welfare when there is no third market. Privately informed trading reduces market makers' cost of holding a position of a given size, but increases the size of the positions they must hold; in this model, these effects have equal and opposite effects on total cost. Therefore, any expenditures on information lead only to wealth transfers and thus are a pure loss. In contrast, if a third market exists, informed trading reduces TC_{31} because it reduces $\hat{\sigma}^2$. Consequently, there is a social return to expenditures on information when there is a third market. Moreover, the third market reduces the private return to information because informed traders have fewer noise traders to profit from. By reducing the private returns to information and increasing the social returns thereto, the third market reduces the deadweight loss attributable to excessive expenditures on private information.[20]

Together these results imply that an open entry third market that free rides on exchange price discovery improves market performance. This may seem counterintuitive as it implies that an externality–the free acquisition of costly trade information by the third market–improves welfare and makes prices more informative. This result obtains because we are in the world of the second best. The "tippiness" of the exchange market leads to a natural

[20]In this analysis information affects only trading costs. It is possible that more informative securities prices improve resource allocation in other ways, such as by leading to superior real investment decisions. See Hirshleifer (1971) for a discussion of these issues. Third markets have an ambiguous effect on such benefits because they have an ambiguous effect on the informativeness of prices when K is endogenous; holding K constant, third markets make prices more informative, but this effect can be offset by a third-market induced decline in informed trading. Also note that the assumption that noise trader demand is completely inelastic understates the potential benefits of third markets. In this case, private information does not impair risk allocation. In contrast, if noise trader demand is not completely inelastic, privately informed trading impedes risk sharing. The existence of a third market that (a) free rides, and (b) is not subject to adverse selection mitigates adverse selection problems and improves risk sharing.

monopoly that restricts the supply of risk bearing to enhance its members' profits. This is inefficient. The externality reduces the costs of enhancing the supply of risk bearing and thereby mitigates the inefficiency.

The foregoing analysis assumes that exchange membership does not change in response to the creation of a third market. In fact, an exchange may respond to the third market by increasing its membership. This is not merely a theoretical possibility. In the 1860s the NYSE faced vigorous competition from the Open Board, the Gold Board, and the curb market. The NYSE responded in 1869 by expanding its membership from 533 to 1060 to include the 527 participants on the Open and Gold Boards (Davis and Neal, 1998).

An exchange faces a tradeoff when considering expansion. Specifically, the analysis of section 2 implies that such an increase in exchange membership reduces its members' profits unless the addition of these new members is sufficient to drive the third market out of business. Note that there is always a $\hat{T}_1 > T_1^*$ such that $x_1(1, \hat{T}_1) \leq x_{31}(T^A - \hat{T}_1)$. That is, since execution costs on exchange are decreasing in T_1, and execution costs on the third market are increasing in T_1 (since increasing T_1 reduces T_3), there is always some critical exchange membership level that is sufficiently large to ensure that exchange execution costs when all noise traders trade there are lower than third market execution costs even assuming information spillovers. Under the assumption that noise trader choices are coordinated to minimize the costs of those who can choose trading venue, this implies that there is always some critical level of exchange membership sufficiently large to drive the third market out of business even when there are information externalities. Since third market execution costs are larger when there are no information externalities, *a* *fortiori* there is some level of exchange membership with associated total risk tolerance $\bar{T}_1$, $\hat{T}_1 > \bar{T}_1 > T_1^*$, sufficient to deny any order flow to the third market absent information externalities.

If the exchange membership finds it profitable to expand to drive out the third market, it is clear that welfare is higher than when the third market does not or cannot exist. This is true because as shown in the appendix total cost for $q_1 = 1$ is decreasing in T_1. Therefore, in the second best world, even the *potential* competition of the third market improves welfare. Indeed, *all* noise traders, not just those who can switch to the third market, are better off if the exchange expands membership to drive out the third market.

The incentive of the exchange to expand membership to drive out the third market depends on q^*. By expanding, the exchange members in $\mathbf{L}^*$ reduce the profits they reap from their "captive" noise traders who cannot utilize the third market. Only if the potential gain in profits obtained by attracting the third market's customers is sufficiently large to offset the loss in profit from "captive" customers will the exchange choose to expand. Therefore, an exchange may not expand if q^* is small, but may expand if q^* is large. Moreover, the incentive to expand depends on whether trading information spills over from the exchange to the third market because a larger expansion is required to eliminate the third market when there are information externalities than when there are not.

When there are spillovers, the exchange expands if and only if

$$\frac{.5t_j\sigma^2 S}{\hat{T}_1^2} > \frac{.5t_j\sigma^2(1-q^*)S}{T_1^{*2}}. \tag{21}$$

The left hand side of this inequality is the profit of exchange member $j \in \mathbf{L}^*$

if the exchange drives out the third market and therefore services 100 percent of the noise trader order flow. The right hand side is this member's profit if the exchange does not expand membership and allows the third market to service fraction q^* of the uninformed order flow. This simplifies to:

$$\frac{1}{\hat{T}_1^2} > \frac{1-q^*}{T_1^{*2}}. \tag{22}$$

If there are no spillovers, the exchange expands if and only if:

$$\frac{1}{\bar{T}_1^2} > \frac{1-q^*}{T_1^{*2}}. \tag{23}$$

There are three cases to consider:

1. The exchange chooses not to expand membership to eliminate the third market when there is no information externality. *A fortiori* the exchange will not expand if there is an information externality.

2. The exchange expands membership to eliminate the third market when there is no information externality, but does not expand membership to eliminate the third market when there is information spillover.

3. The exchange expands membership to eliminate the third market when there is an information externality. *A fortiori* the exchange also expands to eliminate the third market when there is no spillover.

Which outcome occurs depends on q^*. The outcomes are more conveniently analyzed numerically rather than formally.[21]

In the numerical examples, $S = 10$, $\sigma^2 = 3$, and $K = 1$. Market makers are distributed evenly along a line segment $[0, 1]$ at intervals of .0001. The

[21] See the appendix for a formal analysis.

risk tolerance of an intermediary depends on his location on the segment. Specifically, the risk tolerance of a market maker located at $i \in [0,1]$ is $10 - 4i$. In this case $T^A = 8$, and $L^* = .4385$. That is, if all intermediaries $j \in [0, .4385]$ join the exchange, it will account for just over 50 percent of total risk tolerance and thus be immune from competition by another exchange, whereas if only $j \in [0.4384]$ join, an exchange with membership $k \in [.4385, 1]$ could enter and capture the business of all noise traders. Therefore, $\mathbf{L}^* = [0, .4385]$.

First consider the case when $q^* = .1$. In this case, the third market serves a relatively small fraction of the noise traders. The primary exchange must expand its total risk bearing capacity to $\bar{T}_1 = 4.247$ to drive out the third market when there are no spillovers. The profit of member L^* equals 6.86 if the exchange expands membership to this level. If the exchange restricts its membership to $\mathbf{L}^*$, the profit of this member is 6.95. Expanding the exchange to drive out the third market would therefore reduces the profits of all $j \in \mathbf{L}^*$, so when $q^* = .1$, the exchange maintains a membership of $\mathbf{L}^*$ and allows the third market to exist. In this case, the average noise trader execution cost is 8.31 if there are information spillovers and 8.38 if there are not; noise trader execution costs equal 8.40 if there is no third market. Total cost is equal to 3.71 when the third market free rides. Total cost without the third market (and with a non-free riding third market) equals 3.75.

Now consider the case when $q^* = .2$. As before, to drive out the third market when there are no spillovers, the exchange must expand so that $T_1 = \bar{T}_1 = 4.247$. If it does so, the profit of the member L^* is 6.86; if the exchange does not expand to eliminate the third market, the profit of this member when $q^* = .2$ equals 6.18. Thus, when $q^* = .2$ it pays the exchange to expand to eliminate the third market when there are no spillovers. It does not pay to expand when there are spillovers. When $q^* = .2$, $\hat{T}_1 = 4.565$. The profit of member L^* when $T_1 = \hat{T}_1$ is 5.94, which is smaller than the 6.18 this member would earn if the exchange restricts membership to $\mathbf{L}^*$.

In this case, when there are no spillovers and the exchange expands, the noise trader execution costs equal 8.00 and total cost equals 3.53. When there are spillovers and the exchange does not expand, average noise trader execution costs are 8.20 and total cost is 3.67. Thus, when $q^* = .2$ noise trader costs and total cost are actually lower when there are no spillovers than when the third market can free ride because the exchange's incentive to expand membership depends on whether or not there are spillovers.

Finally, when $q^* = .35$ the exchange must choose membership $\hat{T}_1 = 4.62$ to eliminate the third market when it can free ride. The profit of member L^* equals 5.80 when the exchange so expands, whereas this member's profit equals only 5.02 if the exchange restricts membership to $\mathbf{L}^*$. Thus, for q^* sufficiently large, the exchange increases membership to eliminate the third market even when its dealers free ride. In this case, execution costs of noise traders fall to 7.49 and total cost equals 3.25, which is less than the cost of 3.58 incurred when the exchange does not expand, and the cost of 3.75 incurred when there is no third market.[22]

These examples demonstrate several points. First, in all cases, noise

[22] Another possibility arises if q^* is sufficiently large; the exchange may choose to restrict trading to those who can prove they are uninformed and maintain membership at $\mathbf{L}^*$. For large enough q^*, the exchange members' profits are larger when it restricts membership to $\mathbf{L}^*$ and trades only with the demonstrably uninformed than when it expands and trades with everyone.

traders' execution costs and total cost are lower when the open entry third market can exist than when it cannot. Second, when the third market attracts a sufficiently large market share, the primary exchange has an incentive to expand membership to reduce noise trader execution costs and reclaim business from the third market. Thus, even potential third market competition benefits noise traders and improves welfare. Third, the third market's ability to obtain information from the primary market influences the incentives of the exchange to expand. It is possible that welfare is highest when the third market cannot free ride off of the exchange's price discovery. This is not because the free riding is detrimental *per se.* Instead, the ability of the third market to free ride reduces the exchange's incentive to expand. For intermediate values of q^*, the exchange is more likely to expand when there is no free riding than when there is. Since expansion of the exchange provides greater benefits than the third market, eliminating the information externality makes noise traders better off if it induces the exchange to expand. This suggests that if the third market exists when its dealers cannot free ride, then exchanges should be required to disclose price information.

In summary, competition from alternative trading venues–third markets–reduces overall uninformed trading costs and can improve welfare when entry into these venues is unrestricted. Absent competition from third markets, exchanges are too small and risk is allocated inefficiently as a result. Free entry third markets increase the supply of risk bearing capacity and as a consequence reduce the trading costs of the uninformed and improve welfare when third market dealers free ride on the primary market's price. Although externalities of this sort would be suboptimal if the exchange did not inefficiently limit entry, the externalities improve welfare because exchanges have too few members. Therefore, the existence of an information externality is not sufficient to justify restrictions on the dissemination of trading information. In the second best world, information spillovers increase the competition that exchanges face and thereby can improve welfare.

4 The Third Market–Restricted Entry Case

Just as exchange dealers have an incentive to limit membership in order to increase their rents, third market dealers may have an incentive to limit entry. Although historically many third markets have not done so, it is worthwhile to consider the implications of such limits.[23]

At first blush, it would appear that a coalition of third market dealers that is just large enough to undercut by a small amount the trading costs offered on the exchange could enter successfully; results 1 and 2 imply that this coalition could offer less than T_1^* risk tolerance and still peel off the q^* noise traders who can prove they are uninformed. If the primary exchange maintains membership at $\mathbf{L}^*$, it is clear that noise trading costs rise and welfare declines in the face of entry of such a limited third market; the execution costs of the noise traders who must remain on exchange rise, while the execution costs of those who switch to the third market barely fall.

[23]This raises the question of why many third markets have not in fact limited entry. This is most likely due to the fact that exchanges historically have operated physically centralized auction markets from which it is possible to exclude non-members, whereas third markets largely are not and have not been physically centralized. The lack of physical centralization makes it difficult, if not impossible, to exclude traders and limit membership. Some third markets, such as the curb market, were physically centralized but conducted business in public spaces, which precluded exclusion. When the curb market moved indoors with the formation of the American Stock Exchange, it did limit membership.

This outcome cannot be an equilibrium, however. If the third market dealers form a coalition with risk tolerance just sufficient to undercut the exchange's trading cost when its membership is $\mathbf{L}^*$, a small increase in exchange membership is sufficient to deny the third market any business. Expression (21) holds in this case, and the exchange therefore expands.

To survive, therefore, the third market must undercut exchange trading costs by enough to make it unprofitable for the exchange to expand its membership in response. Assume that the third market chooses a membership that offers a risk tolerance $\mathcal{T}_3$ such that the exchange would have to expand membership to $\mathcal{T}_1$ to keep the q^* verifiable noise traders from defecting to the third market. By (21), the exchange has no incentive to expand risk tolerance this much if:

$$\mathcal{T}_1 \geq \frac{T_1^*}{\sqrt{1-q^*}} \tag{24}$$

First consider equilibrium when the third market free rides. Define $\mathcal{T}_1^* = T_1^*/\sqrt{1-q^*}$. Call $\lambda_1(T, q)$ the reciprocal of exchange market depth when it offers risk tolerance T and serves fraction q of the noise traders. Call $\hat{\sigma}^2(T, q)$ the conditional price variance under these conditions. With free riding, if the exchange expands membership to $\mathcal{T}_1^*$, demonstrably uninformed traders defect if the third market members supply total risk tolerance $\mathcal{T}_{31}$ such that:

$$\mathcal{T}_{31} \geq \frac{\hat{\sigma}^2(\mathcal{T}_1^*, 1-q^*)}{\lambda_1(\mathcal{T}_1^*, 1)} \tag{25}$$

Defining $\mathcal{T}_{31}^*$ as the smallest value of $\mathcal{T}_{31}$ that can be supplied by some coalition of third market dealers such that (25) holds with strict inequality, the exchange does not expand to drive out the third market if the latter offers

$\mathcal{T}_{31}^*$ in risk tolerance. In this case aggregate execution costs are:

$$S[(1-q^*)\lambda_1(T_1^*, 1-q^*) + q^* \frac{\hat{\sigma}^2(T_1^*, 1-q^*)}{\mathcal{T}_{31}^*}] \tag{26}$$

Without free riding, verifiable noise traders defect to the third market when the exchange expands risk tolerance to $\mathcal{T}_1^*$ if and only if:

$$\mathcal{T}_{32} \geq \frac{\sigma^2}{\lambda_1(\mathcal{T}_1^*, 1)} \tag{27}$$

If $\mathcal{T}_{32}^*$ is the smallest value of $\mathcal{T}_{32}$ that can be supplied by a coalition of third market dealers such that (27) holds strictly, when the third market offers risk tolerance of $\mathcal{T}_{32}^*$, aggregate execution costs are:

$$S[(1-q^*)\lambda_1(T_1^*, 1-q^*) + q^* \frac{\sigma^2}{\mathcal{T}_{32}^*}] \tag{28}$$

It is possible to show that expected average execution costs on the third market equal $S\lambda_1(\mathcal{T}_1^*, 1)$ in this case. This is an intuitive result. To survive, the third market must offer an execution cost low enough to undercut the exchange by so much that it is uneconomic for it to expand and drive out the third market. By cutting per noise trader costs to $S\lambda_1(\mathcal{T}_1^*, 1)$ the non-free riding third market ensures that the primary exchange will not expand.

Aggregate execution costs are smaller with no free riding when the third market limits entry because $\hat{\sigma}^2(T_1^*, 1-q^*)/\mathcal{T}_{31}^* > \hat{\sigma}^2(\mathcal{T}_1^*, 1-q^*)/\mathcal{T}_{31}^* = \sigma^2/\mathcal{T}_{32}^*$. The effect on execution costs of the additional risk bearing capacity the third market must supply to compete with the exchange when it cannot free ride more than offsets the effect of its information disadvantage.

Third markets with limited entry raise execution costs for some range of q^*. This is most easily seen in the no-free riding case. Without free riding, the first derivative of aggregate execution cost with respect to q^* is:

$$S[-\lambda_1(T_1^*, 1-q^*) - (1-q^*)\frac{\partial\lambda_1(T_1^*, 1-q^*)}{\partial q} + \lambda_1(\mathcal{T}_1^*, 1) + q^*\frac{\partial\lambda_1(\mathcal{T}_1^*, 1)}{\partial T}\frac{\partial\mathcal{T}_1^*}{\partial q^*}]$$

Since (1) $\lambda_1(T_1^*, 1-q^*) = \lambda_1(\mathcal{T}_1^*, 1)$ when $q^* = 0$, and (2) $\partial\lambda_1/\partial q < 0$, this expression is positive at $q^* = 0$. Thus, creation of a third market with limited membership raises aggregate execution costs for some q^*. Since aggregate trading costs with free riding exceed those without free riding, creation of a free riding third market increases execution costs even more.

Limited entry third markets also reduce welfare unless their competition induces the exchange to expand. Note that total cost with free riding is:

$$TC_{31} = \frac{.5\sigma^2(1-q^*)S}{T_1^*} + \frac{.5\hat{\sigma}^2(T_1^*, 1-q^*)q^*S}{\mathcal{T}_{31}^*}$$

Without free riding total cost is:

$$TC_{32} = \frac{.5\sigma^2(1-q^*)S}{T_1^*} + \frac{.5\sigma^2 q^* S}{\mathcal{T}_{32}^*}$$

Since $\mathcal{T}_{32}^* < T_1^*$, total cost with a limited entry third market is larger than when there is no third market. Moreover, $TC_{31} > TC_{32}$. Thus, third markets reduce welfare and free riding reduces welfare further if the third market limits entry and q^* is such that the exchange does not to expand.

Although limited entry third markets reduce welfare for some q^*, this is not true for all q^* because the potential competition from the third market can induce the exchange to undertake a welfare-increasing expansion. Note that $\mathcal{T}_{31}$ and $\mathcal{T}_{32}$ are constrained by the supply of potential market makers. With free riding, if $\mathcal{T}_{31}^* > T_A - \mathcal{T}_1^*$, the exchange always expands membership to prevent entry by the third market. Similarly, if $\mathcal{T}_{32}^* > T_A - \mathcal{T}_1^*$, the exchange expands when there is no free riding. In either case, the expanded exchange offers a total risk tolerance that is less than $\mathcal{T}_1^*(q^*)$ but larger than T_1^*. Thus, potential competition from a third market sometimes induces exchange expansion that reduces execution costs and increases welfare.[24]

Exchange expansion is most likely to occur with large q^* because $\mathcal{T}_1^*$ increases in q^*. That is, for sufficiently large q^* the competitive threat of a limited entry third market induces the exchange to expand, which reduces noise trader costs and increases welfare. Since third markets must be larger to survive when they cannot free ride, as noted in section 3 the exchange is more likely to expand when free riding is precluded than when it is allowed.

Further analysis demonstrates that total execution cost exhibits an inverted-U shape with a discontinuity at the q^* at which the exchange expands, as illustrated in Figure 2. The figure assumes that $S = 10$, $T_A = 8$, and $\sigma^2 = 3$. Given these parameters, although total costs first rise with q^*, there is a crucial q^* (larger with free riding) such that execution costs are smaller than those incurred when there is no third market, $S\lambda_1(T_1^*, 1)$. For a given q^*, the proportionate rise in execution costs is smaller, the larger are S and σ^2, and the smaller is T_1^*. Under these conditions, the exchange's restriction of risk bearing capacity is especially costly, and thus entry of the third market's capacity is especially valuable.

The foregoing implies that third markets that successfully restrict entry

[24]The choice of the third market's risk tolerance is also constrained by the ability of the exchange to implement its own screening mechanism. If the exchange can implement the screening technology at a cost per noise trader of c_e greater than that incurred by the third market, the third market must choose a T_3 that ensures that its per noise trader cost is smaller than $S\hat{\sigma}^2/T_1^* + c_e$. If c_e is small, this would require $T_3 \approx T_1^*$. Thus, third market membership limitations may not overturn the analysis of section 3 if the exchange can implement the screening mechanism at only a small cost disadvantage.

increase execution costs and reduce welfare unless their competitive threat is so great as to cause the exchange to expand. If competition from the third market causes exchange expansion, however, total cost declines. Thus, the welfare effects of limited entry third markets are ambiguous. Note, however, that the third market survives only if it is in the interest of the exchange's dealers to maintain its suboptimally small membership. Thus, any inefficiency resulting from the entry of the third market is attributable to the exchange's membership limits (compounded by the third market's own entry restrictions), not to the inherent inefficiency of third markets *per se*. This situation is similar to that of a cartel. Inefficient producers can enter under a cartel's price umbrella. This creates deadweight losses that are attributable to the cartel. However, as in the free entry case, restricted entry third markets can improve welfare because their competitive threat can cause the exchange to increase membership. Moreover, the analysis of this section reinforces the basic conclusion that limitations on the number of market makers is the basic source of inefficiency in financial markets.

5 Summary and Conclusions

The issue of competition from satellite markets–third markets–has proved perenially vexing to academics and policymakers alike. Conflicting considerations have largely stalemated the debate. On the one hand, the perceived benefits of concentrating trade in securities and derivatives on a single marketplace suggests that third markets are detrimental to liquidity and price discovery, especially if the third market can free ride on the price discovery functions of the primary exchange. On the other hand, such concentration raises the possibility that a monopoly exchange would exercise market power, in which case satellite markets provide a valuable competitive check. Even if the competitive benefits of the third market are granted, it is still necessary to determine whether an exchange's price and order flow information is a public good that should be disclosed to those who compete with exchange.

Any analysis of these issues must be predicated on a firm understanding of the macrostructure of a financial market. The conventional approach of exogenously specifying a market structure is inadequate because it does not answer the fundamental question that lies at the root of the debate over third markets: namely, given the nature of trading in financial instruments, do exchanges exercise market power? In contrast, this article derives financial market macrostructure endogenously from a canonical model of the microstructure of the trading process. Only in this context is it possible to make welfare comparisons of different market structures and rules.

The analysis implies that due to the nature of liquidity, exchanges are natural monopolies that restrict suboptimally the supply of liquidity and risk bearing services. An exchange enhances member profits by limiting its membership to a suboptimally small number. Given this source of inefficiency, a third market benefits uninformed traders in aggregate and improves welfare if entry into the third market is open (as has been the case in many historical instances) and its dealers can free ride off of the exchange's price discovery. When third market dealers successfully limit their numbers, the beneficial effects of third markets on execution costs and welfare are ambiguous. Thus, analysis of the effects of externalities must be undertaken in a second best world in which a deviation from optimality (market power) already exists.

This analysis does not contradict the notion that all else equal, trading of financial instruments should be concentrated on a single market; in the model, welfare would be maximized if third market dealers traded on the exchange. The point is that it is not in the self-interest of exchange members to permit this. Thus, the fundamental source of inefficiency–and the *raison d'être* of the third market–is that the nature of trading in financial instruments permits exchanges to enhance their members' profits by excluding some intermediaries from membership. Free riding results from exclusion, for which the efficient remedy is inclusion. Therefore, cutting the Gordion Knot and opening exchanges to all would improve market efficiency. If that step is deemed too radical, the analysis of this article implies that free riding, free entry third markets should be encouraged because they are a useful antidote to exchange market power.

A Appendix

To see that conditional price variance is increasing in S, recall that

$$\hat{\sigma}^2 = \frac{S\sigma^2}{K^2\beta^2\sigma^2 + S} \tag{29}$$

Thus, the sign of $d\hat{\sigma}^2/dS$ is given by the sign of:

$$S + \sigma^2 K^2 \beta^2 - S(1 + 2\sigma^2 K^2 \beta \frac{d\beta}{dS}) = K^2[\sigma^2\beta^2 - 2S\sigma^2\beta\frac{d\beta}{dS}] \tag{30}$$

The quadratic that defines β is:

$$K\beta^2\sigma^2 + \frac{(K+1)S\sigma^2\beta}{T_1} - S = 0 \tag{31}$$

Therefore:

$$\frac{d\beta}{dS} = \frac{1 - \frac{(K+1)\sigma^2\beta}{T_1}}{2K\beta\sigma^2 + \frac{(K+1)S\sigma^2}{T_1}} \tag{32}$$

Making further substitutions from the quadratic implies:

$$\frac{d\beta}{dS} = \frac{\beta(1 - \frac{(K+1)\sigma^2\beta}{T_1})}{S + K\beta^2\sigma^2} > 0 \tag{33}$$

Thus,

$$2S\beta\frac{d\beta}{dS} = \frac{2S\beta^2(1 - \frac{(K+1)\sigma^2\beta}{T_1})}{S + K\beta^2\sigma^2} \tag{34}$$

This implies:

$$\sigma^2\beta^2 - 2S\sigma^2\beta\frac{d\beta}{dS} = \frac{\beta^2\sigma^2}{S + K\beta^2\sigma^2}[K\beta^2\sigma^2 + \frac{2(K+1)S\sigma^2\beta}{T_1} - S] \tag{35}$$

Since

$$K\beta^2\sigma^2 + \frac{(K+1)S\sigma^2\beta}{T_1} - S = 0 \tag{36}$$

$$\sigma^2\beta^2 - 2S\sigma^2\beta\frac{d\beta}{dS} = \frac{\sigma^2\beta^2}{S + K\beta^2\sigma^2}\frac{(K+1)S\sigma^2\beta}{T_1} > 0 \tag{37}$$

The inequality holds because $\beta > 0$.

Proof of Result 3. Recall that if all noise traders transact on exchange:

$$\lambda_1 = \frac{S\sigma^2}{T_1(S + K^2\beta^2\sigma^2)} + \frac{K\beta\sigma^2}{S + K^2\beta^2\sigma^2} \tag{38}$$

Since λ_1 increases when some of the uninformed transact on the third market, when $T_1 = T_3$, the difference between execution costs on exchange and on the third market when there are no spillovers is no smaller than:

$$(K\beta + \frac{S}{T_1})\frac{\sigma^2}{S + K^2\beta^2\sigma^2} - \frac{\sigma^2}{T_1} \tag{39}$$

This simplifies to:

$$\frac{\sigma^2(K\beta T_1 - S - K^2\beta^2\sigma^2 + S)}{T_1(S + K^2\beta^2\sigma^2)} = \frac{K\beta\sigma^2}{T_1(S + K^2\beta^2\sigma^2)}(T_1 - \sigma^2 K\beta) \tag{40}$$

The quadratic implies:

$$\sigma^2\beta = \frac{T_1(S - K\beta^2\sigma^2)}{(K+1)S} \tag{41}$$

Therefore, the execution cost difference is:

$$\frac{K\beta\sigma^2}{S + K^2\beta^2\sigma^2}[1 - \frac{K(S - K\beta^2\sigma^2)}{(K+1)S}] > 0. \tag{42}$$

Derivation of Total Cost. The total cost of operating the market equals noise trader's execution costs minus informed trader profits minus risk-adjusted market maker profits. Given v and z, exchange execution costs are $z\lambda_1(\beta v + z)$, informed traders' profits are $-K\beta v\lambda_1(K\beta v + z) + Kv^2/(K+1)\lambda_1$ and risk-adjusted market maker profits are:

$$\sum_{j=1}^{L^*}[-\frac{t_j}{T_1^*}(K\beta v + z)(v - \lambda_1(K\beta v + z)) - \frac{.5t_j\hat{\sigma}^2(K\beta v + z)^2}{T_1^{*2}}]. \tag{43}$$

Since $\sum_{j=1}^{L^*} t_j = T_1^*$, this simplifies to:

$$-(K\beta v + z)(v - \lambda_1(K\beta v + z)) - \frac{.5\hat{\sigma}^2(K\beta v + z)^2}{T_1^*}. \tag{44}$$

Substituting for $\hat{\sigma}^2$ and simplifying implies that the total cost of trading on the exchange is:

$$vz + (K^2\beta^2 v^2 + 2K\beta vz + z^2)\frac{.5(1-q)S\sigma^2}{T_1^*(K^2\beta^2\sigma^2 + (1-q)S)} \tag{45}$$

Taking expectations over v and z implies that expected total cost equals:

$$\frac{.5\sigma^2 S(1-q)}{T_1^*} \tag{46}$$

Note that expected total cost equals the cost exchange members incur to bear noise trader order flow risk when they cannot condition their trades on price.

Similar analysis implies that with spillovers, the expected total cost of operating the third market is:

$$\frac{.5\hat{\sigma}^2 qS}{T_3} \tag{47}$$

Without spillovers, the expected third market cost is:

$$\frac{.5\sigma^2 qS}{T_3} \tag{48}$$

Analysis of an exchange's incentive to expand. Recall that $\hat{T}_1(q^*)$ is the exchange size required to eliminate the third market when spillovers exist. This function is defined by the equation:

$$\frac{\hat{\sigma}^2}{T_A - \hat{T}_1} = \frac{\hat{\sigma}^2}{\hat{T}_1} + \frac{\beta(q^*, \hat{T}_1)\hat{\sigma}^2}{S(1-q^*)} \tag{49}$$

If $T_1 > \hat{T}_1$ the third market cannot survive. Eliminating the $\hat{\sigma}^2$ from both sides of the equation, substituting from (11), simplifying and taking a derivative with respect to q^* implies:

$$\frac{d\hat{T}_1}{dq^*} = \frac{.5y^{-.5}/S\sigma(1-q)^2}{\frac{1}{(T_A-\hat{T}_1)^2} + .5y^{-.5}\frac{2\sigma}{\hat{T}_1^3}} > 0 \tag{50}$$

where

$$y = \frac{\sigma^2}{\hat{T}_1^2} + \frac{1}{S(1-q^*)} \tag{51}$$

Thus, the larger is q^*, the more the exchange must expand to eliminate the third market.

The exchange expands if $\mathcal{T}_1(q^*) > \hat{T}_1(q^*)$, where $\mathcal{T}_1(q^*) = T_1^*/\sqrt{1-q^*}$. Note that $\hat{T}_1(q^*) < T_A$. Moreover, the exchange is willing to admit all intermediaries to membership if $q^* > 1 - (\frac{T_1^*}{T_A})^2 \approx .75$. Therefore, there is some $q^* < 1 - (\frac{T_1^*}{T_A})^2$ such that expansion increases the profits of the exchange members in $\mathbf{L}^*$.[25] This critical value of q^* is determined by the intersection of the $\hat{T}_1$ and $\mathcal{T}_1$ loci. Since third market execution costs are higher when there are no spillovers than when there are, the level of exchange membership required to eliminate the third market in the absence of spillovers, $\bar{T}_1(q^*)$, is smaller than $\hat{T}_1(q^*)$. (Due to the assumption of coordinated noise trader choice, $\bar{T}_1$ does not depend on q^*.) Moreover, $\bar{T}_1 > T_1^*$. Therefore, for q^* such that $\mathcal{T}_1 < \bar{T}_1$, the exchange does not expand whether or not there are spillovers. For q^* such that $\hat{T}_1(q^*) > \mathcal{T}_1 > \bar{T}_1$, the exchange expands when there are no spillovers, but does not expand if there are spillovers. When $\mathcal{T}_1(q^*) > \hat{T}_1(q^*)$, the exchange expands when there are spillovers.

[25] Expansion may not occur even if intermediaries in $\mathbf{L}^*$ profit by it because intermediaries $j \notin \mathbf{L}^*$ may earn higher profits on the third market than in the expanded exchange. This is most likely to occur when σ^2 is small. With a small σ^2, $\hat{T}_1$ is large and $\sigma^2/\hat{T}_1$ is small.

References

[1] Admati, A., and P. Pfleiderer. "A Theory of Intraday Patterns: Volume and Price Variability." *Review of Financial Studies* 1 (1988): 3-40.

[2] Admati, A., and P. Pfleiderer. "Sunshine Trading and Financial Market Equilibrium." *Review of Financial Studies* 4 (1991): 443-481.

[3] Arthur, W. *Increasing Returns and Path Dependence in the Economy.* Ann Arbor: University of Michigan Press, 1994.

[4] Battalio, R., B. Hatch, and R. Jennings. "Dimensions of Best Execution for Market Orders: Assessing Differences between the NYSE and the Nasdaq Third Market." Unpublished Manuscript. Kelley School of Business, Indiana University (2000).

[5] Bessembinder, H., and H. Kaufman. "A Cross-Exchange Examination of Trading Costs and Information Flow for NYSE-listed Stocks." *Journal of Financial Economics* 46 (1997): 293-319.

[6] Brown, D., and Z. Zhang. "Market Orders and Market Efficiency." *Journal of Finance* 52 (1997): 277-307.

[7] Chowdhry, B., and V. Nanda. "Multi-Market Trading and Market Liquidity." *Review of Financial Studies* 3 (1991): 483-511.

[8] Coffee, J. "Comment" in A. Lo (ed.) *The Industrial Organization of the Securities Industry.* Chicago: University of Chicago Press, 1996.

[9] David, P. "Clio and the Economics of QWERTY." *American Economic Review* 75 (1985): 332-337.

[10] Davis, L., and L. Neal. "Micro Rules and Macro Outcomes: The Impact of Microstructure on the Efficiency of Security Exchanges, London, New York, and Paris, 1800-1914." *American Economic Review* 88 (1998): 40-45.

[11] Diamond, D., and R. Verrecchia. "Disclosure, Liquidity, and the Cost of Capital." *Journal of Finance* 46 (1991): 1325-1359.

[12] Easley, D., N. Kiefer, and M. O'Hara. "Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow." *Journal of Finance* 51 (1996): 811-833.

[13] Farrell, J., and G. Saloner. "Standardization, Compatibility, and Innovation." *Rand Journal of Economics* 16 (1985): 70-83.

[14] Fudenberg, D., and J. Tirole. "Pricing Under the Threat of Entry by a Sole Supplier of a Network Good." Unpublished Manuscript, Harvard University (1999).

[15] Glosten, L. "Is the Electronic Limit Order Book Inevitable?" *Journal of Finance* 49 (1994): 1127-1161.

[16] Hagerty, H., and R. McDonald. "Brokerage, Market Fragmentation, and Securities Market Regulation" in A. Lo (ed.) *The Industrial Organization of the Securities Industry*. Chicago: University of Chicago Press, 1996.

[17] Hasbrouck, J. "One Security, Many Markets: Determining the Contributions to Price Discovery." *Journal of Finance* 50 (1995): 1175-1199.

[18] Hirsleifer, J. "The Private and Social Value of Information and the Reward to Inventive Activity." *American Economic Review* 61 (1971): 561-574.

[19] Huang, R. and H. Stoll. "Competitive Trading of NYSE Listed Stock: Measurement and Interpretation of Trading Costs." Vanderbilt University Financial Markets Research Center Working Paper 94-13 (1994).

[20] Katz, M. and C. Shapiro. "Technology Adoption in the Presence of Network Externalities." *Journal of Political Economy* 75 (1986): 822-841.

[21] Krugman, P. *Peddling Prosperity.* New York: W. W. Norton, 1994.

[22] Kyle, A. "Continuous Auctions and Insider Trading." *Econometrica* 53 (1985): 317-355.

[23] Leibowitz, S., and S. Margolis. *Winners, Losers, and Microsoft: Competition and Antitrust in High Technology.* Oakland, CA: The Independent Institute, 1999.

[24] Macey, J., and D. Haddock. "Shirking at the SEC: The Failure of the National Market System." *Univeristy of Illinois Law Review* 1985 (1985): 315.

[25] Madhavan, A., and M. Cheng. "In Search of Liquidity: Block Trades in the Upstairs and Downstairs Markets." *Review of Financial Studies* 10 (1997): 175-203.

[26] McInish, T., and R. Wood. "Competition, Fragmentation, and Market Quality" in A. Lo (ed.) *The Industrial Organization of the Securities Industry.* Chicago: University of Chicago Press, 1996.

[27] Miller, G. "Comment" in A. Lo (ed.) *The Industrial Organization of the Securities Industry.* Chicago: University of Chicago Press, 1996.

[28] Mulherin, H., J. Netter, and J. Overdahl. "Prices are Property: The Organization of Financial Exchanges from a Transaction Cost Perspective." *Journal of Law and Economics* 34 (1991): 591-644.

[29] Mulherin, H. "Comment" in A. Lo (ed.) *The Industrial Organization of the Securities Industry.* Chicago: University of Chicago Press, 1996.

[30] Oesterle, D., D. Winslow, and S. Anderson. "The New York Stock Exchange and its Outmoded Specialist System: Can the Exchange Innovate to Survive?" *Journal of Corporation Law* 17 (1992): 223-310.

[31] Oesterle, D. "Comments on the SEC's Market 2000 Report." *Journal of Corporation Law* 19 (1994): 483-551.

[32] O'Hara, M. *Market Microstructure Theory.* Malden, MA: Blackwell Business Publishers, 1997.

[33] Pagano, M. "Trading Volume and Asset Liquidity." *Quarterly Journal of Economics* 104 (1989): 255-274.

[34] Pirrong, C. "The Industrial Organization of Financial Markets" Theory and Evidence." *Journal of Financial Markets* 2 (1999): 329-358.

[35] Seligman, J. *The SEC and the Future of Finance.* New York: Praeger, 1985.

[36] Seppi, D. "Block Trading and Asymmetric Information." *Journal of Finance* 45 (1990): 73-94.

[37] Smith, B., D. Turnbull, and R. White. "Upstairs Market for Principal and Agency Trades: Analysis of Adverse Information and Price Effects." *Journal of Finance* 56 (2001): 1723-1746.

[38] Stoll, H. "Principles of Trading Market Structure." *Journal of Financial Services Research* 6 (1992): 75-107.

[39] Stoll, H. "The Causes and Consequences of the Rise in Third Market and Regional Trading." *The Journal of Corporation Law* 19 (1994): 509-520.

[40] Subrahmanyam, A. "Risk Aversion, Market Liquidity, and Price Efficiency." *Review of Financial Studies* 4 (1991): 417-441.

[41] Telser, L. "Comment" in R. Anderson, (ed.) *The Industrial Organization of Futures Markets.* Lexington, MA: Lexington Press, 1984.

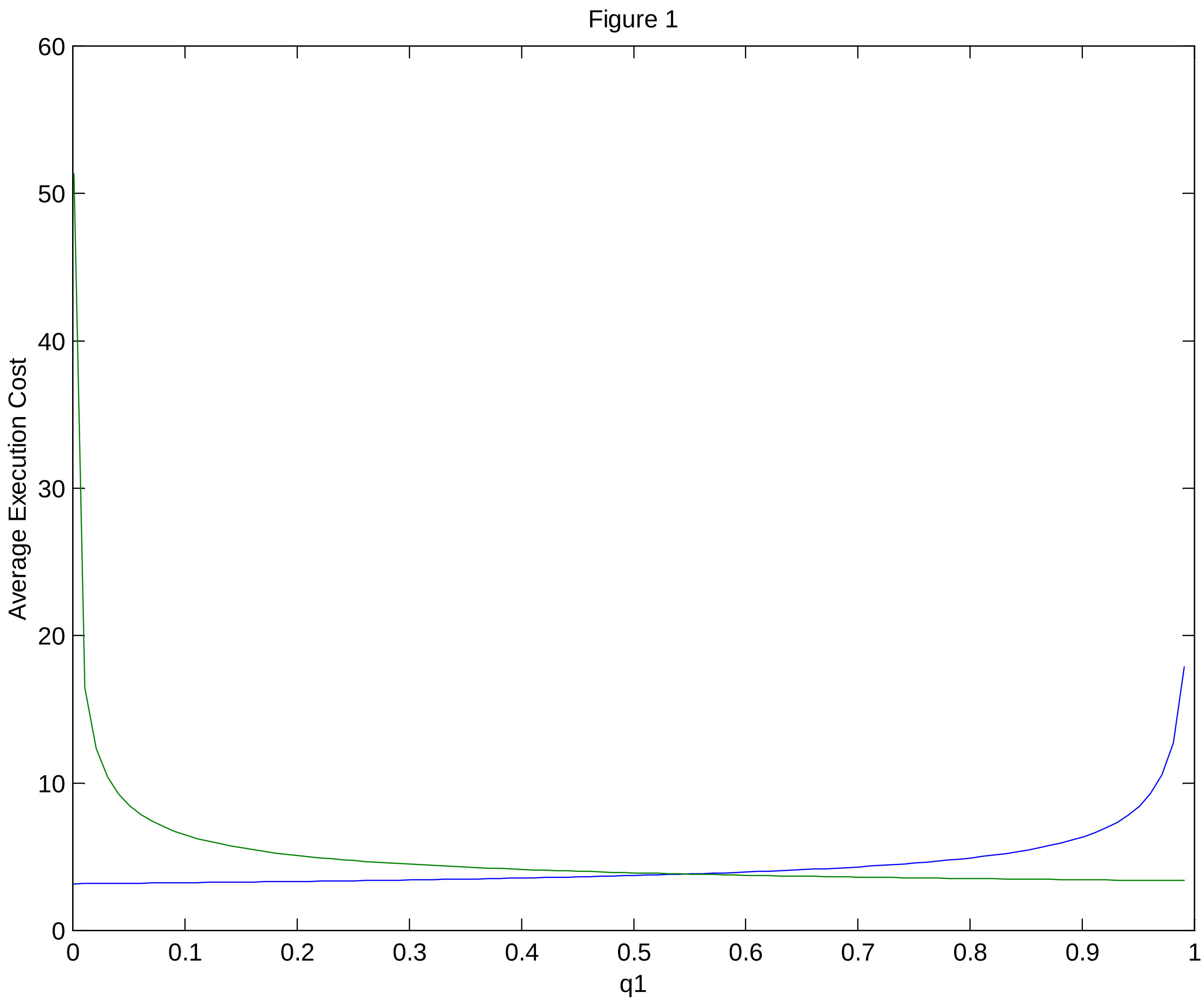
Figure 1
Average Execution Cost
q1
0
10
20
30
40
50
60
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

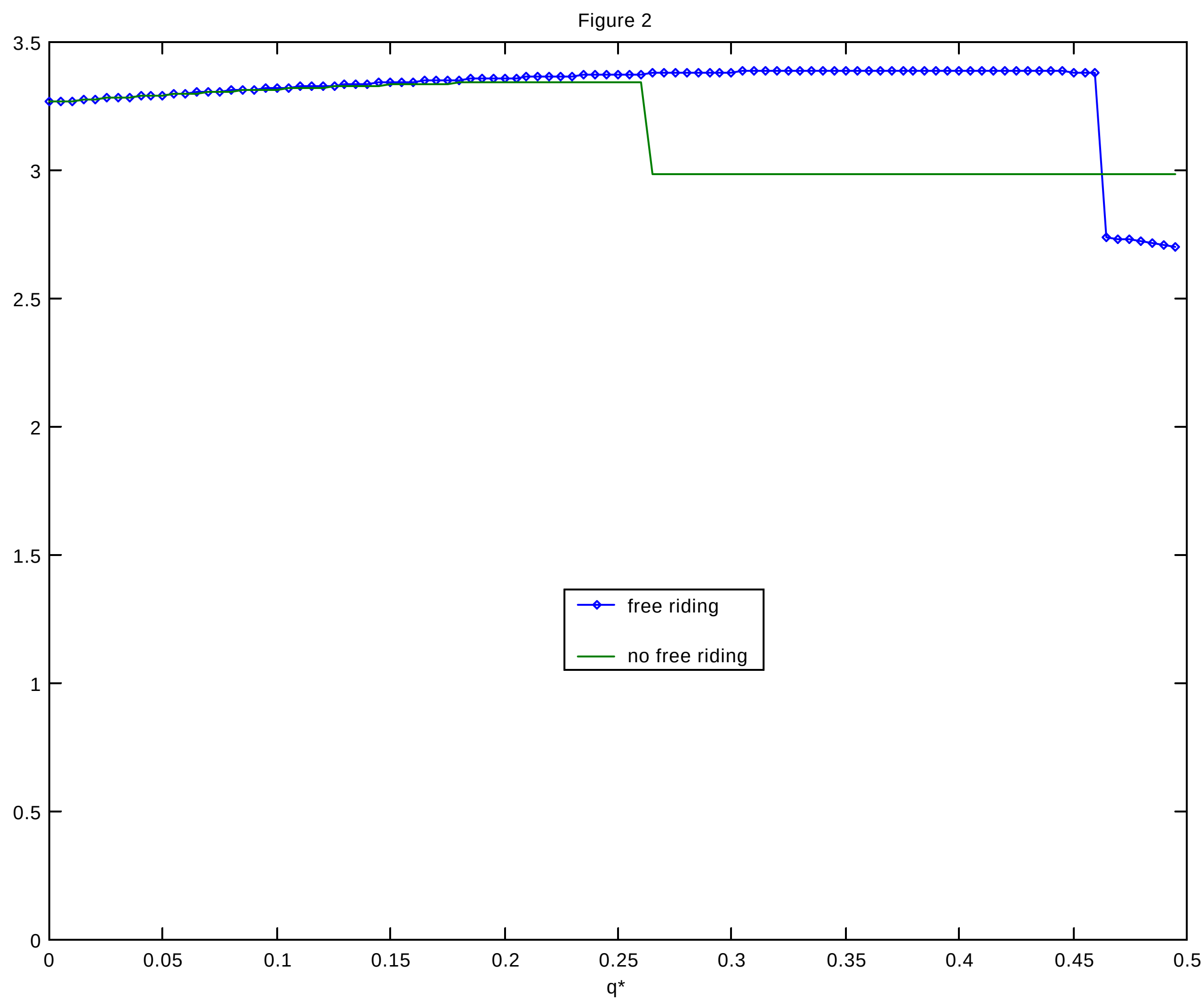
Figure 2
3.5
3
2.5
2
1.5
1
0.5
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
q*
free riding
no free riding