Middle Men versus Market Makers: A Theory of Competitive Exchange

John Rust* and George Hall
Yale University

March 2001

Abstract

What determines how trade in a commodity is divided between privately negotiated transactions via "middle men" (dealer/brokers) in a telephone or "dealer market" versus transactions via "market makers" (specialists) at publicly observable bid/ask prices? To address this question, we extend Spulber's (1996a) search model with buyers, sellers, and price setting dealers to include a fourth type of agent, *market makers*. The result is a model where market microstructure — the division of trade between dealers and market makers — is determined endogenously. In Spulber's model, dealers are the exclusive avenue of exchange, and prices are private in the sense that price quotes can only be obtained through direct contact (e.g. telephone calls) to individual dealers. In contrast a market maker can be conceptualized as operating an exchange that posts publicly observable bid and ask prices. In our model buyers and sellers can either trade with the market maker at the publicly posted bid/ask price or they can search for a better price in the dealer market. We show that the entry of a monopolist market maker can be profitable if it has a lower marginal cost of processing transactions than the least efficient middle man in the equilibrium without market makers. If this is the case the entry of a market maker segments the market; the highest valuation buyers and the lowest cost sellers trade with the market maker and the residual set of intermediate valuation buyers and sellers search for better prices in the dealer market. Dealers act as a "competitive fringe" that undercut the bid/ask spread charged by the monopolist market maker. However less efficient dealers are driven out of business. The remaining dealers are still profitable although the entry of a monopolist market maker significantly reduces their profits and bid-ask spreads. Thus, entry by a marker maker results in uniformly higher surpluses for buyers and sellers and higher trading volumes. When there is free entry into market making and market makers' marginal costs of processing transactions tend to zero, bid-ask spreads converge to zero and a fully efficient Walrasian equilibrium outcome emerges.

---

*Corresponding author. Department of Economics, Yale University, 37 Hillhouse Avenue, P.O. Box 208264 New Haven CT 06520-8264, phone: (203) 432-3569, fax: (203) 432-6323 e-mail: `jrust@gemini.econ.yale.edu`, web page: `http://gemini.econ.yale.edu/jrust` We thank Ian Domowitz and Daniel Spulber for helpful comments. We are especially grateful to an anonymous contact at a U.S. steel service center (i.e. a steel middle man) for providing detailed data on steel prices and inventories that initiated our inquiry into this topic. We also thank the National Science Foundation for financial support.

1 Introduction

What determines how trade in a commodity (or asset) is divided between privately negotiated transactions with "middle men" (dealer/brokers) in a telephone or "dealer market" versus transactions via "market makers" (specialists) at publicly observable bid/ask prices? To address this question we extend Spulber's (1996a) search model with buyers (consumers), sellers (producers), and price setting middle men (also referred to as dealers, intermediaries, or brokers) to include a fourth type of agent, *market makers.* In our model, the market microstructure — as represented by the division of trade between dealers and market makers — is determined endogenously. In the original model without market makers, middle men are assumed to be the exclusive avenue of exchange: every producer wishing to sell a commodity (asset) and every consumer wishing to purchase it is required to transact via middle men rather than trade directly with each other. Transactions in the dealer market occur over a range individually negotiated prices — the outcome of a search process with random matching.

We study the effect of introducing a monopolist market maker on the search equilibrium in the dealer market. The market maker can be conceptualized as operating an exchange that posts publicly observable bid and ask prices. Buyers and sellers now have the option of trading on the exchange at the publicly posted bid/ask price, or searching for a better price in the dealer market. The latter requires a costly, time-consuming process in which a sequence of inquiries are made to randomly selected middle men who set individually determined bid/ask prices. These prices are private in the sense that the only way one can obtain price quotes is via direct contact (e.g. telephone calls) to individual dealers.

We show that entry by a monopolist market maker can be profitable if it has a lower marginal cost of processing transactions than the corresponding cost for the least efficient middle man in the equilibrium without market makers. In this case entry by a market maker segments the market: the highest valuation consumers and the lowest cost producers trade on the exchange and the residual set of intermediate valuation consumers and producers search for better prices in the dealer market. Middle men act as a "competitive fringe" who undercut the bid/ask spread charged by the monopolist market maker. However the entry of the monopolist market maker creates additional competition that results in significant reductions in the bid-ask spreads in the dealer market, uniformly higher consumers' and producers' surpluses, and higher trading volumes.

Since the entry of a market maker narrows bid-ask spreads, this lowers dealers' profits and causes the least efficient middle men to be driven out of business. In the limit as the market maker's marginal cost of processing transactions c tends to zero, there are two possibilities. If the marginal cost k_l of the most efficient middle man is positive, then in the limit as the search cost, δ, approaches zero, all middle men will eventually be driven out of business and the market maker will set a bid-ask spread equal to k_l. If $k_l = 0$, then as $\delta \to 0$, 50% of the middle men who were present in the equilibrium without a market maker will be forced to exit, and the market maker will handle 50% of the volume of trade that would occur in a frictionless Walrasian equilibrium. When there is free entry into market making, the bid-ask spreads are forced down to the marginal cost of processing transactions by the most efficient market maker. If the market maker's marginal cost c is less than the marginal cost k_l of the most efficient middle man, the dealer market will be driven out of business by the free entry of market makers. If $c = 0$, bid-ask spreads are forced to zero and a fully efficient Walrasian equilibrium outcome emerges.

We conclude that middle men and market makers represent complimentary types of exchange institutions. Market makers provide a superior exchange technology; publicly posted bid and ask quotes raise welfare and reduce bid-ask spreads compared to the search equilibrium in the dealer-only market. However, the existence of middle men is remarkably robust, and they serve an important role as a "competitive fringe" that limits the market power of a monopolist market maker. Without the competitive threat of middle men, a monopolist market maker would quote a very large bid-ask spread and consumer and producer surplus would be significantly lower.

The "best" outcome is when there is free entry of both middle men and market makers. The free entry of market makers drives down the bid-ask spread to its lowest possible level short of the frictionless Walrasian equilibrium outcome. This results in the highest possible level of market efficiency and producer and consumer surplus. As long as there are middle men who can process transactions more cheaply than market makers, a small fringe of highly efficient middle men will continue to exist even when there is free entry into market making. The incremental competition from these middle men provide an additional incremental reduction in bid-ask spreads, and a corresponding incremental increase in producer and consumer surplus. The only case where all middle men are driven out of business is when the marginal cost of processing transactions by the most efficient market maker is lower than the corresponding marginal transactions cost of the most efficient middle man.

Our analysis is similar in some respects to a previous paper by Gehrig (1993).[1] Gehrig developed a model in which buyers and sellers of a commodity have the option of trading at publicly posted bid/ask prices on an exchange run by a market maker (which Gehrig refers to as an intermediary), or entering a "search market" in which buyers and sellers are randomly matched and engage in bargaining in an attempt to negotiate a mutually acceptable price. Gehrig's results differ from ours since his model treats the bargaining process as a static (one-shot) random matching game in which buyers and sellers negotiate directly with each other rather than transacting through middle men. Our model builds on the dynamic search model of Spulber (1996a) which allows buyers and sellers who enter the search market to repeatedly contact a sequence of middle men until they eventually succeed in finding acceptable prices at which they can transact.

A number of special aspects of Gehrig's conclusions stem from particular assumptions about the nature of the search and bargaining process. In his model agents have only one opportunity to enter the search market and match and transact. If a match occurs, either the buyer or seller in a match is randomly designated as the "proposer" who quotes a take it or leave it price. These assumptions imply an unavoidable minimum level of inefficiency in the search market even in the limit as the probability of matching λ approaches 1. Part of the inefficiency results from unconsummated trades due to the one-shot nature of the search process: if the proposer quotes a price that the other party does not find acceptable, the match dissolves and neither party has the option to re-enter the search market to try to negotiate another deal. The second source of unavoidable inefficiency is due to "excess entry" into the search market. That is, "extra-marginal' traders who would not trade in a frictionless Walrasian equilibrium are induced to enter the search market due to the fact that entry is costless and there is a small probability that a random match could result in a profitable trade. As a result, a monopolist intermediary can capitalize on these unavoidable inefficiencies in the search market and will charge a positive bid-ask spread even in the limit as the probability of matching in the search market tends to one. Gehrig's model has multiple equilibria, and even specially selected equilibria cannot be Pareto ranked as the level of "search efficiency" (as indexed by the matching probability λ) is varied.

Since our analysis builds on Spulber's search model with competitive intermediaries and sequential search, our conclusions about the properties of the profit-maximizing bid-ask spread

[1]Related but more recent papers by Baye and Morgan (2001) and Pirrong (2000) will be discussed in the concluding section.

chosen by a monopoly market maker differ from Gehrig's. In particular, in our anlysis, if $k_l = 0$, then in the limit as $\delta \to 0$ the search equilibrium becomes arbitrarily efficient and the competition from middle men forces the monopolist market maker's bid-ask down to 0 (if $k_l > c \geq 0$ the monopolist market maker's bid-ask spread converges to k_l as $\delta \to 0$). However in Gehrig's model the market maker's limiting bid-ask spread remains positive even as the probability of a match between buyers and sellers tends to one. In Gehrig's model the monopolist market maker always trades the same quantity regardless of the level of efficiency in the search market (as indexed by the matching probability λ). In our model, the division of trade between middle men and the market maker is endogenously determined: it depends on the relative costs of transacting on the exchange versus the costs and benefits of searching for prices in the dealer market. Additionally, our model has a unique stationary equilibrium for any set of parameter values and this makes it possible to conduct meaningful "comparative steady state" analyses. In particular, the relative efficiency of equilibria for various values of δ can be unambiguously ranked: allocative efficiency (the sum of profits and producers' and consumers' surplus) decreases monotonically as the cost of search increases (i.e. as δ increases). For these reasons we feel that our model provides new insights into the question of why trading of most commodity and assets is divided between two types of institutions: formal exchanges with posted prices and informal "dealer markets" that require active searching and bargaining.

Our work on this topic was motivated by empirical research on intermediation in the market for steel (see Athreya *et. al.* 2001 and Hall and Rust 2000a,b). Interestingly, there are no exchanges or market makers for any major steel products (although as we will see in section 6, two potential market makers, `e-STEEL.com` and `metalsite.com`, have recently entered the steel market). Currently, most transactions occur via middle men known as *steel service centers* (SSCs). These companies purchase steel inventories from steel producers (and other SSCs) and subsequently sell these inventories to retail customers (and other SSCs) at a markup. According to the Steel Service Center Institute (SSCI, 1999), in the U.S. alone there are SSCs in over 5000 separate locations that employ over 88,000 people. A recent report by the SSCI (available online at `http://www.ssci.org/exec_summary.adp`) estimates that in 1997, half the total sales of all the major product lines of steel in the U.S. occurred via SSCs rather than via direct sales from producers to end consumers. The prices charged by SSCs are private information, in the sense that these companies do not advertise or publicly post their prices. Instead, consumers must make telephone calls to individual SSCs to obtain price quotes and negotiate a final transaction price. Analysis of customer level price data for a particular SSC that we have direct contact with (see Athreya, 2001) shows that there is considerable amount of "unexplained" price dispersion in a wide variety of steel products (i.e. dispersion that cannot be accounted for by location of the customer, variations in the production price of steel, quantity discounts, and other observable factors). This evidence clearly refutes the "law of one price" in the market for steel, and is consistent with the hypothesis of equilibrium price dispersion due to costly search.

It is puzzling that there is no commodity market where steel products are traded, even though it is an example of a standardized, homogeneous commodity which would seem to be ideal for trading over an exchange. The large degree of price dispersion in the steel market suggests that there could be substantial gains from the introduction of an exchange. There are formal exchanges for many other less homogeneous commodities such as fish (which are traded in a newly created electronic marketplace `gofish.com`), or wheat and pork bellies (which are traded on the Chicago Mercantile Exchange). There are also exchanges for other metals such as aluminum ingot and precious metals such as gold or silver (the latter are traded on COMEX and the London Metal Exchange, respectively). However the market for plate and coil steel in the U.S. operates exclusively as a dealer or "telephone market" with SSCs serving as competitive middle men that intermediate a substantial share of all trade. There are no centrally posted prices, and there is no central market place at which steel is traded. Instead, consumers must search for the lowest prices by making a series of phone calls to producers or middlemen in their area who are likely to have the type of steel product they are searching for. To the extent this sequence of phone calls constitutes a costly search process, Spulber's (1996a) model provides a rough caricature of how the steel market actually works. One practical goal of this paper is to provide new insights into why no exchanges trade commodities such as steel, and why there are active exchanges for many other assets and commodities. A secondary goal, which will be addressed in section 5, is to consider how changes in communications and transactions technologies resulting from the advent of the world wide web and the "information revolution" will affect the microstructure of trade in a variety of different asset and commodity markets.

Our paper offers a simple analytical framework that provides new insights into how market microstructure is endogenously determined. In particular, we provide conditions under which entry of a monopolist market maker can simultaneously be profitable and result in significant welfare improvements for buyers and sellers by reducing bid/ask spreads. Although the model is extremely stylized, we believe it provides some new insights into why we observe market makers in some markets but not in others (such as the steel market). In particular, our analysis shows that there are three key parameters that determine whether there can be profitable entry by a market maker: the level of search costs δ, the marginal transactions cost of the most efficient middle man k_l, and the marginal transactions cost c of the most efficient market maker. If k_l and δ are sufficiently small, then the equilibrium in the dealer market already closely approximates the fully efficient Walrasian outcome, and so there is little or no scope for profitable entry by a market maker. However even if k_l and δ are relatively high, the market maker's marginal costs of processing transactions c must be sufficiently low to enable profitable entry by a market maker to occur. These costs are the sum of the marginal costs that the market maker incurs to execute a transaction, plus the marginal cost of trading by producers and consumers (e.g. communication costs associated with placing orders on the exchange). The advent of computerized trading and the world wide web have greatly reduced the marginal costs of transacting over exchanges. Unless search and transactions costs in the "telephone market" are declining at a comparable rate there should eventually be room for profitable entry by a market maker into the steel market.

Section 2 reviews Spulber's original model in the case where there are no market makers (exchanges). In Section 3 we introduce the possibility of entry by a monopolist market maker who runs a centralized exchange with publicly posted prices. We derive conditions characterizing when entry is profitable. If entry occurs, we show that a monopolist market maker reduces bid-ask spreads, drives out inefficient middle men, and increases welfare of consumers and producers. Section 4 considers the case where there is free entry of market makers into the commodity exchange business. Free entry drives the bid-ask spread down to the marginal transactions cost c of the most efficient market maker. The free entry equilibrium converges to the fully efficient Walrasian outcome as c tends to zero. However unless the market maker's transactions costs are smaller than the transactions cost of the most efficient middle man, free entry of market makers will not drive the dealer market out of business. Thus, the coexistence of middle men and market makers appears to very robust feature of the microstructure of markets. We conclude in section 5 by offering some observations about the insights and limitations of our simple model as a stylized description of the U.S. steel market. We discuss several other potential explanations for the "puzzle" of why there is no formal centralized exchange for steel products. In particular, we speculate on why the entry of the web-based market maker `e-STEEL.com` has had relatively little impact on the microstructure of exchange in the steel market so far. We compare the outcome in the steel market to the market for fish, where the "seafood exchange" operated by the web-based market maker `gofish.com` seems much more likely to revolutionize the microstructure of trade in this market.

2 Search Equilibrium in the Dealer Market

This section reviews and extends the model of competitive intermediation introduced by Spulber (1996a). We call the market that Spulber analyzes a *dealer market,* since it is populated by a continuum of three different types of agents: producers, consumers, and middle men. The latter type of agents are intermediaries between producers and consumers and are also known as "dealer/brokers" in the financial markets. Producers manufacture a homogeneous commodity that is demanded by consumers. An alternative interpretation is that producers are sellers of an asset that is demanded by buyers, so we will use interchangeably the terms producer/seller and consumer/buyer.

Producers are indexed by their marginal cost of production, w, which is assumed to be distributed uniformly on the $[0,1]$ interval. A producer of type $w \in [0,1]$ can produce a single unit of the commodity at cost w. Similarly, consumers are indexed by their valuations for the commodity, which is also assumed to be uniformly distributed on the $[0,1]$ interval. A consumer of type $v \in [0,1]$ is willing to pay at most v for a single unit of the commodity.

Initially, assume that there is no central exchange or marketplace where the commodity is traded. In particular there are no publicly quoted market bid or ask prices, so the only way to obtain price quotes is via direct contacts with individual middle men. We follow Spulber in assuming that consumers and producers cannot trade with each other directly. Instead, all exchange is intermediated via a continuum of middle men, who serve as dealer/brokers who are willing to buy the commodity from producers and sell it to consumers at a markup. Middle men are constrained to sell only units they have on hand, so the number of units a middle man buys from producers is constrained to equal the number of units it sells to consumers period by period. Middle men indexed by their per period cost of holding the commodity, k, which is distributed uniformly on the interval $[k_l, 1]$. The lower bound k_l is the marginal transactions cost of the most efficient middle man. Since the exchange activity is costly, we assume $k_l \geq 0$. A middle man of type k chooses a pair of stationary bid-ask prices $(p_b(k), p_a(k))$ that maximize its discounted profits, where $p_a(k)$ denotes the ask price at which the middle man is willing to sell to consumers and $p_b(k)$ denotes the bid price at which the middle man is willing to buy from producers.

Consumers and producers are able to contact at most one middle man per period. The middle man they contact is assumed to be a random draw from the population of active middle men. Producers and consumers adopt optimal search strategies in an attempt to find the middle man that offers the best bid or ask price. If middle men adopt stationary pricing strategies, consumers and producers will be searching over stationary distributions of bid and ask prices. Let $F_a(p)$ denote the distribution of ask prices facing consumers. This is the distribution implied by the optimal ask prices $p_a(k)$ for the fraction of middle men that are active in equilibrium. Let $F_b(p)$ denote the distribution of bid prices facing producers. This is the distribution implied by the optimal bid prices $p_b(k)$ for the fraction of middle men that are active in equilibrium. All producers and consumers are assumed to have a common per period discount factor of $\beta = 1/(1+\delta)$, where $\delta \geq 0$ is the per period interest rate. Since δ determines the degree to which deferred gains from trade are discounted, this parameter can be viewed as indexing the cost of search, with higher values of δ corresponding to a more costly search process.

Let $V(p, v)$ denote the present discounted value of an optimal search strategy for a consumer of type v who has randomly contacted a middle man who quotes an ask price of p. The consumer has three choices: 1) do nothing (i.e. do not buy at the ask price p and do not search), 2) accept the middle man's ask price of p, and 3) reject the middle man's ask price of p and continue searching for a better price. These three options are reflected in the Bellman equation for the consumer's problem which is given by

$$V(p, v) = \max \left[0, v - p, \beta \int_{a_l}^{a_u} V(p', v) F_a(dp') \right], \tag{1}$$

where $[a_l, a_u]$ is the support of the distribution of ask prices charged by various middle men. The value of 0 in the Bellman equation corresponds to the option of not searching, not trading, and not consuming, which we assume has a payoff of 0. All consumers with sufficiently low valuations will choose this option: for example any consumer with valuation $v < a_l$ immediately knows that it is fruitless to engage in search, and will never enter the dealer market.

Now consider the remaining high valuation consumers. As is well known, the optimal search strategy for a consumer of type v takes the form of a *reservation price rule:* accept any ask price which is less than the reservation price $r(v)$ where $r(v)$ is the function implicitly defined by the unique solution to

$$v = L(r(v)) = r(v) + \frac{1}{\delta}\int_{a_l}^{r(v)} F_a(p)dp. \tag{2}$$

It is not difficult to see from (2) that $r(v)$ is a strictly increasing function of v. Thus, if a consumer of type v is willing to accept a given ask price p, then all consumers with higher valuations are also willing to accept this ask price as well.

Let $W(p,w)$ denote the present discounted value of an optimal search strategy of a producer of type w facing a bid price of p. The Bellman equation for the producer is given by

$$W(p,w) = \max\left[0, p - w, \beta\int_{b_l}^{b_u} W(p',w)F_b(dw')\right], \tag{3}$$

where $[b_l, b_u]$ is the support of the distribution of bid prices offered by various middle men. The optimal strategy for a producer of type w also takes the form of a reservation price strategy, but in this case it is optimal to accept any bid price p that *exceeds* the reservation price $r(w)$, where $r(x)$ is given by the unique solution to

$$w = Y(r(w)) = r(w) - \frac{1}{\delta}\int_{r(w)}^{b_u} [1 - F_b(w)]dw. \tag{4}$$

Figure 0 graphs the reservation price functions $r(v)$ and $r(w)$ for consumers and producers for a specific (equilibrium) pair of bid and ask distributions F_a and F_b which will be derived shortly. The reservation price function for producers lies uniformly above the 45 degree line, where the latter represents the producer's cost of production. It is easy to see from equation (4) that when the seller's cost w exceeds the upper bound b_u of the distribution F_b of bid prices offered by middle men, there is no point in searching, so all producers with costs $w > b_u$ remain out of the market. In figure 0 $b_u = .39$ is the point at which the reservation price function for producers intersects the 45 degree line. Thus, the set of active producers (i.e. those who choose to search and ultimately deliver a unit of the good to a buyer) is the interval $[0, b_u) = [0, .39)$. Symmetrically, the reservation price function for buyers lies uniformly below their valuations, intersecting it at the lower support point $a_l = .61$ of the distribution of ask prices F_a. Thus, the set of active consumers is the interval $(a_l, 1] = (.61, 1]$: these are the individuals who choose to search and ultimately consume a unit of the good.

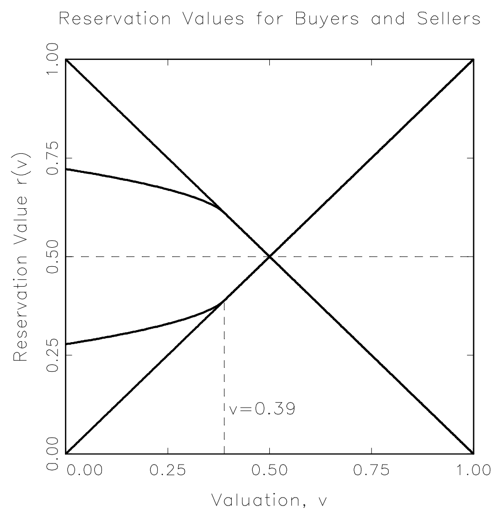

Figure 0: Reservation Prices for Buyers (Consumers) and Sellers (Producers)

The difference between a person's valuation v and their reservation value $r(v)$ is the net value of search. This equals consumers' and producers' expected discounted surplus or "gains from trade". For reference, we plot a horizontal dashed line of height equal to .5. The area in the triangular regions between the 45 degree line and this horizontal line (and to the left of the intersection of the "supply" and "demand" curves) represents the surplus that consumers and producers would achieve in a frictionless Walrasian equilibrium where the equilibrium price for the good equals $p^* = .5$. In that case the consumers' surplus for consumer with valuation v is given by $\max[0, v - .5]$ and the producers' surplus for a producer with cost v is given by $\max[0, .5 - v]$. It is easy to see that the area of the triangular regions, i.e. the surplus achieved by producers and consumers in Walrasian equilibrium, is 1/8 and 1/8, respectively, resulting in a total surplus of 1/4. The area between the horizontal dashed line and the reservation price curves represents the inefficiency of the search equilibrium outcome, i.e. the lost gains from trade to producers and consumers.

Intermediaries maximize expected discounted profits subject to the constraint that the supply and demand for the commodity is equal in every period. This constraint is necessitated by the assumption that middle men do not carry inventories across successive periods. If the middle man sets an ask price of p_a, the quantity of the good demanded in period i, $D_i(p)$, is given by

$$D_i(p_a) = \frac{1}{N} \int_{p_a}^{r_u} [1 - F_a(r)]^i h(r) dr, \tag{5}$$

where N is the number (total mass) of middle men, $r_u \equiv r(1)$ is reservation price of the highest valuation buyer, and $h(r)$ is the distribution of reservation prices in the population. Thus, $h(r)/N$ is the per-firm density of consumers. So $D_i(p_a)$ equals the integral of the product of the probability $[1 - F_a(r)]^{(i-1)}$ of not being successful in search in periods $0, \ldots, i-1$, times the per-firm density of consumers $h(r)/N$, integrated over the region of reservation values $[p_a, r_u]$ corresponding to buyers who would be willing to purchase at price p_a. Using a change of variables, the density $h(r)$ can be derived from the distribution $r(\tilde{v})$ and the fact that consumers' valuations are uniformly distributed on $[0, 1]$:

$$h(r) = L'(r) = 1 + F_a(r)/\delta. \tag{6}$$

Thus, total expected discounted demand is given by

$$\begin{aligned} D(p_a) &= \sum_{i=0}^{\infty} \frac{D_i(p_a)}{(1+\delta)^i} \\ &= \sum_{i=0}^{\infty} \int_{p_a}^{r_u} \frac{[1 - F_a(r)]^i}{(1+\delta)^i} \frac{[\delta + F_a(r)]}{\delta N} dr \\ &= (r_u - p_a) \frac{(1+\delta)}{\delta N}. \end{aligned} \tag{7}$$

By similar reasoning the middle man's expected discounted supply from quoting a stationary bid price p_b is given by

$$S(p_b) = \sum_{i=0}^{\infty} \frac{S_i(p_b)}{(1+\delta)^i} = (p_b - r_l) \frac{(1+\delta)}{\delta N}, \tag{8}$$

where $r_l \equiv r(0)$ is reservation value associated with the lowest cost producer. Given the discounted supply and demand functions, we can write the present discounted profits for a middle man of type k as follows

$$\Pi(p_a, p_b, k) = p_a D(p_a) - (p_b + k) S(p_b) = \frac{(1+\delta)}{\delta N} \left[p_a(r_u - p_a) - (p_b + k)(p_b - r_l) \right]. \tag{9}$$

The middle man's problem is to:

$$\max_{p_a, p_b} \Pi(p_a, p_b, k) \quad \text{subject to: } D_i(p_a) \leq S_i(p_b) \;\; i = 0, 1, \ldots \tag{10}$$

Spulber (1996a) demonstrates that the solution to this problem is given by

$$\begin{aligned} p_a(k) &= (3r_u + r_l + k)/4 \\ p_b(k) &= (r_u + 3r_l - k)/4. \end{aligned} \tag{11}$$

Substituting these solutions back into the profit function, we obtain the following formula for the optimal profits for a middle man of type k:

$$\Pi(p_a(k), p_b(k), k) = \frac{(1+\delta)}{\delta 8N}\left[r_u - r_l - k\right]^2. \tag{12}$$

This implies that the marginal firm is $k_u = r_u - r_l$ so that $N = k_u - k_l = r_u - r_l - k_l$. Thus the number of active middle men are uniformly distributed on the interval $[k_l, k_u]$. Since $p_a(k)$ and $p_b(k)$ are linear functions of k, the equilibrium distribution of bid and ask prices must be uniformly distributed. The equilibrium distribution of ask prices is given by

$$F_a(p) = \frac{p - .75r_u - .25(r_l + k_l)}{.25(r_u - r_l - k_l)}, \tag{13}$$

and its support is the interval $[a_l, a_u] = [.25(r_l + k_l) + .75r_u, r_u]$. Similarly the equilibrium distribution of bid prices is given by

$$F_b(p) = \frac{p - r_l}{.25(r_u - r_l - k_l)}, \tag{14}$$

and its support is the interval $[b_l, b_u] = [r_l, .75r_l + .25(r_u - k_l)]$. Substituting F_a into the equation for the reservation value r_u of the highest valuation consumer $v = 1$ we obtain

$$\begin{aligned} 1 &= r_u + \frac{1}{\delta}\int_{.25(r_l+k_l)+.75r_u}^{r_u} F_a(p)dp \\ &= \frac{(1+8\delta)r_u - r_l - k_l}{8\delta}. \end{aligned} \tag{15}$$

Similarly substituting F_b into the equation for the reservation value r_l for the lowest cost producer $w = 0$ we obtain

$$\begin{aligned} 0 &= r_l - \frac{1}{\delta}\int_{r_l}^{.75r_l+.25(r_u-k_l)} [1 - F_b(p)]dp \\ &= \frac{(1+8\delta)r_l - r_u - k_l}{8\delta}. \end{aligned} \tag{16}$$

Solving these equations for r_l and r_u we obtain

$$r_l = \frac{1 - k_l}{(2+8\delta)}, \quad r_u = \frac{(1+8\delta) + k_l}{(2+8\delta)}. \tag{17}$$

The marginal consumer is the type v_l for which the expected value of searching equals the value of not participating in the market, or 0. Thus we have

$$0 = \beta\int_{a_l}^{a_u} V(p', v_l)F_a(dp') = v_l - r(v_l). \tag{18}$$

Using the equation for the consumer's reservation price (2) we find that

$$v_l = .25(r_l + k_l) + .75r_u = \frac{(1+6\delta) + (1+2\delta)k_l}{(2+8\delta)}. \tag{19}$$

Similarly the marginal producer is the type w_u that satisfies

$$0 = \beta \int_{b_l}^{b_u} W(p', w_u) F_b(dp') = r(w_u) - w_u, \tag{20}$$

and from equation (4) we find that

$$w_u = .75r_l + .25(r_u - k_l) = \frac{(1+2\delta)(1-k_l)}{(2+8\delta)}. \tag{21}$$

Thus, the total demand from consumers in the interval $[v_l, 1]$ equals total supply from producers in the interval $[0, w_u]$ interval. The equilibrium quantity traded is given by

$$q = w_u = 1 - v_l = \frac{(1+2\delta)(1-k_l)}{(2+8\delta)}. \tag{22}$$

Substituting the expressions for r_l and r_u, we see that the distribution of ask prices F_a is uniform on the interval

$$\left[\frac{(1+6\delta) + (1+2\delta)k_l}{(2+8\delta)}, \frac{(1+8\delta) + k_l}{(2+8\delta)}\right]. \tag{23}$$

The distribution of bid prices F_b is uniform on the interval

$$\left[\frac{1-k_l}{(2+8\delta)}, \frac{(1+2\delta)(1-k_l)}{(2+8\delta)}\right]. \tag{24}$$

The number of middle men that are active in equilibrium is

$$N = k_u - k_l = \frac{4\delta(1-k_l)}{(1+4\delta)}. \tag{25}$$

Clearly, the dealer market collapses to autarky in the limit as $k_l \to 1$. Figure 1 illustrates equilibrium outcomes in three cases: a) a monopoly middle man of type $k = 0$, b) the search equilibrium in a competitive dealer market as described above, and c) the Walrasian equilibrium outcome. We see that the search equilibrium with $\delta = .2$ and $k_l = 0$ (middle panel of figure 1) lies between the monopoly outcome and the perfectly competitive Walrasian outcome in terms of prices, quantities, and undiscounted surplus. The Walrasian equilibrium price and quantity are $p^* = .5$ and $q^* = .5$, respectively, resulting in a total surplus of .25. The total quantity eventually traded in the search equilibrium is about 20 percent less, $q = .3889$, and the total (undiscounted) surplus is less: $1/4 - (.5 - q)2\delta/(2 + 8\delta) = .2377$. Note that when $k_l = 0$ the Walrasian equilibrium is the limiting outcome of the search equilibrium as $\delta \to 0$. If $k_l > 0$, then the limiting search equilibrium converges to a degenerate distribution with a bid-ask spread equal to k_l (i.e. $p_a = (1 + k_l)/2$, $p_b = (1 - k_l)/2$ and $Q = (1 - k_l)/2$). Thus, a positive value for k_l represents an unavoidable inefficiency in the dealer market, at least given our assumption that all exchange must be intermediated by middle men.

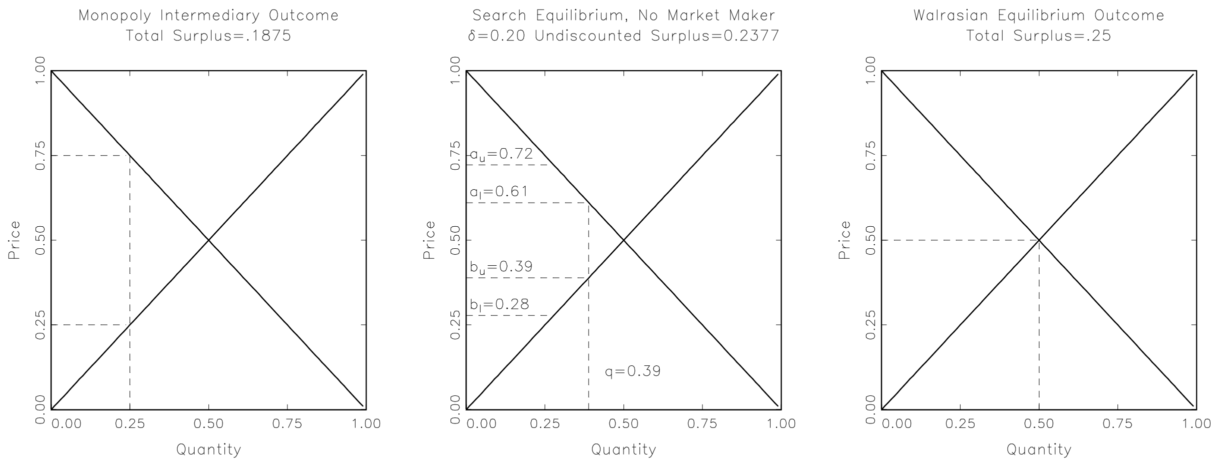

Figure 1: Equilibrium Outcomes in Alternative Market Structures

It is also easy to see that the search equilibrium results in a lower bid-ask spread, higher quantity, and higher surplus than would be attained in an equilibrium with a single monopolist middle man with $k = 0$. The monopoly ask price is $p_a = .75$ and the monopoly bid price is $p_b = .25$ and this results in an equilibrium trade quantity of $q = .25$ and a total surplus of .1875. Note that the search equilibrium outcome does not converge to the monopoly outcome as $\delta \to \infty$. The limiting search equilibrium outcome as $\delta \to \infty$ can be viewed as an equilibrium with a continuum of middle men who behave as "local monopolists". Thus, a middle man of type k sets an ask price equal to $p_a(k) = (3 + k)/4$ and a bid price equal to $p_b(k) = (1 - k)/4$ and transacts a total quantity equal to $q(k) = (1 - k)/4$. The total volume of trade, $\int_0^1 q(k)dk = 1/8$, which is only 1/2 the volume in the case of a single monopoly middle man with cost $k = 0$. The full range $k \in (k_l, 1]$ of inefficient middle men exist in the limiting search equilibrium due to the extreme impatience of the producers and consumers.

Total consumers' surplus S_c is given by

$$S_c = (1 + \delta) \int_0^1 [v - r(v)]dv, \tag{26}$$

and producers' surplus S_p is given by:

$$S_p = (1+\delta)\int_0^1 [r(w) - w]dw. \tag{27}$$

These discounted surplus values will clearly be smaller than the undiscounted surplus (where the latter equals $.25 - 4\delta^2/(2+8\delta)^2$) due to the discounting of the realized surplus from delayed transactions of producers and consumers who searched for multiple periods before they found a bid or ask that they were willing to accept. Total surplus is the sum of producers' surplus, consumers' surplus, and the total discounted profits of middle men, where the latter is given by

$$\int_0^{k_u} \Pi(p_a(k), p_b(k), k)dk = \frac{8\delta(1+\delta)}{3(2+8\delta)^2}. \tag{28}$$

Note that profits converge to 0 as $\delta \to 0$ and converge to 1/6 as $\delta \to \infty$. The latter quantity is equal to the integral of the profits of all the "local monopolist" middle men as k varies over the $[0,1]$ interval. As $\delta \to 0$ it is not difficult to show that $r(v) \to min[v, .5]$, $F_a(p)$ tends to a unit mass at the Walrasian equilibrium price of $p = .5$, and consumers' surplus tends to the Walrasian value of 1/8. As $\delta \to \infty$ we have $r(v) \to v$ so consumers' surplus tends to 0. Similarly, reservation values for producers tend to $r(w) = \max[w, .5]$ as $\delta \to 0$ and $r(w) \to w$ when $\delta \to \infty$. Thus, overall market efficiency tends to the Walrasian value of 1/4 ($1/4 = 1/8 + 1/8 + 0$ with $S_p = 1/8$, $S_c = 1/8$ and total profits of 0, respectively), whereas when $\delta \to \infty$ we have $S_p \to 1/96$, $S_c \to 1/96$. Since total profits tend to 1/6, total surplus is only 3/16, which is only 3/4 of the Walrasian value of 1/4 but the same as the surplus of 3/16 that achieved when there is a single monopolist middle man of type $k = 0$. However the distribution of this surplus is very different in the limiting search equilibrium outcome. When there is a single monopolist middle man, consumer and producer surplus are equal to 1/32 and the monopolist's profits of 1/8 is less than 1/6, which equals the total profits of all middle men $k \in [0,1]$ that are able to survive in the limiting search equilibrium.

3 Search Equilibrium with a Monopoly Market Maker

In this section we extend the Spulber (1996a) model presented in the previous section by introducing a fourth type of agent: a monopolist market maker. Initially we assume there is exactly one potential entrant who could assume the role of a monopolist market maker by introducing a commodity exchange at which any buyer or seller can transact at the publicly announced bid and ask prices of (p_a^m, p_b^m).

The market maker's bid and ask prices are assumed to be costlessly observable. However buyers and sellers may have to incur transactions costs of (c_a, c_b), respectively in order to carry out a transaction if so desired. These costs might include the costs of a telex, fax, or phone call to place an order to buy or sell on the exchange, or the monetary equivalent of any psychic costs associated with logging into a web to execute a transaction. The latter example suggests that the transactions costs incurred by producers and consumers (c_a, c_b) should be very small, at least for individuals who have access to a computer connected to the web and who are computer literate. It is convenient to think of the market maker as setting the bid and ask prices in such as way as to rebate any transactions costs that consumers and producers incur in trading over the exchange, so that $p_a^m - c_a$ is the net income that the market maker earns from selling a unit to a consumer, and $p_b^m + c_b$ is the net cost that the market maker must incur in order to purchase a unit of the good from a producer. It should be clear that as long as (c_a, c_b) are not so large as to place binding constraints on the market maker's pricing decision, there is no loss of generality in the assumption that (p_a^m, p_b^m) are the total costs (inclusive of any transactions costs) that consumers pay and producers receive, respectively. For example, if the market maker adopts a policy of rebating the transaction cost c_a to consumers, then the market maker can obtain a desired net price of p_a^m by charging the consumer the gross of transactions cost amount $p_a^m + c_a$, and subtracting the transactions cost rebate of c_a in order to realize the desired net ask price p_a^m per unit sold.

Once there is a market maker, producers and consumers have the options of a) trading with the market maker at the publicly announced bid and ask prices (p_a^m, p_b^m), or b) searching for a middle man who may offer a better bid or ask price. We begin our analysis by characterizing how the presence of an exchange affects the solution to consumers' and producers' optimal search problems.

Consider a consumer who has not yet chosen to search. The consumer has three options: 1) do nothing (i.e. do not search and do not transact on the exchange), 2) purchase a unit of the commodity in the exchange at price p_a^m, or 3) search for a better price in the dealer market. The consumer's value function is given by

$$V(p_a^m, v) = \max\left[0, v - p_a^m, \beta \int_{a_l}^{a_u} V(p', p_a^m, v) F_a(dp')\right], \tag{29}$$

where $[a_l, a_u]$ is the support of F_a and $V(p, p_a^m, v)$ denotes the value function for a consumer who has chosen to search and has received an ask price of p from a middle man, modeled as a random draw from F_a. Once the consumer has an offer in hand, she has four options: 1) do nothing, 2) accept the offer from the middle man, 3) trade at the ask price posted at the commodity exchange, or 4) obtain another price quote from a new, randomly selected, middle man. The Bellman equation for $V(p, p_a^m, v)$ is given by

$$V(p, p_a^m, v) = \max\left[0, v - p, v - p_a^m, \beta \int_{a_l}^{a_u} V(p', p_a^m, v) F_a(dp')\right]. \tag{30}$$

Theorem 1: *Let a_l denote the lower support of the distribution of ask prices charged by middle men and let $v(a_l, p_a^m)$ be defined by:*

$$v(a_l, p_a^m) = p_a^m + \frac{1}{\delta} \int_{a_l}^{p_a^m} F_a(p) dp. \tag{31}$$

If $v(a_l, p_a^m) \le 1$, then there are three different optimal search/purchase strategies depending on the consumer's type. If $v \in [0, a_l)$ then it is optimal for the consumer not to participate in the market (i.e. don't search and don't trade on the commodity exchange). If $v \in [a_l, v(a_l, p_a^m))$ then it is optimal for the consumer to search for the commodity by contacting middle men. If $v \in (v(a_l, p_a^m), 1]$ then it is optimal for the consumer to purchase the commodity in the commodity exchange at price p_a^m.

We also have symmetric results for producers, namely:

Theorem 2: *Let b_u denote the upper support of the distribution of ask prices charged by middle men and let $w(p_b^m, b_u)$ be defined by:*

$$w(p_b^m, b_u) = p_b^m - \frac{1}{\delta} \int_{p_b^m}^{b_u} [1 - F_b(p)] dp. \tag{32}$$

If $w(p_b^m, b_u) \ge 0$ then there are three different optimal search/purchase strategies depending on the producer's type. If $w \in (b_u, 1]$ then it is optimal for the producer not to participate in the market (i.e. don't search and don't trade on the commodity exchange). If $w \in (w(p_b^m, b_u), b_u]$ then it is optimal for the producer to search for the commodity in the dealer market. If $w \in [0, w(p_b^m, b_u))]$ then it is optimal for the producer to sell the commodity to the market maker at the bid price p_b^m.

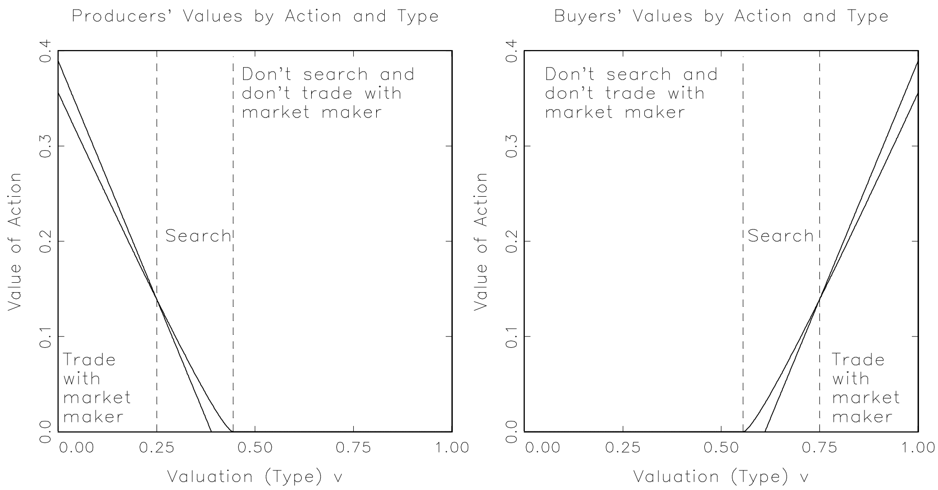

Figure 2: Decision Rules for Producers and Consumers

Figure 2 illustrates Theorems 1 and 2. It plots the net value of search, ($v - r(v)$ for consumers and $r(w) - w$ for producers) and the net values of trading on the exchange at the market maker's bid and ask prices of (p_b^m, p_a^m) given by the lines $p_b^m - w$ and $v - p_a^m$, respectively. We see that in this example that $p_b^m - w > r(w) - w$ for $w \in [0, w(p_b^m, b_u)]$, where in this example we have $w(p_b^m, b_u) = .25$. If follows that all producers in the interval $[0, .25]$ prefer to trade on the exchange rather than search. However for producers in the interval $(.25, b_u]$ we see that $r(w) - w > p_b^m - w$, and thus these producers choose to search for a better price in the dealer market rather than trade on the exchange at price p_b^m. The remaining producers with production costs $w \in (b_u, 1]$ would lose from trading on the exchange or searching, and so these producers do not trade and earn a net surplus of 0. Symmetrical results hold for consumers: buyers with valuations in the interval $v \in [0, a_l]$ do not trade and earn a surplus of 0, buyers with valuations $v \in (a_l, v(a_l, p_a^m)]$ (where in this example $v(a_l, p_a^m) = .75$) choose to search in the dealer market rather than trade on the exchange and earn a net surplus of $v - r(v) > v - p_a^m$, and buyers with valuations in the interval $v \in (.75, 1]$ choose to trade on the exchange and earn a surplus of $v - p_a^m > v - r(v)$. Note that the reservation price functions for consumers and producers are given by the same functions $r(v)$ and $r(w)$ that were defined in equations (2) and (4) of the previous section.

Since all consumers know they have the option of purchasing the commodity from the market maker at price p_a^m, and all producers know they can sell the commodity to the market maker at price p_b^m, it should be clear that no intermediary in the dealer market would be able to sell if they were to charge an ask price p_a higher than the ask price p_a^m posted in the commodity exchange, and similarly no intermediary would be able to purchase the commodity if they offer a lower bid price p_b than the bid price charged by the market maker. We record this observation as:

Theorem 3: *p_b^m is a lower bound on the support of the distribution of bid prices offered by middle men and p_a^m is an upper bound on the support of the distribution of ask prices charged by middle men in the dealer market.*

If the monopolist market maker sets a net ask price of p_a^m and offers a net bid price p_b^m, Theorem 1 implies that the quantity of the commodity that will be demanded by the highest valuation consumers in the interval $(v(a_l, p_a^m), 1]$ will result in a total quantity demanded of:

$$Q_d^m(p_a^m, p_b^m) = 1 - p_a^m - \frac{1}{\delta}\int_{.75p_a^m+.25(p_b^m+k_l)}^{p_a^m} F_a(p)dp. \tag{33}$$

Similarly, the quantity of the commodity supplied by the lowest cost producers in the interval $[0, w(p_b^m, b_u))$ results in a total quantity supplied of:

$$Q_s^m(p_a^m, p_b^m) = p_b^m - \frac{1}{\delta}\int_{p_b^m}^{.75p_b^m+.25(p_a^m-k_l)} [1 - F_b(w)]dw. \tag{34}$$

Applying Spulber's (1996a) analysis, it is not difficult to show that F_a and F_b are uniform distributions given by

$$F_a(p) = \frac{p - .75p_a^m - .25(p_b^m + k_l)}{.25(p_a^m - p_b^m - k_l)}, \tag{35}$$

and

$$F_b(w) = \frac{p - p_b^m}{.25(p_a^m - p_b^m - k_l)}. \tag{36}$$

Substituting these formulas into the formulas for Q_d^m and Q_s^m we obtain the following result:

Theorem 4: *$Q_d^m(p_a^m, p_b^m) = Q_s^m(p_a^m, p_b^m)$ if and only if $p_b^m = 1 - p_a^m$.*

For any given values of p_a^m and p_b^m satisfying this "supply=demand" constraint, the analysis of the previous section can be applied to show that the equilibrium distribution $F_a(p)$ of ask prices charged by intermediaries is uniformly distributed on the interval $[.75p_a^m + .25(p_b^m + k_l), p_a^m]$. Symmetrically the equilibrium distribution of bid prices offered by intermediaries, $F_b(p)$, is uniformly distributed on the interval $[p_b^m, .75p_b^m + .25(p_a^m - k_l)]$.

It follows that once a monopolist market maker enters and quotes public bid and ask prices (p_b^m, p_a^m), the equilibrium set of bid and ask prices offered by middle men in the dealer market adjusts to make p_b^m and p_a^m the lower and upper bounds of the support of bid and ask prices, respectively. When $k_l = 0$ the lower support point of the distribution of ask prices charged by middle men is a convex combination of the bid and ask prices set by the monopolist market maker, with 3/4 of the weight being put on the ask price p_a^m. Symmetrically the upper support point of the distribution of bid prices charged by middle men is also a convex combination of the the bid and ask prices set by the monopolist market maker, but with 3/4 of the weight being put on the market maker's bid price p_b^m. Thus, the intermediaries are acting as a "competitive fringe" that undercut the bid-ask spread set by the monopolist market maker.

The following result verifies that for any suitable choice of p_a^m and p_b^m that equate supply and demand on the exchange, there is a well-defined search equilibrium in the dealer market for those producers and consumers who were not willing to trade immediately on the exchange.

Theorem 5: *For any bid and ask prices quoted by the monopolist that satisfy $p_a^m = 1 - p_b^m$ and:*

$$\frac{1 + k_l}{(2 + 8\delta)} \leq p_a^m \leq \frac{(1 + 8\delta) + k_l}{(2 + 8\delta)}, \tag{37}$$

there exists a unique stationary search equilibrium in the intermediary market. That is, the distributions F_a and F_b given in equations (35) and (36) are unique equilibrium price distributions implied by the profit-maximizing bid-ask prices $(p_a(k), p_b(k))$ for all active intermediaries k. Consumers and producers who decide to participate in the dealer market adopt optimal search strategies based on beliefs that F_b and F_a are the correct distributions of bid and ask prices in the market for intermediaries. All intermediaries are maximizing expected discounted profits given beliefs about their supply and demand functions which are derived from the optimal behavior of consumers and producers. These beliefs are self-fulfilling, and constitute the unique stationary equilibrium in this market.

The constraint on p_a^m in (37) ensures that the market maker's demand Q_d is in the $[0, 1]$ interval. Note that when $k_l = 0$, $1/(2 + 8\delta) = b_l$ the lower bound on the distribution of bid prices in the search equilibrium without a market maker. It should be clear that it would not be possible to balance supply and demand if the monopolist chose an ask price lower than this value (in fact, since $1/(2 + 8\delta) < 1/2$ this price would not even enable the market maker to balance supply and demand and obtain a positive bid/ask spread). The upper bound on p_a^m in (37), $a_u = (1 + 8\delta + k_l)/(2 + 8\delta)$, is the upper bound on the support of ask prices in the search equilibrium without a market maker. No market maker could successfully enter and charge a higher price than the highest possible price in the pre-entry search equilibrium. Note that if

$p_a^m \le a_u$ then $p_b^m = 1 - p_a^m \ge 1 - a_u = (1 - k_l)/(2 + 8\delta) = b_l$, so the constraint implies that the monopolist must charge a bid price that is better than the lowest possible bid price in the pre-entry search equilibrium. This result is another sense in which the intermediaries constitute a "competitive fringe" that provides a check on the power of a monopolist market maker.

Clearly the number of intermediaries that can be supported in equilibrium depends on how aggressively the monopolist prices. The following result shows that the number of middle men operating in the search market is determined by the monopolist market maker's bid-ask spread.

Corollary: *The number of active intermediaries is given by*

$$N = \max(0, p_a^m - p_b^m - k_l). \tag{38}$$

If $k_l = 0$, then as long as the bid-ask spread charged by the monopolist market maker is positive, the dealer market will never be driven out of existence by the entry of the market maker. This will be true even if $p_a^m = a_l$, which is the lowest price that the most efficient profit-maximizing intermediary (with cost $k = 0$) would charge in the search equilibrium without a market maker. What happens is that some of the less efficient intermediaries are driven out of the market, but the remaining more efficient intermediaries are able to set prices below the prices that would exist in the absence of a market maker and still make positive profits. Thus the lower support point of the distribution of ask prices in the equilibrium with a market maker will be lower than the lower bound on the support in the equilibrium without a market maker. Symmetrically, the upper bound of the support of the distribution of bid prices in the presence of a monopoly market maker will be higher than the upper bound of the support of the distribution of bid prices in the equilibrium without the market maker. This implies that producers and consumers are uniformly better off due to the entry of a monopolist market maker.

We now consider the monopolist market maker's optimal pricing problem. Once we solve this we can determine by how much the entry of a monopolist market maker will increase overall market efficiency. By Theorem 5 we know that when $p_b^m = 1 - p_a^m$ we have $Q_d^m = Q_s^m$, we can write the monopolist's problem as follows

$$\max_{p_a^m, p_b^m} \, [p_a^m - p_b^m - c]\, Q_s^m(p_b^m, p_a^m) \quad \text{subject to: } p_b^m = 1 - p_a^m, \quad \frac{(1 + k_l)}{(2 + 8\delta)} \le p_a^m \le \frac{(1 + 8\delta + k_l)}{(2 + 8\delta)}. \tag{39}$$

where $c \ge 0$ denotes the total marginal costs of processing transactions on the commodity exchange. This is the sum of the direct marginal costs of operating the exchange plus the per unit rebate paid to consumers c_a and producers c_b to cover the costs of trading on the exchange. Substituting in the constraint $p_b^m = 1 - p_a^m$ and simplifying, we find that the monopolist's profit function can be reduced to maximizing a quadratic objective function with a single control p_a^m:

$$\Pi^m(p_a^m, c) = (2p_a^m - 1 - c) \max\left[0, 1 - p_a^m - \max[0, (2p_a^m - 1 - k_l)/8\delta]\right]. \tag{40}$$

It is easy to verify that provided $Q_d^m(p_a^m, p_b^m) > 0$ and $(2p_a^m - 1 - k_l) > 0$ the market maker's objective is concave in p_a^m, and thus has a unique optimal solution provided the monopolist's choice of p_a^m does not violate the search equilibrium conditions in the intermediary market. The solution to the monopolist's problem is given in Theorem 6:

Theorem 6: *If the market maker's marginal costs per transaction executed on the commodity exchange satisfies:*

$$0 \le c \le \frac{(4\delta + k_l)}{(4\delta + 1)}, \tag{41}$$

then it is profitable for the market maker to enter the market and post optimal publicly observable optimal bid and ask prices. There are three pricing regimes depending on the value of k_l. The monopoly regime occurs for k_l satisfying:

$$k_l \in \left[\frac{1+c}{2}, 1\right]. \tag{42}$$

In this regime the market maker charges the unconstrained monopoly bid and ask prices given by:

$$\begin{aligned} p_a^m &= \frac{3+c}{4} \\ p_b^m &= \frac{1-c}{4} \end{aligned} \tag{43}$$

The market maker trades the monopoly volume $Q_s^m(p_a^m, p_b^m) = (1-c)/4$ and the dealer market is priced out of existence, i.e. the quantity supplied and demanded in the dealer market is $Q_d^d = Q_s^d = 0$. The limit pricing regime occurs for k_l satisfying:

$$k_l \in \left[\frac{(4\delta+1)(1+c) - 1}{8\delta + 1}, \frac{(1+c)}{2}\right). \tag{44}$$

In this regime the market maker drives the dealer market out of existence (i.e. $Q_d^d = Q_d^d = 0$) by setting bid and ask prices equal to:

$$\begin{aligned} p_a^m &= \frac{3+k_l}{4} \\ p_b^m &= \frac{1-k_l}{4} \end{aligned} \tag{45}$$

and the market maker trades the volume $Q_s^m(p_a^m, p_b^m) = (1 - k_l)/4$. The competitive regime occurs for k_l satisfying:

$$k_l \in \left[0, \frac{(4\delta + 1)(1 + c) - 1}{8\delta + 1}\right), \tag{46}$$

The market maker coexists with a dealer market in this case. The market maker's bid and ask prices are given by:

$$\begin{aligned} p_a^m &= \frac{3 + c}{4} + \frac{(k_l - 1)}{(16\delta + 4)} \\ p_b^m &= \frac{1 - c}{4} + \frac{(1 - k_l)}{(16\delta + 4)}. \end{aligned} \tag{47}$$

The quantity traded by the market maker is given by:

$$Q_d^m(p_a^m, 1 - p_a^m) = Q_s^m(p_a^m, 1 - p_a^m) = \frac{1 - c}{4} + \frac{(1 - k_l)}{(16\delta + 4)} - \frac{1}{8\delta}\left[\frac{1 + c}{2} + \frac{(k_l - 1)}{(8\delta + 2)} - k_l\right]. \tag{48}$$

The equilibrium number of intermediaries in the dealer market is given by:

$$N = k_u - k_l = \frac{1 + c}{2} + \frac{(k_l - 1)}{(8\delta + 2)} - k_l. \tag{49}$$

The equilibrium distribution of ask prices quoted by intermediaries in the dealer market is uniformly distributed on the interval:

$$\left[\frac{5 + c}{8} + \frac{(k_l - 1) + (8\delta + 2)k_l}{(32\delta + 8)}, \frac{3 + c}{4} + \frac{2(k_l - 1)}{(32\delta + 8)}\right]. \tag{50}$$

The equilibrium distribution of bid prices quoted by intermediaries in the residual market is uniformly distributed on the interval:

$$\left[\frac{1 - c}{4} + \frac{2(1 - k_l)}{(32\delta + 8)}, \frac{2 - c}{8} + \frac{(1 - k_l) - k_l(8\delta + 2)}{(32\delta + 8)}\right]. \tag{51}$$

The equilibrium quantity demanded in the dealer market Q_d^d equals the quantity supplied Q_s^d:

$$Q_d^d = Q_s^d = (p_a^m - p_b^m - k_l)\left(\frac{1}{4} + \frac{1}{8\delta}\right) = \left[\frac{1 + c}{8} + \frac{(k_l - 1)}{(32\delta + 8)} - \frac{k_l}{4}\right]\left[1 + \frac{1}{2\delta}\right]. \tag{52}$$

Note that the fundamental condition for the viability of entry by a monopolist market maker, $c \leq (4\delta + k_l)/(4\delta + 1)$, given in equation (41) of Theorem 6 has a very simple interpretation. The upper bound $(4\delta + k_l)/(4\delta + 1)$ equals the efficiency level k_u of the marginal middle man in the equilibrium without a market maker. Thus, the monopolist market maker's per unit transactions cost must be lower than the marginal cost of the marginal middle man in the pre-entry equilibrium if entry is to be feasible.

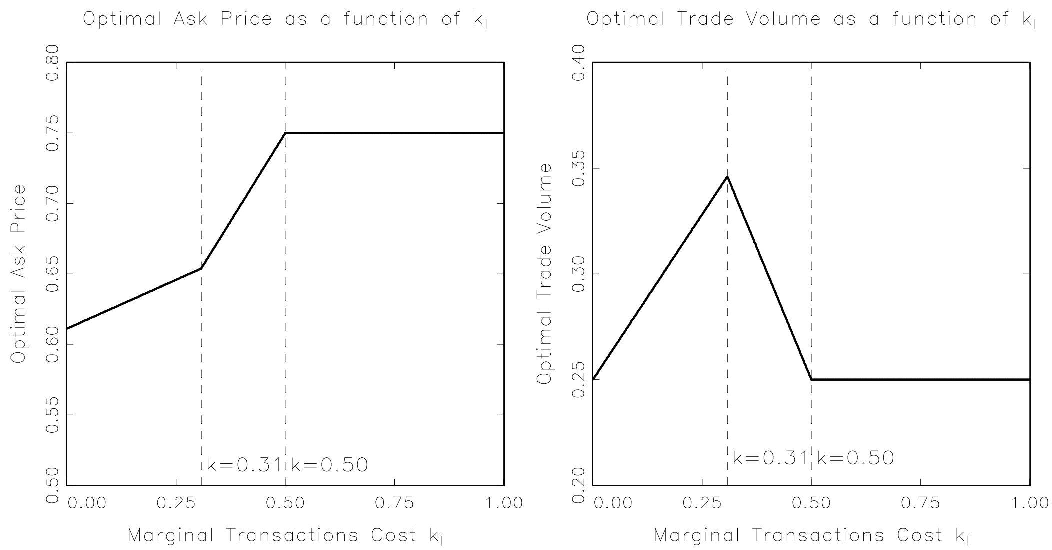

Figure 3: Market Maker's Optimal Trade Volume and Ask Price as a function of k_l

Figure 3 illustrates the market maker's pricing and quantity decisions in the three regimes in the case where $c = 0$ and $\delta = .2$. In this case the cutoff between the competitive and limit pricing regimes is $[(4\delta + 1)(1 + c) - 1]/(8\delta + 1) = .31$ and the cutoff between the limit pricing and monopoly regimes is $(1 + c)/2 = .5$. Note that the market maker's ask price and quantity traded are increasing functions of k_l in the competitive regime. As k_l increases the dealer market gets increasingly inefficient relative to transacting with the market maker, and the market maker exploits this by raising his ask price (and bid-ask spread). The volume of trade in the dealer market decreases monotonically in k_l until at $k_l = .31$ the dealer market vanishes. For $k_l \in [.31, .50)$ the market maker adopts a limit pricing strategy, choosing the largest possible ask price that will not induce entry by middle men. As k_l increases in this region the market maker is able to raise prices, leading to a reduction in quantity traded. When $k_l > .5$, the most efficient middle man is so inefficient relative to the market maker that he no longer fears the possibility of entry and is able to set the bid-ask spread equal to the unconstrained monopoly level of .75 and trade the monopoly quantity of .25. Although we did not plot the market maker's profits as a function of k_l, it should be clear that they are monotonically increasing for $k_l \in [0, (1 + c)/2]$ and constant in the monopoly regime when $k_l > (1 + c)/2$.

Figure 4 compares the equilibria with and without a monopolist market maker in the case where $k_l = 0$, $c = 0$, and $\delta = .2$. The market maker trades a total quantity of $Q^m = .25$ and the dealer maker trades a total quantity of $Q^d = .19 = .44 - .25$, which is less than half the total amount that dealers would have traded in the equilibrium without a market maker. The market maker charges a bid-ask spread of $p_a^m - p_b^m = .22 = .61 - .39$, which equals the bid-ask spread of the most efficient middle man in the equilibrium without a market maker. The net effect of entry is to substantially reduce the average bid-ask spread in the dealer market, from .1666 to .085.

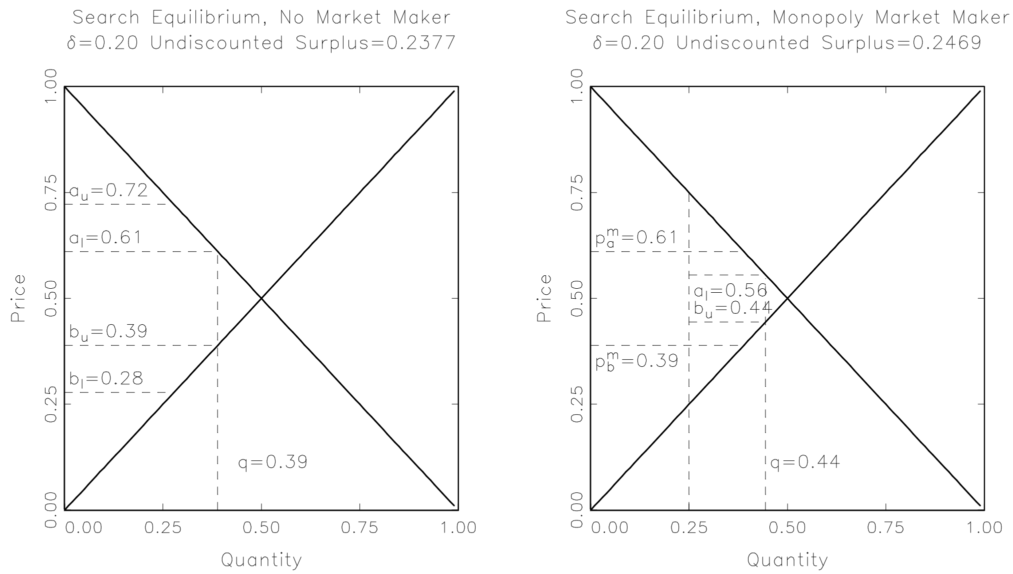

Figure 4: Comparison of Search Equilibria with and Without a Market Maker

Thus, the entry of the market maker lowers bid-ask spreads and raises quantity traded. We can show that overall surplus is higher when a market maker is present, and producers and consumers are uniformly better off. We formalize the latter result in

Theorem 7: *Let (r_p^i, r_c^i), $i = d, m$ denote the reservation price functions for producers (p) and consumers (c) in the search equilibrium with only dealer markets and no market maker $(i = d)$ and in the search equilibrium with a market maker $(i = m)$ (i.e. the competitive regime in Theorem 6). Then we have:*

$$r_c^d(v) \geq r_c^m(v) \quad \text{and} \quad r_p^d(w) \leq r_p^m(w), \tag{53}$$

with strict inequality for $w \in [0, b_u^m)$ and $v \in (a_l^m, 1]$ where b_u^m is the upper bound on the support of the distribution of bid prices in the dealer market and a_l^m is the lower bound on the support of the distribution of ask prices in the post entry search equilibrium in the dealer market.

Note that the same results also hold in the limit pricing regime even though no actual dealer market exists in this case: the threat of entry of middle men is sufficient to restrain the market maker's pricing decisions, keeping them below the unconstrained monopoly solution. So producers and consumers are uniformly better off in this regime as well.

The economic interpretation of Theorem 7 is straightforward: since the entry of the market maker narrows bid/ask spreads and improves the distributions of bid and ask prices for producers and consumers, it follows that consumers can expect better ask prices in the post-entry equilibrium and will therefore have lower reservation values. Similarly producers can expect higher bid prices in the post-entry equilibrium and thus will have higher reservation values. Since $v - r(v)$ is the expected consumer's surplus for a consumer of type v, it follows that no consumer will be worse off, and some consumers will be strictly better off in the post-entry equilibrium. Similarly, since $r(w) - w$ is the expected producer's surplus for a producer of type w, it follows that no producer will be worse off, and some producers will be strictly better off due to the entry of a monopoly market maker.

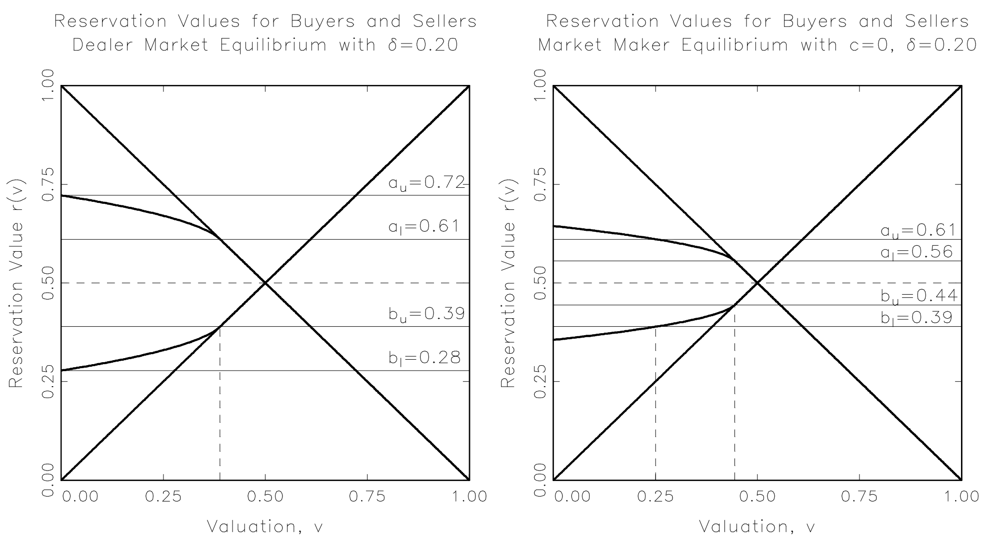

Figure 5: Comparison of Search Equilibria with and Without a Market Maker

These results are illustrated in figure 5. In figure 5 any consumer in the interval $(a_l^m, 1]$ is made strictly better off from the entry of the market maker, where $a_l^m = .56$ is the lower support point of the distribution of ask prices in the equilibrium with a monopolist market maker. The low-valuation consumers located in the interval $[0, a_l^m]$ do not trade, and their welfare is not affected by the entry of the market maker. Note however, that $a_l^d \geq a_l^m$, i.e. entry of the market maker causes the lower support of the distribution of ask prices to fall. Thus, there is a larger set of active consumers in the equilibrium with a market maker, and the consumers who switch from non-participation to search become strictly better off. Symmetric results apply to producers. We record this observation as:

Corollary: *If $c < (4\delta + k_l)/(4\delta + 1)$, entry of a monopolist market maker is strictly Pareto improving: no producer and consumer is worse off and a positive mass of producers and consumers are strictly better off in the post entry equilibrium.*

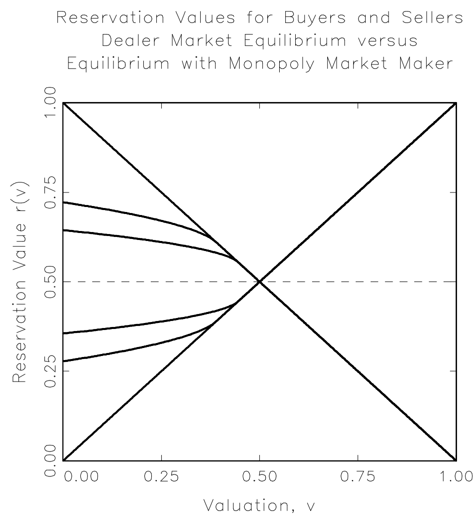

Figure 6: Reservation Values in the Pre and Post Entry Equilibria

Figure 6 summarizes the welfare implications of successful entry by the monopolist market maker when $k_l = 0$ and $\delta = .2$. We see that reservation values for the post-entry equilibrium are higher for producers and lower for consumers, and in each case are closer to the limiting Walrasian equilibrium reservation value of 1/2 represented by the dashed line in figure 6. The area between the dashed line and the reservation price functions represent the lost consumer and producer surplus due to the inefficiency of the search equilibrium. This deadweight loss is lower after the entry by the monopolist market maker, and in this sense, an exchange with posted prices represents a superior institution for conducting trade, at least provided that the cost of running the exchange is not too high. However if the exchange is a monopoly, such as is approximately the case in the NYSE, NASDAQ, and many other securities and commodities exchanges, the existence of the dealer market serves as a "competitive fringe" that prevents the market market from charging too high of a bid/ask spread. Without such competition, the monopoly market outcome (illustrated in the left hand panel of figure 1) results in a much lower total surplus and a worsening of the welfare of consumers and producers relative to the search equilibrium in the dealer market. Thus, middle men and market makers can be thought of as competing exchange institutions. The entry of a monopolist market maker results in a more competitive market by reducing bid-ask spreads and increasing producer and consumer surplus. Conversely as long as k_l is not too large, free entry of middle men into the dealer market provide an important competitive restraint on the market power of a monopolist market maker resulting in lower bid-ask spreads and increased producer and consumer surplus. However if c is sufficiently smaller than k_l or vice versa, we may observe equilibria without middle men or without market makers, respectively.

Corollary: *If $c \in [0, (4\delta + k_l)/(4\delta + 1)]$, then the number of middle men operating in the dealer market in an equilibrium where entry of a market maker is prohibited is greater than the number of middle men operating in the market with a market maker. When $k_l = 0$, the number of middle men in the post-entry equilibrium is never less than 1/2 of the number of middle men that would be operating in an equilibrium where entry by a market maker is prohibited. Equivalently, we have:*

$$\frac{2\delta - (4\delta + 1)k_l}{4\delta + 1} \leq N = \frac{1 + c}{2} + \frac{k_l - 1}{8\delta + 2} - k_l \leq \frac{4\delta(1 - k_l)}{4\delta + 1}. \tag{54}$$

Corollary: *If $c \in [0, (4\delta + k_l)/(4\delta + 1)]$, the total quantity that is ultimately traded in the search market with a monopolist market maker is greater than the total quantity that is ultimately traded in a search market without a monopolist market maker.*

Corollary: *If $c \in [0, (4\delta + k_l)/(4\delta + 1)]$, market efficiency (the sum of producers' and consumers' surplus and the total profits of middle men and market makers) is higher in the search market with a monopolist market maker.*

We conclude by studying the limiting properties of the equilibrium with a market maker as $\delta \to 0$ and $\delta \to \infty$.

Theorem 8: *Suppose that $k_l = 0$ and $c \in [0, 4\delta/(4\delta + 1)]$. In the limit as $\delta \to 0$, the equilibrium prices, quantities, and producer and consumer surpluses tend to the Walrasian equilibrium values. In the limit the market maker handles half of the transactions in the market, and the most efficient intermediaries (with $k = 0$) handle the other half of the transactions. Both charge a limiting bid/ask spread of 0 at the common Walrasian equilibrium price of $p^* = 1/2$.*

Theorem 9: *As $\delta \to \infty$ the equilibrium in the market with a monopolist market maker is more efficient than the limiting outcome in the search market without a market maker. Total quantity traded is higher, total surplus is higher, prices are lower, and fewer high cost intermediaries are operating in the limiting equilibrium when a market maker is present.*

4 Search Equilibrium with Competitive Market Making

This section considers the case where there is free entry into market making. Initially we assume that there are no fixed entry costs. In this case, free entry will result in zero profits to market makers, which implies that the bid-ask spread will be driven down to the marginal cost of executing transactions by the most efficient market maker.

Theorem 10: *Suppose that $c \in [0, (4\delta + k_l)/(4\delta + 1)]$. Then entry by market makers is feasible. In the equilibrium with free entry by competitive market makers, all market makers charge the same bid and ask prices (p_b^c, p_a^c) and no market maker earns positive profits. The equilibrium bid ask spread charged by market makers is given by:*

$$\begin{aligned} p_a^c &= (1+c)/2 \\ p_b^c &= (1-c)/2 \end{aligned} \tag{55}$$

If $k_l > c$, then middle men will be driven out of business by the entry of market makers and the total volume traded by market makers is give by $Q^c(p_a^c, p_b^c) = (1-c)/2$. If $k_l \leq c$, then entry of market makers will not succeed in driving out all middle men. Market makers will trade a total volume equal to $Q^c(p_a^m, p_b^m) = (1-c)/2 - (c-k_l)/(8\delta)$, and the total volume of trade in the dealer market is given by $Q_d^d = Q_s^d = (c - k_l)(1/4 + 1/(8\delta))$.

Corollary: *When $c = 0$, the equilibrium with free entry of market makers is identical to the Walrasian equilibrium outcome, and hence is fully efficient. All middle men are driven out of business in this case.*

If there are fixed costs to entering the market, we would expect that a positive bid-ask spread could emerge even if $c = 0$, but the bid-ask spread would only be large enough so that the present discounted value of profits from market making equals the fixed cost of entry. We do not attempt to analyze this case here because there is some indeterminacy about how many market makers will enter, and how large their respective shares of the market will be. If there is one market maker, the threat of entry by competing market makers should force the monopolist to set a limit price below the optimal bid-ask prices derived in the previous section to the point where a new entrant's expected profits from market making equals the fixed costs of entry. Clearly consumers and producers are uniformly better off as a result of the threat of such entry, even if no entry actually occurs.

5 Applications

In this section we attempt to relate the theoretical results of the preceding sections to the "real world" by pointing out some of the insights and intuitions that the model provides into the microstructure of a number of different commodity and asset markets, including steel, fish, and financial securities. We also compare our analysis and conclusions to two recent related studies. In the conclusion we will discuss some of the limitations of our analysis for understanding the microstructure of these markets.

The theory in this paper appears to be applicable to a wide range of asset and commodity markets: in almost all of these markets we observe analogs of middle men and market makers. Our predictions of the positive welfare benefits from competition between middle men and market makers also seems to be relevant for policymaking, especially in connection with financial markets where there is substantial interest in the question of whether entry of middle men can help reduce the market power of market makers such as the New York Stock Exchange (NYSE). The NYSE is a collection of market makers known as *specialists.* Each specialist is responsible for creating a market (maintaining an inventory and providing liquidity) for one or more individual securities. They do this by holding inventories and posting publicly observable bid and ask prices at which they are willing to transact. However there is also an active "dealer market" which is commonly referred to as the "over the counter market" (OTC). The middle men in the OTC market are known as *broker/dealers.* Pirrong (2000) estimates that the OTC market comprises only 8 percent of the volume and 10 percent of the transactions in NYSE listed securities. The relatively small share of trade conducted by middle men may be due to attempts by the NYSE to maintain its monopoly position in market making by limiting competition from middle men through tactics such as its "Rule 390" that prevents NYSE members from trading in the OTC market. The NYSE has also attempted to prevent or limit entry by other market makers through vigorous lobbying of the Securities and Exchange Commission (SEC). This lobbying has succeeded in creating high regulatory barriers to entry to competing exchanges such as R. Steven Wunsch's computerized Arizona Stock Exchange. However in recent years the advent of "electronic communication networks" (ECNs), competition from computerized foreign exchanges, and a change to a more pro-competitive regulatory regime at the SEC has subjected the NYSE to much more competitive pressure than it has experienced in the past, forcing it to make changes it had previously resisted such as moving towards 24 hour trading and allowing prices to be quoted in decimals rather than in 1/8 increments. However it is not yet clear whether this increased competitive pressure will have its most important impact, namely, forcing NYSE specialists to reduce their bid/ask spreads.

Except for the recent work by Pirrong (2000), there has been relatively little theoretical attention to the question of competition between middle men and market makers in the financial market microstructure literature.[2] Pirrong considers the effect of competition between a monopolist market maker and middle men in a dealer market, which he refers to as the "third market". His analysis focuses on the effects of the effect of entry by middle men on an initial equilibrium with a monopolist market maker, whereas we have focused on the effect of the entry of a monopolist market on an initial search equilibrium where there is free entry and exit of middle men. Pirrong's model also focuses on the issue of differentially informed traders, and the question of whether the creation of a third market would result in free riding on the "price discovery" provided by a monopolist market maker. The NYSE has advanced this latter argument to the SEC as its rationale for the need to limit the formation of third markets and competing exchanges. However Pirrong concludes that "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." (p. 2). We view this as consistent with our result that in the absence of free entry into market making, free entry by middle men into the dealer market provides a significant competitive threat to a monopolist market maker, forcing it to substantially reduce its bid/ask spreads and resulting in a significant welfare gain to buyers and sellers. In our model, entry of competitive middle men will result in undercutting of the market maker's bid-ask spread and will result in additional price dispersion compared to an initial equilibrium with a monopolist market maker and no dealer market. The monopolist market maker might claim that this additional price dispersion is evidence that the middle men are free riding on price discovery and are creating unnecessary price dispersion, but our analysis indicates that any costs associated with the extra price dispersion are outweighed by the benefits of the additional competition from middle men that force the market maker to reduce his bid/ask spread.

There is also a more IO-oriented microstructure literature that Spulber (1999) has surveyed; but apart from the work by Gehrig (1993) that we discussed in the introduction, Spulber's book

[2]For example O'Hara's (1997) book on financial market microstructure contains a detailed discussion of market makers and specialists on the NYSE, and describes briefly the role of broker/dealers, but the book contains no formal analysis of the interaction and competition between specialists and broker/dealers.

(and the overall literature) devotes relatively little attention to the issue of competition between middle men and market makers. The only other related work that we are aware of is a recent paper by Baye and Morgan (2001) that studies the impact of entry of market makers (which they refer to as "information gatekeepers") on an initial equilibrium where geographic dispersion of information and high costs of search give producers monopoly power over their local consumers. Examples of information gatekeepers are web sites such as `Shopper.com` or `Mortgage-quotes.com` that help to centralize dispersed price information and reduce search costs by allowing buyers and sellers to post bid and ask prices on the web. The Baye-Morgan model predicts that in equilibrium, the gatekeeper charges fees that are sufficient to induce all consumers to subscribe to the gatekeeper's service. Even though all consumers subscribe to this service and all posted prices are costlessly observable, the Baye and Morgan model yields the paradoxical result that "price dispersion persists even though in equilibrium all consumers purchase from a firm offering the lowest price." (p. 35). This result is a consequence of their assumption that there are a continuum of consumers but only a finite number of producers. Since any potential pure strategy equilibrium pricing strategy can be broken by undercutting the best alternative price by ϵ, the only Nash equilibrium for firms' pricing game is a mixed strategy equilibrium where advertising and pricing decisions are randomized. This is the source of the equilibrium price dispersion in the Baye and Morgan paper. Baye and Morgan conclude that the entry of such gatekeepers has ambiguous social welfare effects. We will discuss how their analysis and conclusions relates to ours in more detail shortly.

It is reasonable to suppose that the advent of the world wide web could have a significant impact on the microstructure of a wide variety of markets. We think our model offers insights into the effects of this technology and the recent proliferation of "business to business" (B2B) web-based trading sites and exchanges. Unlike the Baye and Morgan study, our analysis suggests that the entry of web-based market makers has had unambiguous positive welfare effects. Computerization and the advent of cheap technology for conducting secure transactions over the world wide web has greatly reduced the cost of making "public broadcasts" of bid/ask prices compared to the pre-web technology that relied on newspapers, faxes, telephones, and telexes. Computerized trading systems have greatly reduced the costs of executing transactions by allowing automated matching of trades and bookkeeping. It has also greatly reduced entry costs for potential market makers since computerized trading systems can now be created at a fraction of the cost of creating traditional human "trading pits" such as the Chicago Board of Trade. In effect, the "information revolution" has dramatically lowered both the cost of entry and the marginal cost of undertaking transactions by market makers. The theory of the previous sections suggests that this should lead to the entry and proliferation of computerized market making services, and a corresponding decline in transactions in traditional dealer markets. By in large, this is indeed what is happening. The rates of growth in the B2B market are nothing short of phenomenal: Forrester Research predicts that sales via computerized market makers will expand fivefold in the next two years, and will account for at least 25% of all sales in 2002, and a total volume of $1.4 trillion in transactions by 2004.

However when we look at individual markets such as steel the story is not quite so simple. Presumably, computer/communication technology and the world wide web have also significantly lowered the cost of search (which is represented in our model by the discount parameter δ). Our theoretical results suggest that the net impact of computerization on various markets depends on whether transaction and entry costs for market makers have declined more rapidly than search costs and transactions cost in the dealer market. To get more insight into this issue will we focus on two particular markets, fish and steel, that have exhibited marked differences in the acceptance and penetration levels by new computerized market maker entrants. Our analysis suggests that the relative costs of search versus transacting online should be a key to explaining the division of exchange between market makers and dealers. However we also find that historical inertia and institutional details such as the particular rules that market makers choose for conducting online transactions can have a big impact on the relative attractiveness of trading via a market maker versus searching in the dealer market. Unfortunately, our model has little to say about the impact of these latter details.

An important conclusion of our analysis of the steel and fish markets is that "transparency" is extremely important for the success of market makers. Web-based market makers such as `gofish.com` appear to be successful because the bid/ask prices and histories of previous transactions on this site are essentially easily accessible public information. On the other hand in the steel market the web-based market maker `e-STEEL.com` has been far less successful. We think this is due in part to the fact that current bid/ask prices and histories of previous transactions are not publicly posted on this site. Instead `e-STEEL.com` is more akin to a "computerized chat room" where private transactions are negotiated. As such, `e-STEEL.com` does not really constitute a market maker in the sense of posting publicly observable bid/ask prices. Instead it appears to be more akin to a computerized extension of the existing dealer or "telephone market" for steel, but where individual deals are negotiated by typing messages into a computer terminal rather than being conducted verbally over the phone. To the extent that most of the producers and consumers in the steel market find it easier to negotiate over the phone than by typing messages on a computer terminal, it might be that `e-STEEL.com` not only fails to perform a role as a market maker, it also constitutes an inferior technology for intermediation in the dealer market compared to the pre-existing telephone technology. For this reason we do not expect `e-STEEL.com` will have a revolutionary impact on the microstructure of exchange in the steel market, at least not in its current form.

It is important to observe at the outset that neither `e-STEEL.com` nor `gofish.com` are attempting to operate as explicit market makers in the same way as a NYSE specialist does, i.e. neither of them hold their own steel or fish inventories and neither of them posts their own bid/ask prices. Instead, these sites operate more like "information gatekeepers" that Baye and Morgan (2001) have analyzed. Both `e-STEEL.com` and `gofish.com` allow buyers and sellers to post the quantity, location, and the price of goods for sale or purchase, but neither site holds any inventories or conducts any transactions themselves and both view themselves as neutral market participants. On both of these sites, posting bid and ask prices are free; however if a transaction is consummated, both sites charge the seller a transaction fee which is a fixed percentage of the sale (0.875% for `e-STEEL.com`, 0.75% for `gofish.com`, with plans to increase the the fee to 1.6% in the near future). `gofish.com` also imposes an initial initiation fee of $1,500 for a buyer, $4,000 for a seller, and $3,000 for a buyer/seller. `e-STEEL` has no membership or initiation fees. The analysis by Baye and Morgan suggests that the `e-STEEL.com` strategy may be the better approach: one of their main conclusions is that "the gatekeeper's profits are maximized in an equilibrium where fees charged for consumer access are set low enough to induce all consumers to subscribe." (p. 35).

We can reinterpret the market maker in our model as an information gatekeeper that charges commissions to buyers and sellers rather than buying and selling on its own account and charging its own bid/ask spreads. In our model, a monopoly gatekeeper that charges a symmetric per unit transaction fee equal to one half of the bid-ask spread charged by a monopolist market maker implements the same equilibrium outcome. Thus, our model predicts that there can be no dispersion in the bid and ask prices posted on the gatekeeper's web site, but there will be price dispersion in the dealer market due to the search/matching frictions discussed in the previous sections. Unlike Baye and Morgan (2001), our results imply that entry of a monopolist gatekeeper has unambiguously positive effect on welfare, resulting in narrower bid/ask spreads. To see the equivalence between gatekeeping and market making in our model, note that when there is a gatekeeper, buyers and sellers who agree to pay a commission are free to post public bid and ask prices on the gatekeeper's web site. It is straightforward to verify that the unique equilibrium outcome of this alternative interpretation of our model is exactly the same as the outcome with a price setting market maker. In particular, suppose the gatekeeper charges a per unit commission or transaction fee of τ to buyers and sellers to post prices on the site. Thus if a buyer posts a bid of p_b and succeeds in transacting, the total per unit cost he actually pays would be $p_b + \tau$. Similarly if a seller posts an ask price of p_a and transacts, the seller's actual per unit proceeds are $p_a - \tau$. With a continuum of buyers and sellers, the only Nash equilibrium outcome is for all buyers to post a common bid price p_b and all sellers to post a common ask price p_a. For any given value of τ, supply and demand for units advertised by the gatekeeper will be equated if and only if $p_a = 1 - p_b$. Consider an equilibrium where $p_a = p_b = 1/2$ and the gatekeeper's commission is one half the bid-ask spread charged by a monopolist market maker, namely $2\tau^* = (p_a^m - p_b^m) = (1+c)/2 + (k_l - 1)/(8\delta + 2)$ where c is the gatekeeper's marginal cost per transaction (i.e. the marginal cost of posting a bid/ask pair on the gatekeeper's web site). In this equilibrium, a seller posts an ask price of 1/2, but net of the commission, the seller receives a per unit sales price equal to $1/2 - \tau^* = p_b^m$, which is the same price he would have obtained from a monopolist market maker. Symmetrical remarks apply to buyers. So unlike Baye and Morgan, we find that the gatekeeper's optimal commission is the same for both buyers and sellers. Further, it is not the case that in equilibrium all buyers will choose to post prices on the gatekeeper's site: some will choose to search for better prices offered by middle men in the dealer market. Our results suggest that it is immaterial whether a market has a monopolist market maker or a monopolist gatekeeper: both lead to exactly the same equilibrium outcome.[3]

Although certain aspects of the "details" of how transactions are carried out are not uniquely determined by our model, it does capture what we regard as the key difference between dealer

[3]Note that the gatekeeper can use a variety of different commission structures to implement the same outcome as a monopolist market maker. For example the gatekeeper might charge 0 fees to buyers who post bids on the site, but a commission equal to $\tau^* = (p_a^m - p_b^m)$ to sellers. In this case the equilibrium outcome would be for all buyers to place a bid equal to p_a^m and all sellers would place asks equal to p_a^m. The net of commission price received by sellers is then the same as the bid price p_b^m that a monopolist market maker would choose.

markets and markets run by market makers or gatekeepers, namely, that price information is publicly available to all market participants. This difference, which we have referred to as "transparency", appears to be a crucial factor behind the very marked differences in the success of `gofish.com` and `e-STEEL.com`. Although both sites charge commissions, `gofish.com` has several features designed to make it easy for buyers and sellers to observe prices, whereas bid, ask and transaction prices are not easily observable on the `e-STEEL.com` site. Both web sites allow buyers and sellers to post prices, and right next to each posted price is a "negotiate" or "counter-offer" button. It is our understanding that these posted prices are much like the list prices for a new car; these prices represent "first offers", not "take or leave it" prices. Negotiation is expected. Nevertheless it is much easier to find out the market price of fish on `gofish.com` than it is to find the market value of steel on `e-STEEL.com`. In addition to allowing posted prices, `gofish.com` runs double auction markets for certain types of fish where real market prices are determined and reports the prices of every transaction that was executed via the site, not only in the auction markets but also for the individually negotiated transactions as well. As a result there is no point in a posting unrealistic ask or bid prices on `gofish.com` since buyers and sellers can readily observe recent transaction prices.

In contrast `e-STEEL.com` does not post historical transaction data and even allows buyers and sellers to limit who can look at their own postings. There is a question on the "frequently asked questions" page of the `e-STEEL.com` web site, "Does e-STEEL create pricing transparency?" The posted answer is that "Since e-STEEL is not an auction, your pricing remains private. e-STEEL preserves your current way of doing business since online negotiations and transactions between you and your trading partners are kept private and secure." This emphasis on privacy makes it difficult to learn much about the current market price of steel from visiting `e-STEEL.com`. For this reason we conclude that `e-STEEL.com` is simply offering an alternative communication channel to the telephone to enable buyers and sellers to privately negotiate. It is not fulfilling the role of a market maker or an information gatekeeper that posts publicly observable bid and ask prices.

In both the steel and fish markets, timely delivery and payment and quality assurances are of first-order importance (as they say in the fish market, "either sell it or smell it"). From our conversations with several steel executives, their primary reluctance to switch from their traditional way of dealing business with known customers over the telephone to using the Internet to buy steel in more anonymous auction markets is the uncertainty about the timing of delivery and the quality of steel that is sold. Several steel executives have emphasized that many sellers of steel do not deliver the goods when promised and that the quality of the steel that shows up can often be lower than promised. A certain fraction of the units in an order for a supposedly homogeneous steel product can actually exhibit a number of different potential non-uniformities or defects, which when combined with informational asymmetries can result in "lemons" type problems. For this reason many steel purchasers may be unwilling to purchase steel from a supplier that they do not know and, in particular, are reluctant to go on a web site to consider purchasing steel from unidentified companies or traders.

This difficultly in learning about the true quality of steel for sale is mentioned in an opinion article in the *American Metal Market* by Arnold Koldenhoven, President of Arrowhead Steel Company (a SSC).[4] Mr. Koldenhoven asks "Will our customers use the Internet to source steel and bypass us?" His response is that "The Internet will not significantly change the way end-users look at their steel suppliers." He states that

> The question, then, is will sites such as Metalsite and E-Steel bring buyers and sellers in such a way as to efficiently distribute steel? The key is information — timely, accurate information that helps us identify the materials we are bidding on. Excess and secondary steel, the offerings we on seeing on the Internet to date, are not so much of a known product. There are scores of possible steel defects that affect applications of material, so our knowing the full descriptions of material is crucial to our making resales. I know Metalsite and their partners are trying to standardize defect descriptions that make up the unique characteristics of each coil. How well they succeed is vital to our success in using the Internet to purchase.

On the selling side these executives are reluctant to sell to people they have no credit information about. The steel executives we talk to complain bitterly about customers that do not pay up, or who are late in paying bills.

These same concerns also exist in the wholesale fish market. But `gofish.com` appears to gets around these problems in two ways. First, `gofish.com` has set up a grading system where buyers have an opportunity to to grade suppliers on every transaction.[5] This system is a useful screening device and a powerful incentive to ensure prompt delivery and quality as promised. They also

[4] Koldenhoven, Arnold "Internet Another Tool for Trading Steel" *American Metal Market*, April 7 1999, p 14.

[5] Such a grading system is not unique to `gofish.com`; for example, `e-bay.com` has similar system.

have a conflict resolution department to resolve disputes. Second, on the selling side `gofish.com` exploits its ties to Seafax, a major provider of credit information on firms in the wholesale fish business. Hence sellers can lookup the credit history on any potential buyer before agreeing to the sale. Furthermore, buyers on `gofish.com` acquire a line of credit with `gofish.com`. Part of the initiation process on `gofish.com` involves a credit check. If the buyer does not pay, `gofish.com` covers the debt. Before an transaction is consummated the seller knows how many dollars of the transaction will be covered by `gofish.com`.

Once again, the contrast with `e-STEEL.com` is stark. `e-STEEL.com`'s response to problem of buyers who don't pay is to give the seller the option of restricting access to their postings through their STEELDIRECT program. That is, a seller on `e-STEEL.com` can list which companies can and cannot observe it's postings. `e-STEEL.com` performs no credit checks on members and provides no information about the past payment/delivery history of its members. In particular a frequently asked question page on the `e-STEEL.com` web site states:

> e-STEEL does not change the claims process; claims continue to be handled between companies. However, all trades are fully documented on the e-STEEL system, facilitating discovery in the case of conflict. e-STEEL strictly monitors compliance with exchange rules, and in the case of misuse penalizes offending members. This ensures that e-STEEL remains a venue for quality industry participants. Sales contracts between the two parties continue to regulate the actual transaction.

The `e-STEEL.com` web site claims that they will provide credit checks, credit lines and financing in the future, but they do not currently do so. While `e-STEEL.com` allows sellers to restrict access to data, in contrast `gofish.com` posts credit history data directly on their web site, another important aspect of its greater transparency.

However the fact that a particular entrant such as `e-STEEL.com` has not been successful still doesn't solve the puzzle of why a "real" market maker hasn't entered the steel market. As we discussed in the introduction, the high degree of unexplained price dispersion in the daily transactions prices for particular steel products revealed in the recent study by Athreya (2001) suggests that there are significant inefficiencies in the existing dealer market for steel. To the extent that our theoretical analysis is any guide, we would expect that successful entry of a market maker could be very profitable and would significantly reduce the amount of price dispersion.

However while successful market maker would probably drive a significant number of the 5,000+ plant locations of steel service centers in the U.S. out of business, the result would be a more concentrated but leaner and more competitive market with significantly lower bid/ask spreads. We now discuss some of the historical inertial and "sociological" factors that might be at play and might lead to a more general resistance to trading through computerized market makers. Our analysis has benefited from our opportunity to learn about these factors first-hand, via direct conversations of executives in the steel service center industry (for further details of our "case study" of this industry and individual middle men operating in it, see Athreya *et. al.* 2001).

From our observation and discussion with middle men in the steel business, the industry is somewhat "old fashioned" in its reliance on the telephone as the preferred communications medium for negotiating transactions. Our impression is relatively few traders are comfortable using computers to conduct transactions. Furthermore, it appears that knowing the identity of a buyer or seller is very important due to the factors discussed above. To the extent a computerized market would enable parties to conceal their true identities, this would be regarded as a step backward. Indeed, we observe many signs of resistance to computerization and centralized market making by executives in the steel service center industry, and perhaps this isn't surprising since our theory predicts that the least efficient middle men will be squeezed out by the entry of a market maker, and the remaining middle men will have lower profits as a result of the reduction in bid/ask spreads. All of these sociological and inertial factors can be thought of as adding to the marginal costs facing a prospective market maker who is considering entering the steel market. Our analysis suggests that if these costs are too high, entry won't be profitable.

One of the steel executives that we have direct contact with is interested in the possibility of computerized market making as a potential line of business for his company, but so far he is very reluctant to step up and post any prices for any of the products he carries on the company's web site. He states that his main fear is that if he posts a price of x, his competitors will observe this and will quote private prices of $x - \epsilon$ and steal most of his business. Our analysis suggests that middle men will in fact undercut the publicly posted bid and ask prices charged by a market maker, but this undercutting will not cause the market maker to lose all or even most of his business. Instead, our model predicts that entry of a market maker will lead to a surge of business by buyers and sellers who prefer to trade at publicly posted prices rather than to engage in costly search. However this argument has not been successful in convincing this executive to enter as a market maker thus far. This executive sees lots of other of impediments to creating a market for steel. As he notes, "the steel business is stupid" by which he means that there are some ingrained practices that could take a long time to change. Perhaps the best and cheapest way to post publicly observable bid and ask prices is via the company's web site, but if most of the buyers and sellers in the steel industry don't use the web or don't know about his site, then his company may not receive the surge in demand that our model predicts. These factors are implicitly accounted for in our model by the private hassle costs of using the web for trading: our analysis suggests that if these costs are sufficiently high, then indeed there will be no room for profitable entry by a market maker. However we expect these transactions costs will drop over time as more and more younger, computer literate traders assume key roles as buyers and sellers in the steel industry. Most of the steel executives we have spoken with believe that sooner or later, web-based market makers will be successful in capturing a large share of the trade in steel products. We conclude that there may eventually be an opportunity for successful entry of a market maker into the steel business, but perhaps not right now.

6 Conclusions

This paper has developed a theory of competitive exchange in which the microstructure of exchange is determined endogenously. We have done this by introducing a fourth type of agent, *market makers* into the equilibrium search model with competitive middle men developed by Spulber (1996a). We have shown that middle men and market makers represent complimentary and competitive exchange institutions: market makers post publicly observable bid and ask prices whereas prices quoted by middle men in the "dealer market" constitute private information that can only be obtained through a costly search process. We have focused on the effect of entry by a monopolist market maker on an initial equilibrium where there is free entry by competitive market makers. We have demonstrated that whenever the market maker's marginal cost of executing transactions is lower than the marginal cost of the least efficient middle man operating in the dealer market, entry by a market maker will be profitable and will result in an Pareto improvement for buyers and sellers (with a positive mass of buyers and sellers becoming strictly better off). This occurs because the market maker succeeds in driving out the least efficient middle men and reducing bid ask spreads. The entry of the market maker creates a segmented market where the highest valuation buyers and lowest cost sellers prefer to trade with the market maker, whereas set of buyers and sellers with intermediate valuations prefer to search for better prices in the dealer market (the remaining set of the lowest valuation buyers and highest cost sellers prefer not to participate in either market). Conversely, entry by competitive middle men can substantially reduce the bid ask spreads that a monopolist market maker would quote in the absence of competition from a dealer market. Thus, while a market maker that quotes publicly posted prices is arguably more efficient type of exchange institution, we find that as long as there are sufficiently low barriers to entry that enable sufficiently efficient middle men to enter the market, the co-existence of a dealer maker with atomistic middle men and an exchange run by one or more market makers is a remarkably a robust outcome. The middle men serve as a "competitive fringe" that limits the market power of a monopolist market maker. Each type of trading institution competes with each other, and the beneficiaries of this competition are the buyers and sellers of the asset or commodity.

Figure 7 summarizes the range of equilibrium outcomes predicted by the model. The type of equilibrium outcome depends on three key parameters: the search cost parameter δ, the marginal transactions cost of the most efficient market maker c, and the marginal transactions cost of the most efficient middle man k_l. The left hand panel displays the four possible equilibrium regimes that occur for various combinations of the (c, k_l) parameters when the search cost is fixed at $\delta = .2$. We see that the model can predict that no market maker will be present if c is sufficiently high relative to k_l (region 1), and that no middle men will exist if k_l is sufficiently large relative to c (regions 3 and 4). Region 2 represents the intermediate slice of (c, k_l) values that permit a coexistence of middle men and market makers. The right hand panel of figure 7 shows the share of trade handled by the market maker. We see that in region 2 this share increases linearly in k_l for any fixed c, or conversely, declines linearly in c in for any fixed k_l. We can conceptualize the steel industry as corresponding to values of (c, k_l) in region 1 where entry by a market maker is currently unprofitable, whereas in the securities industry, entry barriers to middle men and other market makers created by the New York Stock Exchange corresponds to values of (c, k_l) that are close to region 3 where the market maker engages in limit pricing, i.e. it chooses the largest possible bid-ask spread subject to the constraint that this spread is not sufficiently high to encourage significant entry of dealer/brokers into the OTC market.

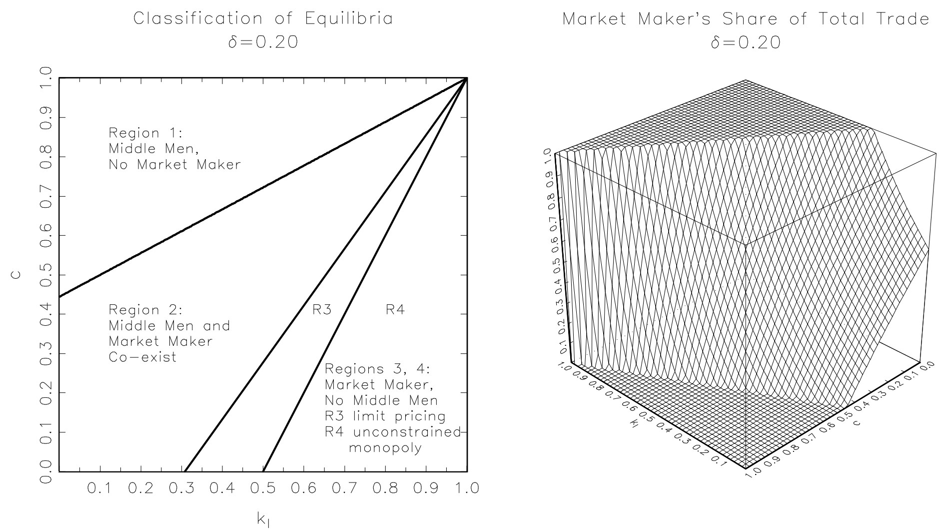

Figure 7: Summary of Equilibria, $\delta = .2$

Although our model is highly simplified and stylized, we believe it provides important insights into the structure of a variety of different real world asset and commodities markets. It also offers a basis for understanding how the "information revolution" might affect the microstructure of these markets, and for thinking about the welfare effects of policies designed to foster competition between different types of exchange institutions.

However we recognize that there are many limitations to our analysis that qualify the types of conclusions we can draw from it. First, in our model, all exchange is conducted either through middle men or market makers; we do not allow for the possibility of direct unintermediated exchange between buyers and sellers (producers and consumers). As we noted in the introduction, only 50% of the volume of trade in steel occurs through middle men, the remaining 50% represents direct exchange between producers and consumers. We would like to extend our model to account for these types of direct transactions. Presumably there are fixed as well as marginal costs of engaging in search. Consumers who buy large quantities of a commodity on a repeated basis may find it optimal to incur fixed setup costs necessary to adopt improved search technologies that enable them to bypass middle men and market makers and purchase the commodity through direct negotiations with producers. Examples include large automobile producers and ship builders who find it more economical to procure steel through their own in-house purchasing and inventory management units than to rely on purchases from middle men.

A second limitation of our analysis is that we constrained the supply and demand for the commodity for both market makers and middle men to be equal in every period. As a result these agents have no inventory holdings in our model. While this may be a reasonable assumption for the fresh fish market, in most other markets intermediaries do hold substantial inventories because supply and demand is stochastic and not guaranteed to match period by period. An important function of market makers (specialists) and middle men (dealer/brokers) is to hold inventory to provide a buffer stock that offers their customers "liquidity" in times when there is an imbalance between supply and demand. In the securities business, liquidity means being able to buy or sell a reasonable quantity of shares on short notice. In the steel market liquidity is also associated with a demand for "immediacy", i.e. a customer can be guaranteed of receiving shipment of an order within a few days of placement.

The flip side of liquidity is the problem of "thin markets". If there are significant fixed costs to setting up an exchange and non-negligible marginal costs to operating it, then unless the volume traded on the exchange is sufficiently large it may not be profitable to operate it. This is a commonly cited reason for why there are no exchanges for highly differentiated commodities. There are thousands of different types of steel and this may be part of the explanation for why markets have not arisen for individual steel products so far. However just as there are specialists who provide liquidity by holding inventories and posting public bid and ask prices for thinly traded stocks, it is still puzzling why there are not market makers who do the same for individual steel products. Middle men do play this role, but the key difference is that middle men in the steel market are unwilling to post publicly observable bid and ask prices. As we discussed in the previously, this is partly due to a fear that their publicly posted bid and ask prices will be undercut by their competitors. While this undercutting is predicted to occur in our model, our model predicts that it is profitable to post publicly observable bid and ask prices because the inflow of business from the highest valuation buyers and cost valuation sellers greatly exceeds any business that would be "stolen" by other middle men who undercut the market maker's posted prices. To better understand these fears, we need to consider a model where posted prices may not be immediately observable to all buyers and sellers in the market. If it isn't the case that everyone observes the publicly posted bid and ask prices, the market maker may not experience the instantaneous surge in demand predicted by our model, and our conclusions about the desirability and positive welfare effects of entry by a market maker might need to be modified.

A final issue has to deal with the possible problems of breakdown in "price discovery" that may occur when there is free entry into market making. When there are sufficiently many market makers quoting their own prices or operating exchanges, there is an issue of whether there will be sufficient arbitrage to guarantee that the prices quoted by all these different market makers will be the same. If not, then consumers will need to engage in a potentially costly search process to find the market maker with the best bid or ask price. In this case it appears that many of the informational gains of having a single market maker who posts a single set of publicly observable bid and ask prices could be lost, and indeed it becomes less clear how a market populated by many competing market makers differs from a dealer market populated by many competing middle men. A deeper analysis would be required to addresses the question of whether market making has elements of "natural monopoly" or whether regulations would be required to create a single gatekeeper who posts all of the bid and ask prices of all the competing market makers. Should the government serve as this "gatekeeper", and if so, isn't this central gatekeeper tantamount to a monopoly market maker one level higher up?

We are presently in the process of econometrically estimating and testing a detailed dynamic model of the inventory and pricing decisions of middle men in the steel market (see Hall and Rust, 1999), and it is likely that this model could be extended to provide a dynamic model of specialist behavior in financial markets (O'Hara and Oldfeld 1986 were among the first to develop theoretical models of specialist behavior). Once this model is complete we plan to build and estimate realistic dynamic models of producers and consumers of steel. These models will allow us to develop a more realistic dynamic spatial equilibrium model of the steel market. If it is possible to develop such a model (it would almost certainly need to be solved numerically rather than analytically) we could use it to assess the profitability and welfare gains from the entry of a market maker. However we believe this first, simple model provides some intuition and guidance about what to look for and expect in a more realistic, but inherently more complex analysis.

References

[1] Athreya, R. (2001) "Price Dispersion in the Wholesale Market for Steel" manuscript, Department of Economics, Yale University.

[2] Athreya, R, T. Bewley, W. Brainard, H. Chan, G. Hall, J. Rust and H. Scarf (2001) "Intermediation in the Steel Market" manuscript, Yale University.

[3] Baye, M. and J. Morgan (2001) "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets" forthcoming, *American Economic Review.*

[4] Gehrig, T. (1993) "Intermediation in Search Markets" *Journal of Economics and Management Strategy* **2-1** 97–120.

[5] Hall, G. and J. Rust (1999) "Econometric Methods for Endogenously Sampled Time Series: The Case of Commodity Price Speculation in the Steel Market" manuscript, Yale University.

[6] Hall, G. and J. Rust (2000a) "An Empirical Model of Inventory Investment by Durable Commodity Intermediaries" *Carnegie-Rochester Conference Series on Public Policy* **52** 171–214.

[7] Hall, G. and J. Rust (2000b) "The (S, s) Rule is an Optimal Trading Strategy in a Class of Commodity Price Speculation Problems" manuscript, Yale University.

[8] O'Hara, M. and G. Oldfield (1986) "The Microeconomics of Market Making" *Journal of Financial and Quantitative Analysis* **21** 361–376.

[9] O'Hara, M. (1997) *Market Microstructure Theory* Basil Blackwell, Oxford.

[10] Pirrong (2000) "Third Markets and the Second Best" manuscript, Washington University.

[11] Spulber, D. (1996a) "Market Making by Price Setting Firms" *Review of Economic Studies* **63** 559–580.

[12] Spulber, D. (1996b) "Market Microstructure and Intermediation" *Journal of Economic Perspectives* **10** 135–152.

[13] Spulber, D. (1999) *Market Microstructure: Intermediaries and the Theory of the Firm* Cambridge University Press.

[14] Steel Service Center Institute (1999) "Steel Market Outlook" Chicago, IL.