Middlemen versus Market Makers: A Theory of Competitive Exchange

John Rust

University of Maryland

George Hall

Yale University

We present a model in which the microstructure of trade in a commodity or asset is endogenously determined. Producers and consumers of a commodity (or buyers and sellers of an asset) who wish to trade can choose between two competing types of intermediaries: "middlemen" (dealer/brokers) and "market makers" (specialists). Market makers post publicly observable bid and ask prices, whereas the prices quoted by different middlemen are private information that can be obtained only through a costly search process. We consider an initial equilibrium with which there are no market makers but there is free entry of middlemen with heterogeneous transactions costs. We characterize conditions under which entry of a single market maker can be profitable even though it is common knowledge that all surviving middlemen will undercut the market maker's publicly posted bid and ask prices in the postentry equilibrium. The market maker's entry induces the surviving middlemen to reduce their bid-ask spreads, and as a result, all producers and consumers who choose to participate in the market enjoy a strict increase in their expected gains from trade. When there is free entry into market making and search and transactions costs tend to zero, bid-ask spreads of all market makers and middlemen are forced to zero, and a fully efficient Walrasian equilibrium outcome emerges.

We received helpful feedback from Ian Domowitz, Bruno Jullien, Deborah Minehart, Daniel Spulber, the editor John Cochrane, and three anonymous referees. We are grateful to an anonymous steel "middleman" who gave us access to confidential data on steel prices and inventories that initiated our inquiry into this topic. This research was funded by National Science Foundation grant 9905145.

[*Journal of Political Economy,* 2003, vol. 111, no. 2]
© 2003 by The University of Chicago. All rights reserved. 0022-3808/2003/11102-0006$10.00

> In the early 1970s, the firm [Cantor Fitzgerald] vaulted to nearly monopolistic pre-eminence in the bond market by doing something no one else had done before: Cantor began posting government bond prices on computer screens for clients, mainly Wall Street bond dealers, bringing transparency to a market where opacity and rumor had been the norm. [Zuckerman, Davis, and McGee 2001, p. A1]

> Why would I want to post my prices on the web? If I did that, my competitors could see what I am charging and would undercut my prices by a few pennies, and I would lose most of my business. [Private communication with a middleman in the steel service center industry]

I. Introduction

We observe considerable differences in the microstructure of trade in various commodities and assets. A significant share of trade in commodities such as wheat and pork bellies and in financial assets such as common stocks and Treasury securities is intermediated by *market makers* (also known as specialists) at publicly posted bid and ask prices. Market makers typically either own or are members of an *exchange* such as the New York Stock Exchange or the Chicago Board of Trade. However, for other commodities such as steel, virtually all trade is conducted by *middlemen* (also known as dealer-brokers in financial markets or "steel service centers" in the steel market) at individually negotiated prices. In the steel market, there are no market makers or exchanges in which valid current bid and ask prices are publicly posted. Instead, transaction prices are private information, forcing traders in the steel market to engage in search and bargaining to find a good price.

It is commonly believed that market makers and exchanges are appropriate for trading highly standardized commodities and assets for which the volume is sufficiently large to produce "thick" and "active" markets. However, most types of steel coil and plate are at least as standardized as wheat or pork bellies, and the volume of trade in these particular steel products is at least as large. A number of potential market makers such as MetalSite, e-STEEL, and Enron have recently attempted to enter the steel market, but so far without success. Enron went bankrupt in December 2001, MetalSite's web site was closed in June 2001 (although it was reopened in November 2001), and e-STEEL changed its name to NewView Technologies in November 2001, reflecting a new focus away from steel. Thus, at present, no market makers handle a significant share of trade in steel. So the first puzzle is to explain why

market makers have been successful in entering and transforming trade in some markets (e.g., bonds) but not in others (e.g., steel).

The second quotation above suggests that even though new information technologies such as the Internet make it nearly costless to post and update bid and ask prices publicly, it might still be unprofitable for potential market makers to enter markets such as steel. If buyers of steel are accustomed to searching, then middlemen can easily respond to the entry of a market maker by slightly undercutting the market maker's prices, stealing most of his business. So the second puzzle is to explain how a market maker such as Cantor Fitzgerald can successfully enter a market such as the bond market if the existing middlemen can respond by undercutting the market maker's publicly posted bid and ask prices.

As a step toward answering these puzzles and understanding the differences in the microstructure of trade that we observe across various markets, we present a simple model in which the share of trade intermediated by middlemen and market makers is endogenously determined. Our model also provides insights into the likely effects of the huge reduction in search and transactions costs resulting from the information revolution and the advent of the World Wide Web. These technologies have facilitated the rapid emergence of market makers operating web-based "B2B exchanges" that intermediate business-to-business trade, threatening the existence of traditional middlemen in these markets.[1]

Our model is an extension of a model of intermediation and search due to Spulber (1996*a*). Spulber's model has three types of agents: buyers (consumers), sellers (producers), and price-setting middlemen. Since middlemen are often called *dealers,* we refer to Spulber's model as an analysis of the *dealer market.* Our main extension is to consider the effect of introducing a fourth type of agent, *market makers.* In Spulber's model, middlemen are assumed to be the exclusive avenue of exchange: every producer wishing to sell a commodity (or asset) and every consumer wishing to purchase it are required to transact via middlemen rather than trade directly with each other. Transactions in the dealer market occur over a range of individually negotiated prices—the outcome of a costly sequential search process.

We study the effect of introducing a monopolist market maker on the search equilibrium in the dealer market. The market maker can be

[1] The rate of growth of B2B markets is 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 percent of all sales in 2002 and a total volume of $1.4 trillion in transactions by 2004. These estimates may be conservative: alternative forecasts quoted in a recent symposium on B2B e-commerce published in the *Journal of Economic Literature* are substantially higher. See, e.g., Bakos (2001) and Lucking-Reiley and Spulber (2001) for even more optimistic forecasts of rapid growth in retail e-commerce.

conceptualized as operating an exchange on which publicly observable bid and ask prices are posted. Producers and consumers now have the option of trading on the exchange at the publicly posted bid and ask prices or searching for a better price in the dealer market. Since the prices quoted by middlemen are not publicly posted, producers and consumers must obtain them by direct contact (e.g., a telephone call), and the delays associated with searching constitute an implicit search cost.

Entry by a monopolist market maker is profitable even if the market maker has a higher per unit transactions cost than most middlemen, provided that it is lower than that of the least efficient middleman operating in the dealer market before the entry of the market maker. If so, the entry of the market maker drives the least efficient middlemen out of business and segments the market: the highest-valuation consumers and the lowest-cost producers (i.e., those with the largest implicit search costs) trade with the market maker, and the residual set of intermediate-valuation consumers and intermediate-cost producers search for better prices in the dealer market. All the surviving middlemen undercut the bid-ask spread charged by the monopolist market maker. Thus 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 consumer and producer surpluses, and higher trading volumes.

Middlemen and market makers represent competing institutions for the intermediation of trade. A market maker offers a superior exchange technology for the highest-valuation buyers and the lowest-cost sellers, and its entry raises welfare and reduces bid-ask spreads compared with the free-entry search equilibrium when all intermediation is done by middlemen. On the other hand, free entry of middlemen provides a "competitive fringe" that limits the market power of a monopolist market maker. Without the competitive threat of middlemen, a monopolist market maker would quote a wider bid-ask spread, and consumer and producer surplus would be significantly lower. However, in some situations the market maker can enter the market, drive all middlemen out of business, and set unconstrained monopoly bid and ask prices in the postentry equilibrium. Even this monopoly outcome results in a strict Pareto improvement relative to the free-entry equilibrium that existed in the dealer market before the entry of the market maker.

The relative share of trade intermediated by middlemen rather than the market maker depends on three parameters: the intertemporal discount rate δ and the per unit transactions costs of the market maker k_m and most efficient middleman $\underline{k}$. For a broad range of parameter values, the market maker coexists with the dealer market. However, if the market maker's per unit transactions cost k_m exceeds that of the *least* efficient

middleman $\bar{k}$ operating in the dealer market before the entry of a market maker, then entry by the market maker is not profitable and only middlemen will exist in equilibrium. Conversely, if the per unit transactions cost of the *most* efficient middleman $\underline{k}$ is sufficiently high relative to the transactions cost of the market maker k_m, then the entry of the market maker drives all middlemen out of business. The dealer market can survive the entry of a market maker even if the transactions cost $\underline{k}$ of the most efficient middleman exceeds the transactions cost k_m of the market maker—provided that it is not too much greater.

Our analysis is similar in some respects to those of Gehrig (1993) and Neeman and Vulkan (2001), although our conclusions are quite different.[2] Gehrig studies a model in which producers and consumers of a commodity have the option of trading at publicly posted bid and ask prices on an exchange run by a market maker (which Gehrig calls an "intermediary") or entering a "search market" in which consumers and producers are randomly matched and engage in bargaining in an attempt to negotiate a mutually acceptable price. Gehrig's model differs from ours primarily in the formulation of the search market: he models the bargaining process in the search market as a static (one-shot) random matching game in which consumers and producers negotiate directly with each other rather than transact through middlemen. In Gehrig's model the market maker always coexists with the search market and charges the same bid-ask spread and trades the same volume regardless of the level of search costs in the search market. Neeman and Vulkan obtain a very different result, namely, that the market maker can never coexist with the search market in equilibrium. In their model, agents have a choice between trading at posted prices with a market maker in a centralized market and engaging in direct negotiations with a randomly chosen producer or consumer in the search market. They prove a result similar to ours, namely, that the highest-valuation consumers and lowest-cost producers prefer to trade with the market maker rather than engage in direct negotiation. However, in their model the entry of a market maker causes a complete unraveling of direct negotiations, and in equilibrium all trade is conducted in the centralized market by the market maker.

In Section II we review a dynamic equilibrium model, introduced by Spulber (1996*a*), of trade with search among competing middlemen. In Section III we consider whether the dealer market equilibrium characterized in Section II can be upset by the entry of a monopolist market

[2] We discuss related papers by Pirrong (2000), Baye and Morgan (2001), Caillaud and Jullien (2001), and Hendershott and Zhang (2001) later in the paper. A number of other studies on the role of intermediaries deserve mention, although we do not explicitly discuss them. An incomplete list includes Garman (1976), Rubinstein and Wolinsky (1987), Yanelle (1989), Yavas (1992), and O'Hara (1995).

maker who runs a centralized exchange with publicly posted prices. In Section IV we consider the case in which there is free entry of market makers, resulting in Bertrand-style competition that forces the bid-ask spread down to the transactions cost of the most efficient market maker. Although our interest in alternative intermediation technologies arose from our observations of the microstructure of trade in the steel market, we believe that our theory provides insights into the microstructure of trade in a wide range of markets. In Section V we discuss three other markets besides steel: the market for goods in Radford's (1945) prisoner of war (POW) camp, the U.S. equity market, and the U.S. Treasury market. In Section VI we offer some concluding remarks and suggestions for further research.

II. Search Equilibrium with Middlemen but No Market Maker

Our point of departure is a simple exchange economy in which the only intermediaries are middlemen. We present a modified version of Spulber's (1996*a*) equilibrium search model with three types of agents: producers, consumers, and middlemen. In this model producers and consumers cannot trade directly with each other. Instead all trade must be intermediated by middlemen. To keep our presentation self-contained, we review Spulber's model in this section before presenting our extension of his model in which we analyze the effect of introducing a fourth type of agent—a market maker. Since middlemen of the type studied in this section are called dealers in a variety of financial and commodity markets, we refer to Spulber's work as an analysis of a competitive *dealer market*.

The dealer market consists of a continuum of heterogeneous producers, consumers, and middlemen. A producer of type v can produce at most one unit of the good at a cost of v. A consumer of type v can consume at most one unit of the good and is willing to pay at most v to consume it. Producers and consumers remain in the market for a random (geometrically distributed) length of time before permanently exiting. Let $\lambda \in (0, 1)$ be the probability that a producer or a consumer exits the market in period t. A consumer or producer may randomly exit before having a chance to consume or sell a unit of the good, respectively. However, if they succeed in trading prior to exiting, the unitary supply-demand assumption implies that these individuals will not make any subsequent transactions after their initial trade.

Suppose that whenever a producer or a consumer exits the market, he or she is replaced by a new producer or consumer who is randomly drawn from $U[0, 1]$, the uniform distribution on the $[0, 1]$ interval. Suppose that at time $t = 0$ the initial distribution of types v of producers and consumers is $U[0, 1]$. Then in all subsequent periods $t =$

$1, 2, 3, \ldots$, the distribution of types will also be $U[0, 1]$. Thus $U[0, 1]$ is the unique invariant distribution of this entry and exit process, and in each subsequent period $t = 1, 2, 3, \ldots$, a fraction λ of the population of producers and consumers exits the market and is replaced by an inflow of an equal fraction of new producers and consumers.

In a dealer market there is no central exchange or marketplace in which the commodity is traded. In particular, there is no advertising or central, publicly accessible site on which middlemen can post bid and ask prices. Instead, the only way for producers and consumers to obtain price quotes is to directly contact individual middlemen. Middlemen are infinitely lived and set a pair of stationary bid and ask prices to maximize their expected discounted profits. There are a continuum of middlemen indexed by k, the marginal costs of executing each trade between a producer and a consumer. Transactions costs k are distributed uniformly over the interval $[\underline{k}, 1]$. The lower bound $\underline{k}$ is the marginal transactions cost of the most efficient middleman. It may not be possible for all potential middlemen to enter the dealer market and make a profit. We shall let $\bar{k}$ denote the transactions cost of the least efficient middleman who participates in the dealer market in equilibrium. Thus profits earned by this marginal middleman $\bar{k}$ are zero, although the more efficient, inframarginal middlemen with $k \in [\underline{k}, \bar{k})$ can earn positive profits in equilibrium. A middleman of type $k \in [\underline{k}, \bar{k}]$ chooses a pair of stationary bid and ask prices $(b(k), a(k))$ that maximizes his expected discounted profits, where $a(k)$ denotes the ask price at which the middleman is willing to sell to consumers, and $b(k)$ denotes the bid price at which the middleman is willing to purchase from producers.

Producers and consumers engage in sequential search. Each period a searcher obtains a single price quote from one middleman, drawn randomly from $U[\underline{k}, \bar{k}]$. Although there is no explicit cost to obtain a price quote, there is an implicit "delay cost" involved in searching for prices. All producers and consumers discount the future using the factor $\rho(1 - \lambda)$. The first term in this composite discount factor, $\rho \in (0, 1)$, reflects the rate of time preference, and the second, $1 - \lambda$, is the "survival probability" that accounts for the possibility of random exit from the market prior to trading (in which case the exiting agent fails to receive any gains from trade).[3] The stationarity of the bid and ask prices charged by middlemen together with the heterogeneity in their transactions costs implies that the realized price quotes obtained by consum-

[3] It is possible to extend the model by including explicit per period search costs γ. However, it is not difficult to show that a stationary equilibrium in the dealer market cannot exist if $\gamma > 0$. Therefore, we restrict $\gamma = 0$ in the analysis that follows, accounting only for the implicit search costs resulting from exiting the market before having an opportunity to execute a transaction (the λ parameter) and the discounting of delayed gains from trade (the ρ parameter).

ers and producers are independently and identically distributed (i.i.d.) draws from the stationary distributions of bid and ask prices charged by middlemen. Let $F_\alpha(a)$ denote the distribution of ask prices facing consumers and $F_\beta(b)$ denote the distribution of bid prices facing producers.

Let $V_c(a, v)$ denote the present discounted value of an optimal search strategy for a consumer of type v who has received a quoted ask price of a from a randomly chosen middleman. The consumer has three choices: (a) do nothing (i.e., do not buy at the ask price a and do not search), (b) accept the middleman's ask price of a, or (c) reject the middleman's ask price of a and continue searching for a better price. These three options are reflected in the following Bellman equation for the consumer's problem:

$$V_c(a, v) = \max\left[0, v - a, \rho(1 - \lambda)\int_{\underline{a}}^{\bar{a}} V_c(a', v)F_\alpha(da')\right], \tag{1}$$

where $[\underline{a}, \bar{a}]$ is the support of the distribution of ask prices charged by middlemen. The value of 0 in the Bellman equation corresponds to the option of not searching, not trading, and not consuming. All consumers with sufficiently low valuations will choose this option. Clearly, any consumer with valuation $v < \underline{a}$ will never search or trade in the dealer market.

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

$$v = r_c(v) + \frac{1}{\delta}\int_{\underline{a}}^{r_c(v)} F_\alpha(a)da, \tag{2}$$

where

$$\delta = \delta(\rho, \lambda) = \frac{1}{\rho(1 - \lambda)} - 1 \tag{3}$$

is the composite exit-adjusted discount rate per period. It is not difficult to see from equation (2) that $r_c(v)$ is a strictly increasing function of v on the interval $(\underline{v}_c, 1)$, where $\underline{v}_c$ is the marginal consumer for whom the gain from entering the dealer market is zero. We have $\underline{v}_c = r_c(\underline{v}_c) \equiv \underline{r}_c = \underline{a}$.

Let $V_p(b, v)$ denote the present discounted value of an optimal search

strategy of a type v producer facing a bid price of b. The Bellman equation for the producer is given by

$$V_p(b, v) = \max\left[0, b - v, \rho(1 - \lambda)\int_{\underline{b}}^{\bar{b}} V_p(b', v)F_\beta(db')\right], \tag{4}$$

where $[\underline{b}, \bar{b}]$ is the support of the distribution of bid prices offered by middlemen. The optimal strategy for a type v producer also takes the form of a reservation price strategy, but in this case it is optimal to accept any bid price b that *exceeds* the reservation price $r_p(v)$, given by the unique solution to

$$v = r_p(v) - \frac{1}{\delta}\int_{r_p(v)}^{\bar{b}} [1 - F_\beta(b)]db. \tag{5}$$

In the producer case, $r_p(v)$ is monotonically increasing over the interval $(0, \bar{v}_p)$, where $\bar{v}_p = r_p(\bar{v}_p) \equiv \bar{r}_p = \bar{b}$ is the marginal producer for whom the expected gain from searching is zero.

Figure 1 graphs the reservation price functions $r_c(v)$ and $r_p(v)$ for an example in which $\delta = .2$ with a specific (equilibrium) pair of bid and ask distributions F_α and F_β, which will be derived shortly. Valuations for buyers are plotted from high to low, whereas sellers' costs are plotted from low to high, resulting in notional "supply" and "demand" curves. However, actual transactions in this market are determined by producers' and consumers' optimal search behavior. Consumers purchase the good from middlemen at prices in the interval $[\underline{a}, \bar{a}] = [.61, .71]$, and producers sell the good to middlemen at prices in the interval $[\underline{b}, \bar{b}] = [.28, .39]$. Note that the reservation price function for producers, $r_p(v)$, lies uniformly above the sellers' supply curve, provided that the producer participates in the dealer market. It is easy to see from equation (5) that when the seller's cost v exceeds the upper bound $\bar{b}$ of the distribution F_β of bid prices offered by middlemen, there is no point in searching. Thus all producers with costs $v > \bar{b} = \bar{v}_p = .39$ remain out of the market. Symmetrically, the reservation price function for buyers lies uniformly below their valuations, intersecting it at the lower support point $\underline{a} = \underline{v}_c = .61$ of the distribution of ask prices F_α. Therefore, the set of active consumers are those for whom $v \in (\underline{v}_c, 1] = (\underline{a}, 1] = (.61, 1]$.

The difference between a consumer's valuation v and his or her reservation value $r_c(v)$ is the net value of search, that is, the expected discounted surplus or "gains from trade." For reference, we plot a horizontal dashed line of height equal to .5 in figure 1. 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)

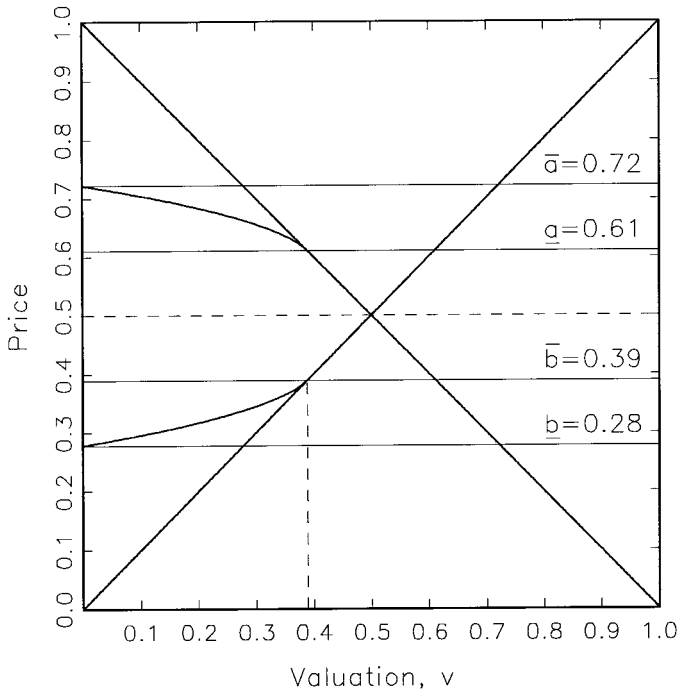

FIG. 1.—Reservation prices for buyers (consumers) and sellers (producers). Along the abscissa, buyers' valuations are plotted from high to low, whereas sellers' valuations (costs) are plotted from low to high.

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 surplus for a consumer with valuation v is given by $\max [0, v - .5]$, and the surplus for a producer with cost v is given by $\max [0, .5 - v]$. The area of the triangular regions, that is, the surplus achieved by producers and consumers in Walrasian equilibrium, is $\frac{1}{8}$ and $\frac{1}{8}$, respectively, resulting in a total surplus of $\frac{1}{4}$. The area between the horizontal dashed line and the reservation price curves represents the inefficiency of the search equilibrium outcome, that is, the lost gains from trade to producers and consumers.

Intermediaries maximize expected discounted profits subject to the constraint that supply and demand for the commodity are equal in every period. This constraint is necessitated by the assumption that middlemen do not carry inventories across successive periods. Suppose that a middleman sets an ask price of a. Let $D_t(a)$ denote the mass of consumers

who were among the initial population at $t = 0$ who purchase the good in period i:

$$D_i(a) = \frac{1 - \underline{v}_c}{N} \int_a^{\bar{r}_c} (1 - \lambda)^i [1 - F_\alpha(r)]^i h(r) dr, \tag{6}$$

where N is the number (total mass) of middlemen, $\bar{r}_c \equiv r_c(1)$ is the reservation price of the highest-valuation buyer, and $h(r)$ is the (conditional) distribution of reservation prices among the fraction $1 - \underline{v}_c$ of the initial population of consumers who chose to participate in the dealer market. Thus $h(r)(1 - \underline{v}_c)/N$ is the per firm density of consumers. So $D_i(a)$ equals the integral of the product of the probability $(1 - \lambda)^i$ of not exiting the market in periods $t = 0, 1, \ldots, i - 1$ times the probability $[1 - F_\alpha(r)]^i$ of not trading in periods $0, \ldots, i - 1$, times the per firm density of consumers $h(r)(1 - \underline{v}_c)/N$, integrated over the region of reservation values $[a, \bar{r}_c]$ corresponding to buyers who are willing to purchase at price a. By a change of variables, the density $h(r)$ can be derived from the distribution $r_c(v)$ and the fact that the valuations of those consumers who participate in the dealer market are uniformly distributed on $[\underline{v}_c, 1]$:

$$h(r) = \left(\frac{1}{1 - \underline{v}_c}\right) \left\{\frac{dr_c}{dv}[r_c^{-1}(r)]\right\}^{-1} = \frac{1 + [F_\alpha(r)/\delta]}{1 - \underline{v}_c}. \tag{7}$$

Let D_i be the share of the initial population that purchases from *any* middleman at time i:

$$D_i = \int_a^{\bar{a}} D_i(a) F_\alpha(da) = \int_{\underline{v}_c}^1 (1 - \lambda)^i [1 - F_\alpha(v)]^i F_\alpha(v) dv. \tag{8}$$

It is not hard to verify that, when $\lambda > 0$, the share of consumers who ultimately purchase the good is less than the share of consumers who choose to search for it in the dealer market:

$$\sum_{i=0}^{\infty} D_i = \int_{\underline{v}_c}^1 \frac{F_\alpha(v)}{1 - (1 - \lambda)[1 - F_\alpha(v)]} dv < 1 - \underline{v}_c. \tag{9}$$

This occurs since some of the consumers who attempted to search for the good ended up exiting the dealer market before they were able to find a sufficiently attractive price. This is part of the deadweight loss involved in the operation of the dealer market. Another component of the deadweight loss due to sequential search and the implied delay in trading and consuming is the discounting of the gains from trade for those transactions that are ultimately realized.

Total expected discounted demand is the expected discounted value of the stream of demands in all future periods by the initial population

at time $t = 0$ as well as the stream of demands from each succeeding generation of new producers and consumers entering the dealer market. The subsequent entrants are drawn from the same $U[0, 1]$ distribution of types as the initial population of producers and consumers at $t = 0$, but their mass is scaled by the factor λ, the fraction of the population entering and exiting each period. Assuming that all middlemen discount future flows using the discount factor ρ, we have

$$\begin{aligned} D(a) &= \sum_{i=0}^{\infty} \rho^i D_i(a) + \lambda \sum_{j=1}^{\infty} \rho^j \sum_{i=0}^{\infty} \rho^i D_i(a) \\ &= \frac{1 - \underline{v}_c}{N} \int_a^{\bar{r}_c} \left\{ \sum_{i=0}^{\infty} \rho^i (1 - \lambda)^i [1 - F_\alpha(r)]^i \right. \\ &\quad \left. + \lambda \sum_{j=1}^{\infty} \rho^j \sum_{i=0}^{\infty} \rho^i (1 - \lambda)^i [1 - F_\alpha(r)]^i \right\} h(r)\, dr \\ &= \frac{1}{N} \int_a^{\bar{r}_c} \left\{ \frac{1 + [\rho\lambda/(1 - \rho)]}{1 - \rho(1 - \lambda)[1 - F_\alpha(r)]} \right\} \left[\frac{\delta(\rho, \lambda) + F_\alpha(r)}{\delta(\rho, \lambda)} \right] dr \\ &= \frac{1}{N} \int_a^{\bar{r}_c} \left\{ \frac{1 + [\rho\lambda/(1 - \rho)]}{1 - \rho(1 - \lambda) + \rho(1 - \lambda) F_\alpha(r)} \right\} \\ &\quad \times \left[\frac{1 - \rho(1 - \lambda) + \rho(1 - \lambda) F_\alpha(r)}{1 - \rho(1 - \lambda)} \right] dr \\ &= \frac{\bar{r}_c - a}{N} \left\{ \frac{1 + [\rho\lambda/(1 - \rho)]}{1 - \rho(1 - \lambda)} \right\}. \end{aligned} \tag{10}$$

By similar reasoning the middleman's expected discounted supply function is

$$S(b) = \frac{b - \underline{r}_p}{N} \left\{ \frac{1 + [\rho\lambda/(1 - \rho)]}{1 - \rho(1 - \lambda)} \right\}, \tag{11}$$

where $\underline{r}_p \equiv r_p(0)$ is the reservation value of the lowest-cost producer. Given the discounted supply and demand functions, the middleman's present discounted value of future trading profits is given by

$$\begin{aligned} \Pi(a, b, k) &= aD(a) - (b + k)S(b) \\ &= \left\{ \frac{1 + [\rho\lambda/(1 - \rho)]}{1 - \rho(1 - \lambda)} \right\} \left[\frac{a(\bar{r}_c - a) - (b + k)(b - \underline{r}_p)}{N} \right]. \end{aligned} \tag{12}$$

The middleman's problem is to

$$\max_{a,b} \Pi(a, b, k) \quad \text{subject to } D(a) \leq S(b). \tag{13}$$

The optimal bid and ask prices are

$$a(k) = \frac{3\bar{r}_c + \underline{r}_p + k}{4}, \quad b(k) = \frac{\bar{r}_c + 3\underline{r}_p - k}{4}. \tag{14}$$

These bid and ask prices also equate supply and demand in every period: $D_i(a(k)) = S_i(b(k))$, $i = 0, 1, 2, \ldots$.

The linearity of $a(k)$ and $b(k)$ in k implies that the distributions of the bid and ask prices F_α and F_β are uniform. To find the support of these distributions, we need to compute $\bar{k}$, the transactions cost of the marginal middleman entering the dealer market in equilibrium. Plugging the solutions in equation (14) back into the profit function (12), we obtain

$$\Pi(a(k), b(k), k) = \left\{\frac{1 + [\rho\lambda/(1-\rho)]}{1 - \rho(1-\lambda)}\right\}\left[\frac{(\bar{r}_c - \underline{r}_p - k)^2}{8N}\right]. \tag{15}$$

Solving for $\Pi(a(k), b(k), k) = 0$, we see that the marginal middleman has a transactions cost of $\bar{k} = \bar{r}_c - \underline{r}_p$, and the number of middlemen operating in equilibrium is $N = \bar{k} - \underline{k} = \bar{r}_c - \underline{r}_p - \underline{k}$. Letting k equal $\underline{k}$ and $\bar{k}$, we obtain supports of the equilibrium distributions of bid and ask prices. The upper and lower support points of the distribution of bid and ask prices are functions of the highest and lowest reservation values of buyers and sellers. This characterization will play a key role in Section III when we analyze how the dealer market is affected by the potential entry of a market maker.

LEMMA 1. If $\underline{k} < 1$, there is a unique stationary equilibrium in the dealer market. Bid prices are uniformly distributed on the interval $[\underline{b}, \bar{b}]$, and ask prices are uniformly distributed on the interval $[\underline{a}, \bar{a}]$, where

$$\bar{a} = \bar{r}_c, \quad \underline{a} = \underline{r}_c = .25(\underline{r}_p + \underline{k}) + .75\bar{r}_c,$$

$$\underline{b} = \underline{r}_p, \quad \bar{b} = \bar{r}_p = .75\underline{r}_p + .25(\bar{r}_c - \underline{k}). \tag{16}$$

To complete the characterization of the dealer market equilibrium, we need to derive expressions for the reservation values of the highest-valuation consumer $\bar{r}_c$ and the lowest-cost seller $\underline{r}_p$ in terms of the underlying preference and technology parameters $(\rho, \lambda, \underline{k})$. Substituting the uniform distribution F_α of ask prices and the expression for its lower

support point $\underline{a}$ into equation (2) for the consumer with the highest reservation value $\bar{r}_c = r_c(1)$, we obtain

$$1 = \bar{r}_c + \frac{1}{\delta}\int_{\underline{a}}^{\bar{r}_c} F_\alpha(a)da$$

$$= \frac{(1 + 8\delta)\bar{r}_c - \underline{r}_p - \underline{k}}{8\delta}. \tag{17}$$

Similarly, substituting F_β and the formula for its upper support point $\bar{b}$ into the reservation price equation (5) for the producer with the lowest reservation value $\underline{r}_p = r_p(0)$, we obtain

$$0 = \underline{r}_p - \frac{1}{\delta}\int_{\underline{r}_p}^{\bar{b}} [1 - F_\beta(b)]db$$

$$= \frac{(1 + 8\delta)\underline{r}_p - \bar{r}_c + \underline{k}}{8\delta}. \tag{18}$$

Solving these equations for $\bar{r}_c$ and $\underline{r}_p$, we obtain $1 = \bar{r}_c + \underline{r}_p$, which implies that

$$\bar{r}_c = \frac{1 + \underline{k} + 8\delta}{2 + 8\delta}, \quad \underline{r}_p = \frac{1 - \underline{k}}{2 + 8\delta}. \tag{19}$$

The fraction of consumers and producers who participate in the dealer market is given by

$$q^* = 1 - \underline{v}_c = \bar{r}_p = \frac{(1 - \underline{k})(1 + 2\delta)}{2 + 8\delta}. \tag{20}$$

The distribution of ask prices F_α is uniform on the interval

$$\left[\frac{1 + 6\delta + \underline{k}(1 + 2\delta)}{2 + 8\delta}, \frac{1 + \underline{k} + 8\delta}{2 + 8\delta}\right]. \tag{21}$$

The distribution of bid prices F_β is uniform on the interval

$$\left[\frac{1 - \underline{k}}{2 + 8\delta}, \frac{(1 - \underline{k})(1 + 2\delta)}{2 + 8\delta}\right]. \tag{22}$$

We can see from these equations that as $\delta \downarrow 0$ and $\underline{k} \downarrow 0$, the distributions of bid and ask prices converge to a degenerate distribution with all mass on the single price $p = .5$, the Walrasian equilibrium value. We can also see from equation (20) that the equilibrium quantity traded also converges to $q^* = .5$, the Walrasian equilibrium quantity. Thus the dealer market equilibrium contains the Walrasian equilibrium as a limiting special case as search and transactions costs tend to zero.

The mass of middlemen who are active in equilibrium is given by

$$N = \bar{k} - \underline{k} = \frac{4\delta(1 - \underline{k})}{1 + 4\delta}. \tag{23}$$

Note that the dealer market collapses to a no-trade equilibrium as $\underline{k} \to 1$. We also have $N \downarrow 0$ in the limit as $\delta \downarrow 0$. In this case the dealer market is not collapsing, but rather converging to a degenerate distribution in which the most efficient middleman handles all trade, setting an ask price of $\bar{a} = \underline{a} = (1 + \underline{k})/2$ and a bid price of $\bar{b} = \underline{b} = (1 - \underline{k})/2$. This is a zero-profit equilibrium, since the bid-ask spread of $\underline{k}$ just offsets the transactions cost of the most efficient middleman. As $\underline{k} \downarrow 0$, the competitive positive transactions cost equilibrium converges to the frictionless Walrasian equilibrium outcome.

We conclude by providing formulas for profits and surplus in this economy. These formulas will be important in the next section, since they enable us to determine the relative efficiency of the dealer market equilibrium compared with an equilibrium with possible entry by a market maker. Total consumer and producer surplus, denoted by S_c and S_p, respectively, are given by

$$\begin{aligned} S_c &= \left(1 + \frac{\lambda\rho}{1 - \rho}\right) \int_0^1 [v - r_c(v)]dv, \\ S_p &= \left(1 + \frac{\lambda\rho}{1 - \rho}\right) \int_0^1 [r_p(v) - v]dv, \end{aligned} \tag{24}$$

where we set $r_c(v) = v$ for the consumers who do not trade and $r_p(v) = v$ for producers who do not trade, so these individuals do not contribute to total surplus. The factor $1 + [\lambda\rho/(1 - \rho)]$ adjusts for the discounted surplus of all current and future consumers and producers who enter the market.

The total discounted profits of all middlemen participating in the dealer market are given by

$$\Pi_d \equiv \int_{\underline{k}}^{\bar{k}} \Pi(a(k), b(k), k)dk = \left(1 + \frac{\rho\lambda}{1 - \rho}\right)\left[\frac{(1 + \delta)N^2}{24\delta}\right]. \tag{25}$$

We see from this formula that total profits of middlemen tend to zero as $\delta \downarrow 0$ (because of convergence to a zero-profit competitive equilibrium) or as $\underline{k} \uparrow 1$ (because of convergence to a no-trade equilibrium).

The total gain from trade from the operation of the dealer market, W^d, is the sum of producer surplus, consumer surplus, and the total discounted profits of middlemen:

$$W^d = S_c + S_p + \Pi_d. \tag{26}$$

In the limiting Walrasian equilibrium, we have $S_c = S_p = .125$ and $\Pi_d = 0$, so the total surplus equals $W^d = .25$, the area under the notional supply and demand curves to the left of the intersection at $q = .5$. When $\delta = .2$, $\lambda = 0$, and $\underline{k} = 0$, we compute $S_c = S_p = .0489$ and $\Pi_d = .0494$, so that the total surplus equals $W^d = .1471$, which is only 59 percent of the maximum possible surplus of .25 that would be achieved in a frictionless Walrasian equilibrium.

III. Search Equilibrium with a Monopoly Market Maker

In this section we extend Spulber's (1996*a*) model of search equilibrium with producers, consumers, and middlemen by introducing a fourth type of agent: a monopolist market maker. Initially we assume that there is at most one potential entrant who could assume the role of a monopolist market maker, quoting publicly observable bid and ask prices (a_m, b_m). If this market maker enters, producers and consumers can choose between trading with the market maker at the publicly posted bid and ask prices (a_m, b_m) and searching for a better bid or ask price in the dealer market.

We begin our analysis by characterizing how the presence of a market maker affects the solution to consumers' and producers' optimal search problems. We then derive conditions under which entry by a market maker is profitable and compute the market maker's optimal pricing strategy. In doing so, we compute the dealer's equilibrium response to the market maker's entry and pricing rules. We then compare equilibrium outcomes for economies with and without a market maker. Finally, we summarize the limiting properties of the model.

A. *The Consumers' and Producers' Decision Rules*

Consider a consumer who has not yet chosen to search. The consumer has three options: (*a*) do nothing, (*b*) purchase a unit of the commodity in the exchange at price a_m, or (*c*) search for a better price in the dealer market. The consumer's value function is given by

$$V_c(a_m, v) = \max\left[0, v - a_m, \rho(1 - \lambda)\int_{\underline{a}}^{\bar{a}} V_c(a', a_m, v)F_\alpha(da')\right], \qquad (27)$$

where $V_c(a, a_m, v)$ denotes the value function for a consumer who has chosen to search and has received an ask price of a from a middleman, modeled as a random draw from F_α. Once the consumer has an offer

in hand, he has the fourth option of accepting the offer from the middleman. The Bellman equation for $V_c(a, a_m, v)$ is given by

$$V_c(a,a_m,v) = \max\left[0,v - a,v - a_m,\rho(1 - \lambda)\int_{\underline{a}}^{\bar{a}} V_c(a',a_m,v)F_\alpha(da')\right]. \quad (28)$$

THEOREM 1. Suppose that an equilibrium exists in which the market maker coexists with middlemen in the dealer market. Let $\underline{a}$ be the lowest ask price in the dealer market after the entry of the market maker. Let $v_c(\underline{a}, a_m)$ be the value of the marginal consumer (with reservation value a_m), who is indifferent between trading with the market maker and trading in the dealer market:

$$v_c(\underline{a}, a_m) = a_m + \frac{1}{\delta}\int_{\underline{a}}^{a_m} F_\alpha(a)da. \quad (29)$$

If $v_c(\underline{a}, a_m) < 1$, then there are three different optimal search-purchase strategies depending on the consumer's type. If $v \in [0, \underline{a})$, then it is not optimal for the consumer to trade with the market maker or to search for a middleman in the dealer market. If $v \in [\underline{a}, v_c(\underline{a}, a_m))$, then it is optimal for the consumer to trade in the dealer market. If $v \in (v_c(\underline{a}, a_m), 1]$, then it is optimal for the consumer to bypass the dealer market and to immediately purchase the good from the market maker at the ask price a_m.

We also have a symmetric result for producers.

THEOREM 2. Suppose that an equilibrium exists in which the market maker coexists with middlemen in the dealer market. Let $v_p(b_m, \bar{b})$ be the value of the marginal producer (with reservation value b_m), who is indifferent between trading with the market maker and trading in the dealer market:

$$v_p(b_m, \bar{b}) = b_m - \frac{1}{\delta}\int_{b_m}^{\bar{b}} [1 - F_\beta(b)]db. \quad (30)$$

If $v_p(b_m, \bar{b}) > 0$, then there are three different optimal search-sell strategies depending on the producer's type. If $v \in (\bar{b}, 1]$, then it is not optimal for the producer to trade with the market maker or to search for a middleman in the dealer market. If $v \in (v_p(b_m, \bar{b}), \bar{b}]$, then it is optimal for the producer to trade in the dealer market. If $v \in [0, v_p(b_m, \bar{b})]$, then it is optimal for the producer to bypass the dealer market and to immediately sell the good to the market maker at the bid price b_m.

Figure 2 illustrates theorems 1 and 2 for the case in which $\delta = 0.2$. The curved line in figure 2*a* is the net value of search for consumers, $v - r_c(v)$, and in figure 2*b* it is the value of search for producers,

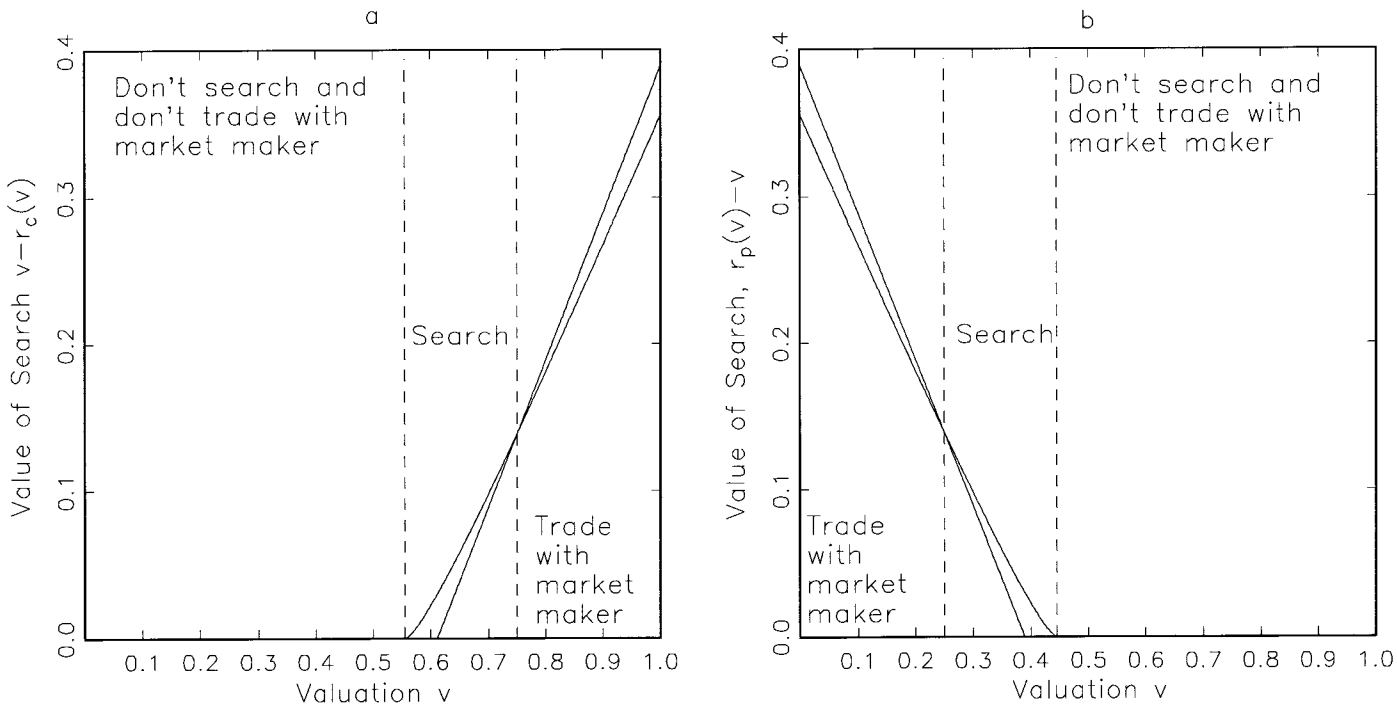

FIG. 2.—Decision rules for producers and consumers. *a*, Value of search for consumers. *b*, Value of search for producers.

$r_p(v) - v$. The straight lines plotted in each panel of figure 2 are net values of trading with the market maker at the publicly posted bid and ask prices (a_m, b_m); that is, they are graphs of the functions $b_m - v$ and $v - a_m$. In this example, $b_m - v > r_p(v) - v$ for $v \in [0, v_p(b_m, \bar{b})]$, where $v_p(b_m, \bar{b}) = .25$. It follows that all producers in the interval [0, .25] prefer to trade with the market maker rather than search for a better price in the dealer market. However, for producers in the interval $(.25, \bar{b}]$, we see that $r_p(v) - v > b_m - v$. Thus these producers choose to search for a better price in the dealer market rather than trade with the market maker at price b_m. The remaining producers with production costs $v \in (\bar{b}, 1]$ would not gain from trading with the market maker or with any middleman in the dealer market, and so these producers do not participate and earn a net surplus of zero. Symmetrical results hold for consumers.

B. The Intermediaries' Entry and Pricing Decisions

In the previous subsection, consumers and producers took prices as given. In this subsection we analyze the market maker's entry and pricing decision. We then derive the response of the middlemen to the market maker's entry. In particular, there are three possible regimes in equilibrium: an unconstrained monopoly regime, a limit-pricing regime, and a competitive regime. In the first two regimes, the market maker drives all the middlemen out of business. In the competitive regime, a set of middlemen coexists with the market maker. In the limit-pricing and competitive regimes, the existence (or potential entry) of middlemen in the market limits the market power of the market maker.

All consumers know that they have the option of purchasing the commodity from the market maker at price a_m, and all producers know that they can sell the commodity to the market maker at price b_m. Clearly, no middleman in the dealer market would be able to sell at an ask price a higher than a_m. Similarly, no middleman would be able to purchase the commodity for less than the bid price charged by the market maker. Thus we have the following theorem.

THEOREM 3. b_m is the lowest reservation price of producers participating in the dealer market and a lower bound on the support of the distribution of bid prices offered by middlemen. a_m is the highest reservation price of consumers participating in the dealer market and an upper bound on the support of the distribution of ask prices charged by middlemen in the dealer market:

$$\underline{b} = \underline{r}_p = b_m, \quad \bar{a} = \bar{r}_c = a_m. \tag{31}$$

Applying the results from Section II, we can easily show that in the

presence of a market maker, F_α and F_β will still be uniform distributions, but with supports given by $[\underline{a}, \bar{a}] = [.75a_m + .25(b_m + \underline{k}), a_m]$ and $[\underline{b}, \bar{b}] = [b_m, .75b_m + .25(a_m - \underline{k})]$. Thus middlemen in the dealer market uniformly undercut the posted bid-ask spread set by the market maker. Even though it is common knowledge that a randomly drawn bid from the dealer market will be less than b_m with probability one, this bid will materialize only after a one-period delay. The highest-valuation buyers and lowest-cost producers prefer to trade immediately at the less favorable prices offered by the market maker rather than incur the search/delay costs involved in trying to find a better price in the dealer market.

If the monopolist market maker sets an ask price of a_m and offers a bid price b_m, theorem 1 implies that the quantity of the commodity that will be demanded by the highest-valuation consumers in the interval $(v_c(\underline{a}, a_m)), 1]$ is

$$Q^d_m(a_m, b_m) = 1 - a_m - \frac{1}{\delta}\int_{\underline{a}}^{a_m} F_\alpha(a)da. \tag{32}$$

Similarly, the quantity supplied by the lowest-cost producers in the interval $[0, v_p(b_m, \bar{b}))$ is

$$Q^s_m(a_m, b_m) = b_m - \frac{1}{\delta}\int_{b_m}^{\bar{b}} [1 - F_\beta(b)]db. \tag{33}$$

THEOREM 4. $Q^d_m(a_m, b_m) = Q^s_m(a_m, b_m)$ for all $\delta > 0$ if and only if $b_m = 1 - a_m$.

Proof. If $a_m = 1 - b_m$, then corresponding symmetry relations hold in the dealer market: $\underline{a} = 1 - \bar{b}$ and $F_\alpha(v) = 1 - F_\beta(v)$. It is easy to see from formulas (32) and (33) that this implies that $Q^d_m = Q^s_m$. Conversely, if $Q^d_m = Q^s_m$ for all $\delta > 0$, then it must also hold in the limit as $\delta \to \infty$. But in the limit $Q^d_m = 1 - a_m$ and $Q^s_m = b_m$. Q.E.D.

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

COROLLARY. If $a_m > \max [.5(1 + k_m), .5(1 + \underline{k})]$, then the mass of middlemen who are active in the dealer market is given by

$$N = \max (0, a_m - b_m - \underline{k}) = 2a_m - 1 - \underline{k}. \tag{34}$$

If $\underline{k} = 0$, 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 $a_m = \underline{a}$, which is the lowest ask price charged by the most efficient middleman (with cost $k = \underline{k} = 0$) in the search equilibrium without a

market maker. The least efficient intermediaries are driven out of the market, and the remaining, more efficient intermediaries are forced to reduce their bid-ask spreads below their pre-entry values in order to avoid taking a loss. Therefore, both the upper and lower supports of the distribution of ask prices in the dealer market fall as a result of the entry of the market maker. Symmetrically, the entry of the market maker increases the lower and upper support points of the distribution of bid prices in the dealer market. This suggests that the entry of the market maker should result in a strict increase in the expected gains from trade for all producers and consumers who participate in the market—regardless of whether they choose to trade with the market maker or search for a middleman in the dealer market.

On one hand, the fact that middlemen uniformly undercut the market maker's posted bid and ask prices could make it unprofitable for the market maker to enter. On the other hand, if the market maker's per unit transactions cost k_m is sufficiently lower than $\underline{k}$, the transactions cost of the most efficient middleman, the entry of the market maker might succeed in driving the entire dealer market out of existence. To determine what will actually happen, we need to derive the market maker's profit function and optimal pricing strategy. By theorem 4 we know that $b_m = 1 - a_m$ implies $Q_m^d = Q_m^s$, and so we can write the market maker's problem as

$$\max_{a_m, b_m} \left(1 + \frac{\lambda\rho}{1-\rho}\right)(a_m - b_m - k_m)Q_m^s(b_m, a_m) \tag{35}$$

subject to

$$b_m = 1 - a_m, \quad \frac{1 + k_m}{2} \leq a_m \leq \frac{1 + 8\delta + \underline{k}}{2 + 8\delta}. \tag{36}$$

When we substitute the constraint $b_m = 1 - a_m$, the market maker's problem reduces to maximizing the following quadratic objective function with respect to the single control variable a_m:

$$\Pi_m(a_m) = \left(1 + \frac{\rho\lambda}{1-\rho}\right)(2a_m - 1 - k_m) \times \left[1 - a_m - \frac{\max(0, 2a_m - 1 - \underline{k})}{8\delta}\right]. \tag{37}$$

The market maker's objective is concave in a_m and thus has a unique optimal solution provided that the market maker's choice of a_m does not violate the search equilibrium conditions in the intermediary market. The solution to the market maker's problem and the resulting configuration of the dealer market are given below.

THEOREM 5. It is profitable for a market maker to enter the dealer market if and only if the market maker's transaction cost k_m satisfies $k_m < \bar{k}_m$, where

$$\bar{k}_m = \frac{\underline{k} + 4\delta}{1 + 4\delta}. \tag{38}$$

If k_m is below the threshold for profitable entry $\bar{k}_m$, there are three possible configurations for the postentry equilibrium depending on the model parameters $(\underline{k}, k_m, \delta(\rho, \lambda))$. There exist thresholds $0 < \underline{k}_l < \underline{k}_u$ such that, if $\underline{k} \in [0, \underline{k}_l)$, the market maker coexists with the most efficient middlemen, but the least efficient middlemen are driven out of business. If $\underline{k} \in [\underline{k}_l, \underline{k}_u)$, the market maker's entry drives all middlemen out of business, and the market maker sets "limit prices" that deter the most efficient middleman from entering. If $\underline{k} \in [\underline{k}_u, 1]$, the market maker sets unconstrained monopoly bid and ask prices.

The fundamental condition for the viability of entry by a monopolist market maker given in equation (38) of theorem 5 has a very simple interpretation. The expression on the right-hand side of equation (38) equals the efficiency level $\underline{k}$ of the marginal middleman 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 middleman in the pre-entry dealer market equilibrium if entry is to be feasible. However, the market maker need not have lower transactions costs than all middlemen in order for entry to be feasible. Even if the market maker's transactions cost is uniformly lower than the transactions costs of all middlemen, the dealer market will not necessarily be driven out of existence.

COROLLARY 5.1. The thresholds $\underline{k}_l$ and $\underline{k}_u$ for the limit-pricing and unconstrained monopoly regimes satisfy $0 < \underline{k}_l < \underline{k}_u$ and are given by

$$\underline{k}_l = \frac{(1 + k_m)(4\delta + 1) - 1}{8\delta + 1},$$

$$\underline{k}_u = \frac{1 + k_m}{2}. \tag{39}$$

We now complete our characterization of the postentry equilibrium by summarizing the quantities traded by the market maker, his bid-ask spreads, the number of middlemen, and the total quantity traded in the dealer market.

COROLLARY 5.2. In the unconstrained monopoly regime, the market maker sets bid and ask prices given by

$$a_m = \frac{3 + k_m}{4}, \quad b_m = \frac{1 - k_m}{4}. \tag{40}$$

The market maker trades the monopoly volume $Q^s_m(a_m, b_m) = b_m$. The dealer market is driven out of existence.

COROLLARY 5.3. In the limit-pricing regime the market maker sets bid and ask prices given by

$$a_m = \frac{1 + \underline{k}}{2}, \quad b_m = \frac{1 - \underline{k}}{2}. \tag{41}$$

The market maker trades the volume $Q^s_m(a_m, b_m) = b_m$. The dealer market is driven out of existence.

COROLLARY 5.4. In the competitive regime the market maker coexists with the dealer market. The market maker's bid and ask prices are given by

$$\begin{aligned} a_m &= \frac{3 + k_m}{4} + \frac{\underline{k} - 1}{16\delta + 4}, \\ b_m &= \frac{1 - k_m}{4} + \frac{1 - \underline{k}}{16\delta + 4}. \end{aligned} \tag{42}$$

The quantity traded by the market maker is given by

$$\begin{aligned} Q^d_m(a_m, 1 - a_m) &= Q^s_m(a_m, 1 - a_m) \\ &= \frac{1 - k_m}{4} + \frac{1 - \underline{k}}{16\delta + 4} - \frac{1}{8\delta}\left(\frac{1 + k_m}{2} + \frac{\underline{k} - 1}{8\delta + 2} - \underline{k}\right). \end{aligned} \tag{43}$$

COROLLARY 5.5. In the competitive regime the postentry equilibrium mass of middlemen operating in the dealer market is given by

$$N = \bar{k} - \underline{k} = 2a_m - 1 - \underline{k} = \frac{1 + k_m}{2} + \frac{\underline{k} - 1}{8\delta + 2} - \underline{k}. \tag{44}$$

COROLLARY 5.6. In the competitive regime the equilibrium distribution of ask prices quoted by middlemen in the dealer market is uniformly distributed on the interval $[\underline{a}, \bar{a}]$, where

$$\begin{aligned} \underline{a} &= \frac{3 + k_m}{8} + \frac{(8\delta + 1) + \underline{k}(8\delta + 3)}{32\delta + 8}, \\ \bar{a} &= \frac{3 + k_m}{4} + \frac{\underline{k} - 1}{16\delta + 4}. \end{aligned} \tag{45}$$

The equilibrium distribution of bid prices is uniformly distributed on the interval $[\underline{b}, \bar{b}]$, where

$$\underline{b} = \frac{1 - k_m}{4} + \frac{1 - \underline{k}}{16\delta + 4},$$

$$\bar{b} = \frac{5 - k_m}{8} - \frac{(8\delta + 1) + \underline{k}(8\delta + 3)}{32\delta + 8}. \tag{46}$$

COROLLARY 5.7. In the competitive regime, the equilibrium mass of producers and consumers who participate in the dealer market is given by

$$Q_d^d = Q_d^s = \bar{b} - \underline{b} + \frac{N}{8\delta}. \tag{47}$$

Figure 3 illustrates the market maker's pricing and quantity decisions in the three regimes in the case in which $k_m = 0$, $\rho = .95$, and $\lambda = .12$. These values imply a discount rate of $\delta(\rho, \lambda) = .2$. In this case the cutoff between the competitive and limit-pricing regimes is $\underline{k}_l = .31$, and the cutoff between the limit-pricing and monopoly regimes is $\underline{k}_u = .5$. Note that the market maker's ask price and quantity traded are increasing functions of $\underline{k}$ in the competitive regime. As $\underline{k}$ increases, the dealer market gets increasingly inefficient relative to transacting with the market maker, and the market maker exploits this by raising the ask price (and bid-ask spread). The volume of trade in the dealer market decreases monotonically in $\underline{k}$ until at $\underline{k} = .31$ the dealer market vanishes. For $\underline{k} \in [.31, .50)$, the market maker adopts a limit-pricing strategy, choosing the largest possible ask price that will not induce entry by middlemen. As $\underline{k}$ increases in this region, the market maker is able to raise prices, leading to a reduction in quantity traded. When $\underline{k} > .5$, the most efficient middleman is so inefficient relative to the market maker that the market maker 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. Thus the market maker's profits increase monotonically for $\underline{k} \in [0, (1 + k_m)/2]$ but are constant in the unconstrained monopoly regime when $\underline{k} > (1 + k_m)/2$.

C. *Equilibria with and without a Monopolist Market Maker*

In this subsection, we compare the equilibrium with a monopolist market maker described in the previous subsection to dealer-market equilibrium described in Section III. Figure 4 compares the equilibria with and without a monopolist market maker in the case in which $k_m = \underline{k} = 0$ and $\delta = .2$. The market maker trades a total quantity of $Q^m = .25$, and the dealer maker trades a total quantity of $Q^d = .19 = .44 -$

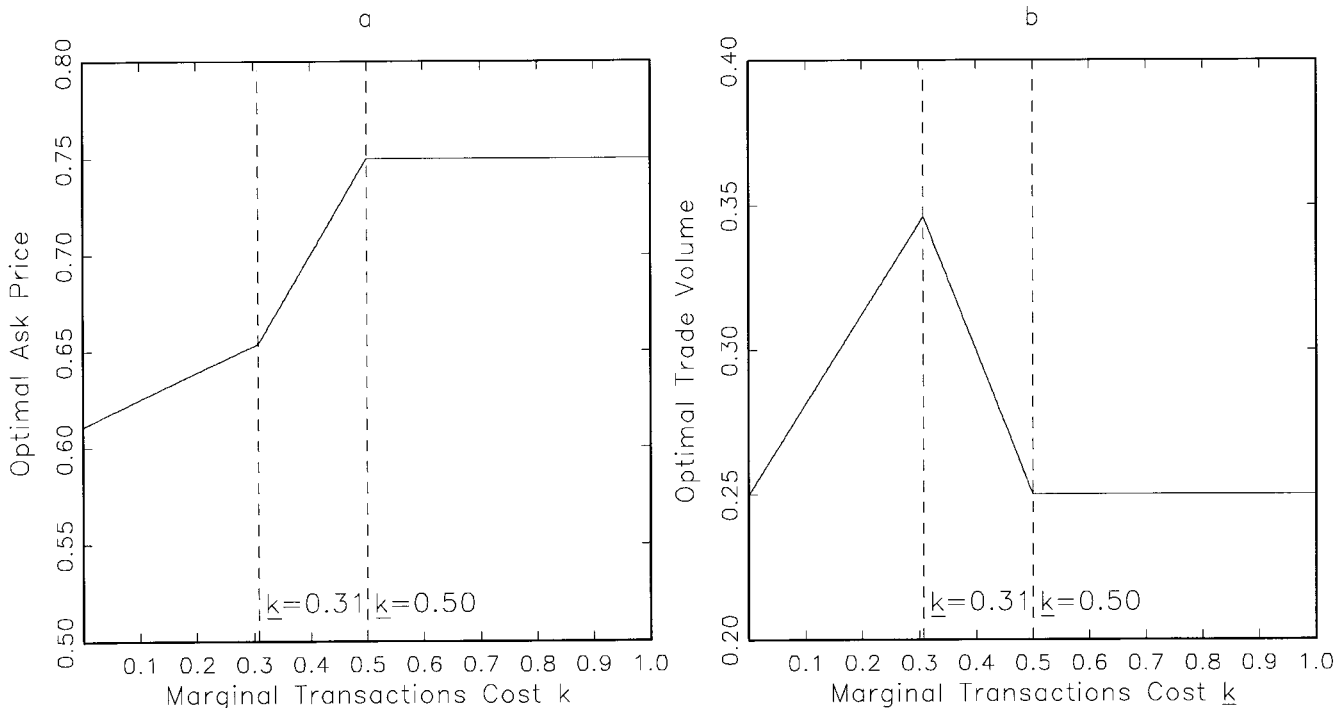

FIG. 3.—Market maker's optimal trade volume and ask price as a function of $\underline{k}$. *a*, Optimal ask price as a function of $\underline{k}$. *b*, Optimal trade volume as a function of $\underline{k}$.

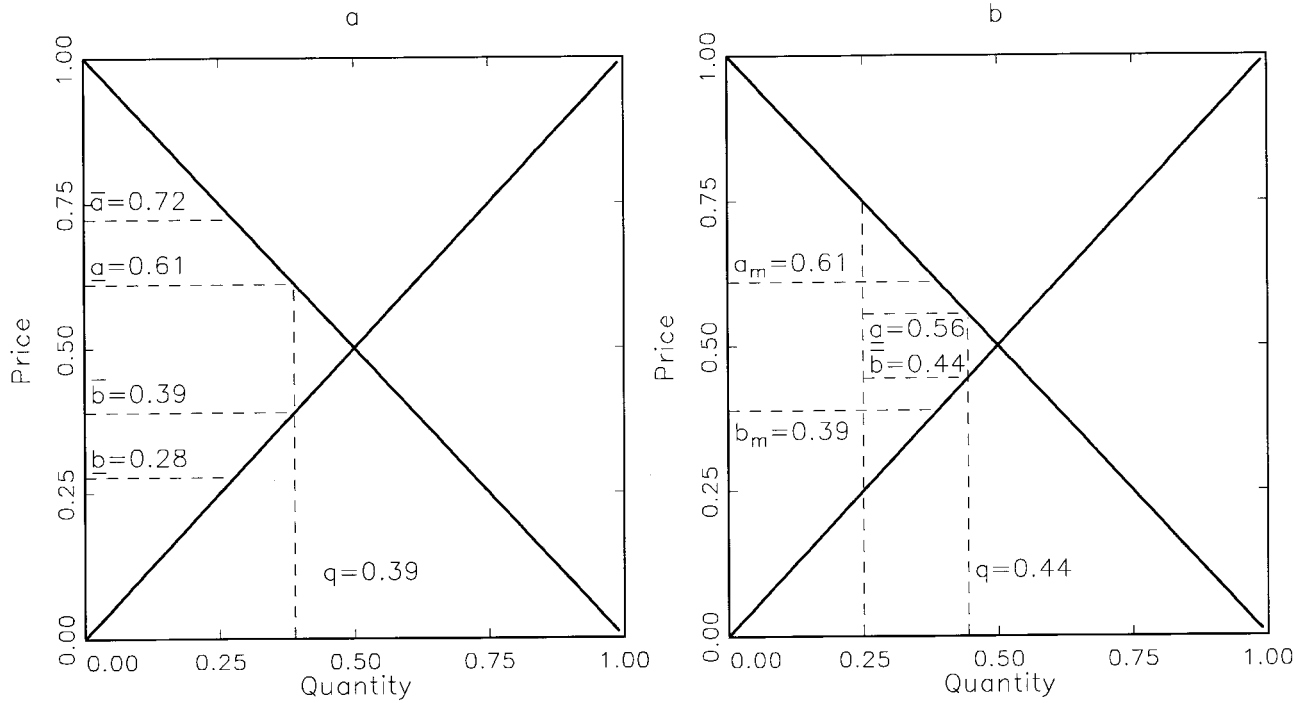

FIG. 4.—Comparison of search equilibria with and without a market maker, $\delta = .2$. *a,* Equilibrium with no market maker. *b,* Equilibrium with a monopoly market maker.

.25, which is less than half the total amount traded in the dealer market prior to the entry of the market maker. The market maker charges a bid-ask spread of $a_m - b_m = .22 = .61 - .39$, which equals the bid-ask spread of the most efficient middleman in the equilibrium without a market maker. The average bid-ask spread in the dealer market nearly halves, from .1666 to .085, following the entry of the market maker. In this case, the market maker's bid and ask prices b_m and a_m equal the highest bid price $\bar{b}_d$ and the lowest ask price $\underline{a}_d$, respectively, in the pre-entry dealer market equilibrium. The following corollary characterizes how the market maker's prices relate to the bid-ask prices prevailing in the pre-entry equilibrium in the dealer market.

COROLLARY 5.8. Assume that $\underline{k} < \underline{k}_m$, so that middlemen coexist with the market maker in the postentry equilibrium. Let $(\underline{b}_d, \bar{b}_d)$ and $(\underline{a}_d, \bar{a}_d)$ be the supports of the pre-entry equilibrium distributions of bid and ask prices, respectively. Then we have

$$a_m < \bar{a}_d, \quad b_m > \underline{b}_d. \tag{48}$$

The prices a_m and b_m can be smaller or larger than $\underline{a}_d$ and $\bar{b}_d$, respectively, depending on the values of $(k_m, \underline{k}, \delta)$. However, if $k_m = \underline{k} = 0$, then

$$a_m = \underline{a}_d, \quad b_m = \bar{b}_d. \tag{49}$$

Proof. From the formula for a_m in equation (42) and the formula for $\bar{a}_d$ in formula (21), it is not hard to show that $a_m < \bar{a}_d$ if and only if $k_m < \bar{k}_m$, where $\bar{k}_m$ is the threshold for which entry by the market maker is profitable given in equation (38). If $k_m = 0 = \underline{k}$, then it is easy to see from equations (42) and (21) that $a_m = \underline{a}_d$. Symmetric arguments establish the results for the case of bids. Q.E.D.

COROLLARY 5.9. Assume that middlemen coexist with the market maker in the postentry equilibrium. Let $\underline{a}_m$ and $\bar{b}_m$ denote the lowest ask price and highest bid price set by middlemen in the dealer market in the postentry equilibrium, respectively. Then

$$\underline{a}_m < \underline{a}_d, \quad \bar{b}_m > \bar{b}_d. \tag{50}$$

If the entry by the market maker eliminates the dealer market in the postentry equilibrium, then we have

$$a_m < \underline{a}_d, \quad b_m > \bar{b}_d. \tag{51}$$

Proof. Using the fact that $a_m = 1 - b_m$ and $\bar{a}_d = 1 - \bar{b}_d$ and the fact that $\underline{a}_m = .75a_m + .25(b_m + \underline{k})$ and $\underline{a}_d = .75\bar{a}_d + .25(\underline{b}_d + \underline{k})$, we see that $a_m < \bar{a}_d$ implies that $\underline{a}_m < \underline{a}_d$. Now consider the case in which the entry of the market maker drives the dealer market out of existence. Consider first the limit-pricing regime. Doing some algebra, we find that $a_m < \underline{a}_d$ if and only if $\underline{k} < 1$. Now consider the unconstrained monopoly re-

gime. The condition $\underline{k} > (1 + k_m)/2$ characterizing the unconstrained monopoly regime implies that $a_m < \underline{a}_d$. As in corollary 5.8, a symmetric argument establishes the results for bids. Q.E.D.

Corollaries 5.8 and 5.9 formalize the notion that entry of the market maker lowers bid-ask spreads. In the competitive regime the market maker's bid-ask spread is strictly smaller than the worst bid-ask spread charged by the highest-cost middleman in the pre-entry equilibrium. Since the market maker's bid-ask spread is necessarily the worst bid-ask spread in the postentry equilibrium, it follows that the market maker's entry has succeeded in reducing both the average and the worst bid-ask spreads in the dealer market. Indeed, the entry of the market maker not only narrows bid-ask spreads but also shifts the entire distribution of ask prices downward. That is, if we let $\underline{a}_i$ and $\bar{a}_i$ denote the support of the distribution of ask prices in the pre-entry dealer market equilibrium ($i = d$) and postentry equilibrium ($i = m$), respectively, then assuming that the market maker and middlemen coexist in the postentry equilibrium, we have the following corollary.

COROLLARY 5.10. If $\underline{k} < \underline{k}_l$ and $k_m < \bar{k}_m$, then

$$\underline{a}_m < \underline{a}_d, \quad \bar{a}_m < \bar{a}_d,$$

$$\underline{b}_m > \underline{b}_d, \quad \bar{b}_m > \bar{b}_d. \tag{52}$$

Since the distributions of bid and ask prices have strictly improved in the postentry equilibrium, it follows that all consumers and producers who participate in the postentry equilibrium are strictly better off as a result of the entry of the market maker. Corollary 5.9 shows that even when the market maker drives the dealer market out of existence, the market maker's ask price a_m is strictly less than the valuation $\underline{a}_d$ of the marginal consumer in the pre-entry equilibrium. Since the reservation price function for buyers in the pre-entry equilibrium is strictly increasing in v with a minimum value of $\underline{a}_d$, it follows that all buyers whose valuations satisfy $v > \underline{a}_m$ are strictly better off in the postentry equilibrium (where $\underline{a}_m$ is the lowest ask price in the dealer market in the postentry equilibrium). We formalize the foregoing discussion as theorem 6.

THEOREM 6. Let (r_c^d, r_p^d) denote the reservation price functions for consumers and producers in the dealer market equilibrium before the entry of a market maker. Assume that $k_m < \bar{k}_m$, so that entry by a market maker is profitable. Let (r_c^m, r_p^m) denote the reservation price functions for consumers and producers in the postentry equilibrium. Then we have

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

with strict inequality for producer valuations in the interval $v \in [0, \bar{b}_m)$ and for consumer valuations in the interval $v \in (\underline{a}_m, 1]$.

Proof. If the entry of the market maker drives the dealer market out of existence, we have $r_c^m(v) = a_m$ for $v \in [a_m, 1]$. It is easy to see from the definition of r_c^d in equation (2) that it is a strictly increasing function of v on the interval $(\underline{a}_d, 1)$ with slope

$$\frac{dr_c^d}{dv}(v) = \frac{1}{1 + [F_\alpha^d(v)/\delta]} > 0, \tag{54}$$

where F_α^d is the distribution of ask prices in the pre-entry dealer market equilibrium. By corollary 5.9 we have $a_m < \underline{a}_d$, so that

$$r_c^d(v) > \underline{a}_d > a_m = r_c^m(v), \quad v \in (a_m, 1]. \tag{55}$$

Now consider the case in which the dealer market coexists with the market maker. In this case, r_c^m will be strictly increasing on the interval $(\underline{a}_m, \bar{a}_m)$ and equal to a_m for $v \in [a_m, 1]$, since these latter consumers trade with the market maker by theorem 5. By corollary 5.9 we have $\underline{a}_m < \underline{a}_d$, so that it is sufficient to show that the slope of r_c^m is strictly less than the slope of r_c^d on the interval $(\underline{a}_m, 1)$ since $r_c^m(v) = r_c^d(v) = v$ for $v \in [0, \underline{a}_m]$. For $v \in (a_m, 1)$, the slope of r_c^m is zero, whereas from formula (54) we can see that the slope of r_c^d is strictly positive for $v \in (\underline{a}_d, 1)$. So we have $0 = dr_c^m/dv < dr_c^d/dv$ for $v \in (\max[\underline{a}_d, a_m], 1)$. For $v \in (\underline{a}_m, \underline{a}_d)$, we have $dr_c^d/dv = 1$, whereas $dr_c^m/dv = 1/\{1 + [F_\alpha^m(v)/\delta]\} < 1$, where F_α^m is the postentry distribution of ask prices. If $\underline{a}_d \geq a_m$, we are done. If $\underline{a}_d < a_m$, we complete the argument by showing that $dr_c^m/dv < dr_c^d/dv$ on the remaining interval $v \in (\underline{a}_d, a_m)$. If $k_m = \underline{k} = 0$, corollary 5.8 implies that $a_m = \underline{a}_d$, and we are done. Corollary 5.10 implies that F_α^d strictly stochastically dominates F_α^m on the interval $(\underline{a}_m, \bar{a}_d)$; that is,

$$F_\alpha^d(v) < F_\alpha^m(v), \quad v \in (\underline{a}_m, \bar{a}_d). \tag{56}$$

Since $r_c^d(v)$ and $r_c^m(v)$ are both strictly increasing functions on the interval $(\underline{a}_d, \bar{a}_m)$ (where $\underline{a}_d > \underline{a}_m$ by corollary 5.9) and since the slope of r_c^m is given by formula (54) but with F_α^m substituted in place of F_α^d, it follows that the strict stochastic dominance condition implies $dr_c^m/dv < dr_c^d/dv$ for $v \in (\underline{a}_m, \bar{a}_d)$. But by corollaries 5.8 and 5.9 we have $\underline{a}_m < \underline{a}_d$ and $a_m < \bar{a}_d$, and so the slope of $r_c^m(v)$ is strictly less than the slope of $r_c^d(v)$ in the remaining interval $(\underline{a}_d, a_m)$. Thus the slope of r_c^m is strictly less than the slope of r_c^d over the entire interval $(\underline{a}_m, 1]$, which implies that $r_c^m(v) < r_c^d(v)$ for $v \in (\underline{a}_m, 1]$ as claimed. Q.E.D.

Figure 5 illustrates the welfare gains resulting from the entry of a market maker. In this example we continue to set $\delta = .2$, $k_m = 0$, and $\underline{k} = 0$. Any consumer in the interval $(\underline{a}_m, 1]$ is made strictly better off from the entry of the market maker, where $\underline{a}_m = .56$ is the lower support point of the distribution of ask prices in the dealer market in the postentry equilibrium. The low-valuation consumers located in the interval

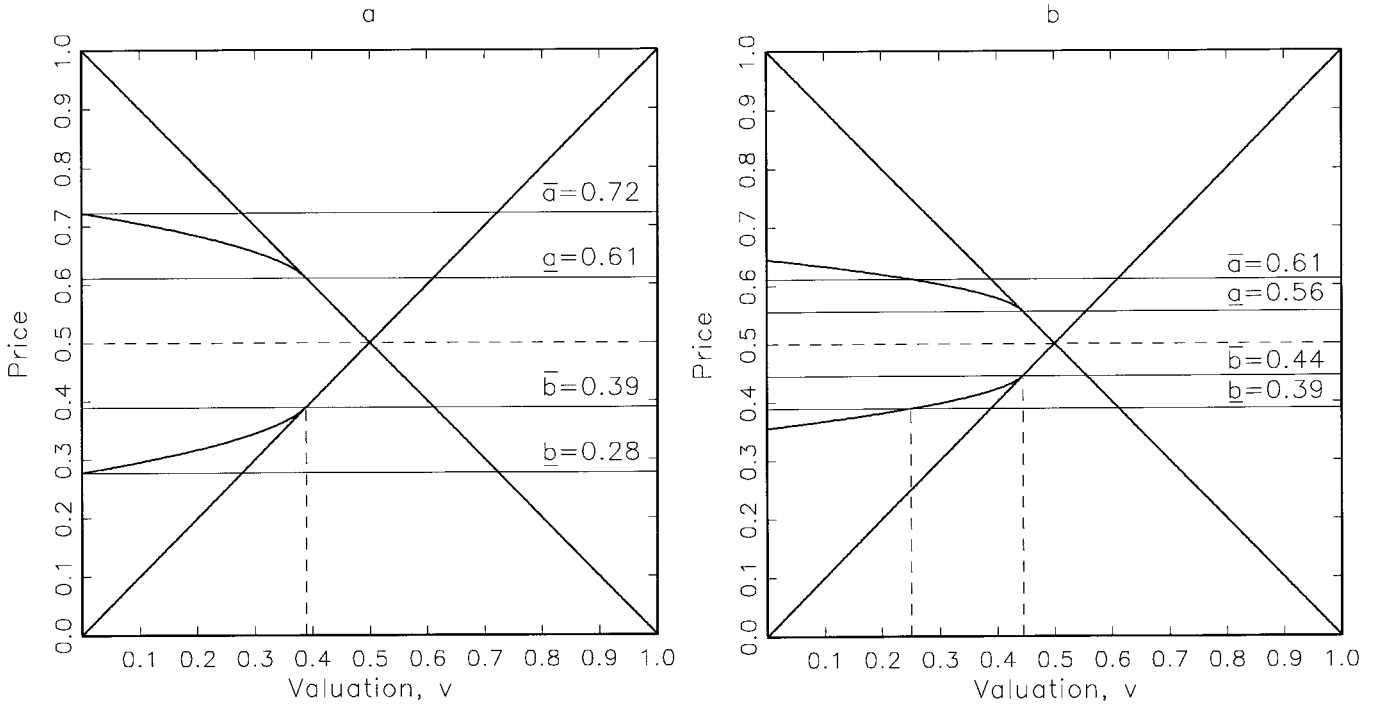

FIG. 5.—Comparison of equilibria before and after the entry of a market maker, $\delta = .2$. Along the abscissa, buyers' valuations are plotted from high to low, whereas sellers' valuations (costs) are plotted from low to high. *a,* Equilibrium with no market maker; discounted surplus .1471. *b,* Equilibrium with a monopoly market maker; discounted surplus .2242.

$[0, \underline{a}_m]$ do not trade, and their welfare is not affected by the entry of the market maker. Clearly, the 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 nonparticipation to search become strictly better off. Symmetric results apply to producers.

The entry of the market maker halves the volume of trade in the dealer market (from .3889 to .1944) and the discounted profits of middlemen and cuts profits to one-fourth of the pre-entry level (from .0494 to .0123). However, the market maker's volume of .25 and profits of .0556 more than make up for the losses incurred in the dealer market, so that overall quantity traded and overall discounted profits increase following the entry of the market maker. Total discounted surplus of consumers (and producers) is .0782 in the postentry equilibrium (i.e., with a market maker) compared with .0489 in the pre-entry equilibrium (i.e., without a market maker). Total discounted surplus (i.e., including the discounted profits of middlemen and the market maker) increases from .1471 in the equilibrium without a market maker to .2242 in the equilibrium after the entry of the market maker. Comparing these total gains from trade with the .25 gains from trade that would be realized in a frictionless Walrasian equilibrium, we see that the entry of the market maker has increased overall market efficiency from 58.8 percent to 89.7 percent.

The rise in total surplus is due largely to the reduction in the deadweight loss of the transactions costs of the less efficient middlemen who were forced out of business by the more efficient market maker (whose marginal transactions cost is zero). However, another source of the reduction in deadweight loss is the reduction in search costs due to the entry of the market maker: in the postentry equilibrium, over half of all trades occur at the bid and ask prices set by the market maker, so that a much smaller fraction of trading is subject to search costs in the postentry equilibrium. We obtain these favorable results even though the market maker is a monopolist: ordinarily monopoly power creates rather than reduces deadweight losses. Provided that the market maker's transactions costs are not too high, publicly posting bid and ask prices represents a superior institution for conducting trade than a search market.

The following theorem shows that entry of the market maker does not always increase total discounted profits: total discounted profits increase in the postentry equilibrium only when k_m is not too high.

THEOREM 7. Let $\Pi_m(k_m, \underline{k})$ denote the total discounted profits of the market maker and $\Pi_d^d(\underline{k})$ and $\Pi_d^m(k_m, \underline{k})$ denote total discounted profits

of all middlemen operating in the dealer market in the pre- and post-entry equilibria, respectively. Let $\underline{k}_m$ be given by

$$\underline{k}_m = \frac{36\delta^2 + 16\underline{k}\delta^2 + 29\underline{k}\delta + 4\underline{k}}{52\delta^2 + 29\delta + 4}. \tag{57}$$

Then $\underline{k} < \bar{k}_m$, where $\bar{k}_m$ is the threshold for profitable entry by a market maker, given in equation (38). If $k_m \in [0, \underline{k}_m)$, the entry of the market maker increases total discounted profits from intermediation; if $k_m \in (\underline{k}_m, \bar{k}_m)$, the entry of the market maker decreases total discounted profits from intermediation (relative to the pre-entry equilibrium value).

Proof. The threshold $\underline{k}_m$ is the (smallest) solution to the equation

$$\Pi_m(k_m, \underline{k}) + \Pi_d^m(k_m, \underline{k}) = \Pi_d^d(\underline{k}). \tag{58}$$

When we ignore the factor $1 + [\lambda\rho/(1 - \rho)]$, which is common to Π_m, Π_d^m, and Π_d^d, and we substitute the formula for N in the pre-entry equilibrium in equation (23) into the formula for Π_d^d in (25), we obtain

$$\Pi_d^d = \left(\frac{1 + \delta}{24\delta}\right)\left[\frac{4\delta(1 - \underline{k})}{1 + 4\delta}\right]^2. \tag{59}$$

Similarly, substituting the formula for N in the postentry equilibrium from corollary 5.5, we obtain

$$\Pi_d^m = \left(\frac{1 + \delta}{24\delta}\right)\left[\frac{1 + k_m}{2} - \frac{\underline{k}(8\delta + 1) + 1}{8\delta + 2}\right]^2. \tag{60}$$

Corollary 5.4 implies that in the competitive regime

$$\Pi_m = \left(\frac{1 - k_m}{2} + \frac{\underline{k} - 1}{8\delta + 2}\right)\left[\frac{1 - k_m}{4} + \frac{1 - \underline{k}}{16\delta + 4} - \frac{1 + k_m}{16\delta} + \frac{\underline{k}(8\delta + 1) + 1}{8\delta(8\delta + 2)}\right]. \tag{61}$$

Substituting these expressions, we see that equation (58) is quadratic in k_m and thus has two roots. One root is $\bar{k}_m$, since $\Pi_m(\bar{k}_m, \underline{k}) = 0$ and $\Pi_d^m(\bar{k}_m, \underline{k}) = \Pi_d^d(\underline{k})$ by theorem 5. It is not hard to show that the other root satisfies $0 < \underline{k}_m < \bar{k}_m$. Further using the formulas above, we can show that $\Pi_m(k_m, \underline{k}) + \Pi_d^m(k_m, \underline{k})$ is a strictly convex function of k_m on the interval $(0, \bar{k}_m)$. It follows that entry of the market maker increases total discounted profits if $k_m \in [0, \underline{k}_m)$ and decreases total profits if $k_m \in (\underline{k}_m, \bar{k}_m)$. Q.E.D.

Figure 6 illustrates theorem 7 by plotting the profit functions Π_m, Π_d^m, and Π_d^d as functions of k_m for $\underline{k} = 0$ and $\underline{k} = .6$. Profits of the market

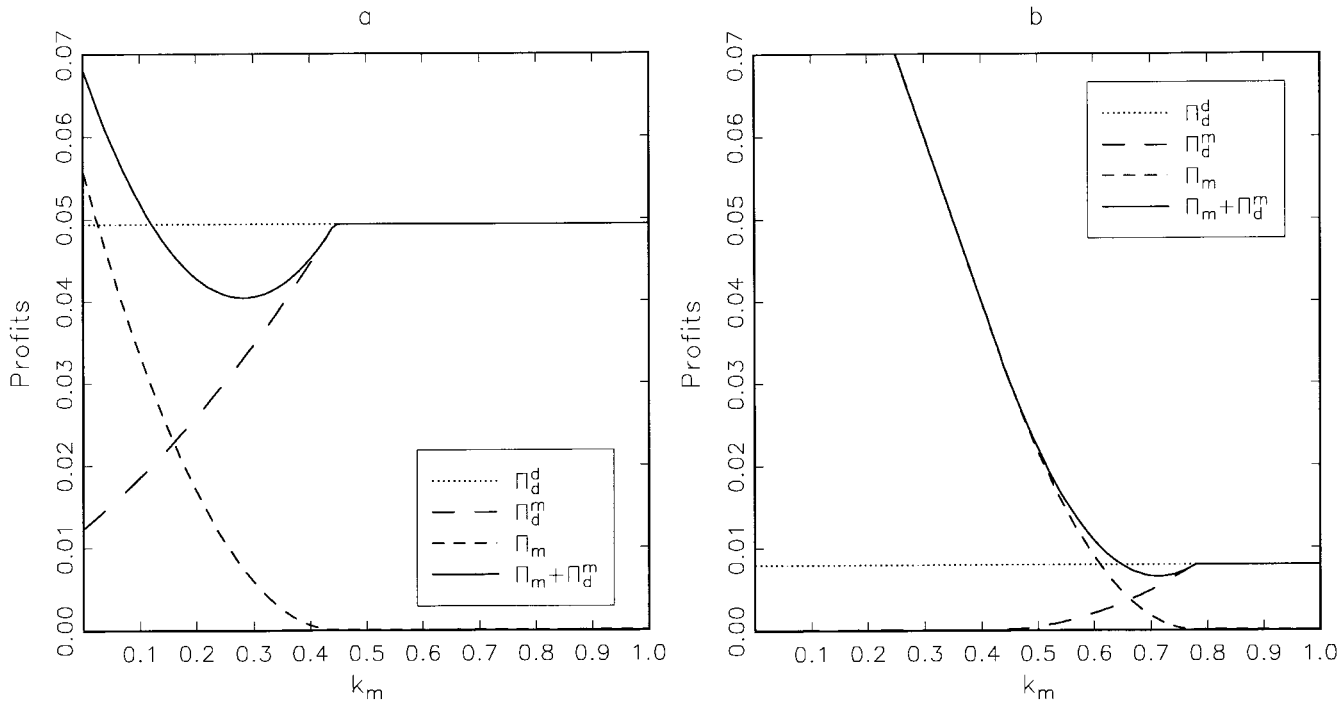

FIG. 6.—The effect of entry on total discounted profits for various k_m when (*a*) $\underline{k} = 0$ and (*b*) $\underline{k} = .6$

maker decline monotonically in k_m whereas the profits in the dealer market increase monotonically in k_m until $k_m > \bar{k}_m = .44$. Above this point, entry by the market maker is no longer profitable, so $\Pi_d^m(k_m, \underline{k}) = \Pi_d^d(k_m, \underline{k})$ for $k_m > \bar{k}_m$. When $\underline{k} = 0$, we have $\underline{k}_m = .12$, so entry of the market maker increases total discounted profits for k_m in the interval $[0, .12)$ and decreases total profits for k_m in the interval $(.12, .44)$. Figure 6*b* plots the profit functions in the case in which $\underline{k} = .6$. In this case the larger value of $\underline{k}$ has increased the lower threshold $\underline{k}_m$ to .65 and the upper threshold $\bar{k}_m$ to .77. Thus the range of k_m for which entry by the market maker increases total discounted profits from intermediation is a monotonically decreasing function of $\underline{k}$, as is evident in the formula for $\underline{k}_m$ given in equation (57).

COROLLARY 7.1. If $\underline{k}_l < \underline{k}_m$, that is, if the entry of the market maker drives the dealer market out of existence, the market maker's total discounted profits exceed the total discounted profits of middlemen in the pre-entry dealer market equilibrium.

Since the entry of the market maker strictly increases both consumer and producer surplus but can decrease total discounted profits from intermediation when $k_m > \underline{k}_m$, the effect of entry on total discounted gains from trade (i.e., the sum of discounted surplus and discounted profits) is unclear. Figure 7 plots a decomposition of total surplus in the two cases $\underline{k} = 0$ and $\underline{k} = .6$. In each case, total gains from trade increase following the entry of the market maker. However, we also see that total surplus is a monotonically declining function of k_m until $k_m > \bar{k}_m$, at which point entry by the market maker is no longer profitable. In figure 7 we indicate the various equilibrium regimes by the symbols R_1 (for the unconstrained monopoly regime), R_2 (for the limit-pricing regime), R_3 (for the competitive regime in which the middleman and dealer market coexist), and R_4 (for the case in which $k_m > \bar{k}_m$, so that entry by the market maker is not profitable). Note that surplus is constant as a function of k_m in region R_2 since, in this regime, the market maker's bid and ask prices are functions of $\underline{k}$, not k_m.

Figure 7 illustrates how the competition between middlemen and the market maker affects the division of the gains from trade between producers, consumers, and intermediaries. Consider the case in which the market maker has transactions cost $k_m = 0$. When the most efficient middleman also has transactions cost $\underline{k} = 0$, the dealer market and the market maker coexist and generate total gains from trade of .2242, or nearly 90 percent of the Walrasian surplus of .25. Producers and consumers realize a surplus of .1564, or nearly 70 percent of the total discounted gains from trade. However, when $\underline{k} = .6$, producers and consumers realize a surplus of only .0625, which is only one-third of the total gains from trade of .1875. The market maker obtains the lion's share of the surplus, earning the unconstrained monopoly profit of .125.

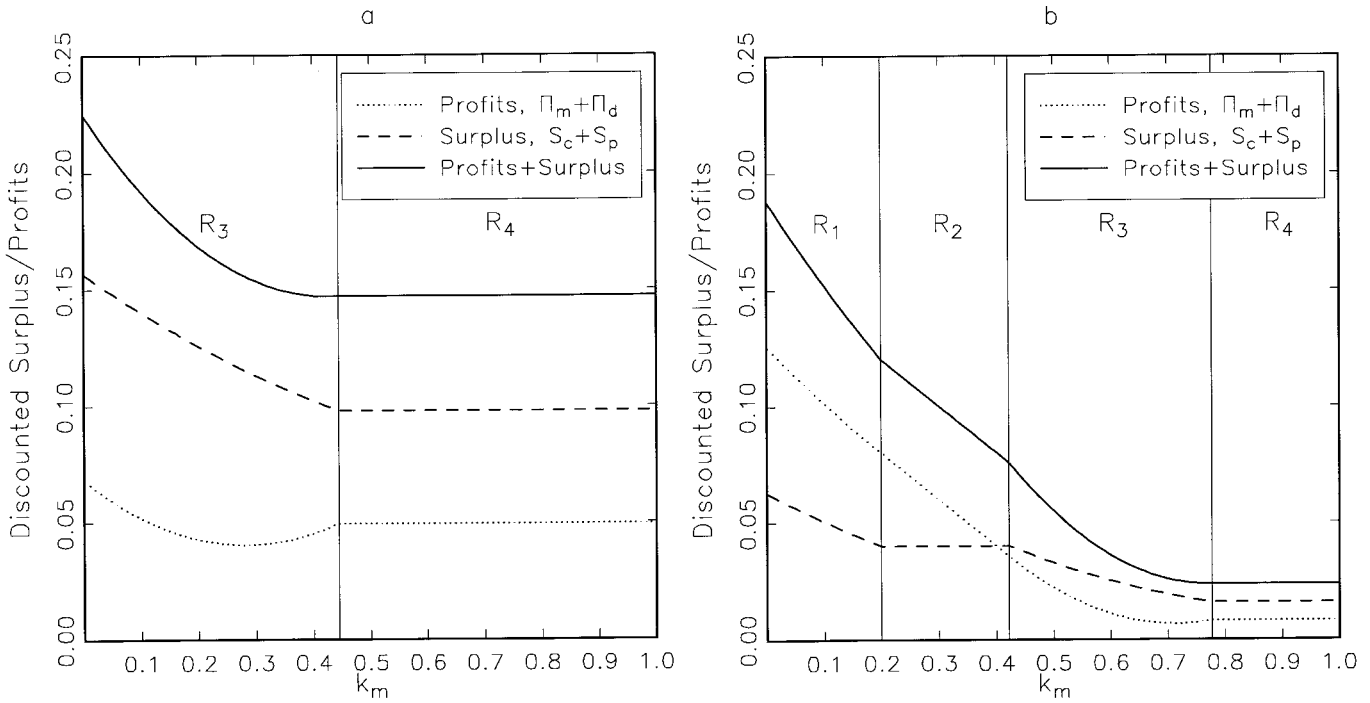

FIG. 7.—Decomposition of total gains from trade for various k_m. *a*, Welfare gains from entry of a market maker as a function of k_m with $\underline{k} = 0$. *b*, Welfare gains from entry of a market maker as a function of k_m with $\underline{k} = .6$.

Thus the market maker is made better off, and producers and consumers are made correspondingly worse off, by any measure that limits entry or increases the transactions costs of middlemen.

Similarly, middlemen can increase their profits at the expense of consumers and producers by promulgating measures that artificially limit entry or increase the transactions cost of the market maker. For example, figure 7*a* indicates that if middlemen were successful in passing a law preventing the entry of market makers, total discounted surplus of producers and consumers would fall significantly. For example, in the presence of a market maker with $k_m = 0$, total producer and consumer surplus is .1564, whereas if entry of a market maker is prohibited, total producer and consumer surplus would fall by 38 percent to .0977.

Although all our numerical results indicate that total gains from trade are a monotonically decreasing function of k_m for $k_m < \bar{k}_m$ (which implies that the entry of a market maker always increases total gains from trade for any $k_m < \underline{k}$), we have not yet been able to prove this analytically. Below we simply state our conjecture that entry by the market maker is always welfare-improving.

CONJECTURE 1. Let W^d denote total market surplus in the pre-entry dealer market equilibrium, given in formula (26). Let W^m denote total market surplus in the postentry equilibrium, given by

$$W^m = S_c^m + S_p^m + \Pi_d^m + \Pi_m, \tag{62}$$

where S_c^m and S_p^m are total discounted consumer and producer surplus in the postentry equilibrium, respectively. Then W^m is a strictly monotonically decreasing function of k_m for $k_m < \bar{k}_m$. In particular, for any $k_m < \bar{k}_m$, we have

$$W^m > W^d. \tag{63}$$

THEOREM 8. If $k_m \in [0, \bar{k}_m)$, then the mass of middlemen operating in the pre-entry dealer market equilibrium is greater than the mass of middlemen operating in the postentry dealer market equilibrium. When $\underline{k} = 0$, the number of middlemen in the postentry equilibrium is never less than half of the number of middlemen that would be operating in an equilibrium in which entry by a market maker is prohibited.

Proof. Let N^d be the mass of middlemen in the pre-entry dealer market equilibrium and N^m be the mass of middlemen in the postentry dealer market equilibrium. We have

$$N^m = \frac{1 + k_m}{2} + \frac{\underline{k} - 1}{8\delta + 2} - \underline{k} \leq \frac{4\delta(1 - \underline{k})}{4\delta + 1} = N^d. \tag{64}$$

Q.E.D.

THEOREM 9. If $k_m \in [0, \bar{k}_m)$, the total mass of producers and consumers

who participate in the market is strictly larger in the postentry equilibrium than in the pre-entry equilibrium.

Proof. The mass of consumers who participate in the dealer market prior to the market maker's entry is

$$Q_d^d = \bar{b}_d = \frac{(1 - \underline{k})(1 + 2\delta)}{8\delta + 2}. \tag{65}$$

By theorem 5, when $\underline{k} > \underline{k}_u$, the market maker drives the dealer market out of existence, so that $Q_d^m = 0$. Thus the total mass of consumers who participate are those who trade with the market maker, Q_m. By corollary 5.2 we have $Q_m = (1 - k_m)/4$ in the unconstrained monopoly regime. It is easy to see that $Q_m > Q_d^d$ for $\underline{k} \in [\underline{k}_u, 1]$. In the limit-pricing regime we still have $Q_d^m = 0$, but the monopoly quantity is given by $Q_m = (1 - \underline{k})/2$. Since $Q_m > Q_d^d$ at $\underline{k} = \underline{k}_u$, it suffices to show that the slope of Q_m is steeper than the slope of Q_d^d for $\underline{k} \in (\underline{k}_l, \underline{k}_u)$. The slope of Q_m with respect to $\underline{k}$ is $-\frac{1}{2}$, which is less (i.e., steeper) than the slope of Q_d^d with respect to $\underline{k}$, $-(1 + 2\delta)/(2 + 8\delta)$. For the final case, $\underline{k} \in [0, \underline{k}_l)$, we proceed similarly. We need to show that the slope of $Q_m + Q_d^m$ is steeper than Q_d^d in $\underline{k}$. The slope of the former is $-(8\delta + 3)/(32\delta + 8)$, and it can be readily verified that this is less (i.e., steeper) than $-(2 + 2\delta)/(2 + 8\delta)$, the slope of Q_d^d. Q.E.D.

D. *Limiting Properties*

Finally, we summarize the limiting properties of the postentry equilibrium as $\delta \downarrow 0$, $k_m \downarrow 0$, and $\underline{k} \downarrow 0$.

THEOREM 10. In the limit, as $\delta \downarrow 0$, $\underline{k} \downarrow 0$, and $k_m \downarrow 0$, 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 middleman (with $k = \underline{k} = 0$) handles the rest. Both charge a limiting bid-ask spread of zero at the common Walrasian equilibrium price of $p^* = \frac{1}{2}$.

This result implies that the coexistence of middlemen and market makers should be fairly robust in a world of steadily declining search and transactions costs because of technological improvements in information and communications technologies. However, the conceptual distinctions between middlemen and market makers start to blur in the limit since the prices charged by all surviving middlemen are virtually the same as the bid and ask prices charged by the market maker. We view theorem 10 as a characterization of "efficient markets": when search and transactions costs are small, bid-ask spreads are very narrow, so that there is not much difference between trading with a market maker and trading in the dealer market. For this reason there is approximately a

50-50 split in trade between these two competing institutions. However, in the limit, it does not matter whether the microstructure of trade involves only middlemen or a combination of middlemen and a market maker. As we saw in Section II, we have convergence to a Walrasian equilibrium even in the absence of entry by a market maker provided that $\delta \downarrow 0$ and $\underline{k} \downarrow 0$.

The entry of a market maker is much more important when $\underline{k}$ or δ is large. Conversely, if a market maker is able to obtain unconstrained monopoly power by creating artificial barriers to entry by middlemen or other, competing market makers, then large efficiency and welfare gains can be achieved by breaking down these barriers and subjecting the market maker to competitive pressure to reduce bid-ask spreads. In the next section we shall consider the case in which the market maker faces competition from other potential market makers as well as from middlemen.

IV. Search Equilibrium with Competitive Market Making

Now suppose that there is free entry into market making. If there are no fixed entry costs and entry occurs simultaneously, Bertrand-style price competition will ensue among competing market makers. Since all producers and consumers can costlessly observe the bid and ask prices quoted by alternative market makers, all trade will occur with the market maker that offers the best bid and ask prices. Bertrand price cutting will result in two possible outcomes, depending on the heterogeneity in transactions costs. If there is more than one market maker that has the smallest per unit transactions cost, price competition will result in a zero-profit equilibrium in which the bid-ask spread equals the per unit transactions cost of the most efficient market maker. The division of trade among the competing most efficient market makers is indeterminate. If the most efficient market maker has a strictly lower per unit transactions cost than the next most efficient market maker, there may be a positive-profit monopoly market maker equilibrium in which the most efficient market maker deters the entry of other market makers by charging a bid-ask spread equal to the per unit transactions cost of the next most efficient potential entrant.

If entry into market making occurs sequentially and there are fixed costs associated with entry, the outcome depends on the order in which potential market makers arrive and enter the market. If the first potential entrant can enter profitably, he will do so according to the theory of the previous section. When the next potential entrant into market making arrives, the existing market maker will limit price, setting a bid-ask spread such that the net profits for any new entrant would be zero. If a potential entrant had sufficiently low fixed costs of entry and suf-

ficiently low per unit transactions costs, then it is possible that the incumbent would not be able to limit price and still make a profit. In this case, the incumbent would be displaced by the most efficient potential entrant, and so forth.

It should be clear from the analysis in the previous section that any additional competition among market makers, in addition to the competition that already exists between middlemen and market makers, will result in further narrowing of bid-ask spreads and additional welfare gains for producers and consumers. Total profits earned by middlemen and market makers will generally fall as a result of additional competition from other potential market makers, but this fall in profits will be offset by an increase in surplus that will be reaped by producers and consumers.

The stationary equilibrium model used in this paper is an unsatisfactory framework in which to study the issue of competition between market makers. A more realistic analysis would have to consider the large fixed costs and frictions involved in becoming an "established" market maker. If there are many potential market makers and consumers must incur positive search costs to discover the bid and ask prices charged by competing market makers, then there is the possibility of "market fragmentation." In this case, competing market makers may not be much different from competing middlemen: producers and consumers will not know which market maker is offering the best bid-ask spread without undertaking a sequential search process to obtain the best price quotes. Thus most of the benefits of having a single marketplace in which credible bid and ask prices are quoted could be lost. Further, if there are elements of increasing returns to scale and network externalities associated with having larger populations of traders (which enable a market maker to have a "thicker" and more "continuous" market), there may be "natural monopoly" elements to market making. In our framework these sorts of increasing returns could be modeled by assuming that the market maker's per unit transactions cost is a decreasing function of the total volume traded. In such an environment the equilibrium outcome may be indeterminate, or there may be multiple equilibria.

Caillaud and Jullien (2001) analyzed the "chicken and egg" problems arising from the network externalities involved in competition between market makers (or "matchmakers") in a different framework. These issues are worthy of further exploration, but they require a more complicated model than we have employed here. A more realistic model will have to account for various types of nonstationarities and network externalities in order to yield a more satisfactory dynamic analysis of competition in market making in which issues of market fragmentation and natural monopoly can be addressed.

V. Applications

The theory in this paper is applicable to a wide range of asset and commodity markets. In this section we look at four markets: the market for commodities in a World War II POW camp, the U.S. stock market, the U.S. Treasury bill market, and the U.S. steel market. In the first three markets we observe the coexistence of both middlemen and market makers. In the fourth we observe only middlemen despite recent attempts by Enron and two potential web-based market makers, e-STEEL and MetalSite, to enter the market. In light of the quotation at the beginning of the article, we discuss why the brokerage firm Cantor Fitzgerald was able to become a market maker in the U.S. Treasury bond market whereas the two steel dot-coms have not become market makers.

A. The Commodities Market in a POW Camp

Perhaps one of the best-known examples (at least among economists) of the coexistence of middlemen and market makers is the World War II POW camp described by Radford (1945). In this camp, prisoners traded a variety of commodities among themselves: canned milk, jam, biscuits, and chocolate. In the absence of fiat money, cigarettes became a form of currency. Trade was facilitated by "Exchange and Mart notice boards" on which bid and ask prices for different goods were posted. When a deal was consummated, the posting was crossed out. Radford notes that "the public and semi-permanent record of transactions led to cigarette prices being well known and thus tending to equality throughout the camp" (p. 191). In addition to the Exchange and Mart, middlemen were active in the camp, although they were viewed with disdain:

> Despite the fact that his very existence was proof to the contrary, the middleman was held to be redundant in view of the existence of the Shop and the Exchange and Mart. … And middlemen as a group were blamed for reducing prices. Opinion not withstanding, most people dealt with a middleman, whether consciously or unconsciously, at some time or another. [P. 199]

In this camp the Exchange and Mart served the role as a market maker. Both current and historical bid and ask prices were publicly and costlessly observable. As our model predicts, these middlemen did help reduce ask prices. Radford does not mention whether middlemen had any effect on bid prices.

B. The Stock Market

In the contemporary market for U.S. equities there is substantial interest in the question of whether entry of middlemen can reduce the bid-ask spread 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 in one or more individual securities. Each holds inventories, posts publicly observable bid and ask prices, and reports a history of past transaction prices.[4] However, there is also an active set of middlemen commonly referred to as the "over-the-counter" (OTC) market. Pirrong (2000) estimates that the OTC market accounts for only 8 percent of the volume and 10 percent of the transactions in NYSE-listed securities. Despite this small market share, the OTC market does appear to play the role of the competitive fringe as in our model. Although the evidence is mixed, several empirical studies, such as Battalio (1997), document that the quoted bid-ask spread for NYSE-listed securities tightens when "third-market" dealer/brokers enter and compete against the NYSE to execute trades.

The NYSE also faces increasing competition from other potential market makers such as R. Steven Wunsch's computerized Arizona Stock Exchange (http://www.azx.com) and the Cincinnati Stock Exchange. In recent years the advent of "electronic communication networks," competition from computerized foreign exchanges, and a change to a more procompetitive regulatory regime at the Securities and Exchange Commission (SEC) have subjected the NYSE to much more competitive pressure than it has experienced in the past.[5] This has forced the NYSE to make changes it had previously resisted such as repealing "Rule 390," moving toward 24-hour trading, and allowing prices to be quoted in decimals rather than in one-eighth increments.[6] Battalio, Greene, and Jennings (1997) studied the effect of a set of rule changes in the early 1990s that made it easier to trade NYSE-listed securities on regional exchanges. They found that, after the rule changes, bid-ask spreads decreased for about two-thirds of the securities in their sample.

Although the narrowing of bid-ask spreads that Battalio (1997) and Battalio et al. (1997) find in response to increased competition is consistent with our model, it runs counter to the intuition of a set of models in the financial intermediation literature (see, e.g., Easley, Kiefer, and O'Hara 1996; Fong, Madhavan, and Swan 1999). These papers empha-

[4] Typically a specialist on the NYSE posts prices for only relatively small transactions. Almost all large transactions are negotiated "upstairs" via middlemen; only after the deal is consummated is the transaction price publicly posted.

[5] See http://www.island.com for an example of an electronic communication network with posted transactable prices and completely open order books.

[6] The NYSE's Rule 390, which prevented member firms from trading with middlemen in the OTC market, was repealed on May 5, 2000.

size the potential adverse selection problem associated with competition between intermediaries. These models generally consider two types of traders: informed and uninformed. If the middlemen are able to selectively trade with only the uniformed traders (e.g., by accepting only small orders), then the market maker is left trading with only the informed traders. This "cream skimming" by the middlemen leaves the market maker at an informational disadvantage against the informed traders. In response, the market maker must widen his bid-ask spread in order not to systematically lose money.

Within this adverse selection literature, Pirrong's (2000) model is perhaps the closest to ours. He considers the effect of competition between a monopolist market maker and a "third market" consisting of middlemen. He focuses on the effects of entry by middlemen on an initial equilibrium with a monopolist market maker, whereas we focus on the effect of the entry of a monopolist market maker on an initial search equilibrium in which there is free entry and exit of middlemen. Pirrong's model has differently informed traders and studies whether the creation of a third market results 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).

Our model is silent on this adverse selection problem and, as noted above, does not address potential network externalities associated with market making. In our model, free entry by middlemen provides a significant competitive threat to a monopolist market maker, forcing it to substantially reduce its bid-ask spreads. This results in a significant welfare gain to both buyers and sellers. Since middlemen undercut the market maker's quoted prices, their entry increases price dispersion compared with the initial equilibrium. From the monopolist market maker's point of view, this additional price dispersion is unnecessary and is evidence that the middlemen are free-riding on its price discovery; however, from the point of view of producers and consumers, any additional costs associated with the extra price dispersion are outweighed by the benefits of the reduction in the market maker's bid-ask spreads.

C. The Treasury Bill Market

Thirty years ago the secondary market for U.S. Treasury securities was dominated by middlemen. Trades between dealers were made through a small set of interdealer brokerage firms. Trades were conducted over the telephone, and dealers did not know the prices other dealers re-

ceived, nor were customers (e.g., pension funds) shown the prices dealers faced. However, in 1972, Cantor Fitzgerald, a brokerage firm, began allowing dealers to see transactable bid and ask prices on computer screens. Dealers could now see the prices other dealers received. As the quotation at the beginning of the article states, by making its prices transparent, Cantor Fitzgerald "vaulted to nearly monopolistic preeminence in the bond market." In 1990, in response to calls from the SEC and customers for greater market transparency, several of the major brokers in the interdealer market formed a joint venture, GovPX, to consolidate interdealer data. Once consolidated, these data are transmitted in real time to customers through vendors such as Bloomberg. In our view the Treasury market has been transformed from primarily a dealer market like the one described in Section II to a market with competitive market making like the one described in Section IV.

Today trades in the interdealer market are made through brokers either by telephone or over an *electronic transaction system* (ETS). Like the market maker in our model, ETSs allow dealers to post transactable prices and quantities and execute trades electronically. The Bond Market Association (2001) reports that there are currently 33 ETSs active in the U.S. Treasury market, but the two largest ETSs, eSpeed and BrokerTec Global, dominate the market.[7] The Bond Market Association reports that although the ETSs are capturing a large share of trades in the most liquid issues, trades in less liquid issues still take place primarily over the telephone. We are told by Treasury market participants that the telephone market is helpful in getting information about the depth of the market that is not always available on an ETS. Although our model does not explicitly account for liquidity, one can stretch the model's intuition by interpreting a decrease in liquidity as an increase in the cost of carrying out a transaction. This would suggest that liquid securities are associated with low k_m's relative to $\underline{k}$ whereas the reverse holds true for less liquid securities.

Although precise data on the fraction of interdealer trades executed electronically are unavailable, one (albeit noisy) way to measure the movement away from telephone trading and toward ETSs is to view the fraction of total interdealer trades not captured by GovPX. Neither eSpeed nor BrokerTec reports its trades to GovPX; therefore, the trades recorded by GovPX are dominated by the telephone brokers. All interdealer broker trades, including those made through the ETSs, are reported to the Federal Reserve Bank of New York. Fleming (2001) reports that GovPX's coverage of the total market has fallen in recent years from 65 percent in 1997 to 57 percent in 1998 and 52 percent in

[7] Cantor Fitzgerald owns 55 percent of eSpeed. BrokerTec is run and owned by a consortium of large Wall Street firms.

1999. Although certainly not conclusive, this evidence suggests that the ETSs' market share is increasing.

Although our model predicts that the rise in electronic trading should lead to a decrease in bid-ask spreads in the interdealer market, bid-ask spreads are generally higher today than they were in 1997. In particular, bid-ask spreads widened substantially in 1998 during the Russian crisis and the near collapse of Long Term Capital Management and have not returned to precrisis levels. We do not have access to the high-quality transactions data necessary to tease out the effect of the ETSs on bid-ask spreads, taking into account other broad market events (e.g., the reduction in U.S. Treasury debt during the second half of the Clinton administration and the Treasury's buyback program).

D. The Steel Market

Just as ETSs have transformed the U.S. Treasury market for securities from one dominated by middlemen to one dominated by market makers, it is not hard to imagine similar conversions happening in other markets. In particular, we believe that the U.S. steel market is a likely candidate for such a transformation. The model of a dealer market presented in Section III provides a reasonable caricature of the current state of the U.S. steel market. It is a highly competitive and unconcentrated market in which over 5,000 "steel service centers" play the role of middlemen between buyers and sellers of steel products. In 1998, three potential market makers—Enron, e-STEEL, and MetalSite—entered the market. None has yet been successful in garnering a significant share of transactions.[8] To be clear, neither e-STEEL nor MetalSite claims to be an explicit market maker: neither holds its own inventories, and neither posts its own bid and ask prices. Instead, these sites operate like the Exchange and Mart in the POW camp discussed above or like the "information gatekeepers" studied by Baye and Morgan (2001). Information gatekeepers are web sites, such as Shopper.com or Mortgage-quotes.com, that centralize dispersed price information and reduce search costs by allowing buyers and sellers to post bid and ask prices on the web.

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

[8] Enron declared bankruptcy in December 2001. Although both e-STEEL and MetalSite are still in business, neither firm is capturing a significant share of transactions. MetalSite shut down its web site and all trading in June 2001, but relaunched it in late November 2001. e-STEEL is now NewView and has switched its focus to licensing interenterprise software. We, like many of Enron's shareholders, do not fully understand what Enron was able to do in the steel market. So we focus our attention on the other two firms.

spreads. Suppose that the gatekeeper charges a per unit commission or transaction fee τ to buyers and sellers to post prices on the site. Thus if a buyer posts a bid of b and succeeds in transacting, the total per unit cost he actually pays would be $b + \tau$. Similarly, if a seller posts an ask price of a and transacts, the seller's actual per unit proceeds are $a - \tau$. With a continuum of buyers and sellers, the only Nash equilibrium outcome is for all buyers to post a common bid price b and all sellers to post a common ask price a. For any given value of τ, supply and demand for units advertised by the gatekeeper will be equated if and only if $a = 1 - b$. Consider an equilibrium in which $a = b = \frac{1}{2}$ and the gatekeeper's commission is one-half the bid-ask spread charged by a monopolist market maker, namely,

$$2\tau^* = a_m - b_m = \frac{1 + k_m}{2} + \frac{\underline{k} - 1}{8\delta + 2},$$

where k_m 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 $\frac{1}{2}$ but, net of the commission, receives a per unit sales price equal to $\frac{1}{2} - \tau^* = b_m$, which is the same price the seller would have obtained from a monopolist market maker. Symmetrical remarks apply to buyers. It is immaterial whether a market has a monopolist market maker or a monopolist gatekeeper: both lead to exactly the same equilibrium outcome.[9] 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 as a result of the search frictions.

The crucial difference between the two steel dot-coms and the ETSs in the U.S. Treasury market is price transparency. In the Treasury market, dealers must post transactable, "take-it-or-leave it" prices. The ETSs in the interdealer Treasury market are designed so that no negotiation over price occurs (although negotiation over quantity may still occur). Posted prices are available to all subscribers, and the history of past transactions is made public. In contrast, e-STEEL and MetalSite are designed as "computerized chat rooms" in which private transactions are negotiated. Both web sites allow buyers and sellers to post prices, but next to each posted price is a "negotiate" or "counteroffer" button. From our discussions with steel middlemen, we learned that these posted prices were much like the list price for a new car; they represent first

[9] 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 nothing to buyers who post bids on the site but a commission equal to $\tau^* = a_m - b_m$ to sellers. In this case the equilibrium outcome would be for all buyers to place bids equal to a_m and all sellers to place asks equal to a_m. The price received by sellers net of commission is then the same as the bid price b_m that a monopolist market maker would choose.

offers, not take-it-or-leave-it prices. Negotiation is expected. Consequently, these two web sites are more akin to a computerized extension of the existing dealer or telephone market for steel, but individual deals are negotiated by typing messages into a computer terminal rather than conducted over the telephone. To the extent that most producers and consumers in the steel market find it easier to negotiate verbally by telephone than by typing messages over a computer terminal, e-STEEL and MetalSite not only fail to perform the role of market maker but may indeed constitute an inferior technology for intermediation in the dealer market compared with the pre-existing telephone technology.

Furthermore, neither e-STEEL nor MetalSite posts historical transaction data, and both allow buyers and sellers to limit who can view their own postings. There is a question on the "frequently asked questions" page of the e-STEEL web site: "Does e-STEEL create pricing transparency?" The posted answer is "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 about the current market price of steel from visiting e-STEEL. We conclude that these two web sites are simply offering an alternative communication channel to the telephone to enable buyers and sellers to negotiate privately. Neither fulfills the role of a market maker or an information gatekeeper that posts publicly observable and transactable bid and ask prices. This may be part of the reason that these firms have not been successful in gaining a significant share of transactions in steel.

VI. 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 middlemen introduced by Spulber (1996*a*). Middlemen and market makers represent complementary and competitive exchange institutions: market makers post publicly observable bid and ask prices, whereas prices quoted by middlemen in the dealer market constitute private information that can be obtained only through a costly search process. We have focused on the effect of entry by a monopolist market maker on an initial equilibrium in which there is free entry by competitive middlemen.

Figure 8 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 per unit transactions cost of the market maker k_m, and the per unit transactions cost of the most

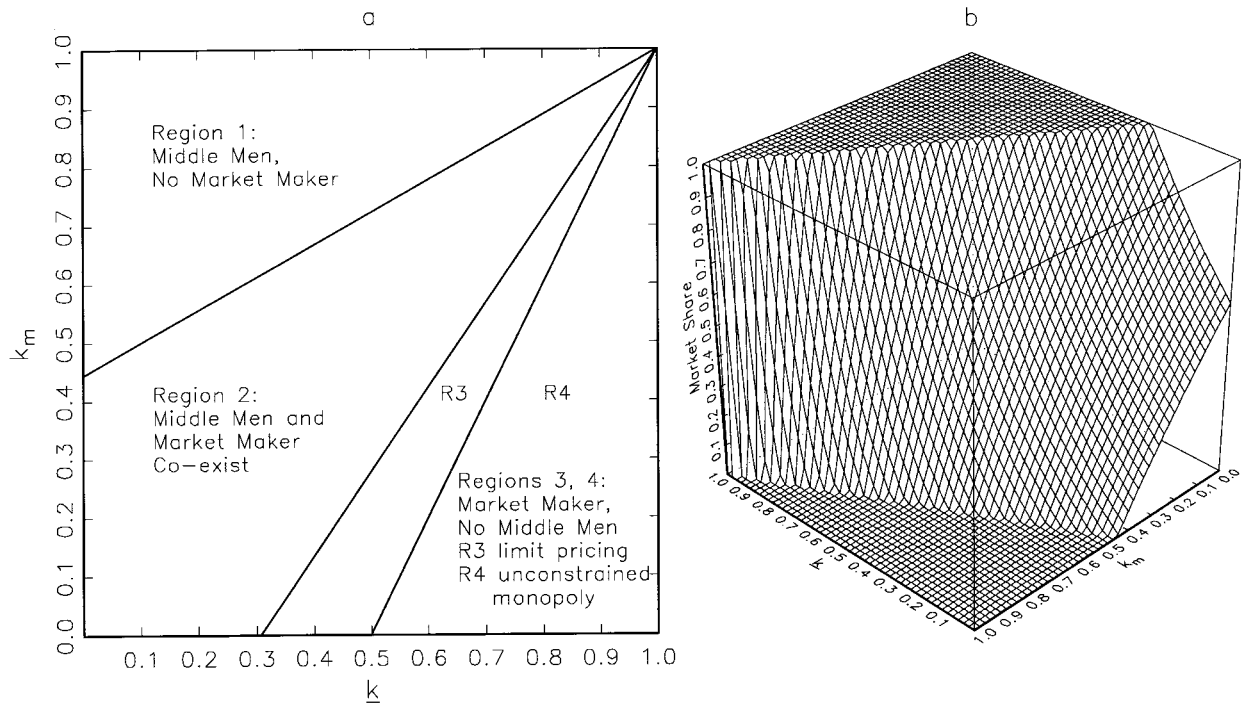

FIG. 8.—Summary of equilibria, $\delta = .2$ *a,* Classification of equilibria. *b,* Market maker's share of total trade.

efficient middleman $\underline{k}$. Figure 8*a* displays the four possible equilibrium regimes that occur for various combinations of the $(k_m, \underline{k})$ parameters when the search cost is fixed at $\delta = .2$. The model predicts that no market maker will be present if k_m is sufficiently high relative to $\underline{k}$ (region 1) and that no middlemen will exist if $\underline{k}$ is sufficiently large relative to k_m (regions 3 and 4). Region 2 represents the intermediate slice of $(k_m, \underline{k})$ values that permit the coexistence of middlemen and market makers. Figure 8*b* shows the share of trade handled by the market maker. In region 2, this share increases linearly in $\underline{k}$ for any fixed k_m or, conversely, declines linearly in k_m for any fixed $\underline{k}$. We can conceptualize the steel industry as corresponding to values of $(k_m, \underline{k})$ in region 1, where entry by a market maker is currently unprofitable. In the securities industry, entry barriers to middlemen and other market makers created by the NYSE correspond to values of $(k_m, \underline{k})$ that are close to region 3, where the market maker engages in limit pricing; that is, 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.

We recognize that there are several limitations to our analysis that qualify the types of conclusions we can draw from it. First, as noted in Sections IV and V, our model does not account for information asymmetries or network externalities that could affect our conclusions about whether entry by middlemen or competing market makers always benefits buyers and sellers in the market. A richer analysis would be required to determine whether some intermediaries might free-ride on the price discovery provided by a market maker, so that market fragmentation could occur and raise traders' search costs and reduce welfare. If there are network externalities in addition to the information problems, there may be conditions under which market making has elements of natural monopoly.

A second limitation is that we assumed that all exchange must be intermediated by either a middleman or a market maker. As we noted in the Introduction, only half of the volume of trade in steel occurs through middlemen; the rest does not occur through market makers but through direct transactions between producers and consumers. Hendershott and Zhang (2001) study an extension of Spulber's (1996*a*) model in which a monopolist producer can sell directly to consumers or through a middleman. Direct sales involve lower search costs than intermediated sales. In equilibrium, the market segments. In a result analogous to ours, high-valuation consumers purchase directly from the producer, and intermediate-valuation consumers choose to search for better prices in the dealer market.

A third limitation of our analysis is that we constrained supply and demand for the commodity for both market makers and middlemen to

be equal in every period. As a result, these agents have no inventory holdings in our model. An important function of intermediaries is to hold inventory to provide a buffer stock that offers their customers liquidity when there is an imbalance between supply and demand (see Spulber 1996*b*). 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" so that a customer can be guaranteed of receiving shipment of an order within a few days of placement. Lacking inventories and stockouts, this model cannot be used to analyze the important role of intermediaries in providing liquidity.

Although our model is highly simplified and stylized, it provides insights into the organization of a variety of different asset and commodities markets. Consider the two puzzles raised in the Introduction. One puzzle was to explain how entry could be profitable even if middlemen uniformly undercut the market maker's publicly posted bid and ask prices. Our explanation is that even though it is common knowledge that middlemen offer better prices, the highest-valuation buyers and lowest-cost sellers still find it optimal to trade immediately at these prices rather than incur the search costs involved in trying to find a better price in the dealer market.

The other puzzle was to explain why market makers intermediate a significant share of trade in financial assets such as bonds and stocks, but virtually none of the trade in steel. Our model suggests that an explanation for this puzzle is that transactions costs for market making are high for commodities such as steel but low for financial assets such as bonds. However, this explanation may seem tautological. Why would transactions costs associated with market making be so high for commodities such as steel and so low for financial assets such as bonds?

Recall that k_m can be thought of as including a "rebate" to buyers and sellers to offset any transactions costs involved in transacting with the market maker. In financial assets such as bonds and agricultural commodities such as wheat and pork bellies, buyers and sellers (or producers and consumers) may be relatively sophisticated and may be used to conducting transactions through a market maker at a central exchange. Thus their "hassle" or transactions costs may be fairly low, resulting in a low value of k_m relative to $\underline{k}$. However, in the steel market, traders have little experience in conducting transactions over an exchange. They may perceive relatively high transactions costs to doing business with a market maker compared with their local steel service center. Thus a new entrant to the steel market may face a different culture than the initial entrants to the bond or wheat markets, and this difference in culture could translate into a higher effective transactions cost k_m as explained in the beginning of Section III.

However, the increasing penetration of computers and the World Wide Web seems to be gradually changing the culture in the steel market, lowering k_m and creating the possibility for profitable entry. As we noted, there have recently been several unsuccessful attempts at entry by potential market makers. This may indicate that technology and associated cultural changes are reducing k_m relative to $\underline{k}$, moving the market toward the boundary between region 1 (where entry by a market maker is unprofitable) and region 2 (where entry is profitable). From our own observation of the steel market, we think that it is only a matter of time before successful entry does occur. Also, some of the problems experienced by the first wave of entrants into the steel market may have been due to the general effects of the dot-com crash of 2001–2, and some of these entrants (such as Enron) might have succeeded in entering the steel market had it not been for problems in their other lines of business.

More generally, our model provides new insights into how the information revolution could affect the microstructure of a variety of different markets. We expect that improvements in computing and communications technologies will tend to drive all three parameters δ, $\underline{k}$, and k_m toward zero. In this case we expect that most markets will ultimately be in a configuration near the origin in region 2 of figure 8. Our theoretical results predict that middlemen and market makers will coexist, with each handling approximately half of the total volume of trade.

References

Bakos, Yannis. "The Emerging Landscape for Retail E-Commerce." *J. Econ. Perspectives* 15 (Winter 2001): 69–80.

Battalio, Robert H. "Third Market Broker-Dealers: Cost Competitors or Cream Skimmers?" *J. Finance* 52 (March 1997): 341–52.

Battalio, Robert H.; Greene, Jason; and Jennings, Robert. "Do Competing Specialists and Preferencing Dealers Affect Market Quality?" *Rev. Financial Studies* 10 (Winter 1997): 969–93.

Baye, Michael R., and Morgan, John. "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets." *A.E.R.* 91 (June 2001): 454–74.

Bond Market Association. "eCommerce in the Fixed-Income Markets: The 2001 Review of Electronic Transaction Systems." Manuscript. Washington: Bond Market Assoc., December 2001.

Caillaud, Bernard, and Jullien, Bruno. "Chicken and Egg: Competing Matchmakers." Manuscript. Toulouse, France: Univ. Toulouse, 2001.

Easley, David; Kiefer, Nicholas M.; and O'Hara, Maureen. "Cream-Skimming or Profit Sharing? The Curious Role of Purchased Order Flow." *J. Finance* 51 (July 1996): 811–33.

Fleming, Michael. "Measuring Treasury Market Liquidity." Manuscript. New York: Fed. Reserve Bank, June 2001.

Fong, Kingsley; Madhavan, Ananth; and Swan, Peter. "Why Do Securities Markets Fragment?" Manuscript. Los Angeles: Univ. Southern California, November 1999.

Garman, Mark B. "Market Microstructure." *J. Financial Econ.* 3 (June 1976): 257–75.

Gehrig, Thomas. "Intermediation in Search Markets." *J. Econ. and Management Strategy* 2 (Spring 1993): 97–120.

Hendershott, Terrence, and Zhang, Jie. "A Model of Direct and Intermediated Sales." Manuscript. Berkeley: Univ. California, November 2001.

Lucking-Reiley, David, and Spulber, Daniel F. "Business-to-Business Electronic Commerce." *J. Econ. Perspectives* 15 (Winter 2001): 55–68.

Neeman, Zvika, and Vulkan, Nir. "Markets versus Negotiations: The Predominance of Centralized Markets." Working paper. Boston: Boston Univ., 2001.

O'Hara, Maureen. *Market Microstructure Theory.* Oxford: Blackwell, 1995.

Pirrong, Craig. "Third Markets and the Second Best." Manuscript. St. Louis: Washington Univ., November 2000.

Radford, R. A. "The Economic Organisation of a P.O.W. Camp." *Economica* 12 (November 1945): 189–201.

Rubinstein, Ariel, and Wolinsky, Asher. "Middlemen." *Q.J.E.* 102 (August 1987): 581–93.

Spulber, Daniel F. "Market Making by Price-Setting Firms." *Rev. Econ. Studies* 63 (October 1996): 559–80. (*a*)

———. "Market Microstructure and Intermediation." *J. Econ. Perspectives* 10 (Summer 1996): 135–52. (*b*)

Yanelle, Marie-Odile. "The Strategic Analysis of Intermediation." *European Econ. Rev.* 33 (March 1989): 294–301.

Yavas, Abdullah. "Marketmakers versus Matchmakers." *J. Financial Intermediation* 2 (March 1992): 33–58.

Zuckerman, Gregory; Davis, Ann; and McGee, Suzanne. "Before and After: Why Cantor Fitzgerald Can Never Re-create What It Once Was." *Wall Street J.* (October 26, 2001), p. A1.