WORKING PAPER SERIES

Market Structure and Inefficiency in the Foreign Exchange Market

Mark D. Flood

Working Paper 1991-001D
http://research.stlouisfed.org/wp/1991/91-001.pdf

PUBLISHED: Journal of International Money and Finance,
April 1994.

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
411 Locust Street
St. Louis, MO 63102

MARKET STRUCTURE AND INEFFICIENCY IN THE FOREIGN EXCHANGE MARKET

Mark D. Flood*

91-001D

First Draft: November, 1990
Revised: January, 1993

Federal Reserve Bank
P. O. Box 442
St. Louis, MO 63166
(314) 444-8584

All comments are very welcome.

* Economist, Federal Reserve Bank of St. Louis. I thank Mark Flannery, Mustafa Gültekin, Kalman Cohen, Richard McEnally, Harvey Wagner, James Burnham, Richard Lyons, David Hsieh, Stanley Black, two anonymous referees, and participants at an NBER conference on exchange rate regimes for many helpful comments. All remaining errors pertain to the author.

MARKET STRUCTURE AND INEFFICIENCY IN THE FOREIGN EXCHANGE MARKET

ABSTRACT:

This paper investigates the intradaily operational efficiency of the U. S. foreign exchange market by conducting computer simulation experiments with market structure (the numbers of market-makers, brokers and customers). The results indicate significant operational inefficiencies which can be explained by temporary inventory imbalances inherent in a decentralized market. The results also suggest that much of this inefficiency could be alleviated through a centralization of price information.

KEYWORDS:

Foreign exchange, industrial organization, market microstructure, market efficiency, simulation, experimental methodology.

JEL Classification:

C88, C99, F33, G14, L19

"It is by no means clear what object there could be in exchanging one bone for another."

- Adam Smith (1937), p. 13.

This paper addresses the short-term performance of the interbank foreign exchange market as a problem in industrial engineering. This study is motivated by the general question of the origin of microstructures. Why, for example, do foreign exchange futures contracts trade in an open-outcry pit system on organized exchanges, while foreign exchange forward contracts trade in the decentralized and unregulated interbank dealer market? While it seems unlikely, *prima facie*, that both microstructures are identically efficient in all respects, there may be differences between futures and forward contracts that make open-outcry pit trading optimal for the former and a decentralized dealer market best for the latter. On the other hand, there may be other factors, such as economies of scale or large fixed costs of transition, that impede institutional change even when the status quo is not globally optimal. Because of this possibility for obstacles to institutional change, it is not obvious that a given microstructure will always attain the theoretical optimum for a specific measure of efficiency.

In general, when a class of inefficiencies can be eliminated by individuals acting unilaterally in their own self-interest, we should expect such inefficiencies to be negligible; for example, arbitrage opportunities should be rare. On the other hand, when relative efficiency gains require the orchestration of multilateral action in the common interest, the implications are not so immediate; for example, the "long overdue reforms" of the London Stock Exchange's Big Bang were preceded by years of debate over the sometimes conflicting interests of investors, jobbers, brokers and regulators (*Euromoney*, 1986, p. 3). In the context of the U. S. foreign exchange market we address these microstructural issues indirectly, by examining market structure: the absolute and relative proportions of market-makers, brokers and customers constituting the market. Ultimately, we wish to learn whether

the market performs its immediate functions -- establishing a market-clearing price and matching buyers of currency with sellers -- in a cost-minimizing fashion.

A thorough understanding of the microstructure of the foreign exchange market is of interest at two levels. At a theoretical level, the combination of a fully decentralized interbank direct market with the quasi-centralized interbank brokered market composes a particular microstructure of considerable importance. It is found in various forms in the secondary market for U. S. government securities, in the federal funds market, as well as in the foreign exchange market. Stoll (1978) and Ho and Stoll (1983) examine the equilibrium number of dealers and dealer behavior in a fairly general setting. Garbade (1978) and Garbade and Silber (1976) consider the Treasuries market. Bossaerts and Hillion (1991) provide a valuable application of theoretical microstructural analysis to the foreign exchange market. Flood (1991) offers an overview of microstructural theories in the context of the foreign exchange market.

At a practical level, a thorough microstructural understanding can offer insights into the operational efficiency of the foreign exchange market.[1] Daily trading volume in the U. S. market alone was $128.9 billion in April 1989. Annualizing, this rounds out at $32.4 trillion per year, or roughly twice world GDP.[2] More significantly, 95 percent of spot volume involved the shifting of funds among commercial bank market-makers, while only 5 percent was devoted to customer transactions.[3] The literature has produced at least two explanations for this concentration of trading in the interbank market. Frankel and Froot (1990), for example, suggest that commercial banks engage actively in exchange rate speculation.[4] Burnham (1991) concludes instead that this topheavy distribution of trading volume largely results from decentralization of exchange. The simulations here support the latter view.

Numerous recent articles provide empirical evidence on the intradaily behavior of exchange rates. The basic issues that emerge revolve around the statistical distribution of prices: the unconditional distribution of returns has excess kurtosis relative to the normal, volatility is clustered in time, and returns appear to exhibit negative serial correlation. Wasserfallen and

Zimmerman (1985), using Swiss franc (CHF) data from nine days in 1978-80 in the Swiss market, analyze returns over intervals ranging from one to ten minutes, and find evidence of various non-normalities, including skewness and excess kurtosis. They suggest that time-varying parameters may be to blame. Wasserfallen (1989), with CHF returns over five-minute intervals for 231 days in 1983 in the Swiss market, finds evidence of skewness and excess kurtosis, but little evidence of negative serial correlation; he also implicates time-varying parameters. He observes that the spread (in this case for Union Bank of Switzerland) varies positively, although infrequently, with price volatility.

Baillie and Bollerslev (1990) address one aspect of the time-varying parameters issue -- volatility clustering -- directly. Using round-the-clock hourly data for four currencies (GBP, DEM, CHF, and JPY), they find that a seasonal GARCH specification with dummies for hours of the day and for days following vacations fits the data well. Periods of high volatility within the day tend to correspond to opening of trading in the world's major market centers, especially London and New York. They suggest (p. 578) that the intradaily pattern of volatility clustering may be caused by trading volume.[5] Bollerslev and Domowitz (1992) confirm the presence of volatility clustering in a sample of round-the-clock, screen-based DEM rates at five-minute intervals over nearly three months in 1989. Goodhart and Demos (1990) describe in detail patterns in Reuters submission rates in this data set.

Goodhart and Figliuoli (1991), analyzing a set of screen-based DEM rates at frequencies as high as once per minute from three days in 1987 in the London market, find statistically significant negative serial correlation in high-frequency ask prices. They also observe that dealer spreads in their sample are largely unaffected by market conditions.[6] Bollerslev and Domowitz similarly observe that trading activity has a negligible effect on the conditional mean spread size; they also note cross-sectional differences in spread-setting behavior between large and small institutions. In contrast, Melvin and Tan (1992), using daily South African Rand rates, conclude that spreads are indeed sensitive to market conditions; in particular, instances of social unrest are associated with wider spreads. In the model used here,

spreads are affected by inventory; thus, while trading activity per se does not affect the spread, anything that disrupts inventory, such as a news event, will also affect spreads.

In a series of papers, Ito (1987), Ito and Roley (1987, 1988), Engle, Ito and Lin (1990), and Ito, Engle and Lin (1992) also examine the effect of news on the behavior of exchange rates intraday. Each of these five papers uses some subset of a series of JPY rates recorded five times daily in the New York and Tokyo markets between 1979 and 1988. Hakkio and Pearce (1985) use a similar data set, but consider seven bilateral exchange rates while restricting their attention to the New York market. The evidence of Ito and his various co-authors generally supports the "meteor showers" hypothesis, in which volatility afflicts markets in different countries sequentially as the earth turns, while rejecting the "heat wave" hypothesis, in which turbulence is restricted to the same country on successive days. The simulation model here restricts attention to intradaily trading in the New York market, and thus does not directly address global patterns in volatility clustering.

The paper proceeds as follows. Section I outlines a detailed model of the U. S. foreign exchange market, involving heterogeneous market-makers, brokers and customers. Appendix B provides both a brief defense of the simulation methodology and a description of the model's validation against real-world data. In section II, five measures of market inefficiency are defined. The experimental technique employed, described in section III, is a full-factorial response surface design over four experimental variables: the number of market-makers, the number of brokers, the number of customers, and the rate of unanticipated news arrival. The response variables are the five inefficiency measures. Section IV presents the experimental results, which reveal a significant and complex relationship between market inefficiency and market composition. Much of the inefficiency present can be explained by temporary inventory imbalances resulting from the decentralized nature of the market.

I. An Outline of the Model

The market consists of a set of foreign exchange traders connected by computer and telephone network.[7] The market combines two distinct modes of exchange. The customer and interbank direct markets are decentralized, continuous, open-bid, double-auction markets. The brokered market is a quasi-centralized, continuous, limit-book, single-auction market. The temporal setting for the simulations is intraday; the trading day is approximated as a sequence of discrete time periods, subscripted $t \in \{1,\ldots,T\}$. There are three types of participants in the model: market-makers, brokers, and customers. Market-makers, subscripted $m \in \{1,\ldots,M\}$, represent dealers for commercial and investment banks, trading for their own accounts and subject to in-house position constraints. Brokers, subscripted $b \in \{1,\ldots,B\}$, collect limit orders from the market-makers into private limit order books, in turn making a market by acting as a liaison for those same market-makers; brokers do not trade for their own account. Customers, subscripted $c \in \{1,\ldots,C\}$, represent companies and individuals trading to satisfy an exogenous demand for foreign exchange (export financing, for example). The subscripts $i,j \in \{1,\ldots,N\}$ denote arbitrary market participants ($N \equiv M + B + C$). The market-makers, brokers and customers together compose the basis for the set of price information. The agents are allowed to interact in the following four ways:

(1) Market-maker calls market-maker
(2) Market-maker calls broker
(3) Broker calls market-maker
(4) Customer calls market-maker

News events occur throughout the day at randomly determined times. There are two types of news, or fundamental information: anticipated and unanticipated events, arriving according to Poisson processes with parameters Ω_a and Ω_u, respectively. This dichotomy captures the disruptive impact of news on expectations and the propensity of market-makers to assume speculative currency inventories in anticipation of future events. For anticipated events (e. g., a congressional vote or a weekly money supply announcement), the timing of the news arrival is known in advance, thus providing an opportunity for speculation. Market-makers know that change impends, and they hold idiosyncratic beliefs about the post-event state of the world. If current

prices deviate from a market-maker's beliefs about post-event prices, he believes he can profit from this discrepancy, and he acts on this belief by taking a speculative inventory position. Upon the arrival of news, whether anticipated or not, all customers and market-makers are assigned new prior price expectations, and market-makers establish new desired inventories.

A. Market-makers

Market-makers are central to the model, because they are involved in every transaction, and because interdealer trades (direct or brokered) compose the overwhelming majority of activity in the market. Market-makers are assumed to try to call out each period (it is preferable for a market-maker to call out, because calling out provides more price information and control over inventory than does receiving a call and making a market). Incoming calls take precedence, however, so that if a market-maker (or broker) receives a call, he is preempted and cannot call out in that period. The decision regarding whom to call each period and whether the call goes through is modeled as a random matching of agents, subject to the constraint that each pairing must be one of the four allowable types listed above. Given a specific pairing, the market-maker must decide on price and/or quantity. If he receives a call from a customer or market-maker, he makes a market by quoting bid and ask prices. If he receives a call from a broker, he decides on the price and quantity of a limit order. If he places a call to a broker or market-maker, he must decide whether and how much to buy or sell. All such contacts reveal information to the market-maker.

The market-maker's pricing decision must be consistent with the realities of the marketplace. In particular, the presence of other well-informed and well-capitalized market-makers implies an imperative of arbitrage avoidance: market-makers must strive first for a price consensus (i. e., overlapping spreads). A quoted spread that is "off the market" serves only to saddle the market-maker with a large inventory that could have been had at a better price. The importance of consensus is conveyed in the words of one market-maker:[8]

> "Ninety percent of what we do is based on perception. It doesn't matter if that perception is right or wrong or real. It only matters that other people in the market believe it. I may know it's crazy. I may think it's wrong. But I lose my shirt by ignoring it. ... I can't afford to be five steps ahead of everybody else in the market. That's suicide."

Accordingly, if a market-maker believes that the current price quotes of other market-makers do not adequately reflect fundamental information, then attempts to profit from this belief should take the form of speculative inventory positions, not off-market quotes.

Market-maker m maintains an appraisal function, A_m, summarizing all her prior beliefs and information on the current status of the market; it is her assessment of the current market consensus price. In establishing an appraisal, she does not consider fundamentals directly, but rather only as she expects fundamentals to alter the quotes of other traders. In addition, she has a preferred inventory position, $I^*_{mt} \equiv I^*_m(\Theta_t)$, that reflects her speculative beliefs -- based on fundamental information, Θ_t -- about future movements in the market consensus valuation. Thus, a market-maker's behavior is determined first by a desire to give a bid-ask quote that will avoid arbitrage, and, secondly, by a desire to profit through speculation from general movements in the exchange rate. More formally, the relations determining market-maker m's bid-ask quote, $Q_{mt} \equiv (\beta_{mt}, \alpha_{mt})$, can be expressed as a function, g_m, of the appraisal and of the deviation of inventory from its preferred level, $I_{mt} - I^*_{mt}$:

$$Q_{mt} = g_m[A_m(Q|\Theta_t, Q^*_{mt}), I_{mt} - I^*_m(\Theta_t)] \ , \tag{1}$$

where Q is a TxN matix of past, current and future prices, and Q^*_{mt} is that subset of Q that has been observed by market-maker m up throught time t. Two alternative specifications (dubbed "cautious" and "incautious") of $g_m(\cdot)$ are used, in order to check the robustness of the results.[9] Each market-maker's appraisal is endogenous in the sense that it incorporates price information as it is observed. The special characteristics of incautious and cautious market-makers are described next.

(1) Incautious -- In determining her appraisal, A_m, the incautious market-maker distills all of her information down to a single best guess about the relative values of the two currencies. The mean of her posterior marginal distribution is: $E^{n}_{S}(\mu|k'',v'',n'',w'') = k''$, and her appraisal is: $A_m = \exp(k'')$

$= \exp[(n'k'+nk)/(n'+n)]$; this is the price which, if quoted, would equalize the subjective probabilities of buying and selling.[10] If a call comes in from another market-maker or a customer, the quoted spread is this appraisal, adjusted for inventory considerations. Specifically, given a desired inventory level and a tick size, z, for the market, the spread for a market-maker within one round lot of her desired position is one tick wide and set such that her appraisal falls within the spread. The spread widens as the absolute difference between actual and desired inventories increases, and bid and ask quotes are shaded up (down) to alleviate a shortfall (excess) of actual inventories relative to desired, as described in table 1. **[TABLE 1 APPROXIMATELY HERE]** β° and α° are determined by rounding the appraisal down, and up respectively, to the nearest tick.[11]

Desired inventories depend only on fundamental information, and not on observed price information. Following each news arrival, desired inventories are drawn anew at random for each market-maker. Desired inventories are subject to banks' in-house position constraints, both daylight and overnight. The overnight constraint is assumed to be zero for all market-makers, implying a desired inventory of zero following the final anticipated event of the day. The daylight position constraint, P, is the upper bound on the absolute value of the desired inventory for all market-makers. For the incautious market-maker, desired inventories are drawn from the uniform distribution: $I_m^*(\Theta_t) \sim U\{-P,\ldots,P\}$. It should be emphasized that there is no correlation between the new prior expectations parameters and the new desired inventory. Although speculation implies that market-makers believe their counterparts have misestimated the ultimate impact of news on prices, a correlation between new priors and new desired inventories would imply that market-makers believe that this misestimation occurs in a consistent and predictable (on average) manner. There is no reason to believe that rational expectations are violated in this way. For the same reason, there is no serial correlation in the random draws of desired inventory levels, which track a market-maker's speculative beliefs about the direction in which the new market consensus price will move.[12]

(2) Cautious -- The cautious market-maker uses more information from his

subjective price distribution in establishing an appraisal. Instead of setting the desired-inventory spread at the one-tick minimum, the market-maker who is at his desired inventory position sets his (minimum) spread to enclose some proportion, δ, of his subjective distribution centered around the median. Formally, market-maker m's bid and ask prices at a balanced inventory level (i. e., $I_{mt}=I^*_{mt}$) are set around the mean of his subjective log-transformed marginal price distribution as follows:

$$\beta^\circ_{mt} = \max_k \left\{ kz\colon F_{SI1}[ln(kz)] \le (1-\delta)/2 \right\} , \tag{2}$$

and:

$$\alpha^\circ_{mt} = \min_k \left\{ kz\colon F_{SI1}[ln(kz)] \ge (1-\delta)/2 \right\} , \tag{3}$$

where k is a positive integer, and $F_{SI1}(\cdot)$ is the incomplete first moment of his subjective (general Student normalized) distribution. Adjustments to this minimum spread are then made, as described in table 1, to account for deviations of inventory from its desired level, where α° and β° are now given by equations (2) and (3). For cautious market-makers, new desired inventory levels are generated as deviates from the "triangular" distribution: $I^*_m(\Theta_t) \sim \Delta\{-P,\ldots,P\}$.[13]

B. Brokers and Customers

Brokers operate according to a relatively simple algorithm. Because brokers do not take positions, profiting instead only through brokerage commissions, it is unnecessary to consider their beliefs or expectations. Rather, brokers single-mindedly pursue the accumulation of orders for their limit order books. New limit orders can cross against existing limit orders on any brokerage book; if they do not cross a book, they are posted on the book of the broker requesting the order. When this pursuit is interrupted by an incoming call from a market-maker, the broker quotes him the inside spread from her own book, along with the amounts available at each price. The decision to buy, sell, or pass is then up to the market-maker. If a trade is concluded, accounting ensues to adjust the limit order book and the market-makers' inventory positions, and to assess the brokerage fees. A market-maker

treats the information in a broker's quotation in the same way as he does a quote from another market-maker.[14]

A limit order remains on a brokerage book until it is hit, until news arrives, or until it goes stale. Upon the arrival of news, all existing limit orders are removed from the brokerage books. Staleness is included in the model to approximate the fact that market-makers do not leave limit orders standing indefinitely, because of the tendency of the market consensus price to shift. The number of brokers is separate from the number of brokerage books. Each broker is assigned to a brokerage firm with a single book. Adding brokers to the market increases the number of brokers canvassing the market-makers for limit orders to be entered into a fixed number of books.

The behavior of customers is exogenous to the model in the sense that the behavior of other agents does not affect customers' behavior. In other words, customers consider fundamentals only, being preoccupied with real trade and not being concerned with short-term exchange market dynamics. Thus, quoted exchange rates do not alter customers' beliefs about the relative value of two currencies.[15] Customers provide an exogenous order flow to the market-makers by contacting market-makers and submitting market orders when the quoted bid (ask) price is above (below) the customer's reservation valuation. The number of customers, C, thus affects the rate of order flow. Customers' reservation valuations are assigned from a random draw following each news arrival. Each customer in the model is active in the market in every period. This occurs without loss of generality, because customers in the simulation model are interchangeable: their expecations are unaffected by the prices they observe in the market.

II. Measures of Inefficiency

Market efficiency has several aspects. The measures defined here assess the market's ability to distribute decentralized price information, to allocate inventories, as well as to exhibit price efficiency. The are real costs associated with these inefficiencies. Misallocations of currency inventories generally arise when orders fail to go to the best-price market-

maker, or when arbitrage opportunities or predictable price patterns are exploited. Operational opportunity costs (e. g., back-office costs) are incurred when, *ceteris paribus*, the number of transactions involved in a reallocation of currency exceeds the minimum number of transactions required.

The law of one price dictates that simultaneous price quotes should coincide: that is, all bid-ask spreads should overlap. If they do not, the violation of the law clearly counts as an inefficiency. This is reflected in the first three measures, which assess the consistency of decentralized prices. The fourth measure assesses the allocational efficiency of the market. Finally, weak-form price efficiency implies certain statistical relationships which are independent of the specification of a "true price". The most notable of these is the requirement that a series of prices from an efficient market not be serially correlated, embodied in the fifth inefficiency measure. All five measures, Λ_k (k=1,...,5), have been aggregated over the trading day, so that intraday seasonalities do not affect their interpretation. All the measures are normalized to allow comparisons between markets of different sizes, all are non-negative, and all are equal to zero in a perfectly efficient market.

(1) Arbitrage opportunities -- An "orderly market" is defined here as a market in which all open bid quotes are strictly less than all ask prices at a given point in time. It follows that the presence of an arbitrage opportunity implies a disorderly market. If arbitrage opportunities exist, it implies that markets are failing to transmit all available price information instantaneously.[16] The extent of arbitrage opportunities is measured by the number of arbitrage opportunities available, aggregated over the trading day and normalized by the number of market-makers:

$$\Lambda_1 \equiv \frac{1}{M} \sum_{t=1}^{T} \Lambda_{1t} \equiv \frac{1}{M} \sum_{t=1}^{T} \left\{ K\{\beta_{jt}: \beta_{jt} > \min_i(\alpha_{it})\} + K\{\alpha_{jt}: \alpha_{jt} < \max_i(\beta_{it})\} \right\} \qquad (4)$$

where $K\{\beta_{jt}\}$ is the cardinality of the set of bids in period t, and $K\{\alpha_{jt}\}$ is the cardinality of the set of asks.

(2) Price dispersion -- A more subtle inefficiency related to arbitrage

opportunities is measured by price dispersion. If a market-maker does not quote the inside (i. e., market) spread, but is nonetheless able to conclude deals at those quotes, then there is an informational inefficiency, because the market-maker's counterparties are not finding the best available quotes. Λ_2 measures the total volume of deals concluded away from the inside spread in a single time period; this inefficient volume is then normalized by the total volume for the day, to facilitate comparisons between markets of different sizes:

$$\Lambda_2 \equiv \frac{1}{V} \sum_{t=1}^{T} \sum_{m \in H_t} \left\{ \{q_{mt}: q_{mt}>0,\ \beta_{mt}<\max_i(\beta_{it})\} - \{q_{mt}: q_{mt}<0,\ \alpha_{mt}>\min_i(\alpha_{it})\} \right\} \tag{5}$$

where V is the total volume for the day:

$$V \equiv \sum_{t=1}^{T} \sum_{m \in H_t} |q_{mt}| \ , \tag{6}$$

and where H_t is the set of market-makers whose quotes or limit orders were hit in period t, and q_{mt} is the quantity transacted by market-maker m in period t (positive for a purchase, negative for a sale). Note that it is only necessary to consider market-makers, since they are the only agents providing specific prices to the market.

(3) Adjustment interval -- A measure of dynamic efficiency related to both the arbitrage/dispersion measures just defined, and to semi-strong-form informational efficiency, is the amount of time it takes the market to converge to price consensus from a disequilibrium starting point, where disequilibrium is defined as the presence of arbitrage opportunities. In evaluating the adjustment interval, no further news arrivals are allowed after an initial shock to the model. This presumes that the price adjustment process in the market is stable, so that convergence is assured. Thus, the length of time until the elimination of the last arbitrage opportunity is measured:

$$\Lambda_3 \equiv \min_t \left\{ t: \Lambda_{1t+k}=0 \ \forall\ k \in \{0,1,\ldots,T-t\} \right\}, \tag{7}$$

where Λ_{1t} is measure one at period t, defined in equation (4).

(4) Misallocation -- To measure suboptimal portfolio allocation as an

inefficiency, one must either identify the optimum, or find some other secondary characteristic (such as the equation of marginal rates of return across investors) that marks allocational efficiency. The simulation model readily provides the optimum, namely the point at which all market-makers hold their desired inventories.[17] Suboptimality is measured as the distance (under the absolute norm) from this optimum:

$$\Lambda_4 \equiv \frac{1}{M} \sum_{t=1}^{T} \sum_{m=1}^{M} |I_{mt} - I^*_{mt}| \tag{8}$$

where I_{mt} is market-maker m's actual position at time t, and I is her desired position.

(5) Serial correlation -- Predictable intertemporal price patterns generally reveal an inefficiency by the standard rational expectations argument.[18] For present purposes, predictable patterns are at least a violation of the fair-game properties of a weak-form efficient market. Depending on the extent of the pattern, it may also be sufficient to overwhelm spreads and transaction costs and provide profitable trading opportunities. First-order autocorrelations in the quotes of market-makers and brokers are used here to measure the extent of such patterns.[19] One source of such correlations might be price shifts by market-makers to compensate for inventory imbalances induced by having previous quotes hit (sometimes called "inventory bounce").

The presence of price patterns is measured by aggregating the serial correlation of quotes made by market-makers and brokers:

$$\Lambda_5 \equiv \frac{1}{M+B} \sum_{j=1}^{M+B} \left| \frac{\frac{1}{T-2} \sum_{t=3}^{T} \left[(p_{jt} - p_{jt-1}) \cdot (p_{jt-1} - p_{jt-2}) \right]}{\frac{1}{T-1} \sum_{t=2}^{T} \left[p_{jt} - p_{jt-1} \right]^2} \right| , \tag{9}$$

where

$$p_{jt} \equiv \frac{ln(\alpha_{jt}) + ln(\beta_{jt})}{2} , \tag{10}$$

and M and B are the numbers of market-makers and brokers, respectively, in the market. Λ_5 is simply the aggregation (over agents) of first-order autocorrelations in changes in the mean log-transformed quotes of market-

makers and brokers. Absolute values are taken so that positive and negative correlations do not cancel in the aggregation. The aggregate is normalized by the number of agents, M+B, to facilitate comparisons between markets of different sizes. The measure is designed to ignore bid-ask bounce while capturing inventory bounce.

III. Experimental design

At this point, an acceptable computer model of the foreign exchange market and a set of inefficiency measures are taken as given. The next step is to see how the market performs under circumstances that represent deviations from a base case, defined as the status quo in the real foreign exchange market. To approximate the status quo in the real foreign exchange market, the computer model was subjected to a systematic validation. In addition, the two different behavioral specifications for market-makers (cautious and incautious) are used to gauge the robustness of the model to such differences.

Given the validated model, a randomized full-factorial experimental response-surface design over the four experimental variables: M, B, C, and Ω_u is used.[20] In the response-surface design, the simulation model is treated as a black box that takes a vector of variables as inputs, and generates the five measures of inefficiency as outputs. The inputs are the four experimental variables. The aim is to predict the values of the inefficiency measures, given the values of the experimental variables. In practical terms, this entails fitting five regression equations -- the response surfaces -- with the experimental variables appearing as independent variables in the regressions. The regression coefficients and the associated covariance matrix contain the information of interest.

In particular, the program is assumed to be a smooth vector-valued function of four variables. This function is not explicitly defined, but rather is approximated by a second-order Taylor series in the experimental variables:

$$\Lambda_k^r = \gamma_{0k} + \gamma_{mk}M^r + \gamma_{bk}B^r + \gamma_{ck}C^r + \gamma_{uk}\Omega_u^r + \gamma_{mmk}M^rM^r + \gamma_{bbk}B^rB^r + \gamma_{cck}C^rC^r + \gamma_{uuk}\Omega_u^r\Omega_u^r + \gamma_{mbk}M^rB^r + \gamma_{mck}M^rC^r + \gamma_{muk}M^r\Omega_u^r + \gamma_{bck}B^rC^r + \gamma_{buk}B^r\Omega_u^r + \gamma_{cuk}C^r\Omega_u^r + \epsilon_k^r , \tag{11}$$

where Λ_k^r is the value of the kth inefficiency measure, $k \in \{1,2,4,5\}$, from the rth simulation run, $\gamma_k \equiv (\gamma_{0k}, \gamma_{mk}, \ldots, \gamma_{cuk})$ is the vector of polynomial coefficients for the approximating function, ϵ_k^r is a random disturbance for the rth simulation run, and M^r, B^r, C^r and Ω_u^r are the values of the experimental variables from the rth simulation run.[21]

To measure the adjustment interval, the rates of anticipated and unanticipated news arrival must be set to zero, because the measure would otherwise be almost completely determined by the timing of the last news arrival rather than by the inherent efficiency of the market structure. Thus, the approximating function for the adjustment interval becomes:

$$\Lambda_3^r = \gamma_{03} + \gamma_{m3}M^r + \gamma_{b3}B^r + \gamma_{c3}C^r + \gamma_{mm3}M^rM^r + \gamma_{bb3}B^rB^r + \gamma_{cc3}C^rC^r + \gamma_{mb3}M^rB^r + \gamma_{mc3}M^rC^r + \gamma_{bc3}B^rC^r + \epsilon_3^r , \tag{12}$$

where the terms are defined analogously to those above. The absence of news arrival times also means that this measure assesses only the market's ability to process pure price information, in the absence of fundamental information.

The approximating functions are estimated by generating 300 data points for each of the inefficiency measures, simultaneously varying the experimental variables at random over a convex region. In particular, the values of the experimental variables are drawn from uniform distributions over the ranges shown in table 2. **[TABLE 2 APPROXIMATELY HERE]** For four of the inefficiency measures (i. e., all but Λ_3, the adjustment interval), the same set of simulation runs is used; the inefficiency of the market is measured four different ways at the end of each day. For the adjustment interval, all parameters are set at their validated levels, except for the news arrival rates Ω_u and Ω_a, which are set to zero, and the three experimental variables, M, B, and C, which vary uniformly over the ranges in table 2. Data are generated for both the incautious and cautious models. OLS regressions on the data for measures Λ_k, $k \in \{1,2,4,5\}$, produce the coefficients $\gamma_k \equiv (\gamma_{0k}, \gamma_{mk}, \ldots, \gamma_{cuk})$ in the corresponding columns of tables 3 and 4. **[TABLE 3**

APPROXIMATELY HERE] [TABLE 4 APPROXIMATELY HERE] For the third measure, Tobit regressions are required, because, in many cases, prices failed to converge to the exact no-arbitrage condition within 1000 trading periods. The regression coefficients for Λ_3 presented in tables 3 and 4 are the non-normalized coefficients of the underlying model.

The question of interest is now the impact of a change in the experimental variables on each of the inefficiency measures. Although this can be gauged to some extent by the magnitude and significance of individual regression coefficients, it is more meaningful to consider directly the partial derivatives of the inefficiency measures with respect to the experimental variables. Thus, for example, the impact of the number of market-makers on the number of arbitrage opportunities is reckoned as:

$$\frac{\partial E(\Lambda_1)}{\partial M} = \hat{\gamma}_{m1} + 2\overline{M}\hat{\gamma}_{mm1} + \overline{B}\hat{\gamma}_{mb1} + \overline{C}\hat{\gamma}_{mc1} + \overline{\Omega}_u\hat{\gamma}_{mu1} \quad , \tag{13}$$

where the bars over the experimental variables indicate that these are the particular values of those variables at which the partial derivative is evaluated, and the hats over the coefficients indicate that these are the fitted values from the regression. The statistical significance of this relationship is measured as the probability of attaining a particular value for the partial derivative under the null hypothesis that the experimental variables have no impact on the measures of inefficiency. Specifically, F-statistics for these linear restrictions on the regression coefficients are calculated from the variance-covariance matrix under the maintained assumption of normality in the error terms. Similarly, the impact of experimental variables in the Tobit regressions on Λ_3 are reckoned as partial derivatives. The relationship of interest here is the partial derivative of the expectation of the underlying (i. e., uncensored) inefficiency measure with respect to a change in an experimental variable.[22] For example, the partial derivative of the expected adjustment interval, $E(\Lambda_3)$, with respect to the number of market-makers is simply:

$$\frac{\partial E(\Lambda_3)}{\partial M} = \hat{\gamma}_{m3} + 2\overline{M}\hat{\gamma}_{mm3} + \overline{B}\hat{\gamma}_{mb3} + \overline{C}\hat{\gamma}_{mc3} \ . \tag{14}$$

The statistical significance of such deviations is calculated from the

variance-covariance matrix as an F-statistic on the linear restriction as before.

These partial derivatives are evaluated at two different values of the vector of experimental variables, to ensure the robustness of the conclusions. The two points chosen are the population mean from the experimental region given in table 2, and the status quo levels of the experimental variables, given by their validated levels. The results of these tests of the direction and significance of the impact of the experimental variables on the measures of inefficiency are presented in tables 5 and 6.

IV. Results of the experiments

A. The issues

Before considering the experimental results, it is useful to review the issues that make them relevant. The first three experimental variables (M, B and C) could imply normative policy implications. If there were some market composition minimizing the inefficiency measures, and the market is not currently achieving that optimum, then it would be in the general interest exogeneously to impose barriers to entry or exit, to force the market toward the optimal composition. Such a direct approach falters here on the standard vector optimization problem: none of the four experimental variables is uniformly good or bad across all five inefficiency measures. Establishing an objective function for inefficiency minimization would require some implicit or explicit prioritization of the different types of inefficiency. Such a weighting is not attempted here. A more holistic analysis of the results proves fruitful, however. It is suggested that a single microstructural fact -- the decentralized nature of the market -- accounts for most of the significant effects reported here.

It is not clear *a priori* whether the market should be larger, smaller, or remain at the current size. The traditional argument is for centralization of exchange; at the extreme, it implies natural monopolies for market-makers and brokers. The opposite stand, that interdealer competition promotes

efficiency, is also defensible.[23] Finally, a resort to naive institutional Darwinism argues that the status quo must be best. The purpose of experimenting with M, B and C is to find out which of these arguments is correct. The sources of disagreement are revealed on closer examination. Increasing the number of market-makers obviates the possibility of monopoly profits while simultaneously increasing the market's ability to process information, since there are more individuals with differing perspectives attempting to evaluate the same commodity. On the other hand, the decentralized nature of the market implies that price dispersion and arbitrage opportunities should increase with the number of market-makers providing quotes. Similarly, increased customer order flows may tend to stabilize the market, by forcing exchange rates into line with some real-trade valuation, or they may tend to roil the market, as they effect shocks to individual market-makers' inventories. Finally, more brokers should tend to centralize price information and thus reduce search costs, but, because brokered limit orders do not expire immediately, the same increase in limit orders may induce price dynamics which deviate from the martingale model.

The rate of unanticipated news arrival, Ω_u, differs from the other experimental variables in that it is not a factor in market size or composition. It is included to gauge the market's ability to absorb and process new information. Engle, Ito and Lin (1990), for example, suggest that "traders with heterogeneous priors and private information may take some hours of trading, after a shock, to have expectational differences resolved" (pp. 525-6). Although it is unclear that a resolution of expectational differences, in the sense of a homogeneous posterior distribution over future prices, is in any way desirable, a homogeneous subjective distribution over current prices is a practical prerequisite for market efficiency. That is, arbitrage opportunities are inconsistent with efficiency.

Tables 5 and 6 reveal some interesting general characteristics of the market. **[TABLE 5 APPROXIMATELY HERE]** **[TABLE 6 APPROXIMATELY HERE]** Considering only the statistically significant partial derivatives, none of the experimental variables is uniformly good or bad across all five inefficiency measures.[24] A comforting result is the consistency of signs

and magnitudes across the tables. With a few exceptions, partial derivatives which are significantly positive (or negative) for one behavioral specification and evaluation point exhibit the same property under the other three arrangements; in no case is a partial derivative significantly positive in one table and significantly negative in another. Thus, the inefficiency results are robust across behavioral specifications. It also becomes convenient to summarize the signs on the partial derivatives, as in table 7. **[TABLE 7 APPROXIMATELY HERE]** The impacts of the experimental factors on the adjustment interval, Λ_3, are generally less significant than the other results, a fact that is at least partly explained by the nature of the censored sampling process: *ceteris paribus*, a censored sample must contain less information than an uncensored sample.[25]

B. The impact of market-makers

The impact of the number of market-makers is mixed. Increasing the number of market-makers increases the number of arbitrage opportunities. This is not attributable directly to the relative proliferation of quotes associated with an increase in the number of market-makers, because Λ_1 has been normalized by the number of market-makers. Instead, this measure is assessing a second-order effect, as additional market-makers create more price confusion than they resolve at the margin. Interestingly, although additional market-makers increase the per-dealer instance of arbitrage opportunities, they also tend to achieve a price consensus more rapidly (Λ_3), although the statistical significance is low. At the same time, additional market-makers tend to reduce the degree of price dispersion. Recalling the definition of Λ_2, this means that an increase in the number of market-makers makes it more likely that a randomly selected transaction will take place at the market price. In other words, as more market-makers enter the market, they drive up volume more quickly than they increase the number of suboptimal transactions. This can be explained in terms of greater liquidity: as the number of market-makers increases, the less likely it is that any one of them has an inventory discrepancy in a given period, because it is easier for them to find a market-

making counterparty to resolve inventory imbalances. Thus, the less likely it is that a market-maker will transact at off-market rates simply to resolve inventory imbalances. In other words, the more nearly allocational efficiency is achieved, the more willing are market-makers to shop for better prices. This explanation is also consistent with the impact of market-makers on the measure of suboptimal allocation; market-makers improve the allocation of inventories. Thus, the negative impact of market-makers on the degree of suboptimal allocation indicates that market-makers do indeed create liquidity. Lastly, liquidity, along with its implicit improved allocation of inventories, can be held responsible for the reduction in serial correlations associated with an increase in the number of market-makers: as increasing liquidity associated with additional market-makers smooths out the average misallocation of inventory, "inventory bounce" in prices tends to abate.

Already these results reveal some of the intricacy of the relationship between market structure, microstructure and inefficiency in the foreign exchange market. In particular, the significant impacts of the number of market-makers on inefficiency are consistent with a single causal explanation of market activity. This fact is alluded to by Burnham (1991), who notes that, when a market-maker is hit with an undesired inventory position:

> "[The market-maker] now seeks to restore its own equilibrium by going to another marketmaker or the broker market for a two-way price. A game of 'hot potato' has begun. ... It is this search process for a counterparty who is willing to accept a new currency position that accounts for a good deal of the volume in the foreign exchange market" (p. 135).

Thus, the large volume of interbank trading is not primarily speculative in nature, but rather represents the tedious task of passing undesired positions along until they happen upon a market-maker whose inventory discrepancy they neutralize. This same influence is seen in many of the other significant experimental effects.

C. The impact of brokers

Brokerage is beneficial inasmuch as it reduces arbitrage opportunities through the centralization of prices. At the same time, however, brokerage tends to exacerbate price dispersion. These two seemingly conflicting results

are related, however, because of the way arbitrage opportunities are counted in Λ_1. If a market-maker submits to a broker a bid (ask) quote above (below) the market ask (bid), two things may happen; either this would-be arbitrage opportunity crosses the broker's book, or it does not. If it crosses, it transacts immediately with an opposing limit order on the book, and it is not counted as an arbitrage opportunity in Λ_1; however, it may take out an off-market price on the other side of the book, adding to Λ_2. If it does not cross, it goes onto the book and remains an arbitrage opportunity in all subsequent periods until either the market shifts, or it is removed from the book by news, staleness, or transaction. There is a tendency for brokered limit orders to be away from the market spread, because limit orders do not change when the market spread moves. The level of brokerage does not have a statistically significant effect on Λ_3, the length of the adjustment to a price consensus. This may be due in part to the lower power of the Tobit results. Brokerage does tend to improve the allocation of inventories, Λ_4, although its impact is only mildly statistically significant. Adding brokers increases the level of serial correlation. This can be attributed to the presence of aging limit orders, which tend to make brokered prices mean-reverting as older quotes are removed. In sum, there is a trade-off involved in brokerage: it reduces arbitrage and improves the allocation of inventories through a partial centralization of price information, but it introduces serial correlation and price dispersion as limit orders held on the books are left behind by shifts in the market spread.[26]

D. The impact of customers

Customers are largely detrimental to market efficiency (in the same way that students prevent a university from running smoothly).[27] The expected level of arbitrage opportunities, the adjustment interval, and suboptimal allocation are all significantly increased by an increase in the number of customers active in the market. Even the significant improvement in price dispersion results from an increase in the normalization factor, total market volume, rather than any decrease in transactions occurring away from the

market spread. To the contrary, such transactions significantly increase in absolute terms with the number of customers. In hindsight, this is not too surprising. Again here, there is an identifiable causal link between customer orders and inefficiency. When customers transact, market-maker inventory disruptions result immediately, increasing Λ_4. A proximate effect of these inventory disruptions is higher volume, implying lower levels of price dispersion per unit volume, as measured by Λ_2. Another implication is price shading, which yields both more arbitrage opportunities, Λ_1, and consequently longer adjustment intervals, Λ_3. Whereas additional market-makers tend simultaneously to raise the number of arbitrage opportunities and lower the adjustment interval, the same effect cannot hold for customers. Additional market-makers actively work toward a price consensus, thus lowering Λ_3, but additional customers do not: there is no feedback from market activity into customer appraisals. The "hot potato" story is thus consistent with all four of these impacts. The expected level of serial correlation tends to decline with increases in customer participation, but this can be attributed to an increase in the variance in returns, rather than a decrease in serial covariance. In other words, customers make the price series noisier without bringing it any closer to the martingale property.

E. The impact of news arrival

The rate of unanticipated news arrival has a mixed effect on the market. While one might expect news to have a disruptive effect, and while this is in fact the case for Λ_1 and Λ_5, news is not uniformly disruptive. Although news significantly increases the expected number of arbitrage opportunities, there is no matching increase in price dispersion. The mild negative impact of news on price dispersion implies that as news arrives more frequently, total trading volume increases more rapidly than off-market volume. In comparing the degree of price dispersion to the number arbitrage opportunities, note that, as defined here, even the minimum level of price dispersion (Λ_2=0) is consistent with any number of arbitrage opportunities; as long as an arbitrageur acts on the best opportunity available, the arbitrage will not be

counted as price dispersion. The significantly negative impact of news arrivals on Λ_4 is an indirect result of differences between customer and market-maker appraisals. Given the uniform generation of desired inventories for numerous market-makers, the average desired inventory level across market-makers should be close to zero at all times. After a news arrival, market-maker and customer appraisals are generated from two different log-t distributions. On average over many news arrivals, these two distributions are the same, but after any given news event the mean of market-maker appraisals and the mean of customer appraisals will generally differ. Thus, a given event will be followed by an accumulating aggregate imbalance in market-maker inventories, as a preponderance of customers appear on one side of the market, say selling currency, until market-makers' spreads shift to surround customer appraisals. More frequent news arrivals tend to combine such periods of net selling with countermanding periods net buying, thus bringing the accumulated aggregate inventory imbalance closer to its long-run average level of zero. A similar dynamic is seen in the impact of news on serial correlation. Market-maker's prices converge toward a new consensus following a news event, implying positive autocorrelation over this convergence interval. The more news events there are, the greater is the proportion of the trading day characterized by such periods of convergence.

V. Conclusions

Among the experimental results is a single factor that can be held responsible for most of the statistically significant impacts: the sequence of transactions, set in motion by an inventory disruption, as the undesired position is passed along from one market-maker to another, producing temporary misallocations of currency inventories. This "hot-potato" trading has implications for price efficiency as well. First, there is inventory bounce in quoted prices that accompanies the passing along of an undesired inventory position. Second, there is a disinformation effect on prices, as individual market-makers may mistake a "hot potato" for a general shift in supply or demand. In both cases, the resulting price adjustments can yield arbitrage

opportunities, price dispersion and serial correlation in price changes.

Given this, the theoretical implications are fairly clear. From this perspective, the traditional view of market centralization as the key to operational efficiency is correct. One important advantage of centralization should be the ability to route orders directly to those who want them. If, for example, two customers call the market simultaneously, one wanting to buy yen for dollars, the other wanting dollars for yen, and both willing to transact at a common price, an operationally efficient market should pair these two off immediately, rather than letting two inventory positions meander aimlessly through a forest of market-makers until they happen to coincide. At the very least, such circuitous routing of currency positions produces an unnecessary multiplication of trading volume. Bilateral netting agreements and proposed multilateral payment netting schemes indicate that the implicit inefficiency is economically significant.[28]

Existing microstructures in other markets suggest institutional possibilities for centralization in this context. A radical possibility would be centralization of the market-making function in a monopolistic dealer, in the spirit of the stock exchange specialist. An idealized specialist system, in which the dealer participates in every transaction either through the book or for her own account, would clearly eliminate arbitrage opportunities and price dispersion. Moreover, it would collapse a "hot-potato" sequence of trades into a single transaction, in effect netting countermanding inventory imbalances across market-makers. Indeed, considering only back-office costs and the costs of financing inventory positions, economies of scale imply that market-making should be a natural monopoly in the current microstructure. On the other hand, previous theoretical work concludes that competition among market-makers is desireable -- Stoll (1978), for example, demonstrates that equilibrium market structure in a dealer market is generally not a monopoly -- while the empirical observation of consistent growth in the number of market-makers argues strongly against a natural monopoly for dealer services. Moreover, the capitalization required for a monopolistic foreign exchange market-maker to establish creditworthiness would likely be prohibitive.

A more realistic possibility is a centralization of price information

only. This could be achieved by collecting live direct and brokered quotes on an electronic network (similar to the NASDAQ system on the OTC stock market). Such a system would eliminate any arbitrage opportunities and price dispersion as defined here. Indeed, such systems have already begun to appear.[29] Centralization of quotes would obviate price dispersion, arbitrage opportunities and the adjustment interval, as measured here. Moreover, by encouraging price priority of order execution, such a system would facilitate the routing of currency positions to market-makers who want them, thus reducing the number of "hot-potato" trades. The results here indicate that a full centralization of price quotes should enhance the operational efficiency of the market and lead to a significant *ceteris paribus* reduction in trading volume. Adoption of such a fundamental microstructural change would necessarily entail considerable coordination between market-makers, brokers and regulators, however.

Notes

1. At the extreme, microstructural problems have been held to destroy markets. Miller (1986), pp. 169-70, for example, holds microstructure responsible for the failure of the inflation (CPI) futures contract at the Coffee, Sugar and Cocoa Exchange.

2. Foreign exchange trading volume is taken from FRB-NY (1989); the annual figure is based on 251 trading days per year. The corresponding figure for global foreign exchange volume is $160.64 trillion (BIS 1990b). World GDP ≈ $16 trillion is a rough approximation; this represents an extrapolation from UN (1989) estimates for 1986, using the 1986 growth rates in GDP since 1986, and assuming that per capita GDP for Hungary and Poland was representative of Eastern Europe.

3. See Flood (1991), p. 54.

4. More specifically, they state (p. 182) that, "The tremendous volume of foreign exchange trading is another piece of evidence that reinforces the idea of heterogeneous expectations, since it takes differences among market participants to explain why they trade."

5. While they do not rule out such an argument, the results in the present paper do offer an alternative explanation. In particular, the heavy trading volume and the relatively large price volatility that follow the morning opening of trading may both be symptoms of an underlying common cause, namely a wave of customer orders that disrupt the market as local customers begin their day.

6. This result may be related to the fact that their screen-based prices are not live quotes, but rather for "information purposes" only. Electronic trading screens (e. g., Reuters and Telerate) are not considered in the simulation model here. While such quotes doubtless contain some information

-- carefully constructed time-series statistics may be unbiased, for example -- their usefulness for trading decisions is limited. Reuters displays only one quote at a time, while Telerate has only five. More importantly, these posted quotes are not treated as binding offers to transact by market participants, but rather as "for information purposes". Posted quotes must be confirmed and can be repudiated at the whim of the posting bank, at little cost to their ability to deal in the future (compare this to the reputational cost to a bank that refuses to deal when called on the phone). As a result, posted quotes can lag the market significantly. For example, it is not uncommon to find quotes on Telerate as much as thirty minutes old, when in fact the lifetime of quotes (i. e., the length of time that a telephone-based bid or ask price would be considered a binding offer) should be measured in seconds [see, e. g., Burnham (1990), p. 12].

7. Because of its detail, only an outline of the simulation model is provided here. Details of the Bayesian learning model for market-makers are provided in appendix A. A full specification of the model, including flowcharts and copies of the Fortran simulation routines, are available upon request.

8. James Horhorst, as quoted by Mossberg (1988), p. 29R. Mr. Hohorst directed foreign exchange trading in North America for Manufacturers Hanover.

9. A total of twelve different behavioral specifications were tried. Because the conclusions were robust to all these variations, we report only two representative specifications here.' In addition, an adaptive expectations model was tried; it exhibited extreme cobweb-type instability, and thus was considered an unsatisfactory approximation to the real market.

10. The notation is defined in appendix A.

11. This is a heuristic approach based on Amihud and Mendelson (1980). Their derivation is optimal for a monopolistic market-maker facing a stationary stream of market buy and sell orders arriving according to a Poisson process. However, the two basic forces at work (narrowing the spread as the desired

inventory level is approached, and shading the quote in the direction of a desired inventory movement) are more general. See Burnham (1991), pp. 133-6, for additional analysis of the factors affecting the spread. In private correspondence, he suggests that market-makers widen their spreads in the time immediately prior to an anticipated news arrival, to discourage a last-minute shock to inventory. This aspect of trading is not incorporated in the model.

The bid price is not allowed to exceed $\beta^{\circ}+z = \alpha^{\circ}$, because this would be inconsistent with a trader who believed the market to be at $(\beta^{\circ},\alpha^{\circ})$ and who wanted to avoid being arbitraged. Similarly, the ask price is not allowed to fall below $\alpha^{\circ}-z = \beta^{\circ}$. Spreads larger than five ticks are not allowed, because, at some point, the dealer is perceived to be failing in her role as a market-maker -- a spread of $(0,\infty)$ is equivalent to quitting the market -- and would be asked to narrow her spread.

The spread-setting algorithm in table 1 requires some clarification. Prices in the U. S. market for DEM are quoted European style (i. e., DEM/USD). Spreads are concentrated at ten pfennigs per hundred dollars, with five- and twenty-pfennig spreads also frequent but less common. Intermediate values (e. g., .0009 DEM/USD) occur very rarely, so that .0005 DEM/USD becomes, at a first approximatation, a *de facto* tick size. The failure of market-makers to exploit finer gradations of the spread remains an unanswered puzzle; Goodhart and Figliuoli (1991), pp. 28-29, attribute it to a form of bounded rationality dubbed the "round number syndrome". In the model here, however, all prices and spreads are stated in American terms (USD/DEM), because the intradaily bid-ask spot price data used to validate the simulation model were obtained from the Philadelphia stock exchange (PHLX), which prices its foreign currency options American style (e. g., the quote 1.8910-20 would appear on the PHLX tape as "5288 5285"). This conversion compresses spreads, and the empirical distribution of spreads when prices are converted to American terms is unimodal. The mode is two ticks (i. e., .0002 USD/DEM; see also appendix B), the mean is .000266 USD/DEM, and more than 95% of the observed spreads are five ticks or less.

12. This approach, including the lack of serial correlation in changes in desired inventory, is consistent with Ho and Stoll (1983, pp. 1066-67). They define dealer m's inventory discrepancy ("net unwanted inventory", in their parlance, p. 1067) as $I_m - \epsilon_m/(R\sigma^2)$, where m's desired inventory is $\epsilon_m/(R\sigma^2) \equiv I_m^*$, and where σ^2 and R are both exogenous constants. The one difference between their approach and that used here is that, for Ho and Stoll, the sum across all dealers of ϵ_m -- and thus of $I_m^* \equiv \epsilon_m/(R\sigma^2)$ -- is always zero by construction: ϵ_m is the deviation of m's opinion of the true price from the average such opinions. In the simulations here, on the other hand, the sum across market-makers of the I_m^* is not identically zero, but rather zero in expectation. Note that, in both specifications, desired inventory is unaffected by price information: ϵ_m changes only with a change in fundamentals.

13. Given two uniform U(0,1) random variables, u_1 and u_2, a triangular integer random variable, $I_{mt}^* \sim \Delta\{-P,\ldots,P\}$, is generated as: $I_{mt}^* = [\![u_1(P+1)]\!] + [\![u_2(P+1)]\!] - P$, where $[\![\cdot]\!]$ is the greatest integer function and P is some positive integer. For example, subtracting 7 from a roll of two standard six-sided dice would yield a $\Delta\{-5,\ldots,5\}$ deviate.

14. This is a simplification. Owing to the difference in the way brokers' quotes and market-makers' quotes arise, it may be reasonable for market-makers to interpret their information differently. Precisely modelling such a difference in interpretation would be exceedingly complex, however.

15. Burnham (1991), pp. 136-7, confirms that customers who regularly engage in arbitrage or short-term position-taking will be required by their correspondent banks to reciprocate by making a market themselves. Thus, by definition, customers are precluded from considering short-term dynamics.

16. Taylor (1989) finds significant and persistent violations of covered interest parity in the London foreign exchange market. Similarly, Rhee and Chang (1992) find rare instances of covered interest parity violations, and more significant opportunities for "one-way arbitrage".

17. This point is always feasible in the model, because desired inventories are bounded by the position constraints, while customers provide an essentially unbounded pool of currency on both sides of the market. This is noteworthy, because it contradicts the familiar axiom of nonsatiation, and thus allows a measure of allocational inefficiency that differs from the standard notion of Pareto-efficiency.

18. There are some limits to the application of this argument. Roll (1984), for example, shows that the existence of a double-auction protocol is sufficient to induce negative serial correlation in changes in transaction prices (sometimes called "bid-ask bounce") in a stable efficient market. Fama (1970), on the other hand, argues that correlation patterns in security prices are usually small enough to be swamped by the spread and transaction costs, so that profitable trading algorithms are still not possible, even if pure price efficiency is not achieved. Cohen, Maier, Schwartz, and Whitcomb (1978) also cite this argument, and they provide several other reasons why "serial correlation patterns in transaction returns provide only limited evidence regarding market efficiency" (pp. 725-6). Goodhart and Figliuoli (1991) and Wasserfallen (1989) find evidence of significant negative serial correlation in exchange rates.

19. This clearly does not cover all possible predictable patterns. See Fama (1970), p. 392, note 10, for a simple counterexample. The market efficiency coefficient of Hasbrouck and Schwartz (1988) uses a weighted average of correlation coefficients at several lags. Their construction requires an assumption of Brownian motion, however, which allows the aggregation of lagged autocorrelations; see Schwartz (1988), pp. 361-2.

20. For a discussion of experimental designs in computer simulations, see Naylor, Burdick and Sasser (1967). For a description of response-surface designs see Khuri and Cornell (1987).

21. The choice of a second-order Taylor series to approximate the true response surface is not entirely arbitrary. Implicit in its use are the

presumptions that true response functions are smooth, and that the effect of lower order terms dominates that of higher order ones. Both assumptions are reasonable implications of what Simon (1963) calls the "principle of continuity of approximation" (pp. 230-1). The Taylor series is truncated at the second-order terms to mitigate multicollinearity problems among the numerous terms in same experimental variables.

22. The partial derivatives calculated from the Tobit regressions of the third inefficiency measure are slightly tricky, because a number of such partial derivatives can be calculated from the regression results. For example, one can calculate the partial derivative for the underlying model, or the partial derivative conditional on the dependent variable being at a non-limiting value; see Maddala (1983), p. 160.

23. For example, Taylor (1989), p. 388, argues: "The second finding is quite intuitive and uncontentious. One would expect the efficiency of the international foreign exchange and capital markets to have increased over time with increases in the number market participants ..." He also attributes part of the improvement in efficiency to advances in information technology. Stoll (1978) addresses directly the relative advantages of monopolistic and competitive dealer microstructures. For statements of the traditional view, see Stigler (1964), p. 129, and Garbade (1978), p. 497.

24. In interpreting the results, keep in mind that inefficiency is being measured: a negative partial derivative means an increase in the experimental variable tends to improve market performance. This difference in signs on partial derivatives for the same experimental variable is consistent with an equilibrium market structure, which should balance trade-offs between different types of inefficiency.

25. For the incautious model, 178 of the 300 observations (59 percent) are at the limiting value of 1000; for the cautious model the figure is 182 (61 percent). Normalized coefficients, as described by Maddala (1983), pp. 156-8, can be obtained by dividing the non-normalized coefficients (and their

standard errors) by the standard error of the regression estimate, which is 473.6 for the incautious case, and 645.5 for the cautious case.

26. Brooks (1985), p. 25, describes submitting a limit order as "sticking out the chin so as to be acquainted with the moment that the fight starts."

27. In interpreting the numbers in tables 5 and 6, keep in mind that the number of customers, C, is the number of customers active in the market in every period, not the total number of customer orders per day.

28. See BIS (1989, 1990a) and Duncan (1991). Netting procedures lower processing and accounting costs, and they can reduce counterparty exposure risks. Among current proposals, FXNET is a communications network that allows bilateral netting of payment orders; ECHO and NACHO are clearing house organizations that provide for multilateral netting of payments and counterparty substitution of the clearing house into all trades.

29. Existing and proposed services are either for informational purposes only (e. g., Quotron's FXQuote and Telerate pages 263-4), or they perform a brokerage function, that is, they accept only limit orders, rather than market-makers' spreads (e. g., Reuters's Dealing 2000). The efficiency of prices provided for informational purposes only is not so relevant, because there is no penalty for posting an off-market price. It is noteworthy that Reuters's Phase II was postponed precisely because it would provide live quotes -- participants could not agree on liability in case of a computer failure.

Sources Cited

Amihud, Yakov, and Haim Mendelson, 1980, Dealership Market: Market-Making with Inventory, *Journal of Financial Economics*, 8, 31-53.

Baillie, Richard T., and Bollerslev, Tim, 1990, Intra-Day and Inter-Market Volatility in Foreign Exchange Rates, *Review of Economic Studies*, 58, 565-585.

Bank for International Settlements (BIS), 1989, Report on Netting Schemes (BIS, Basle, Switzerland).

_____, 1990a, Report of the Committee on Interbank Netting Schemes of the Central Banks of the Group of Ten Countries (BIS, Basle, Switzerland).

_____, 1990b, Survey of Foreign Exchange Market Activity (BIS, Basle, Switzerland).

Banks, Jerry, and John S. Carson, II, 1984, Discrete Event System Simulation (Prentice-Hall, Englewood Cliffs).

Bollerslev, Tim, and Ian Domowitz, 1992, Trading Patterns and Prices in the Interbank Foreign Exchange Market, *Journal of Finance*, forthcoming.

Bossaerts, Peter, and Pierre Hillion, 1991, Market Microstructure Effects of Government Intervention in the Foreign Exchange Market, *Review of Financial Studies*, 513-541.

Brooks, J. A., 1985, The dealing room and the dealer, in R. Weisweiller, ed.: Managing a Foreign Exchange Department (Woodhead-Faulkner, Cambridge).

Burnham, James B., 1991, Current Structure and Recent Developments in Foreign Exchange Markets, in S. Khoury, ed.: Recent Developments in International Banking and Finance, Volumes IV and V (Elsevier North Holland, Amsterdam, The Netherlands).

Clarkson, Geoffrey P. E., and Herbert A. Simon, 1982, Simulation of Individual and Group Behavior, in H. A. Simon, ed.: Models of Bounded Rationality, Volume 2, Behavioral Economics and Business Organization (MIT Press, Cambridge, Mass.).

Cohen, Kalman J., Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb, 1978, Limit Orders, Market Structure, and the Returns Generating Process, *Journal of Finance*, 33, 723-736.

Conroy, Robert M., and Robert L. Winkler, 1981, Informational Differences between Limit and Market Orders for a Market Maker, *Journal of Financial and Quantitative Analysis*, 16, 703-724.

Duncan, G. M., 1991, Echo Netting: Multilateral FX Netting in Europe, Discussion Document (Barclays Bank, London, England).

Engle, Robert F., Takatoshi Ito, and Wen-Ling Lin, 1990, Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market, *Econometrica*, 58, 525-542.

Euromoney, 1986, Flinging open the doors of change, Aug. 1986 supplement, 2-6.

Fama, Eugene F., 1970, Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance*, 25, 383-417.

Federal Reserve Bank of New York (FRB-NY), 1989, Summary of Results of U. S. Foreign Exchange Market Turnover Survey Conducted in April 1989, FRB-NY Release, Sep. 13, 1989.

Flood, Mark D., 1991, Microstructure Theory and the Foreign Exchange Market, *Review* (Federal Reserve Bank of St. Louis), 73(6), 52-70.

Frankel Jeffrey A., and Kenneth A. Froot, 1990, Chartists, Fundamentalists, and Trading in the Foreign Exchange Market, *American Economic Review*, 80, 181-185.

Garbade, Kenneth D., 1978, The Effect of Interdealer Brokerage on the Transactional Characteristics of Dealer Markets, *Journal of Business*, 51, 477-498.

Garbade, Kenneth D., and William L. Silber, 1976, Price Dispersion in the Government Securities Market, *Journal of Political Economy*, 84, 721-740.

Goodhart, Charles A. E. and Antonis Demos, 1990, Reuters Screen Images of the Foreign Exchange Market: The Deutschemark/Dollar Spot Rate, *Journal of International Securities Markets* 4, 333-348.

Goodhart, Charles A. E. and L. Figliuoli, 1991, Every minute counts in financial markets, *Journal of International Money and Finance*, 10, 23-52.

Goodhart, Charles A. E. and M. Giugale, 1988, From Hour to Hour in the Foreign Exchange Market, working paper (London School of Economics, London).

Grant, Malcom J., and F. Michael Rabinowitz, 1987, Representing Individual Differences in Computer Models: Theoretical Considerations and Parameter Estimation, *Computers in Human Behavior*, 3, 211-224.

Hakkio, Craig S., and Douglas K. Pearce, 1985, The Reaction of Exchange Rates to Economic News, *Economic Inquiry*, 23, 621-636.

Hasbrouck, Joel and Robert A. Schwartz, 1988, Liquidity and execution costs in equity markets: How to define, measure, and compare them, *Journal of Portfolio Management*, Spring, 10-16.

Ho, Thomas S. Y., and Hans R. Stoll, 1983, The Dynamics of Dealer Markets Under Competition, *Journal of Finance*, 38, 1053-1074.

Ito, Takatoshi, 1987, The Intradaily Exchange Rate Dynamics and Monetary Policies after the Group of Five Agreement, *Journal of the Japanese and International Economies*, 1, 275-298.

Ito, Takatoshi, Robert F. Engle, and Wen-Ling Lin, 1992, Where does the meteor shower come from? The role of stochastic policy coordination, *Journal of International Economics*, 32, 221-240.

Ito, Takatoshi, and V. Vance Roley, 1987, News from the U. S. and Japan: Which Moves the Yen/Dollar Exchange Rate?, *Journal of Monetary Economics*, 19, 255-277.

_____, 1988, Intraday Yen/Dollar Exchange Rate Movements: News or Noise?, working paper RWP 88-07 (Federal Reserve Bank, Kansas City).

Khuri, André I., and John A. Cornell, 1987, Response Surfaces: Designs and Analyses (Marcel Dekker, New York, NY).

Maddala, G. S., 1983, Limited Dependent and Qualitative Variables in Econometrics (Cambridge University Press, Cambridge, England).

Melvin, Michael, and Kok-Hui Tan, 1992, Foreign Exchange Market Bid-Ask Spreads and the Market Price of Social Unrest, working paper (Arizona

State University, Tempe).

Mossberg, Walter S., 1988, Making Book on the Buck, *Wall Street Journal*, Sep. 23, 1R ff.

Naylor, Thomas H., Donald S. Burdick, and W. Earl Sasser, 1967, Computer Simulation Experiments with Economic Systems: The Problem of Experimental Design, *Journal of the American Statistical Association*, 62, 1315-1337.

Philadelphia Stock Exchange (PHLX), 1989, Foreign currency options pricing history tape (PHLX, Philadelphia, PA).

Praetz, Peter D., 1972, The Distribution of Share Price Changes, *Journal of Business*, 45, 49-55.

Raiffa, Howard, and Robert Schlaifer, 1961, Applied Statistical Decision Theory (Harvard University Division of Research, Boston, MA).

Rhee, S. Ghon, and Rosita P. Chang, 1992, Intra-Day Arbitrage Opportunities in Foreign Exchange and Eurocurrency Markets, *Journal of Finance*, 47, 363-379.

Roll, Richard, 1984, A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market, *Journal of Finance*, 39, 1127-1139.

Schwartz, Robert A., 1988, Equity Markets: Structure, Trading, and Performance (Harper and Row, New York, NY).

Simon, Herbert A., 1963, Problems of Methodology -- Discussion, *American Economic Review*, 53, 229-231.

Smith, Adam, 1937, An Inquiry into the Nature and Causes of the Wealth of Nations (Modern Library, New York, NY).

Stigler, George J., 1964, Public Regulation of the Securities Markets, *Journal of Business*, 37, 117-142.

Stoll, Hans R., 1978, The Supply of Dealer Services in Securities Markets, *Journal of Finance*, 33, 1133-1151.

Taylor, Mark P., 1989, Covered Interest Arbitage and Market Turbulence, *The Economic Journal*, 99, 376-391.

United Nations Conference on Trade and Development (UN), 1989, Handbook of International Trade and Development Statistics 1988 (ISBN 92-1-012025-6) (United Nations, New York, NY).

Wasserfallen, Walter, 1989, Flexible exchange rates: A closer look, *Journal of Monetary Economics*, 23, 511-521.

Wasserfallen, Walter, and Heinz Zimmerman, 1985, The Behavior of Intra-Daily Exchange Rates, *Journal of Banking and Finance*, 9, 55-72.

TABLE 1 Market-maker quote function

$$Q_{mt} = g_m(A_m, I_{mt} - I^*_{mt})$$

Inventory discrepancy $(I_{mt} - I^*_{mt})$	Quote $(\beta_{mt}, \alpha_{mt})$
≤ -5	$(\beta^\circ_{mt}+z, \alpha^\circ_{mt}+5z)$
-4	$(\beta^\circ_{mt}+z, \alpha^\circ_{mt}+4z)$
-3	$(\beta^\circ_{mt}+z, \alpha^\circ_{mt}+3z)$
-2	$(\beta^\circ_{mt}+z, \alpha^\circ_{mt}+2z)$
-1	$(\beta^\circ_{mt}, \alpha^\circ_{mt}+z)$
0	$(\beta^\circ_{mt}, \alpha^\circ_{mt})$
$+1$	$(\beta^\circ_{mt}-z, \alpha^\circ_{mt})$
$+2$	$(\beta^\circ_{mt}-2z, \alpha^\circ_{mt}-z)$
$+3$	$(\beta^\circ_{mt}-3z, \alpha^\circ_{mt}-z)$
$+4$	$(\beta^\circ_{mt}-4z, \alpha^\circ_{mt}-z)$
$\geq +5$	$(\beta^\circ_{mt}-5z, \alpha^\circ_{mt}-z)$

Note: For the incautious model, the base prices, β° and α°, are set by rounding the appraisal down and up, respectively, to the nearest tick. For the cautious model, β° and α° are set according to equations (2) and (3).

TABLE 2 Convex Design Region for the Experimental Variables

Experimental Variable	Range of Values	Inefficiency Measures
M	{11,...,500}	1,2,3,4,5
B	{15,...,250}	1,2,3,4,5
C	{1,...,50}	1,2,3,4,5
Ω_u	(0,.2]	1,2,4,5

TABLE 3 Regressions: Incautious Model

	Λ_1	Λ_2	Λ_3	Λ_4	Λ_5
Intercept	406.	0.378	522.	2610.	0.208
	(46.0)	(0.0367)	(343.)	(347.)	(0.0418)
$\hat{\gamma}_m$	1.97	-0.00258	-0.0238	-0.284	-0.000184
	(0.169)	(0.000135)	(1.33)	(1.28)	(0.000154)
$\hat{\gamma}_b$	-4.50	0.00150	0.384	-1.384	0.00105
	(0.359)	(0.000286)	(2.92)	(2.70)	(0.000326)
$\hat{\gamma}_c$	9.40	-0.00327	42.5	61.0	-0.00503
	(1.57)	(0.00125)	(11.5)	(11.8)	(0.00142)
$\hat{\gamma}_u$	957.	-0.0796	---	-5950.	0.558
	(368.)	(0.293)	---	(2770.)	(0.334)
$\hat{\gamma}_{mm}$	-0.00120	38.6 E-7	-0.00174	-0.00491	-1.59 E-7
	(0.000237)	(1.85 E-7)	(0.00191)	(0.00175)	(2.10 E-7)
$\hat{\gamma}_{bb}$	0.00552	-13.1 E-7	0.00350	-0.00845	-9.70 E-7
	(0.00102)	(8.12 E-7)	(0.00817)	(0.00768)	(9.25 E-7)
$\hat{\gamma}_{cc}$	-0.00822	2.19 E-5	-0.699	-0.554	3.48 E-5
	(0.0237)	(1.89 E-5)	(0.176)	(0.178)	(2.15 E-5)
$\hat{\gamma}_{uu}$	367.	-0.999	---	5140.	-2.56
	(1420.)	(1.13)	---	(10700.)	(1.29)
$\hat{\gamma}_{mb}$	0.00307	-26.1 E-7	-0.00219	0.00775	-11.4 E-7
	(0.000449)	(3.58 E-7)	(0.00401)	(0.00338)	(4.07 E-7)
$\hat{\gamma}_{mc}$	-0.0116	5.08 E-6	0.0273	-0.0176	7.20 E-6
	(0.00212)	(1.69 E-6)	(0.0159)	(0.0160)	(1.92 E-6)
$\hat{\gamma}_{mu}$	-1.69	0.000888	---	4.88	0.000237
	(0.551)	(0.000439)	---	(4.15)	(0.000500)
$\hat{\gamma}_{bc}$	-0.0127	-.599 E-6	0.00514	-0.0326	-.867 E-6
	(0.00444)	(3.54 E-6)	(0.0329)	(0.0335)	(4.03 E-6)
$\hat{\gamma}_{bu}$	-0.921	-0.000114	---	8.76	0.000983
	(1.12)	(0.000888)	---	(8.40)	(0.00101)
$\hat{\gamma}_{cu}$	-9.14	0.000376	---	8.73	0.00561
	(5.17)	(0.00412)	---	(38.9)	(0.00469)
F-statistic	332.	131.	---	17.8	20.4
R^2	0.942	0.866	---	0.466	0.500
R^2 (adj.)	0.939	0.859	---	0.440	0.476
$\hat{\sigma}$	71.9	0.0573	474.	541.	0.0652

(standard errors in parentheses)

TABLE 4 Regressions: Cautious Model

	Λ_1	Λ_2	Λ_3	Λ_4	Λ_5
Intercept	154.	0.373	556.	4910.	0.0782
	(45.0)	(0.0313)	(373.)	(470.)	(0.0426)
$\hat{\gamma}_m$	2.03	-0.00202	2.69	-4.58	-0.000227
	(0.154)	(0.000107)	(1.46)	(1.61)	(0.000146)
$\hat{\gamma}_b$	-3.11	0.000271	-1.79	1.38	0.00210
	(0.319)	(0.000222)	(3.70)	(3.33)	(0.000302)
$\hat{\gamma}_c$	7.90	-0.00270	35.0	64.8	-0.000507
	(1.35)	(0.000940)	(14.3)	(14.1)	(0.00128)
$\hat{\gamma}_u$	1340.	-0.321	---	-14400.	0.945
	(346.)	(0.240)	---	(3610.)	(0.328)
$\hat{\gamma}_{mm}$	-0.00198	27.5 E-7	-0.00743	0.00111	2.91 E-7
	(0.000220)	(1.53 E-7)	(0.00236)	(0.00230)	(2.09 E-7)
$\hat{\gamma}_{bb}$	0.00350	2.39 E-7	0.000315	-0.00726	-22.5 E-7
	(0.000889)	(6.18 E-7)	(0.0112)	(0.00929)	(8.43 E-7)
$\hat{\gamma}_{cc}$	-0.0586	2.17 E-5	-0.494	-0.588	-.434 E-5
	(0.0193)	(1.34 E-5)	(0.236)	(0.202)	(1.83 E-5)
$\hat{\gamma}_{uu}$	-4010.	0.368	---	36800.	-0.430
	(1270.)	(0.885)	---	(13300.)	(1.21)
$\hat{\gamma}_{mb}$	0.00266	-5.01 E-7	0.00286	0.00464	-23.1 E-7
	(0.000380)	(2.64 E-7)	(0.00446)	(0.00397)	(3.60 E-7)
$\hat{\gamma}_{mc}$	-0.00995	4.16 E-6	0.0215	-0.0423	3.84 E-6
	(0.00180)	(1.25 E-6)	(0.0199)	(0.0188)	(1.71 E-6)
$\hat{\gamma}_{mu}$	1.47	0.000607	---	20.7	-0.000450
	(0.495)	(0.000344)	---	(5.16)	(0.000469)
$\hat{\gamma}_{bc}$	-0.000883	-.881 E-6	0.0361	0.0309	-5.74 E-6
	(0.00389)	(2.70 E-6)	(0.0415)	(0.0406)	(3.68 E-6)
$\hat{\gamma}_{bu}$	-3.66	-8.25 E-5	---	-26.4	-0.000563
	(0.997)	(69.4 E-5)	---	(10.4)	(0.000946)
$\hat{\gamma}_{cu}$	7.33	0.00125	---	53.3	-0.00524
	(4.58)	(0.00319)	---	(47.9)	(0.00435)
F-statistic	278.	76.1	---	23.0	35.6
R^2	0.932	0.789	---	0.531	0.636
R^2 (adj.)	0.928	0.779	---	0.508	0.618
$\hat{\sigma}$	62.1	0.0431	645.	648.	0.0589

(standard errors in parentheses)

TABLE 5 Impact of the experimental variables on the inefficiency measures, stated as partial derivatives and elasticities.

Behavioral spec.: incautious

Inefficiency Measure	Impact of M	Impact of B	Impact of C	Impact of Ω_u
	Evaluated at:	population means		
$\partial E(\Lambda_1)/\partial\cdot$	1.30	-2.67	3.42	245.
Elasticity	0.577	-0.615	0.152	.0427
	(1660.)*	(1570.)*	(123.)*	(9.43)†
$\partial E(\Lambda_2)/\partial\cdot$	-0.000734	0.000461	-0.000895	-0.0581
Elasticity	-7.75	2.52	-0.944	-0.241
	(838.)*	(73.3)*	(13.3)*	(0.837)
$\partial E(\Lambda_3)/\partial\cdot$	-0.574	0.791	37.4	---
Elasticity	-0.105	0.145	6.87	---
	(4.13)†	(2.24)	(37.4)*	---
$\partial E(\Lambda_4)/\partial\cdot$	-1.36	-1.59	24.9	-2290.
Elasticity	-0.121	-0.0727	0.220	-0.0795
	(32.4)*	(9.76)†	(114.)*	(14.6)*
$\partial E(\Lambda_5)/\partial\cdot$	-0.000209	0.000579	-0.000968	0.379
Elasticity	-0.226	0.325	-0.105	0.161
	(52.2)*	(89.2)*	(11.9)*	(27.4)*
	Evaluated at:	status quo		
$\partial E(\Lambda_1)/\partial\cdot$	1.69	-3.10	5.49	611.
Elasticity	0.604	-0.739	0.0915	0.124
	(460.)*	(445.)*	(31.5)*	(16.7)*
$\partial E(\Lambda_2)/\partial\cdot$	-0.00157	0.000833	-0.00223	-0.126
Elasticity	-1.52	0.536	-0.100	-0.0696
	(624.)*	(50.5)*	(8.17)†	(1.12)
$\partial E(\Lambda_3)/\partial\cdot$	-0.506	0.878	14.5	---
Elasticity	-0.102	0.177	2.93	---
	(0.790)	(0.392)	(25.5)*	---
$\partial E(\Lambda_4)/\partial\cdot$	-0.472	-1.39	48.1	-3400.
Elasticity	-0.0296	-0.0580	0.141	-0.121
	(0.631)	(1.57)	(42.6)*	(9.09)†
$\partial E(\Lambda_5)/\partial\cdot$	-0.000275	0.000763	-0.00307	0.293
Elasticity	-0.152	0.281	-0.0790	0.0922
	(14.7)*	(32.6)*	(11.9)*	(4.65)†

F-statistics appear in parentheses
* -- significantly different from zero at .001 level
† -- significantly different from zero at .100 level

TABLE 6 Impact of the experimental variables on the inefficiency measures, stated as partial derivatives and elasticities.

Behavioral spec.: cautious

Inefficiency Measure	Impact of M	Impact of B	Impact of C	Impact of Ω_u
	Evaluated at:	population means		
$\partial E(\Lambda_1)/\partial\cdot$	1.26	-1.89	2.98	601.
Elasticity	0.670	-0.521	0.159	0.125
	(2110.)*	(1040.)*	(138.)*	(81.1)*
$\partial E(\Lambda_2)/\partial\cdot$	-0.000534	0.000175	-0.000524	-0.0718
Elasticity	-10.5	1.78	-1.03	-0.554
	(786.)*	(18.4)*	(8.81)†	(2.40)
$\partial E(\Lambda_3)/\partial\cdot$	-0.186	-0.0533	20.1	---
Elasticity	-0.0327	-0.00935	3.52	---
	(0.347)	(0.00522)	(41.0)*	---
$\partial E(\Lambda_4)/\partial\cdot$	-2.40	-1.20	33.4	-3920.
Elasticity	-0.140	-0.0363	0.195	-0.0898
	(70.5)*	(3.84)†	(159.)*	(31.7)*
$\partial E(\Lambda_5)/\partial\cdot$	-0.000331	0.000716	-0.00103	0.536
Elasticity	-0.342	0.382	-0.106	0.217
	(163.)*	(166.)*	(18.3)*	(71.9)*
	Evaluated at:	status quo		
$\partial E(\Lambda_1)/\partial\cdot$	1.83	-2.31	6.66	402.
Elasticity	1.07	-0.971	0.0779	0.150
	(568.)*	(301.)*	(57.3)*	(16.5)*
$\partial E(\Lambda_2)/\partial\cdot$	-0.00120	0.000237	-0.00192	-0.165
Elasticity	-1.41	0.201	-0.0453	-0.124
	(534.)*	(6.08)†	(9.75)†	(3.19)†
$\partial E(\Lambda_3)/\partial\cdot$	0.830	-1.18	39.2	---
Elasticity	0.162	-0.230	7.62	---
	(1.31)	(0.699)	(13.0)*	---
$\partial E(\Lambda_4)/\partial\cdot$	-1.88	-1.93	63.4	-6970.
Elasticity	-0.0790	-0.0585	0.0533	-0.188
	(8.95)†	(1.02)	(47.0)*	(27.2)*
$\partial E(\Lambda_5)/\partial\cdot$	-0.000422	0.00120	-0.00108	0.718
Elasticity	-0.222	0.454	-0.0114	0.242
	(30.1)*	(91.7)*	(1.76)	(31.2)*

F-statistics appear in parentheses
* -- significantly different from zero at .001 level
† -- significantly different from zero at .100 level

TABLE 7 Signs on the Partial Derivatives of the Inefficiency Measures with Respect to the Experimental Variables

Measure		Experimental Variable M	B	C	Ω_u
Λ_1	arbitrage opportunities	+	-	+	+
Λ_2	price dispersion	-	+	-	(-)
Λ_3	adjustment interval	(-)	0	+	n.a.
Λ_4	suboptimal allocation	(-)	(-)	+	-
Λ_5	serial correlation	-	+	(-)	+

Note: The zero entry indicates the impact on the inefficiency measure is not significantly different from zero anywhere in table 5 or 6. The five signs in parentheses indicate that the impact is not statistically significant in all cases.

TABLE B1 Output Vectors Corresponding to the Validated Levels of the Input-Ouput Parameters

	Benchmark Value	Incautious Model	Cautious Model
Mean Price	0.5252	0.5252	0.5319
Mean Spread	0.000266	0.000266	0.000246
Modal Price Variance	8.00×10^{-08}	8.01×10^{-08}	8.02×10^{-08}
Modal Serial Correlation	-0.16	-0.0045	-0.0042
Market-maker Volume	15317×10^{6}	15318×10^{6}	12352×10^{6}
Broker Volume	11759×10^{6}	11761×10^{6}	10478×10^{6}
Customer Volume	1119×10^{6}	1119×10^{6}	1121×10^{6}

APPENDIX A: A Bayesian Learning Model for Market-makers

To endogenize quoted prices in the model, market-makers must take the quoted rates of others into account, as they are observed. The particular technique employed is a Bayesian learning model. Market-makers believe that current quotes can be represented as draws from a lognormal distribution with random parameters μ and h.[1] Each market-maker attempts to learn through sampling (i. e., observing the quotes of others) what the actual parameters μ and h are. The analysis is simplified by performing a logarithmic transformation on observed prices, so that learning is seen to take place in the transformed space, where a normal distribution can be used, the Bayesian properties of which are well known. Exploiting this, the uncertain parameters μ and h of this normal process are modeled as draws from a joint (subjective) normal-gamma distribution, which is natural-conjugate for an independent normal process with both mean and variance unknown. Thus, each market-maker's joint posterior density is also normal-gamma. More importantly, the marginal posterior distribution of μ is a generalized t-distribution.[2] This latter is the distribution of interest, because the market-maker wants to produce the best estimate of the median of the (untransformed) price distribution -- the price which, when quoted, equalizes the subjective probability of being hit on

---

[1] This follows Raiffa and Schlaifer (1961) in defining the normal density in terms of its precision, $h \equiv 1/\sigma^2$, rather than its variance, σ^2. For a related learning model, see Conroy and Winkler (1981).

[2] There is some empirical support for this formulation. For example, Praetz (1972) finds that a scaled log t outperforms, among others, stable Paretian distributions in fitting stock price data. For daily changes in foreign exchange rates, Boothe and Glassman (1987) also reject the lognormal in favor of either a t-distribution or a mixture of two lognormals. Both studies, of course, are measuring the empirical properties of a time series of prices, whereas we are interested in a theoretical model of market-makers' beliefs about the cross-sectional distribution of prices at a point in time.

the buy side with the probability of being hit on the sell side.[3] The mean of the marginal distribution is the market-maker's best guess about the mean of the "actual" log-transformed price distribution. Because of the symmetry of the t-distribution, its mean is its midpoint, and thus corresponds to the median of the (untransformed) subjective price distribution when the transformation is reversed.[4]

More formally, the following statistics for the sample of "prices" observed by market-maker m can be defined:

$$n_m \equiv \text{no. of observations}$$

$$w_m \equiv n_m - 1$$

$$k_m \equiv \frac{1}{n_m} \sum_{i=1}^{n_m} x_{im} \tag{A1}$$

$$v_m \equiv \frac{1}{n_m - 1} \sum_{i=1}^{n_m} (x_{im} - k_m)^2 ,$$

which together form a sufficient statistic for the entire sample under the distributional assumptions. The x_{im}'s are market-maker m's "price observations" (described below). Corresponding prior statistics n'_m, w'_m, k'_m, and v'_m are assigned at random for each market-maker at the beginning of the day and following each subsequent news arrival.[5] Following the subsequent observation of prices in the market, the sample information is combined with the prior statistics to form posterior statistics:

$$n''_m = n'_m + n_m$$

---

[3] This strategy applies to a market-maker with a balanced inventory. Cf. Conroy & Winkler's (1981) equation 1; they avoid the issue of inventory entirely. A heuristic based on Amihud and Mendelson (1980) is used to adjust this formulation to account for inventory imbalances.

[4] The correspondence between the median of the price distribution and the mean of the distribution of log-transformed prices is a consequence of the integrand transformation rule; see Raiffa & Schlaifer (1961), pp. 212-13.

[5] The random assignment of behavioral parameters in computer models is well-accepted; see Grant and Rabinowitz (1987), esp. pp. 213-215.

$$w''_m = w'_m + w_m + 1$$

$$k''_m = \frac{n'_m k'_m + n_m k_m}{n'_m + n_m} \tag{A2}$$

$$v''_m = \frac{(w'_m v'_m + n'_m k'^2_m) + (w_m v_m + n_m k^2_m) - n''_m k''^2_m}{w'_m + w_m + 1} .$$

The joint posterior distribution of μ and h is then a normal-gamma distribution defined by these posterior statistics:

$$f''_{\phi\Gamma}(\ \mu, h \mid k''_m,\ v''_m,\ n''_m,\ w''_m) \equiv f''_{\phi}(\ \mu \mid k''_m,\ hn''_m) \cdot f''_{\Gamma}(\ h \mid v''_m,\ w''_m) \ . \tag{A3}$$

As a practical matter, most of these calculations can be bypassed, because the market-maker is explicitly interested only in the posterior marginal density of μ, which is given by the general Student normalized density:

$$f''_S(\ \mu \mid k''_m,\ n''_m/v''_m,\ w''_m) = \int_0^\infty f''_{\phi\Gamma}(\ \mu, h \mid k''_m,\ v''_m,\ n''_m,\ w''_m) \cdot dh \ . \tag{A4}$$

The following algorithm transforms the observed behavior of other market agents into x_{im}'s required by the definition of the sample statistics. A spread, $(\beta_{im}, \alpha_{im})$, quoted by broker or market-maker i and observed by m, is converted to the arithmetic mean of the spread, after taking logarithms:

$$x_{im} = [\mathit{ln}(\beta_{im}) + \mathit{ln}(\alpha_{im})]/2 \ . \tag{A5}$$

Buy, sell or pass orders, whether market orders from a customer or responses by another market-maker to a quoted spread, are converted to the appropriate incomplete first moment, normalized by the appropriate probability. For example, if market-maker m quotes an ask price α_m, and makes a sale to i at this price, she presumes that the purchaser's appraisal is above the price asked. To incorporate this limited price information into her subjective distribution, she imputes a valuation x_{im} from her observation:

$$x_{im} = \frac{k'_m - F'_{SI1}[\ \mathit{ln}(\alpha_m) \mid k'_m,\ n'_m/v'_m,\ w'_m]}{1 - F'_S[\ \mathit{ln}(\alpha_m) \mid k'_m,\ n'_m/v'_m,\ w'_m]} , \tag{A6}$$

where $F'_{SI1}(\cdot)$ is the incomplete first moment of the marginal (general Student)

distribution (prior to this observation), and $F'_S(\cdot)$ is the cumulative general Student distribution. This quotient is simply m's prior conditional expectation of $ln(A_i)$, given that $ln(A_i) > ln(\alpha_m)$, where A_i is purchaser i's appraisal. Similarly, a sale to the market-maker who quotes a bid price β_m implies a valuation:

$$x_i = \frac{F'_{SI1}[\ ln(\beta_m)\ \mid k'_m,\ n'_m/v'_m,\ w'_m]}{F'_S[\ ln(\beta_m)\ \mid k'_m,\ n'_m/v'_m,\ w'_m]}\ , \tag{A7}$$

and a pass, given a spread of β_m and α_m, becomes (suppressing the prior parameters in the notation):

$$x_i = \frac{F'_{SI1}[\ ln(\alpha_m)\ \mid\cdot] - F'_{SI1}[\ \ln(\beta_m)\ \mid\cdot]}{F'_S[\ ln(\alpha_m)\ \mid\cdot] - F'_S[\ \ln(\beta_m)\ \mid\cdot]}\ . \tag{A8}$$

Finally, The random assignment of priors proceeds in two stages. Given the jth news event, heterogeneous expectations implies that each market-maker will respond differently. In particular, for each market-maker, v'_m and w'_m are drawn from the exponential distributions $f_e(v'_m|\Omega_{vj})$ and $f_e(w'_m|\Omega_{wj})$, respectively, and k'_m is drawn from a normal $f_\phi(k'_m|\mu=\mu_{kj},h=1/v_{kj})$ distribution.[6] To model the fact that each news event is different from the previous event, the four parameters, Ω_{vj}, Ω_{wj}, μ_{kj} and v_{kj}, are themselves random draws. Exponential distributions, $f_e(1/\Omega_{vj}|\Omega_v)$, $f_e(1/\Omega_{wj}|\Omega_w)$ and $f_e(v_{kj}|\Omega_k)$, produce Ω_{vj}, Ω_{wj} and v_{kj}, respectively, and the normal $f_\phi(\mu_{kj}|\mu_\mu,h_\mu)$ distribution produces μ_{kj}. Values for the five parameters, Ω_v, Ω_w, Ω_k, μ_μ and h_μ, are determined in the validations.

---

[6] To make the prior statistics consistent, the redundant statistic n'_m is set to $n'_m = w'_m + 1$. Furthermore, w'_m is augmented by adding two to the draw from the exponential distribution, to provide sufficient prior degrees of freedom for the market-maker to conjecture a full prior distribution.

APPENDIX B: Validation of the simulation routine

The central advantages of the simulation methodology in this context are threefold. First, as with any theoretical construction, the simulation model here allows controlled experimentation. That is, one can assign (and systematically vary) parameter values and examine the resulting properties of the model. Of special interest is experimentation with counterfactual scenarios (e. g., what happens when the number of market-makers is reduced tenfold, while the number of customers simultaneously increases tenfold?). The ability of existing data and standard empirical techniques to address counterfactuals is limited in obvious ways. Second, the data availability problem for this market is nontrivial. For example, even a transaction-by-transaction data set for all market participants, if such a thing were available, would be insufficient for the measures defined here. Measures 1, 2, 3 and 5 require all live quotes, including and especially those quotes that did not produce transactions; measure 4 requires knowledge of every market-maker's actual and desired inventories. Third, the trade-off between realism and tractability that usually constrains theoretical models is obviated (or at least mitigated) by recourse to simulation. Tractability as a modelling criterion can be subordinated to realism, making possible a much richer and more realistic specification.[1] A corollary proposition is that, if one is to be confident of the simulation's results, one must first be assured of the

---

[1] In the words of Clarkson and Simon (1982), p. 359: "Apart from its normative uses, simulation is a peculiarly attractive method for describing and explaining the decision-making process at a microeconomic level. Its first, and most obvious, advantage is the same one that has led to its wide use in management science - it allows a degree of complexity to be handled that would be unthinkable if inference could be drawn from the model only by standard analytical techniques."

model's realism. Validation is the implementation of this realism requirement.[2]

Seven of the parameters of the simulation program correspond in a straightforward way to observable real-world counterparts; these parameters are assigned values equal to those of their real-world equivalents:

(1) Brokerage fees: Burnham (1990, p. 30) states, "In a spot trade in a major currency, both sides might be charged $10 per million by the broker." Kubarych (1983, p. 14) identifies a higher rate: "Commissions are negotiable and vary according to the currency, but generally amount to less than 1/100 percent of the selling price, or about $25 per $1 million". Using the benchmark exchange rate of .5252 USD/DEM, brokerage fees in the simulations are set at an intermediate value of .00001 USD/DEM.

(2) Tick size: The tick size, or minimum price increment, for the market is gotten from the PHLX tape (1989), which records bid and ask prices from Telerate. The PHLX converts Telerate quotes from European style (i. e., DEM/USD) to American style (USD/DEM), because foreign currency options are priced American style. The minimum price increment for the German mark is .0001 USD/DEM; see also note 11.

(3) Number of market-makers: The number of market-makers is taken from FRB-NY (1989). The survey lists 148 commercial banks and 14 non-bank institutions as respondents (162 total). This involves overcounting, as some banks report their branches separately, when not all branches are actively engaged as market-makers -- some maintain a trading desk only as a correspondent service

---

[2] The procedure used here is adapted from Banks and Carson (1984).

for bank customers. Indeed, FRB-NY (1989), pp. 1 and 3, cites trading volume for "127 banking institutions" and "14 nonbank financial institutions" (141 total), noting the discrepancy between this count and the list of respondents. For the simulations, an intermediate number of 150 is chosen.

(4) Number of brokerage books: The number of brokerage books for the simulations (14) is taken from FRB-NY (1989), which lists 14 foreign currency brokerage firms. The number of books is distinguished from the number of brokers servicing those books, which is determined by the input-output validations described below.

(5) Number of trading intervals: Subdividing the trading day amounts to selecting the average length of a trading interval, bearing in mind that trading in the New York market occurs at a much faster pace in the morning, when the European markets are still active, than in the afternoon. The New York market is most active for the six hours from 8:30 AM to 2:30 PM Eastern time, as evidenced by the trading hours on the PHLX foreign currency options exchange [see also Bollerslev and Domowitz (1992), and Goodhart and Demos (1990)]. For the simulations, a round figure of 1000 periods per trading day is used; this corresponds to approximately 22 seconds per contact on average.

(6) Maximum deal size: The maximum deal size is stated in terms of round lots in the foreign currency. Kubarych (1983, p. 12) observes that "Lately, customary amounts have been in the $1-5 million range, although for active banks they can be higher." For the German mark, with an exchange rate of roughly .50 USD/DEM, this implies a maximum deal size, in terms of marks, of roughly DEM 10 million.

(7) Validation sampling rate: The input-output validations (described below) require that time series of prices correspond in its construction to the PHLX tape. The PHLX tape has, on average, 75.3 observations between 8:30 AM and 2:30 PM per day over the period 1 March 1989 to 31 May 1989. Observations on the tape are time-stamped with the hour and minute, and only the first observation on the tape for a given minute was used. This sampling process is approximated via Monte Carlo simulation. Exponential sampling intervals are generated for a range of process intensities, extra simultaneous (after rounding to the minute) observations are deleted, and the resulting number of observations are recorded as a function of the sampling rate. The sampling rate which approximates 75 observations per day is .08.

The remaining parameters are continuous inputs parameters that could not be validated as data assumptions, for lack of real-world counterparts; these are subjected to a response surface analysis. For this, the computer program is treated as a vector-valued function on a continuous parameter space. The first step is to define the elements of the program's output vector. Since these are to be compared to reality, they are chosen to correspond to a benchmark vector of summary statistics, the values of which are known for the real world. The summary statistics selected are basic magnitudes representing price levels, dynamic price behavior, and aggregate quantities traded.

All numbers are for the USD-DEM spot market. To insure a consistent set of numbers for the validations, all of the validation techniques restrict themselves to data for the German mark from the spring of 1989. The data for price statistics are taken from the PHLX (1989) tape, which is filtered as follows. Observations must be between 1 March 1989 and 31 May 1989, inclusive, and between 8:30 AM and 2:30 PM, inclusive. Observations implying a nonpositive spread or a spread larger than ten times the modal spread of

.0002 USD/DEM (i. e., spreads greater than .002 USD/DEM) are assumed to be typos and are deleted from the sample. If two or more observations on the tape have the same date and time, only the first such observation is used. The filtered data set consists of two time series: the bid rate and the corresponding ask rate most recently submitted on the Telerate wire at the specified time.

The volume statistics are taken from FRB-NY (1989). The data in the 1989 survey (in contrast to earlier surveys) are broken down in such a way that it is possible approximately to undo the double-counting of transactions. All non-brokered transactions between two domestic parties are assumed to be double-counted, and are divided by two. For brokered transactions, volume figures are taken from the brokers' report, which is not subject to the double-counting problem, because there is only one broker per transaction. Brokered transactions between two foreign parties are omitted as non-domestic trading. The seven elements of the benchmark vector, V^*, are as follows:

V_1^*, mean price: given the sample described above, prices, p_t, are calculated as the exponentiated midpoint of the natural logarithms of the bid and ask rates. That is, $p_t = \exp(p_t)$, where p_t is defined as in equation (10). The benchmark mean price of .525166 USD/DEM is the arithmetic average of all such prices in the sample.

V_2^*, mean spread: given the sample described above, spreads are the difference between the ask and bid rates. The benchmark mean spread of .000266 USD/DEM is the arithmetic average of all such spreads in the sample.

V_3^*, modal variance in price changes: for each day in the sample described above, the daily variance in price changes, V_3^+, is:

$$V_3^+ \equiv \frac{1}{\upsilon-1} \sum_{t=2}^{\upsilon} (p_t - p_{t-1})^2 , \tag{B1}$$

where υ is the number of observations in the sample for that day. A histogram is created by grouping these daily variances together into 27 bins. The variances ranged from .00000002986 to .0000005260 $(USD/DEM)^2$. The histogram is bimodal, with ten observations for each of the two ranges (.00000006,.00000008] and (.00000008,.00000010]. The benchmark modal variance in price changes is .00000008 $(USD/DEM)^2$.

V_4^*, modal serial correlation: prices are calculated as above. The daily covariance in price changes for each day on the tape, V_4^+, is:

$$V_4^+ \equiv \frac{1}{\upsilon-2} \sum_{t=3}^{\upsilon} (p_t - p_{t-1})(p_{t-1} - p_{t-2}) . \tag{B2}$$

The daily serial correlation is the daily covariance divided by the daily variance. Grouping serial correlations into 15 bins produces a histogram of daily serial correlations. The correlations ranged from -.3413 to .2118. The mode of the histogram is the bin (-.180,-.140], with 14 observations. The benchmark modal serial correlation in price changes is set at -.160.

V_5^*, market-maker volume: after adding together figures for commercial banks and non-bank financial institutions [FRB-NY (1989), tables II(a) and II(b)], and adjusting for double-counting, the volume of direct interbank spot USD-DEM transactions in the U. S. market for April 1989 was USD 306,332,000,000. There were twenty business days in April, 1989. The benchmark daily direct domestic market-maker volume is the monthly volume divided by twenty: USD 15,316,600,000.

V_6^*, broker volume: The volume figures were read directly from the brokers'

summary report [FRB-NY (1989), table II(c)], with one exception. Brokered transactions between two banks abroad were ignored. The brokered domestic spot USD-DEM volume for April, 1989 was USD 235,183,000,000. On a daily basis, this is USD 11,759,150,000.

V_7^*, customer volume: volume figures for commercial banks and non-bank financial institutions are added together [FRB-NY (1989), tables II(a) and II(b)]. All non-brokered transactions not concluded with another bank or non-bank financial institution were assumed to be with customers. The domestic spot USD-DEM customer volume for April, 1989 was USD 22,381,000,000. On a daily basis, this is USD 1,119,050,000.

These seven statistics form the benchmark vector, $V^*=(V_1^*,\ldots,V_7^*)$, which the computer model attempts to approximate. Corresponding statistics are generated by the simulation program. For the statistics on price dynamics, the simulation data are sampled at the exponential rate determined above (0.08), and the price statistics are calculated using the procedures for V_1^* through V_4^*.

The next step is to approximate the intractable functional form of the computer program by a Taylor series in the input-output parameters. A first-order approximation is used to simplify the search for the best parameter vector. Exploring the parameter space by assigning parameter values, $\rho^k=(\rho_1^k,\ldots,\rho_{13}^k)$, at random over a compact subset of the parameter space, the simulation generates a series of seven-dimensional output vectors, $V^k=(V_1^k,\ldots,V_7^k)$, where ρ^k and V^k are the kth input and output vectors, respectively ($k \in \{1,\ldots,100\}$). Seven ordinary least-squares regressions are run to fit each element of the program's output vector to the first-order Taylor series of the input vector. That is, the following models are fitted:

$$V_j^k = \psi_{j0}^k + \psi_{j1}^k \rho_1 + \cdots + \psi_{j13}^k \rho_{13} + \epsilon_j^k \ , \quad j \in \{1,\ldots,7\} \tag{B3}$$

where ϵ_j^k is a white-noise error term. The fitted regression equations define the response surface. The result is seven polynomials expressing the elements of the output vector as functions of the input parameters:

$$V(\hat{\psi}_j,\rho) \equiv \hat{\psi}_{j0} + \hat{\psi}_{j1}\rho_1 + \cdots + \hat{\psi}_{j13}\rho_{13} \tag{B4}$$

where $\hat{\psi}_j \equiv (\hat{\psi}_{j0},\ldots,\hat{\psi}_{j13})$ and $\rho \equiv (\rho_1,\ldots,\rho_{13})$. The seven functions, $V(\hat{\psi}_j,\rho)$, are henceforth treated as a substitute for the actual computer program.

The object is now to choose values for the elements of the parameter vector, ρ, such that the substitute functions, $V(\hat{\psi}_j,\rho)$, most closely approximate the real-world benchmark. It is possible to combine the seven-dimensional benchmark vector, $V^*=(V_1^*,\ldots,V_7^*)$, with the corresponding vector-valued output from the simulation runs, $V=(V(\hat{\psi}_1,\rho),\ldots,V(\hat{\psi}_7,\rho))$, into a scalar by calculating $\Delta(\rho)$, the square of the Euclidean distance between the two vectors, after scaling by the benchmark levels:

$$\Delta(\rho) \equiv \sum_{j=1}^{7} \left[\frac{V_j^* - V(\hat{\psi}_j,\rho)}{(V_j^*/\omega_j)} \right]^2 , \tag{B5}$$

where ω_j is a confidence weighting assigned to regression j, designed to give more emphasis in the optimization to those regressions for which the fit is better. The weights used here are:

$$\omega_j \equiv \frac{f_j}{\frac{1}{7}\sum_{j=1}^{7} f_j} \ , \qquad j \in \{1,\ldots,7\} \ , \tag{B6}$$

where f_j is the F-statistic from the jth regression. Note that for equal weighting (i. e., $\omega_j = 1 \ \forall \ j$), the distance formula (B5) becomes simply the sum of squared percentage deviations from the benchmark.

The distance is treated as a function only of the parameters, ρ. By optimizing this over the parameter space, that is, by calculating

$\rho^* \equiv \text{argmin}\ \Delta(\rho)$ -- such that ρ^* is contained in the interior of the estimation range -- the parameter set that best approximates reality (in the sense just defined) is achieved. Making the following four substitutions in (B5) and rearranging:

$$\gamma_{0j} \equiv \frac{V_j - \hat{\psi}_{0j}}{(V_j^*/\omega_j)}, \qquad j \in \{1,\ldots,7\},$$

$$\gamma_{ij} \equiv \frac{-\hat{\psi}_{ij}}{(V_j^*/\omega_j)}, \qquad i \in \{1,\ldots,13\},\ j \in \{1,\ldots,7\}, \tag{B7}$$

$$\rho_0 \equiv 1, \text{ and}$$

$$H_{ik} \equiv \sum_{j=1}^{7} \gamma_{ij}\gamma_{kj} \qquad i,k \in \{0,\ldots,13\},\ j \in \{1,\ldots,7\},$$

the distance function reduces to:

$$\Delta(\rho) \equiv \sum_{i=0}^{13} \sum_{k=0}^{13} H_{ik}\rho_i\rho_k \ . \tag{B8}$$

The reduced form coefficients, H_{ik}, together compose the Hessian matrix for the objective function. Indeed, the function is a pure quadratic form, and is completely determined by the Hessian. Given this, and given that the objective function is bounded from below at zero, the objective function is positive definite everywhere, and the existence of a unique global minimum is assured.

As a check on the validations, the equations (B4) are evaluated at the calibrated parameter values. The results appear in table B1. **[TABLE B1 APPROXIMATELY HERE]** With the exception of the modal serial correlation in price changes, the linear approximations (B4) are able to track the benchmark values quite well. The serial correlations, V_4, followed a noisy process, consequently produced low F-statistics in the regressions (B3), and hence received low weights (B6) in the optimization.