No.65

Foreign Exchange Market Maker's Optimal Spread with Heterogeneous Expectations

Ryosuke Wada

June 2000

Department of Economics
Otaru University of Commerce

Foreign Exchange Market Maker's Optimal Spread with Heterogeneous Expectations

April 1,2000

Name : Ryosuke Wada
Address : Otaru University of Commerce
Dept of Economics
3-5-21 Midori, Otaru
Hokkaido, 047-8501
Japan
Phone: 81-134-23-5109
Fax: 81-134-27-5213
e-mail: rwada@res.otaru-uc.ac.jp

Foreign Exchange Market Maker's Optimal Spread with Heterogeneous Expectations

I. Introduction

Trading volume and bid/ask spread have not shown any systematic relationships in foreign exchange market. The spread does not change in a way that existing models predict from the trading volume. Our paper aims at constructing a model with new analytical tools and using it to explain the varying correlation between the volume and the spread. Heterogeneity of expectations is a key element. We characterize it by a distribution function of a specific point on expected time path on which actions hinge. We formulate a market maker's optimization problem and derive his optimal spread. We show this optimal bid/ask spread increases, ceteris paribus, as the expectations become more heterogeneous among the FX dealers. The market maker widens his spread, because by doing so he can exploit the higher degree of the expectation's heterogeneity. Attributing the change in the spread to the heterogeneity of expectations contrasts our model to the existing literature. According to its models, the larger spread is to protect open position from higher risks of volatile periods.

It can be shown that, as the expectations become more heterogeneous the volatility increases as well,Wada(2000). Hence the spread and the volatility will show a positive correlation. On the other hand, the spread and the volume will not show a clear correlation. This lack of correlation is obtained by identifying sources of the trading volume. The trading volume is determined by the fundamentals' daily demand and supply as well as by frequencies of expectation renewals. These two elements can vary daily independently and their effects may be amplified or canceled out depending on the day. The correlation becomes obscure between the volume and the spread.

We apply our model on seemingly incoherent relationships between the spread and the trading volume as are reported in Bollerslev and Domowitz(1993). The existing literature, Admati and Pfleiderer (1988) and Subrahmanyam (1991), for example, provide either positive or negative correlation. Therefore their models can not explain the observed pattern of changing correlation. If we find effects of the heterogeneous expectations on both of the spread and the volume, the changing correlation is not puzzling.

In the following, we explain such seemingly incoherent changing patterns using our model in section II. A construction of the market maker's expected profit maximization problem and derivation of its solution are presented in Appendix 1 and 2. The optimal solution takes a form of an optimal spread. This optimal spread is larger as FX dealers' expectations become more heterogeneous. This is the key relationship to explain the empirical observations.

II. Implications of the Model

A. Major Results on Empirical Observations

Bollerslev and Domowitz (1993) report lack of systematic relationship between spread and "market activity" in FX market. The market activity is meant to be a renewal of quotes. Its number of times could be a proxy for the trading volume. Their findings indicate firstly that the spread is larger in a period before lunchtime than in a period that follows even if these two periods may be similarly active. Secondly, if we compare lunchtime and late afternoon, lunchtime has fewer activities and much larger spread. Thus the spread and the volume do not show any systematic relationships empirically.

The heterogeneity of expectations is the key element to reconcile the above observations. Firstly, our paper identifies banks' retail transactions as one of the sources of uncertainty in inter-bank transactions. Since the dealers adjust their positions eventually after they have retail transactions, these retail transactions give rise to the inter-bank transactions. The retail transactions matter. So the dealers try to estimate an intraday pattern of the aggregate retail transactions. Their estimates will be heterogeneous. A degree of the heterogeneity varies depending on time and date. Their estimates are more heterogeneous in the morning when banks' retail transactions are taking place than the afternoon when the retail transactions are finishing. Hence, the expectations are more heterogeneous and the spread is larger in the morning, even though the morning and the afternoon may have the similar trading volumes.

In inactive periods such as lunchtime, weekends and days before holidays, competitions among market makers decrease. Many of them voluntarily drop out from the auction process. At the same time, the expected number of retail transactions and hence that of the inter-bank transactions also decrease. If we compare the above inactive periods, then the lunchtime is followed by the most uncertain period. The expectations are the most heterogeneous among those periods. Hence, the lunchtime has the largest spread.

B. Model's Characteristics

A process of price formation in the foreign exchange market poses a couple of theoretical difficulties which existing literature has not provided satisfactory models. One of the difficulties is **continuous auction**. In a continuous auction, there is not a specific length of time to define demand and supply and it is not clear which price is equalibrium among the series of realized transaction prices. It is hard to apply usual equalibrium analysis on the continuous auction. As a tool to formulate the continuous auction as in Garman (1976) and Amihud and Mendelson (1980), our model defines **expected numbers of buyers and sellers per unit time** in stead of demand and supply. In the continuous auction, it is not obvious how to compare heterogeneous expectations. There is not a specific point of time to do it.We identify **the first peak or bottom on the expected time path** as a key value to compare heterogeneous expectations. The key value is the first local extremum on the expected time path. We characterize the heterogeneity of the expectations by a distribution function for this expected extremum. Taking advantage of as many such extremums as possible increases the expected profit. Dealer's action to seek capital gain hinges on it. Therefore the distribution function describes distributions of actions.

Participants in a FX market are dealers. Some of them act as market makers. For given values for bid and ask, sellers and buyers randomly arrive at the market maker and trade with him. These sellers and buyers are other dealers. Their expected numbers of arrivals are determined by the distribution of the reservation prices and the competitions among the market makers. Each time **an arrival** occurs, the market maker's position changes. Revenue or cost is incurred by the arrival. The market maker's position follows a **continuous time Markov process**. The levels of his position constitute "**states**". The arrival causes a transition between the states. The transition brings about *"a reward"*. The market maker is not always passive in that he can initiate transactions with other market makers to adjust his position. Therefore the process is controlled. This process is modeled as a **continuous time controlled Markov process with reward** as in Yushikevich(1977). The arrival process of sellers and buyers is characterized by an "**infinitesimal matrix**", which is a continuous time counterpart to a Markov transition probability matrix. Elements in the infinitesimal matrix are the expected numbers of arrivals for given values of bid and ask. Values of these elements are determined by the distribution of the reservation prices and by the number of competitors. He tries to maximize the expected daily profits by choosing the quotes and hence the spread. The model shows that this optimal spread increases, ceteris paribus, as the expectations become more heterogeneous.

Garman (1976), Amihud and Mendelson (1980) and our model share the model specifications such as **price sensitive random arrivals**. We improve these preceding models by identifying sources of the price sensitive random arrivals.

Intra-day flow of the excess demand for FX by the economy's fundamentals will not be balanced. The dealers as a whole adjust net total positions and absorb the excess demand. The transaction price makes the excess demand just being absorbed. The dealers try to figure out patterns of the excess demand and to estimate forthcoming peaks and bottoms of the transaction prices. The heterogeneity of expectation comes from differences in opinions with regard to the arrival patterns of retail transactions and the distribution of the expected peaks and bottoms.

C. Characteristics of Optimal Spread

1. Intensity Curves of Buyers and Sellers

In Appendix 1, ingredients of the market maker's expected profit maximization are discussed. And in Appendix 2, from the arrival intensity buyers and sellers as functions of the spread, the necessary condition of the optimal spread is derived. We discuss the implications of the optimal spread.

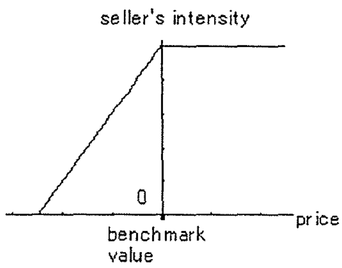

Figure 1: Seller's Intensity Curve

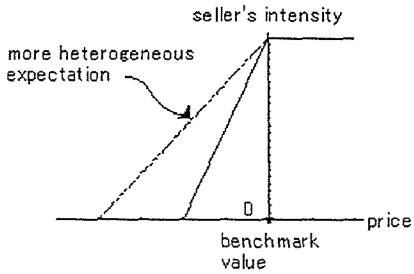

Figure 2: Effect of Heterogeneous Expectations

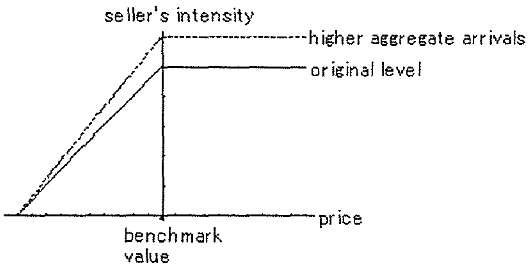

Figure3: Effect of Higher Aggregate Arrivals

The arrival intensity is the expected number of arrivals for a given value of quoted price **Benchmark value** is a weighted average of all the indications and the reported last transaction prices. Dealers use the benchmark value as an index to signify the current transaction price. When one of the dealers contacts the market maker and ask for prices for an immediate execution, probability that the dealer accept quoted price depends on closeness to the benchmark value. Closer to the benchmark value, higher the probability that a transaction takes place. Arrival intensity is higher (Figuere1). For a given closeness, if the dealers' expectations are more heterogeneous, then the market maker sees higher probability of realizing transactions. So the higher degree of expectation heterogeneity results in higher intensity. However, the highest intensity is unchanged, if the total arrivals are unchanged (Figure 2). Then if the aggregate arrivals increase while the heterogeneity of expectations stay the same, then arrival intensity for the market maker increase. However the least competitive price to have arrivals stay the same (Figure 3). The change in a degree of competitions among the market makers has the same effect as Figure 3. Even if the aggregate arrivals over the entire market stay the same, the potential arrivals to an market maker who is still ready to trade.

The aggregate arrivals increase as the frequency of dealers' expectation renewals increases. And also if the retail arrivals increase, the aggregate arrivals increase. The market maker faces such buyer's and seller's intensity curves. He tries to maximize expected daily profits. For simplicity, we assume that the market maker chooses a pair of buying and selling prices with the same arrival intensities(Figure 4)

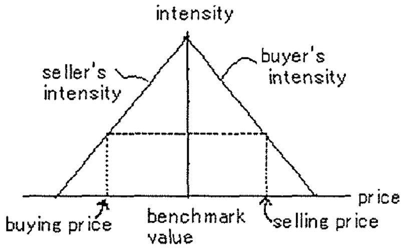

Figure 4 Intensity Curves and Choice of spread

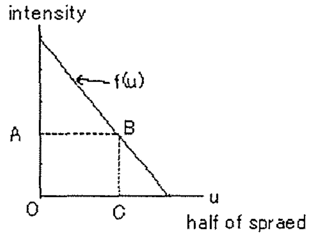

Figure 5 Determination of Optimal Spread

The highest intensities of buyer's and seller's curve need not be the same. If we assume the same retail arrivals, then the market maker's problem becomes choosing the profit maximizing spread.

2. Necessary Condition for Optimal Spread

As benchmark value moves, the intensity curves shift for a given value of quoted price. The market maker revises his buying and selling prices accordingly. We assume that the market maker chooses a price pair with the same arrival intensities. The solution of the expected profit maximization problem becomes choosing the optimal spread. Let u be a half of the market maker's bid/ask spread. Then we express the arrival intensity as a function of u. Let f(u) be that function (Figure 5). Using this function, the expected profit maximization is solved. The solution is derived in Appendix 2. The necessary condition for the optimal spread is given by $u f' + f = 0$. This result is obtained for an asymptotic case such as time goes to infinity. This is not unreasonable situation for the FX market. It is not unusual that several transactions take palace in one minute. If we consider second as unit of time, then it is reasonable to apply the asymptotic result on the intra-day auction process. The necessary condition for the asymptotic case coincides that for the static case. Since the value of f(u) is an expected number of arrivals, rectangle ABCO in Figure 5 is the expected profit for an infinitesimal amount of time. The necessary condition tells that choose the value of u such that the area of ABCO will be maximum.

3.Comparative Statics of the Optimal Spread

effect of heterogeneity of expectation: A slope of f(u) becomes less steep. The optimal value of u becomes larger. Hence the optimal spread is larger.

effect of aggregate arrivals: The slope of f(u) becomes steeper while an intercept with the horizontal axis is fixed. The optimal *u* increase and hence the spread increases. The aggregate arrivals are generated by two sources. The first is renewals of dealers' expectations. The second is retail transactions with customers. The second source is demand and supply from the macro economy. Arrivals from the first source are called "heterogeneity arrivals" and the second "retail arrivals".

4. Application on Intra-Day Pattern of Spread

(1) Morning and Afternoon

The retail arrivals are supposed to show clear intra-day patterns. Their number is large in the morning toward noon lunchtime. Then there is a pause in the lunchtime. After the lunchtime, it surges again and tapers off toward the end of business hours.

The random arrivals create fluctuation of transaction prices. The larger expected number of arrivals is associated with larger variance. (This is one of the mathematical characteristics of Poisson process. We assume Poisson process for the arrival processes.) Hence, the trading volume and volatility has a positive correlation. Since the larger fluctuation is anticipated, a degree of expectation heterogeneity is higher. Hence, the optimal spread is larger for larger trading volume. Even if the trading volumes are same before and after the lunchtime, a degree of the expectation heterogeneity may be different. In the afternoon the retail transactions come in a more predictable way because of communications between the customers and dealers. The heterogeneity is lower in the afternoon. As a result, even if the trading volumes are the same, the optimal spread is smaller in the afternoon.

(2) Lunchtime

At lunchtime, many dealers voluntarily drop out from the auction process. Retail transactions become low. If an market maker is still ready to trade, buyer's and seller's intensity curves for him shift as follows. They shift up in a manner of Figure3 due to the decreased level of competition. The optimal spread is wider. Then the intensity curves shift down due to decrease in retail arrivals. The spread is narrowed. If the effect of decreased competitions is larger, then the optimal spread is wider.

D. Conclusions

Auctions in FX market are continuous time. Transactions take place asynchronically. Equalibrium analysis is difficult to apply. We introduce expected number of arrivals of buyers and sellers for an infinitesimal amount of time. These expected numbers are called arrival intensities. The arrival intensities play a role of demand and supply in the continuous auction. We construct "buyer's and seller's intensity curves". The market maker faces these curves. For given arrival curves, he tries to maximize his expected profit. The expected profits are maximized with respect to bid/ask spread. We obtain the necessary condition for the optimal spread.

The optimal spread increases as the dealers' expectations become more heterogeneous. Even if the trading volumes are similar, the optimal spread will be larger if the degree of the heterogeneity is higher. If the degree of competition among the market makers is lower, the optimal spread is larger. Using these implications of the model, we can resolve the seemingly incoherent changes of correlation between the trading volume and the spread.

Appendix 1. Construction of the Optimization Problem

A. Environments of the Market

1. Two Groups of Dealers

Price" means spot foreign exchange rate. Auction participants are foreign exchange dealers. There are two groups of dealers. The first group consists of market makers. They quote their own prices and stand ready to trade at them. The second group consists of those who do not quote own prices. Transactions take place between the first and the second groups or within the first group. All the dealers can assume open position. Sizes of the open positions are subject to exogenouly imposed constraints. The market makers are allowed to assume larger open positions. The constraints are two transaction units for the market makers and one unit for other dealers. We call those in the second group with single unit constraint "S-dealer", when distinction is necessary for clarity. The FX market is wholesale market. All of the dealers have retail customers and trade with them as well. All the dealers are risk neutral. If S-dealers recognize possibility of capital gain, his position must be one unit open. Contrary, the market makers' positions do not necessarily reflect their view. Their positions are exposed random fluctuations in order to seek profits from bid/ask spread.

2. Continuous Auctions

Auction process goes on continuous time. A market maker quotes a pair of his buying and selling prices, whenever other dealers ask for quotations. These quoted prices are for immediate executions. If one of the prices is good enough to the inquiring dealers, then transactions take place. Transactions occur asynchronously. When a transaction takes place, we say "an arrival of buyer or seller" occurs and count one arrival. The FX market opens in the morning and ends late afternoon. The dealers' daily profits are evaluated at the end of business hours. When retail customers sleep, our market is closed.

B. Ingredients for Market Maker's Optimization Problem

1.Choice variables

The market maker acts to maximize his expected daily profits. His choice variables are a pair of buying and selling prices, and his position. His expected profit maximization problem can be viewed as an inventory control problem. The size of the open position corresponds to an inventory. He can choose only either a price pair or inventory level. He can control either the prices or the inventory at one time.

He gives out the pair of buying and selling prices together when inquired. This is the practice called "two way quotation" and a rule of the market. The inquiring dealer decides which side to trade. The market maker could differentiate competitiveness of his prices so that a particular side is more likely to be chosen. Still he does not have complete control over his own position. It is so unless he quotes absurd prices. Thus if he chooses prices then he cannot control his inventory. If he wants to adjust inventory right away as he wishes, he can do so by trading with other market maker. In that case, the price applied is someone else's. Thus if he chooses his inventory, he cannot control the price.

2. A Little Lagged Inventory Control

The market maker's position is subject to a constraint. Meanwhile, when he has an arrival, a constraint may be violated. In order to reconcile such inconsistent situations, we distinguish "a desired " value and an actual value. And we allow the actual value to be different from the desired value momentarily. The constraint is defined on the desired level of inventory. The market maker chooses actions so that the desired position satisfies the constraint. If the constraint is violated due to the random arrivals, he immediately trades with other market makers. Then the constraint becomes just binding.

3. Buyer's and Seller's Intensity Curves

For a given value of quoted price, an expected number of arrivals of buyers or sellers can be defined. The expected number of arrivals for an infinitesimal amount of time is called "arrival intensity". The arrival intensity is price sensitive, in a very much the same way as demand and supply curves. We call relationships between the price and the arrival intensity "an intensity curve". We have "buyer's intensity curve" and "seller's intensity curve". The market maker daces two intensity curves and chooses actions.

C. Generation of Arrivals of buyers and sellers

1.Inside and Outside Sources of Arrivals

It was defined that an arrival of buyer or seller" implies that a transaction take place with a market maker. Two sources generate these arrivals, one within the market and the other outside. The first source is dealers' revises of heterogeneous expectations. The second source is retail transactions. The second source is, in another words, demand

and supply from the macro economy.

2.Generations of Arrivals by Expectation Renewals

All the dealers have heterogeneous expectations and revise them from time to time. S-dealers assume open positions according to their expectations. When they revise expectations, they try to adjust their positions accordingly. Thus the expectation renewals generate arrivals.

The difference in expectations can be expressed as the difference in the first peak or bottom in an expected time path of transaction prices. Possible capital gain hinges on such extemums. Becoming a buyer or seller, depends on them. The difference with this regard results in the difference in actions. The heterogeneity of expectations takes a form of distribution of such extremums. We call a distribution function for it "heterogeneity distribution". We call arrivals generated by the expectation renewals "heterogeneity arrivals". For a given quoted pair of prices, the heterogeneity distribution determines probability that the next heterogeneity arrival is a buyer or seller. The total numbers of the heterogeneity arrivals are generated by the expectation renewals. These arrivals are sorted into buyers and sellers according to the heterogeneity distribution.

4. Arrival Generation by Retail Transactions

The dealers engage in retail transactions with their customers as well as wholesale transaction in the market. Dealers are ready to trade during the business hours whenever customers want. Meanwhile S-dealers must have constructed the desired levels of positions according to their own expectations. The randomly arriving retail transactions disrupt the already constructed positions. Had this occurred, the dealers would counterbalance retail transactions to recover the desired positions. They trade in the market. Thus a sequence of the retail transactions changes into a sequence of the arrivals in the market. We call the arrivals generated by the retail transactions "retail arrivals".

5.Price Insensitive Aggregate Retail Arrivals

Arrival process is defined for the entire market and for the individual market makers. We call retail arrivals aggregated over the entire market "aggregate retail arrivals". It is assumed that the dealers' customers do not respond to intra-day price movements. Hence the aggregate retail arrivals are price insensitive. Since the intra-day price movements do not influence retail customers, daily demand and supply from them would not be exactly equal, except for by chance. Therefore the daily accumulative numbers of the buyers and the sellers would not be equal. The dealers absorb the difference. The aggregate of the entire dealers' position would be open overnight.

D. Constructing Intensity Curves

1.Poisson Processes

We assume that the retail arrivals and the heterogeneity arrivals constitute Poisson processes. It means that the number of arrivals follows Poisson distribution, that its expected number for a given time interval is proportionate to a length of the interval and that its variance is equal to the expected value. The buyer's and seller's retail arrivals constitute distinct two Poisson processes. The aggregate retail arrivals are price insensitive.

The heterogeneity arrivals, buyers and sellers combined together, constitute a unique Poisson process. We call these heterogeneity arrivals aggregated over the entire market "aggregate heterogeneity arrivals". The aggregate heterogeneity arrivals are also price insensitive. This insensitiveness comes from that the frequency of the expectation renewals, not a price level, determines the number of the aggregate heterogeneity arrivals. Thus, there are three Poisson processes for the aggregate arrivals and none of them are influenced by the price movements.

2. Price Sensitive Process for Individual Market Makers

A fraction of the aggregate arrival processes reaches individual market makers. More precisely speaking, it is a fraction of a sum of Poisson processes. The fraction of the combined aggregate processes appears as two arrival processes of buyers and sellers. The fraction changes as the quoted price changes. Due to competitions, the arrival processes for a given market maker become price sensitive.

3. Price Search

The market makers always post "indications". These indications are intended to be approximate values. An exact price applied in each transaction is known only when the dealers inquire the market makers about them. Besides the indications are not updated continuously. Even if one dealer tries an extensive search, some of the quotations may change before he completes the search. Thus, it is not clear what is the best price available at a given moment. The dealers shop around but their searches are not complete. They accept prices randomly. The probability of the acceptance depends on how good the price is compare with some benchmark value.

4. Benchmark Value

The dealers always monitor the market makers' indications. And information vendors distribute real time transaction prices to the dealers. These reported prices come from multiple sources and are anonymous. We assume that dealers use a weighted average of the indications and the last reported prices as a benchmark value. It is weighted average of pair prices of the all the indications and the last reported transaction prices. The weights are up to the individual dealers. The benchmark value signifies the current price to the dealers. The market makers use the benchmark value to set own prices. All the dealers use benchmark value to calculate retail prices. The retail prices are benchmark value plus or minus fixed margin.

5. Arrival Intensity Curves

The market makers have prices for immediate execution. These prices can change any moment. Hence they are not necessarily same as their indications. It is not clear what are the best prices in the market for a given moment. So dealers use the benchmark value as a substitute for them. When dealers search for the best price for the immediate execution, they accept the market maker's price if they judge it is close enough to the benchmark value. For a given benchmark value, it looks as if the inquiring dealers accept the price with some probability. This probability decreases as the distance between the benchmark value and the market maker's price. This decreasing schedule itself becomes a relationship between the arrival intensity and the quoted price. "The buyer's intensity curve" is defined to the right of the benchmark value and is downward sloping. "The seller's intensity curve" is defined to the left of the benchmark value and upward sloping.

D. Simplified Inventory Control Problem

1. Equal Intensity for Buyers and Sellers

We consider only rather simple case such that a chosen pair has the same arrival intensity. It is assumed that the market maker chooses price pair with equal arrival intensities. It means that when he has open position, he does not differentiate the arrival intensities between buyers and sellers so that the position is more likely to move in a particular direction. It is assumed that market makers take other market makers' indications as given. The indications are truly available price pair when they are revised and quoted. The revisions of the indications occur with random intervals of time. We do not consider the cases such that the market makers manipulate own indications so as to influence the dealers' expectations.

2. Necessary Condition for Optimality

The market maker shifts around his buying and selling prices while keeping their arriving intensities equal. Then the arrival intensity curves are derived as functions of spread. Let u be the half of the spread. And let $q = f(u)$ be intensity curve as a function of u. In the appendix the necessary condition for the expected profit maximization is obtained. It is given by $f' u + f = 0$.

Appendix 2. Derivation of Optimal Spread

1. Profits and Its Expected Value

A market maker wants to maximize his expected profit for a time interval $[0, T]$. This maximization problem can be formulated as follows. Random variables change values from time to time in continuous time. We use integrals to express traded quantities. $Z_1^*(t)$ and $Z_2^*(t)$ are accumulative values that he bought and sold by time t. Hence $dZ_1(t)$ and $dZ_2(t)$ are quantities traded at each transaction at time t. S_1^* and S_2^* are buying and selling prices applied at time t. Using integration notation, $\int_0^T S_1^*(t)\, dE[\, Z_1^*(t)$ means total quantity bought for a time interval $[0, T]$. Similarly $\int_0^T S_2^*(t)\, dE[\, Z_2^*(t)\,]$ means total quantity sold during the interval $[0, T]$. $Z(t)$ is inventory level at time t. Positive value means long position. W_T expresses the position's value at the end of the day. $\mathcal{F}_{t_0}$ is information. Conditional expectation on profits for interval $[0, T]$ is to be maximized. He can choose S_1^* and S_2^* and, when necessary, $Z(t)$ is adjusted. He is risk neutral.

The stared variables are desired values of the original variables. It is allowed that the actual values are different from the desired values temporarily. The market maker's position is exposed random shock of transactions while he has to satisfy the constraint on the position. Temporary discrepancy is allowed. Had it occurred, he trades with other market maker and makes his position just binding.

2. Omitting Retail Transaction

The market maker has retail transactions too. Profits from the retail transactions become implicit in the maximization problem. His retail prices and wholesale prices are always different by constant margin. As the number of the retail transactions increases, his profits from the margin increase. This is exogenously given process. Profits as market maker is realized when he trades with other dealers.

$$\max_{S_1^*, S_2^*} E\Big[\int_0^T S_2^*(t)\, dZ_2^*(t) - \int_0^T S_1^*(t)\, dZ_1^*(t) + W_T(\ Z(t)\ )\ |\ \mathcal{F}_{t_0}\Big]$$
$$= \max_{S_1^*, S_2^*}\Big\{\int_0^T S_2^*(t)\, dE[\,Z_2^*(t)\,] - \int_0^T S_1^*(t)\, dE[\,Z_1^*(t)\,]$$
$$+ E\Big[W_T(\ Z(t)\ )\ |\ \mathcal{F}_{t_0}\Big]\Big\}. \quad (1)$$

subject to $|\,Z^*(t)\,| \leq 2,$ for $0 \leq t \leq T.$

where $W_T(\,Z(t)\,)$ is final reward at time T , having $Z(t)$ of position at time t.

3. Model Specifications

This is a model of a controlled, continuous time Markov process with finite states. This model is presented in Yushkevich (1977). We look for an optimal policy among the time invariant policies. The position denoted as $Z(t)$ is the "state" variable. The arrival processes of buyers and sellers are Poisson processes. These random arrivals cause transitions of the state. As a transaction takes place, the state jumps from a given value to another. We introduce two simplifying assumptions.(A-1) Each arrival's quantity is either one or two units. The transaction incurs cash flow. The jump of the state has an associated "reward". The position is subject to the constraint of 2 units. We consider simple policy such that (A-2) The market maker always chooses price pair with the same arrival intensities. It means that the next arrival is equally likely to be a buyer or a seller. We construct an **"infinitesimal matrix "**. This is a continuous time counterpart of Markov transition matrix. Let Q be the infinitesimal matrix. The market maker chooses prices so that buyer's and seller's arrival intensities are the same. Let q be that value of the intensity.

The quantity associated with each arrival is either one or two units. We denote probabilities that the quantity is one unit and two units by v_1 and v_2 . State space is $\{1, 2, 3, 4, 5\}$. Let S_i denote state i. These state are defined to correspond to the inventory levels $\{2, 1, 0, -1, -2\}$. If buyer of one unit arrives, the inventory decreases by 1 and state moves down by 1. Let Q be the infinitesimal matrix of the process. Using the notation defined above, Q is written as

$$Q = qA \equiv q \begin{pmatrix} -1 & v_1 & v_2 & 0 & 0 \\ 1 & -2 & v_1 & v_2 & 0 \\ v_2 & v_1 & -2 & v_1 & v_2 \\ 0 & v_2 & v_1 & -2 & 1 \\ 0 & 0 & v_2 & v_1 & -1 \end{pmatrix} \tag{2}$$

The raw number is the state before the transition and the column number is the state after that. The interpretation of the matrix is that element $Q_{i,i+1}$ means the expected number of seller's arrival who trades one unit when the state is S_i. The element $Q_{j,j}$ means expected number of arrivals combining buyers and sellers together when the state is S_j. It is expected number of arrivals to force the market maker to leave the state S_j. It has negative value. The probability to stay at j decreases by $Q_{j,j}$ for an infinitesimal amount of time. For $j = 1, 5$, the element's value is $-q$. When the state is at S_1 and S_5, the constraint is already has been just binding. If the position goes out the constraint, he would cancel out excess quantity by counterbalancing transactions right away. So the states 1 has virtually only buyers and state 5 has only sellers. The value of the element is half.

4. Omitting Expected Capital Gains

Since we consider the simple case such that the market maker chooses only the price pair with the same arrival intensities. If he wants to maintain equal arrival intensities, he has to keep shifting around the price pair as the benchmark value, i.e. the current price level, fluctuates. S_1^* and S_2^* have to shift around. Then the market maker has capital gains and losses. We separately calculate profits from the bid/ask spread from the net capital gains. Then it can be shown that we can eliminate the expected capital gains from the expected profit maximization problem. If we separate the capital gains and bid/ask profits, the profits can be written as follows.

$$\Sigma_{k=1}^{N} |I_k| u + I_k \xi_k \tag{3}$$

where N is the number of transaction, ξ_k is the average of bid/ask when kth transaction takes place and I_k indexes whether it is purchase or sale. For simplicity, suppose each quantity is one unit. $I_k = 1$ if it is sale and $I_k = -1$ if it is purchase. ξ_k is approximately equal to the benchmark value. Since we consider the case the arrival intensities of buyers and sellers are kept equal, $I_k = 1$ with probability $\frac{1}{2}$ and hence $E[I_k] = 0$. The if we take expected value of equation (3), then it is given by $uN + \Sigma_{k=1}^{N} E[I_k \xi_k]$. For each k, $E[I_k \xi k] = E[I_k]E[\xi_k] = 0$ holds. The capital gains are expected to be zero, when the dealer maintains equal arrival intensities for buying and selling prices and if he starts out with no inventory. For simplicity, we assume that his initial inventory is zero. Then we can eliminate capital gains from the expected profit maximization problem. As a result, finding a pair of S_1^* and S_2^* is the same as finding the optimal bid-ask spread $2u^*$.

5. Differential Equation of Reward

When the inventory increases by a seller's arrival, we call the transition of state "upward jump". Each jump is associated with a reward. A jump from S_1 to S_2 means one unit of sale. If the state is S_2, the expected reward from the buyer's arrival is given by $v_1 + 2v_2$, . An upward jump caused by seller's arrival incurs cost. We do not have to assign a negative reward to the upward jump, however. Except for the last downward jumps at the end of the day, the downward jump already signifies a profit, not just revenue. Let R be the 5×1 vector and denote a vector of the expected rewards of jumps, i.e.,

$$R \equiv 2uq \begin{pmatrix} v_1 + 2v_2 \\ v_1 + 2v_2 \\ v_1 + 2v_2 \\ 1 \\ 0 \end{pmatrix}. \tag{4}$$

Let the 5×1 vector $W(t)$ denote the expected profit for an time interval $[t, T]$ where T is the end of the day. The ith element of $W(t)$ is the expected profit for the case that at time t the state is S_i. The expected profit satisfies the following differential equation.

We can calculate the expected profit as the solution of the following differential equation (Yushkevich 1977, Corollary 3.2 and Supplementary Remarks 5.),

$$W'(t) = -R - QW(t) \tag{5}$$

with the boundary condition $W(T) = W_T$. (4) is an example of Bellman's equation. The elements of R and Q are constant. Since the sum of the elements in each row of Q is zero, Q is singular. The solution of (5) is given by

$$W(0) = \Big(\int_0^T e^{sQ} ds\Big) R + W_T \tag{6}$$

where $e^{sQ} \equiv I + \frac{sQ}{1!} + \frac{(sQ)^2}{2!} + \frac{(sQ)^3}{3!} + \ldots + \frac{(sQ)^k}{k!} + \ldots$ and I is an identity matrix. Since Q is not invertible, the expression for $\int_0^T e^{sQ} ds$ cannot be simplified. $\int_0^T e^{sQ} ds = T\Big(I + \frac{TQ}{2!} + \frac{(TQ)^2}{3!} + \frac{(TQ)^3}{4!} + \ldots + \frac{(TQ)^{k-1}}{k!} + \ldots\Big)$.

6. Intensity Curves as Functions of Spread

The common arrival intensity of buyers and sellers is q . The value of the arrival intensity q depends on u. We construct the intensity curve as follows. When the dealers ask for market maker's quotations, they compare the quoted price with the benchmark value. The benchmark value is a weighted average of all the indications and reported last transaction prices. Therefore the arrival intensity monotonically decreases as the distance increases between the benchmark value and the quoted price. Then we can derive a relationship between this distance and the arrival intensity. We can do this for both buyers and sellers. We have two intensity curves. They are functions of the distance defined as above. Next, fix the value of the intensity and, using intensity curves find the values of distance. Take a sum of two values of distance. It is the bid/ask spread. If you divide it by 2, then the answer is u. Then for a given value of u, there is unique value of arrival intensity. Thus we have arrival intensity curves as functions of the spread or half of the spread u.

7. Necessary Condition for Optimal Spread

As shown in (3), R is also a function of u. The expected profit is given by equation (6) and this equation is maximized with respect to u. Next we find optimal spread. Let $q = f(u)$ be an intensity curve as a function of u, one half of the spread. Substitute $f(u)$ and (4) into (6) and differentiate with respect to u. Since $\left(\int_0^T e^{sQ} ds \right) R = T\left(qI + \frac{q(TqA)}{2!} + \frac{q(TqA)^2}{3!} + \ldots + \frac{q(TqA)^{k-1}}{k!} + \ldots \right) u2V$, where $V \equiv \frac{1}{2uq} R$ and $Q = qA$, the derivative of the first term of equation (6) is given by

$$\frac{d}{du}\Big[\int_0^T e^{sQ} ds R \Big]$$
$$= T\Big\{ Iq' + (TqA)q' + \frac{(TqA)^2 q'}{2!} + \frac{(TqA)^3 q'}{3!} + \ldots + \frac{(TqA)^{k-1} q'}{(k-1)!} + \ldots \Big\} u2V$$
$$+ T\Big\{ +qI + \frac{q(TqA)}{2!} + \frac{q(TqA)^2}{3!} + \ldots + \frac{q(TqA)^{k-1}}{k!} + \ldots \Big\} 2V$$
$$= \Big\{ Te^{TQ} q'u + q\Big(\int_0^T e^{sQ} ds \Big) \Big\} 2V. \qquad (4-13)$$

The optimal value of u makes (7) equal to zero. Te^{TQ} and $\int_0^T e^{sQ} ds$ become proportional to T, as $T \to \infty$. The proof is given later. Then for $T = \infty$. equation (7) $= 0$ holds if $q'u + q = 0$. Thus the necessary condition for the optimal spread u^* for the asymptotic case is given by

$$f'u + f = 0. \qquad (8)$$

8. proof of asymptotic proportionality

e^{tQ} is the probability distribution of the state in which Z would stay at time t, starting at time 0 from one of the states. The row gives the starting state and the column gives the state at time t. e^{tQ} converges to a stationary distribution, as $t \to \infty$. Let C be the stationary distribution which is associated with Q , i.e., $C = \lim_{t \to \infty} e^{tQ}$. We can show that $\frac{1}{t} \int_0^T e^{sQ} ds$ also converges to C. For a given $\epsilon > 0$, , there exists τ such that $\|e^{tQ} - C\| < \epsilon$ for $t > \tau$. $\| \cdot \|$ represents the maximum of the absolute values of the elements in a matrix. Using the facts that $\frac{1}{t} \int_0^t e^{sQ} ds = \frac{1}{t} \int_0^\tau e^{sQ} ds + \frac{1}{t} \int_\tau^t e^{sQ} ds$ and $\frac{1}{t} \int_0^t e^{sQ} ds - C = \frac{1}{t} \int_0^t (e^{sQ} - C) ds$, we have $\lim_{t \to \infty} \frac{1}{t} \int_0^t (e^{sQ} - C) ds =$

$\lim_{t\to\infty} \frac{1}{t}\int_0^{\tau}(e^{sQ}-C)ds + \lim_{t\to\infty} \frac{1}{t}\int_{\tau}^{t}(e^{sQ}-C)ds$. Because $\lim_{t\to\infty} ||\frac{1}{t}\int_0^{\tau}(e^{sQ}-C)ds|| = 0$ and $\lim_{t\to\infty} ||\frac{1}{t}\int_{\tau}^{t}(e^{sQ}-C)ds|| \leq \lim_{t\to\infty} \frac{1}{t}\int_{\tau}^{t}||(e^{sQ}-C)||ds \leq \epsilon$, we have $||\frac{1}{t}\int_0^t e^{sQ}ds - C|| \leq \epsilon$ and

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t e^{sQ}ds = C = \lim_{t\to\infty} e^{tQ}ds \tag{9}$$

Eequality (9) shows Te^{TQ} and $\int_0^T e^{sQ}ds$ become proportional to T, as $T \to \infty$ Divide (7) by T, let $T \to \infty$, and in (9) replace t by T, then (7) becomes zero, if $q'u + q = 0$.

Reference

Admati, A.R. and P. Pfleiderer, 1988, "A theory of intraday pattern: Volume and price variability," *Review of Financial Studies* 1,3-40

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

Bollerslev, Tim and Domowitz, Ian, 1993, "Trading Patterns and Prices in the Interbank Foreign Exchange Market," *The Journal of Finance*, Vol.48,No4, 1421-1443

Garman, Mark B, 1976 "Market Microstructure," *Journal of Financial Economics* 3, 257-275

Subrahmanyan, A., 1991, "Risk aversion, market liquidity, and price efficiency," *Review of Financial Studies* 4, 417-441

A.A. Yushikevich, 1977, "Controlled Markov Models with Countable State Space and Continuous Time", *The Theory of Probability and Its Applications*, Vol.22, No.2, 215-235

Wada, Ryosuke ,2000, "Microstructure Model of Trading Volume and Volatility in FX market", mimeo

This Discussion Paper Series is published by the Center for Business Creation(changed from the Institute of Economic Research on April 1999) and integrates two old ones published separately by the Department of Economics and the Department of Commerce.

Discussion Paper Series
Institute of Economic Research
Otaru University of Commerce

No.	Title	Author/s	Date
1.	ﾎｰｷﾝｽﾞ=ｻｲﾓﾝの条件に関する諸説の統合について	ﾀﾞｽｸﾞﾌﾟﾀ,ﾃﾞｨﾊﾟﾝｶｰ	Jul.1992
2.	Motivation and Causal Inferences in the Budgetary Control	Yoshihiro Naka	Aug.1992
3.	Проблемы управления рабочей силой на предприятиях Далънего Востока(социологические аспекты)	Анатолий Михайловичн Шкурки	Nov.1992
4.	Dynamic Tax Incidence in a Finite Horizon Model	Jun-ichi Itaya	Jan.1993
5.	Business Cycles with Asset Price Bubbles and the Role of Monetary Policy	Hiroshi Shibuya	Jun.1993
6.	Continuous Double-Sided Auctions in Foreign Exchange Markets	Ryosuke Wada	Aug.1993
7.	The Existence of Ramsey Equilibrium with Consumption Externality	Sadao Kanaya & Tomoichi Shinotsuka	Sep.1993
8.	Money, Neutrality of Consumption Taxes, and Growth in Intertemporal Optimizing Models	Jun-ichi Itaya	Nov.1993
9.	Product Returns in the Japanese Distribution System:A Case Study of a Japanese Wholesaler's Return Reduction Efforts	Jeffery Alan Brunson	Mar.1994
10.	Dynamics, Consistent Conjectures and Heterogeneous Agents in the Private Provision of Public Goods	Jun-ichi Itaya & Dipankar Dasgupta	Jun.1994
11.	Intra-industry Investment and Imperfect Markets A Geometric approach in Simple General Equilibrium	Laixun Zhao	Oct.1994
12.	Sit-Down to Split:Flint GM Workers in 1937-1939	Satoshi Takata	Dec.1994
13.	The Complementarity between Endogenous Protection and Direct foreign Investment	Laixun Zhao	Feb.1995
14.	Consumption Taxation and Tax Prepayment approach in Dynamic General equilibrium Models with Consumer Durables	Jun-ichi Itaya	Mar.1995
15.	Regulatory System and Supervision of the Financial Institutions in Japan	Osamu Ito	May 1995
16.	Financial Restructuring and the U. S. Regulatory Framework	Jane W. D'Arista	May 1995
17.	The Legacy of the Bubble Economy in Japan:Declining cross Shareholding and Capital Formation	Hiroo Hojo	May 1995
18.	Stockownership in the U. S.:Capital Formation and Regulation	Marshall E. Blume	May 1995
19.	International Joint Ventures and Endogenous Protection a Political-Economy Approach	Laixun Zhao	Nov.1995
20.	GM社をめぐるｱﾒﾘｶ労働史研究：ﾌｧｲﾝとｴｯﾂﾞﾌｫｰｽの現場像の吟味	高田聡	Feb.1996
21.	卸売業の経営と戦略－－卸売流通研究会ﾋｱﾘﾝｸﾞ調査録(1):日用雑貨卸売企業	卸売流通研究会（代表　高宮城朝則）	Apr.1996
22.	卸売業の経営と戦略－－卸売流通研究会ﾋｱﾘﾝｸﾞ調査録(2):食品・酒類卸売企業	卸売流通研究会（代表　高宮城朝則）	Apr.1996

23.	A Note on the Impacts of Price Shocks on Wage in Unionized Economies	Laixun Zhao	May 1996
24.	Transfer Pricing and the Nature of the subsidiary firm	Laixun Zhao	Jun.1996
25.	The Incidence of a Tax on Pure in an Altruistic Overlapping Generations Economy	Jun-ichi Itaya	Sep.1996
26.	'Small Government' in the 21st Century	Hiroshi Shibuya	Sep.1996
27.	Characteristics and Reforms of Public Health Insurance System in Japan	Takashi Nakahama	Sep.1996
28.	The Role of Local Governments in Urban Development Policy	Yoshinori Akiyama	Sep.1996
29.	Optimal Taxation and the Private Provision of Public Goods	Jun-ichi Itaya & David de Meza & Gareth D. Myles	Oct.1996
30.	Comparison of Agricultural Policy in the U. S. and the Japan	Toshikazu Tateiwa	Oct.1996
31.	US Health Insurance:Types, Patterns of Coverage and Constraints to Reform	Dwayne A. Banks	Oct.1996
32.	International Capital Flows and National Macroeconomic Policies	Jane W. D'Arista	Oct.1996
33.	Financial Liberalization and Securitization in Housing Finance and the Changing Roles of the Government	Syn-ya Imura	Oct.1996
34.	Social Efficiency and the 'Market Revolution' in US Housing Finance	Gary Dymski & Dorene Isenberg	Oct.1996
35.	Government Expenditure and the Balance of Payments:Budget Deficit, Financial Integration, and Economic Diplomacy	Hiroshi Shibuya	Nov.1996
36.	A History of PBGC and Its Roles	C. David Gustafson	Nov.1996
37.	Dynamic Provision of Public Goods as Environmental Externalities	Toshihiro Ihori & Jun-ichi Itaya	Mar.1997
38.	A Comparative Static Analysis of the Balanced Budget Incidence in the Presence of Sector-Specific Unemployment	Koh Sumino	Mar.1997
39.	An Econometric Study of Trade Creation and Trade Diversion in the EEC,LAFTA and CMEA:A Simple Application of the Gravity Model	Masahiro Endoh	Apr.1997
40.	A Dynamic Model of Fiscal Reconstruction	Toshihiro Ihori & Jun-ichi Itaya	Apr.1997
41.	The Japanese Way of Solving Financial Institution Failures	Osamu Ito	Jul.1997
42.	The Federal Role in Community Development in the U.S. :Evolution vs. Devolution	Jane Knodell	Oct.1997
43.	Rent-Seeking Behavior in the War of Attrition	Jun-ichi Itaya & Hiroyuki Sano	Oct.1997
44.	ｻﾊﾘﾝ石油・ｶﾞｽ開発ﾌﾟﾛｼﾞｪｸﾄと北海道経済の活性化 第1号	北東ｱｼﾞｱｰｻﾊﾘﾝ研究会	May 1998
45.	購買部門の戦略性と企業間連携について	伊藤 一	Jun.1998
46.	The Formation of Customs Unions and the Effect on Government Policy Objectives	Masahiro Endoh	Jul.1998
47.	The Transition of Postwar Asia-Pacific Trade Relations	Masahiro Endoh	Jul.1998
48.	地域型ﾍﾞﾝﾁｬｰ支援ｼｽﾃﾑの研究 Ｉ－道内製造業系ﾍﾞﾝﾁｬｰ企業のｹｰｽｽﾀﾃﾞｨｰ	地域経済社会ｼｽﾃﾑ研究会 日本開発銀行札幌支店	Jul.1998

49. Fiscal Reconstruction Policy and Free Riding Behavior of Interest Groups	Toshihiro Ihori & Jun-ichi Itaya	Aug.1998
50. Quellen zum Markwesen des Osnabrücker Landes im Niedersächsischen Staatsarchiv Osnabrück(mit Schwerpunkt :Verfassung,Hölting,Siedlung und Konflikten im 17.und 18.Jahrhundert)	Susumu Hirai	Sep.1998
51. Equity and Continuity with a Continuum of Generations	Tomoichi Shinotsuka	Dec.1998
52. Public Resources Allocation and Election System	Akihiko Kawaura	Mar.1999

Discussion Paper Series
Center for Business Creation
Otaru University of Commerce

53. 消費者の価格ﾌﾟﾛﾓｰｼｮﾝ反応への影響を考慮した広効果測定告ﾓﾃﾞﾙの構築	奥瀬喜之	Jun.1999
54. 地域型ﾍﾞﾝﾁｬｰ支援ｼｽﾃﾑの研究II-地域型ﾍﾞﾝﾁｬｰ･ｲﾝｷｭﾍﾞｰｼｮﾝの設計-	小樽商科大学ﾋﾞｼﾞﾈｽ創造ｾﾝﾀｰ & 日本開発銀行札幌支店	Jul.1999
55. ｻﾊﾘﾝ石油・ｶﾞｽ開発ﾌﾟﾛｼﾞｪｸﾄと北海道経済の活性化 第2号	北東ｱｼﾞｱｰｻﾊﾘﾝ研究会	May 1999
56. 石鹸洗剤ﾒｰｶｰにおけるﾏｰｹﾃｨﾝｸﾞ･ﾁｬﾈﾙ行動の変遷	高宮城朝則	
57. 長期的取引関係における資源蓄積と展開	近藤公彦&坂川裕司	Dec.1999
58. Exernalities:A Pigovian Tax vs. A Lavor Tax	Ko Sumino	Dec.1999
59. A New Dimension of Service Quality:An Empirical Study in Japan.	Makoto Matsuo & Carolus Praet & Yoshiyuki Okuse	Dec.1999
60. Aftermath of the Flint Sit-Down Strike:Grass-Roots Unionism and African-American Workers, 1937-1939	Satoshi Takata	Mar.2000
61. Tariff induced dumping in the intermediate-good market	Chisato Shibayama	Apr.2000
62. Deregulatoion, Monitoring and Ownership structure:A Case Study of Japanese Banks	Akihiko Kawaura	Apr.2000
63. ｻﾊﾘﾝ石油・ｶﾞｽ開発ﾌﾟﾛｼﾞｪｸﾄと北海道経済の活性化 第2号	北東ｱｼﾞｱｰｻﾊﾘﾝ研究会	Apr.2000
64. A Cooperative and Competitive Organizational Culture, Innovation, and Performance: An Empirical Study of Japanese Sales Departments	Makoto Matsuo	May 2000
65. Foreign Exchange Market Maker's Optimal Spread with Heterogeneous Expectations	Ryosuke Wada	Jun.2000

Discussion Paper Series
Department of Economics
Otaru University of Commerce

No.1-16　　Feb.1985-Oct.1991

Discussion Paper Series
Department of Commerce
Otaru University of Commerce

No.1-2　　Apr.1985-May 1989

Center for Business Creation, Otaru University of Commerce
3-5-21, Midori, Otaru, Hokkaido 047-8501, Japan　Tel +81-134-27-5290　Fax +81-134-27-5293
E-mail:cbc@office.otaru-uc.ac.jp

小樽商科大学ビジネス創造センター
〒047-8501 北海道小樽市緑３丁目５番２１号　Tel 0134-27-5290　Fax 0134-27-5293
E-mail:cbc@office.otaru-uc.ac.jp