System Dynamics of Market Making and Liquidity

by

Boris Raykin

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degrees of

Bachelor of Science in Electrical Engineering and Computer Science

and Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

May 26, 1998

The author hereby grants to M.I.T. permission to reproduce and
distribute publicly paper and electronic copies of this thesis
and to grant others the right to do so.

Author_____
Boris Raykin
Department of Electrical Engineering and Computer Science
May 26, 1998

Certified by_____
James H. Hines, Jr.
Senior Lecturer, MIT Sloan School of Management
Thesis Supervisor

Accepted by_____
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

System Dynamics of Market Making and Liquidity
by
Boris Raykin

Submitted to the
Department of Electrical Engineering and Computer Science

May 26, 1998

In Partial Fulfillment of the Requirements for the Degrees of
Bachelor of Science in Electrical Engineering and Computer Science
and Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

The quality of a securities market is largely determined by the level of liquidity it provides to investors. This study examines the dynamics of liquidity provision by specialists on the New York Stock Exchange. It explores the effects of regulation, competition, adverse selection, and network externalities on transaction costs and liquidity in the market for a stock. The focus is on the causes of illiquidity in certain stocks, such as some of the non-U.S. securities on the NYSE. The system dynamics modeling methodology is used to build causal feedback and simulation models of the problem domain. The models are applied to derive better understanding of the causes for illiquidity and to recommend solutions for this problem. The models can also serve as a foundation for further study of market maker behavior and as tools for testing alternative liquidity-enhancing policies.

Thesis Supervisor: James H. Hines, Jr.
Title: Senior Lecturer, MIT Sloan School of Management

Chapter 1

Introduction

1.1 Motivation

Major equity markets, such as the New York Stock Exchange, the NASDAQ and the London Stock Exchange, compete with each other for new listings of shares and for the volume of transactions. The competition among today's markets is based largely upon the level of transaction efficiency that each market provides to its clients - the investor community. A market is transactionally efficient if it contains sufficient liquidity and has low transaction costs (Smidt, 1971). Liquidity is a comprehensive measure of how easily one can buy or sell shares of stock on a particular exchange. Simply put, the more active investors there are in the market and the more shares they trade in a day, the more liquid the exchange. Large share volume, high trading frequency and a large number of limit orders close to the current market price are all considered signals of market liquidity. Transaction costs are all of the expenses that are incurred by an investor during a transaction: commissions for all agents that represent the investor in the transaction, the bid-ask spread and adverse market impact costs. Markets that can succeed in reducing aggregate execution costs can expect more business in the form of order flow.

Markets for some stocks are naturally liquid. If a large number of investors regularly trade the shares of a stock, then finding a counterparty to a trade in that stock will not be difficult. What happens when an investor wants to transact a stock that is relatively unknown and trades very infrequently? Waiting for a counterparty to arrive may be very costly - one becomes exposed to the risks of unfavorable news arriving in the interim. Moreover, the sale or purchase of stock may need to be immediate, such as if it is motivated by the need for cash or short-term informational advantages. In that case, the investor may be willing to pay a premium for the immediacy of a transaction.

To improve overall liquidity, most equity markets contain some mechanism for the provision of immediacy. Mechanisms of that type are often called market-making mechanisms. Market

makers are traders that always stand ready to buy and sell securities at posted bid and ask prices. For providing this useful function they charge a fee in the form of the spread between the bid and the ask prices. The market-making arrangements of different markets vary quite significantly. Some exchanges have multiple competing market makers for each security, while others allow only a single market maker per security. Some let investors bypass the market-makers and trade with one another, while in others a market maker is a participant in every trade.

On the New York Stock Exchange (NYSE, the Big Board), the largest (by capitalization and dollar volume of trading) equity exchange in the world, the market-making function is delegated to exchange members called specialists. In addition to being the liquidity providers of last resort, these individuals have a number of key responsibilities, which make them central figures in the trading of every stock (Stoll, 1985). For example, specialists utilize their own order flow to stabilize trade imbalances that occur in the market for a stock. The behavior of NYSE specialists is a key determinant of how efficiently a stock trades. It is especially important for stocks with pathologically low liquidity levels, where the services of specialists are most necessary.

1.2 Purpose

This study aims to provide a comprehensive analysis of the issues behind the trading behavior of market makers and its linkage to the transactional efficiency of the market. This behavior is driven by a variety of factors including profit motives, regulatory pressures and the desire to form working relationships with clients. In order to be able to gain some depth of analysis, we focus our attention on the behavior of New York Stock Exchange specialists. We focus on those aspects of specialist behavior that affect the efficiency of the markets she makes. This choice of focus motivates a deeper look at a particular class of specialty stocks, namely non-U.S. securities. A large number of non-U.S. stocks on the NYSE (and other U.S. equity markets), trade less frequently than one would expect from security attributes, such as firm size and name recognition. Thus, the involvement of the specialists with these stocks is high and quite critical to the quality of market. An analysis of specialist activity in non-U.S. stocks will shed light on the nature of liquidity provision by the current specialist system.

As part of a broader analysis of liquidity, a closer look at the relationship between liquidity and the bid-ask spreads is called for. An inverse relationship between spreads and any measure of liquidity has been very well established in the literature (Demsetz, 1968; Stoll, 1985). For

3

example, non-U.S. stocks on the NYSE have both lower volumes and higher spreads than their U.S. counterparts. Policymakers often use this relationship to try to obtain better liquidity by forcing down the spreads. We want to study the effects of such a policy more closely and find out whether it has any adverse side-affects.

1.3 Methods

The methods of inquiry used in the study have evolved with the understanding of the system being studied. The project started out with an empirical analysis of a data sample collected to study the efficiency of markets in NYSE-listed non-U.S. securities. We were fortunate to have access to specialist trading and inventory position data. In our analysis of the data we revealed factors that influence the specialist participation rate, inventory commitment and the bid-ask spreads charged by the specialists. Data analysis was performed using statistical and regression techniques. Regression was used primarily in order to control for factors exogenous to the specialist and to identify the effects of specialist decision-making.

The regression analysis stage was successful at helping us extract important behavioral patterns from the data set. That stage also revealed some of the limitations of a strictly data-oriented approach of examining the system. Such an approach cannot accommodate some of the key issues of the market-making system. Amongst these issues are the relationships that the specialists form with the trading crowd on the floor and the perception of regulatory pressures by the specialist community, Some of these questions are inherently qualitative in nature, for others the acquisition of statistical evidence is not possible. In any circumstance, we found that we needed tools that could deal with the "soft" qualitative variables, as well as hard data in order to proceed with the inquiry.

Another important property of the system that we observed was complex feedback dynamics that governed the interaction between the specialists and other market participants. The standard cross-sectional regression approach that is prevalent in the empirical microstructure literature is not very well suited for the analysis of the dynamic nature of key feedback interactions. Thus, the explanatory power of standard cross-sectional models is quite limited in explaining specialist behavior. This study demands a modeling methodology able to capture the mutual interdependence of elements of the market-making system and clarify how feedback affects specialist's decision-making.

4

The system dynamics approach is extremely well fit for problems with a complex dynamic content and comprised of both quantitative and qualitative components. System dynamic models allow experimentation with the parameters and structural elements of the system for the purposes of normative analysis. In our case, that capability will be useful in the design of improved regulatory mechanisms to evoke desired behavior from the specialists.

1.4 Document Layout

The rest of the thesis document is organized as follows. The next section gives some background on the internal structure of the New York Stock Exchange and a primer on the competition to exchange specialists that is essential to the understanding of the discussion in the rest of the document. It similarly accustoms the reader to the terminology of the field of system dynamics. Section 3 presents a conceptual model of the problem of liquidity provision on the NYSE. It starts by uncovering the current patterns of trade volumes, spreads and regulation in Non-U.S. securities and goes on to identify the dynamic relationships between these key variables. Section 4 presents a numerical model of a selected subset of the model in Section 3. Section 5 summarizes the insights from the conceptual and numerical models and concludes.

Chapter 2

Background

This chapter is designed to give the reader some background information and perspective on the issues being dealt with in the modeling sections that follow. Section 2.1 familiarizes the reader with the terminology and concepts of the field of exchange efficiency. Section 2.2 examines prior works in the analysis of market making and liquidity, which serve as bases for this study, and clarifies the place of our research in the literature.

2.1 Background on Exchange Structure and Efficiency Measures

The discussion of issues in market making and liquidity is inextricable from the context of the structure and operation of the stock market. The first subsection below gives background on the New York Stock Exchange, the marketplace where the trading and market making discussed here actually take place. The current structure and rules of trading on the NYSE are key factors behind the liquidity services it is able to provide. Thus, understanding this structure is necessary to be able to understand the dynamics of market making and liquidity. We attempt to cover material that is outside of the scope of general finance knowledge and that is specifically necessary for the understanding and interpretation of the models and their discussion. In the next subsection we cover the competitive forces that affect the NYSE specialist and influence his decisions and trading behavior. The last subsection explains the terminology of transaction costs and measures used to characterize various aspects of exchange efficiency

2.1.1 The Structure and Operation of the New York Stock Exchange

The New York Stock Exchange (NYSE, the Big Board) is a double auction market where individuals are allowed to trade with one another directly by submitting orders to buy and sell shares of particular stocks. In addition, investors not being able to find a trade partner are able to

use the services of a market maker, called 'the specialist', who provides a reliable last resort of liquidity. Specialists stand ready to both buy and sell securities at prices they post. The price at which the specialist is willing to buy is called a 'bid' and the price at which she will sell is called an 'ask'. These two prices together form the specialist's 'quote'. The difference between the bid and ask prices is called the 'spread' (also called bid-ask spread). The spread compensates the specialist for providing liquidity through trades that she may not have engaged in as an ordinary rational investor. Each stock is assigned to exactly one specialist on the floor of the NYSE (unlike the NASDAQ market-maker system, which has multiple market makers for every security). Specialists are employed by private specialist firms, but are regulated by the NYSE itself and the Securities Exchange Commission (SEC).

Investors can submit different types of orders that allow them to base order execution on the prevailing market conditions. The predominance of all orders falls either into the 'market' or 'limit' order category. Designating an order as 'market' means that it should be executed unconditionally at the prevailing market price and as soon as possible. Limit orders, on the other hand, are contingent upon the prevailing price and execute only if the market price is at or more favorable than the specified limit order price. For example, a limit order to Buy 100 Shares of IBM at $120 per share, would execute if and only if the market price falls to *or below* $120. The specialist enters arriving limit orders into her Limit Order Book (LOB, order book, or specialist's book).

Orders have two possible routes they can take down to the floor of the Big Board. An investor Some orders are submitted by upstairs brokers electronically and arrive, via a system called SuperDOT, directly to the specialist's order book. Others are called in by phone onto the floor and are forwarded to designated floor brokers, who deliver it to the specialist post - a place frequented by other brokers interested in trading the stock specified in the order. Limit orders that arrive through the floor brokers are usually left with the specialist to be placed into the LOB. For more information on the details on NYSE market structure, rules and procedures please see Hasbrouck, Sofianos, and Sosebee (1993).

Companies that aspire to have shares of their stock traded or 'listed' on the NYSE are required to pass certain rather stringent financial and performance requirements. When shares get listed they are assigned to one of the specialist firms by the Exchange's Allocation Committee (specialists apply to get a new listing and the top candidates based on recent performance are chosen by the Committee and are given to the listing company to choose from). The stocks that

have been assigned to a particular specialist for dealership and "fair and orderly market" purposes are referred to as 'specialty' stocks of that specialist.

The NYSE lists shares of both domestic and foreign-based companies. A few of the Non-U.S. companies, mainly Canadian, decide to list regular or 'ordinary' shares of ownership, while others prefer to use instruments called ADR's. ADR's, an abbreviation for American Depository Receipts, are receipts for the shares of a non-U.S. corporation held by a U.S. bank (called the "depository"). The receipts entitle the registered holder to all of the benefits of ownership (except voting in some cases), including dividends and capital gains. ADR's are a popular way of listing foreign shares both on U.S. exchanges and the over-the-counter market. They are denominated in dollars and have the same settlement and clearing procedures as domestic shares. This makes them a very attractive vehicle for U.S. investors seeking international risk exposure.

Three levels of ADR programs exist. These levels differ in the amount of disclosure and compliance to the U.S. accounting standards that is required of the foreign corporation. Level 1 ADR's have minimal requirements and cannot be traded on an organized exchange. Level 2 and Level 3 ADR's are exchange traded and are registered with the Securities and Exchange Commission. They both comply to the U.S. general accounting standards, but Level 3 calls for higher levels of disclosure and U.S. investor relations initiatives than Level 2. Companies interested in raising capital on a U.S. exchange are required to reach Level 3 status. For more information on ADR's please refer to Velli (1994).

2.1.2 The Competition to the NYSE Specialist

Although there can be only one specialist assigned to a stock on the NYSE floor and she is the only designated market-maker in the stock, a profusion of competitive mechanisms in the provision of liquidity services exist in NYSE-listed securities. Below we list the various competitors to the specialist that provide alternative sources of liquidity and means to avoid paying the spread set by the specialist (although carry costs of their own). These mechanisms together with the regulatory pressures and rewards for relationship building, practically rule out monopolistic behavior by the specialists.

Competition to the specialist is both from within the NYSE and from agents in the broader markets. The following are the substitutes to specialist services inside the NYSE:

- Limit order traders avoid paying the spread and compete with the specialist for trades with market orders.

- Floor brokers are representatives of one or several investors, who form the trading crowd around the specialist on the trading floor. They are able to transact securities amongst themselves without any participation of the specialist. The order flow that is routed via the floor brokers has a much lower chance of paying the spread than that which comes in through the SuperDOT electronic system. Floor brokers have a lot more hands-on control over the execution of orders and the search for the counterparty.

- Upstairs block positioners offer services to institutional investors to find counterparties to very large trades. Block positioners have a network of clients who they call to solicit order flow when a block trade needs to be made. After all such parties are found, the trade is consummated and reported on the floor of the NYSE, usually at a price somewhere within the current quote.

There is also a number of sources of order flow competition that are external to the NYSE. Serious external competition to the NYSE is a relatively new phenomenon for the 206 year old Exchange. In mid-1970's the Congress and the SEC have established a National Market System (NMS) designed to bring new players into the securities exchange business. This gave the then infant NASDAQ market and United States regional exchanges a new role and capabilities to compete with the Big Board. The multiplicity of off-Board competitors also has a lot to do with advances in information and financial technologies. Below we list the main types of competitors to both the NYSE itself and its specialists:

- The specialist competes with other specialists on the 5 U.S. regional exchanges and the American Stock Exchange (AMEX), who are able to compete on a quote basis through a computer network called the Intermarket Trade System (ITS) that links all of the exchanges.

- The specialist also faces contention from third-market NASD-affiliated dealers in NYSE-securities. A majority of stocks listed in the last 15 years are not covered by NYSE Rule 390 that prohibits off-board trading of issues and is eligible for the third market.

- There are currently 4 Electronic Communication Networks (ECN's) registered with the SEC and subsequently able to trade in shares of NYSE-listed securities. These networks are proprietary trading systems that mainly connect large institutional investors and aid in

transacting large blocks of stock. ECN's provide another way for investors to avoid using the liquidity services of a specialist.

- The Chicago Board of Options Exchange (CBOE), the AMEX and the Philadelphia Stock Exchange all list option contracts on NYSE-listed stocks. Exchange-traded futures contracts exist on common stock indices, such as the S&P 500 and the Dow Jones Industrial Average. These contracts allow an investor to achieve leveraged exposure to individual stocks or stock portfolios, without any dealings with the NYSE specialist in these stocks.

- The NYSE faces increasing competition from foreign markets in both its domestic and non-U.S. listings. The ADR's and Non-U.S. ordinary securities listed on the NYSE are often also listed elsewhere, frequently, but not always in the country of origin of the stock. The market structure, level of development, and efficiency of the foreign markets that the security trades on has a significant influence on the amount of order flow that is seen by the NYSE. Emerging market stocks often have underdeveloped stock markets that have difficulty competing with the NYSE's market structure, technology and expertise. Thus, it often happens that the NYSE becomes the primary market for transactions in these securities fairly quickly after listing. On the other hand, developed market companies have home stock markets that match or even exceed the level of sophistication of the NYSE. In this case the NYSE may not be able to reach primary market status and have to operate as a secondary source of liquidity. Some domestic securities are also traded in foreign markets either as ordinary securities or in the form of American or Global Depository Receipts. However there is largely no contention to the NYSE's market share in these issues.

2.1.3 Efficiency and Transaction Cost Measures

Below we cover some of the terms used to characterize efficiency and costs associated with executing orders on an exchange. We make frequent use of these terms and concepts in our models of liquidity provision.

'Continuity' measures the stability of a market for a particular stock by looking at the changes in trade price for consecutive trades. Continuity is a very desirable property because it keeps investors from panicking about sharp price changes. Gradual price moves give investors enough time to respond with their orders. Additionally, the quality of markets and usefulness of equity derivatives is largely determined by the continuity in the price of the underlying stock. Furthermore, continuity reduces slippage, which is a measure of how much price moves from the

10

time an investor submits an order to the time it can be executed. High bid-ask spreads reduce continuity because a buy market order trades at the ask and a sell market order at the bid. If two of these orders arrive consecutively the change from one trade to another will be determined largely by the magnitude of the spread. This phenomenon is what's called the "bid-ask bounce".

'Immediacy' is simply the ability to quickly find counterparty to a trade. In a pure auction market, where investors trade with one another, investors depend on a critical mass of market participants to always be present in the market to achieve immediate execution if it is necessary. However, in markets where trading frequency is low, investors often have to wait for long periods of time for a trade partner and risk the arrival of unfavorable news in the interim. In order to give investors the choice to avoid such situations, the NYSE provides specialists always standing ready to buy or sell securities at posted prices.

'Depth' is a measure of trade size that the market can absorb without incurring significant changes in price or 'market impact costs' (see below) for the investor. Depth on the floor of the exchange is provided by the orders in the public limit order book, the trading crowd and the specialist quote. The quantity of shares supplied and demanded by these parties is what determines how easily one can trade a given amount of shares. Institutional investors often turn to upstairs block positioners, who have a network of clients with very large positions in a given stock, to buy or sell really large quantities.

A 'market impact cost' is the price paid by investors who decide to trade in large share quantities at once. When a large buy market order hits the limit order book, for example, it is likely to start 'consuming' smaller sell orders on the other side of the book. The combined sizes of these counterparty orders at prices close the current ask price is what determines how much price movement the buy order will see as it is being filled. If the order book is 'shallow', fewer orders will exist to counteract the movement of price. A 'deep' order book has a lot of orders around the current market quote and can absorb larger trade sizes.

2.2 Related Work

Below we provide a brief overview of key studies from each subfield of research on stock exchange specialists and exchange efficiency. We then try to establish how our work fits into the body of literature and what is new about the perspective it provides.

Market microstructure theory concerns itself with the economic analysis of exchange processes. Unlike a majority of other economic disciplines that focus on equilibrium behavior of systems, microstructure research looks closely at situations of disequilibrium between the supply and demand of securities. This emphasis has been inspired largely by the early findings of Demsetz (1968).

Demsetz set out to study how prices are determined in securities markets and the effect of transaction costs on price formation. He was soon able to understand that the time dimension of supply and demand of securities affects market prices. His reasoning was as follows. In real markets, such as the NYSE, there are both explicit and implicit execution costs. An example of explicit costs is a per share transaction cost levied by an exchange. An example of an implicit cost, is the price one has to pay in order to execute his order immediately. Whereas over long time the number of buyers and sellers is the same, there is no reason for why this number is the same at any particular time. Thus, at a time t there might exist an imbalance of demand and supply and a market-clearing price will not exist. For those traders who can wait for a balance to be achieved, the temporary imbalance does not present a problem. However, those who want to trade immediately might want to raise (lower) price to lure more sellers (buyers), until an "immediate" equilibrium arises. Thus, we arrive at two equilibrium prices: one for those willing to wait and one for seekers of immediacy. The difference between these two prices represents the market price for immediacy.

The implications of these conclusions are multiple. First, the notion of a market equilibrium is problematic. In the analysis above, we arrive at two demand curves and two supply curves in the market, that reflect groups of traders with two distinct time frames. Whether one wants to buy or sell becomes a factor in the price that she is able to attain. Second, the specific structure of the market matters for price formation. The price of immediacy (the bid-ask spread) above depends on the number of traders present in the market and, thus, on the volume of trading. This theoretical finding has been confirmed by Demsetz's empirical analysis of NYSE trading, where he found that the size of the spread depended significantly on the trading volume.

Demsetz's work is considered to be the cornerstone of securities market microstructure research. He was able to show both theoretically and empirically that both transaction costs and market structure can affect the price formation process and motivated the need to study the details of how markets operate. His model of exchange mechanisms, however, had missed a fairly important notion - the interplay of price-setting mechanisms with trader behavior. Since the

trading mechanism matters for price formation, it will have an effect on how traders will formulate their order decisions. Thus, investor order flow is not exogenous to the trading mechanism.

This last observation had been made by Garman (1976), who used it to restate Demsetz's original question as: how do market prices arise given the nature of the order flow and the market-clearing mechanism (market structure and rules)? Garman abstracted the desires and optimization problems of individual traders into a set of buy and sell order flows. These flows produce temporary trade imbalances that have to be absorbed by the market-clearing mechanism. The problem as restated resembles stochastic matching problems in the fields of inventory control and insurance. Uncertainties in the order flow can result in inventory control problems for specialists or dealers. Garman has actually proved by using a form of the "Gambler's Ruin" argument that specialist's disregard of her inventory is bound to lead to her bankruptcy and the disruption of the operation of the market. Thus, market makers must formulate policies of dealing with price and inventory uncertainties.

This last view of the market and market makers places special importance on optimal inventory management strategies of exchange specialists. A number of works that followed had studied these strategies and their effect on price formation. Amihud and Mendelson (1980), Stoll (1978), and Ho & Stoll (1981, 1983) pursue this area of research and come up with a model of the dealer's pricing behavior based on the need to control inventory. The picture painted by these studies is that of a specialist as a typical dealer, selling at a discount when inventory is high (by lowering the position of her quote) and raising prices (quotes) when it is low. They treat the bid-ask spread as largely the price of risks that the specialist assumes by taking on undesirable inventory at times of order flow imbalance, the risk that the inventory may cause losses due to unfavorable patterns in future order flow.

Another strand of literature (Copeland and Galai, 1983; Glosten and Milgrom, 1985; Easley and O'hara, 1987) looks at the differences between order flow from uninformed and superiorly informed traders and the hazards the latter play for specialist. The information about the 'fair' price of a stock is incorporated into the price through the orders of traders who first uncover it. Dealers trading with these superiorly informed traders incur sure losses since the price will move for the informed trader and against the dealer. One way the dealer can deal with this problem is to raise the bid-ask spread both to make it unprofitable for some of the information traders to trade and to recoup losses from others. Thus, the inventory holding risks are complemented by

13

the information risks in determining the dealer spread and trading behavior. The game theoretic framework is especially well-suited to the analysis of competitive behavior with asymmetric information. This framework has been extensively employed by market microstructure studies of the effects of informed traders. For more on the field of theoretical market microstructure please refer to O'hara (1995) or Cohen et al. (1986).

The importance of the specialist for microstructure research, as shown above, in addition to the key role she has in the NYSE market both inspired empirical analyses of facets of specialist behavior and performance. Since specialist inventories are key determinants of the bid-ask spread and the quote movements, some attention has been paid to them in the empirical literature (Hasbrouck and Sofianos, 1993). Madhavan and Sofianos (1994) actually do not find evidence that specialists manage inventories by adjusting their quotes but rather by selectively trading to reduce or increase inventory as desired. The main contribution of their work, however, is to elucidate the patterns in specialist participation or trading involvement across stocks and time. They find that specialists have much higher participation rates in stocks with low daily share volumes, making the market for these stocks look more like a dealer market. Specialists have much lower percentage involvement in more active instruments, where investors are able to provide enough liquidity for themselves. Sofianos (1995) sheds light on specialist's total revenues. The paper defines a dichotomy of revenues based on their origin. Revenues are either obtained through spreads or through positioning. The first measure is mostly related to the size of the spread, while the latter to the change in the position of the midquote (the average of the bid and the ask) from the time the specialist buys to the time she sells.

Papers on market making often discuss the impact of the specialist or market maker on the efficiency and costs of trading in the market. The analysis of exchange efficiency has often been instigated by perceptions of unfairness in contemporary market mechanisms and the desire to change regulatory policy. A number of very influential works have been produced in the early to mid 1970's, after the SEC has conducted an Institutional Investor Study (U.S. Securities and Exchange Commission, 1971) and was considering the reformation of the securities industry. The literature of that period debates the relative merits of single and multiple market-maker systems and examines closely the practices and impact of NYSE specialists. An important by-product of discussing the relative merits of various proposals for securities industry reform was the identification of the types of costs investors incur when executing their orders and the development of terminology necessary to address aspects of exchange efficiency. Prominent

14

studies of that era include Smidt (1971) and Black (1971). These studies have proved to be very influential and led to eventual reform of the securities and exchange industry and to the creation of the National Market System in 1975.

One publication, namely "Making Markets" by Mitchell Abolafia (1996), stands out from the rest in its genuine portrayal of the reality and the social aspects of market making. It describes the author's personal experience interviewing and observing market makers on a number of securities exchanges, including the NYSE. Through very well chosen quotations and remarks, the author succeeds in accurately portraying the social and economic motives for the real behavior of market makers. It is especially helpful in understanding the interplay of specialists' profit motives and a culture of restraint that has recently been established in response to regulatory pressures. This work is a great inspiration and an invaluable source for our models and discussion. It is truly unique and irreplaceable by any amount of theoretical discussion on the subject.

This study is motivated by some of the shortcomings of strictly empirical approaches and combines the empirical aspects of some studies below with a system dynamics model of the market-making process. It is unique in its use of system dynamics to formalize and study the feedback interactions between market makers and other participants. It is also unique in the analysis of the specialists' dealing in non-U.S. stocks.

Chapter 3

The Causal Loop Model

In the next two chapters we present the core of this project - the conceptual and simulation models of liquidity provision on the NYSE. The conceptual model extracts entities in the system, conceptualizes them in the form of model variables and establishes the dynamic relationships between these variables. A single dynamic relationship between two variables specifies what happens to the second variable if the first one changes. As one starts defining the dynamic relationships for all variables in the system, cycles of influence, called causal loops, begin to emerge. The causal loops generate the endogenous behavior of the model and represent the natural operation of the system. They help us identify the reasons for why the system behaves the way it does and to discover alternative modes of operation and means of reaching them.

The layout of the chapter is as follows. We first identify that the problem we are trying to analyze is that of illiquidity of non-U.S. issues on NYSE. We characterize the problem by describing the dynamics of the key variables. We then decide on the sectors of the system that will be conceptualized and modeled and proceed on modeling these sectors and guiding the reader through the model.

3.1 Issue Statement

The problem we are focusing on in this causal loop model is the illiquidity of certain securities listed on the New York Stock Exchange, particularly the shares of stock of non-U.S. companies. Although a number of the non-U.S. companies on the NYSE have volumes and trading frequencies that are comparable to domestic issues, the average non-U.S. stock has considerably lower levels of daily share volume and daily number of trades than an average U.S. stock. The lower average is particularly due to very low levels of activity in stocks of companies from developed countries and companies from the Asian continent. In addition to lower activity levels, an average non-U.S. stock also carries higher bid-ask spreads and, thus, execution costs.

3.2 Reference Modes

The issue statement above identifies a number of key variables of interest, whose behavior we find problematic and are looking to improve. Below we list these key variables and auxiliary variables important for future discussion and verbally describe their problematic behavior.

(1) The average daily trading share volume of non-U.S. securities on the NYSE, especially the volume of securities of developed countries and countries from the Asian continent, are almost twice as low on average as their U.S. counterparts. Both the U.S. and Non-U.S. volumes have been growing rapidly in the last few years, but the gap between them has not been reduced significantly.

(2) The daily number of trades for an average non-U.S. share is also lower than for the domestic counterpart. As a matter of fact, the number of trades for non-U.S. shares is usually around 45% of that for U.S. shares.

(3) The bid-ask spreads for non-U.S. stocks are higher than those of a U.S. control group. Both the non-U.S. and U.S. spreads over the last year and a half have been decreasing due to increased activity and competition brought about by a number of regulatory and operational changes. However, the gap between these two groups of stocks is not going down.

(4) The levels of regulatory pressures are high, but rather steady.

3.3 Choice of System Sectors

The causal loop model consists of several separate pieces that describe the influence of particular sectors of the system on the dynamic behavior of key variables of interest. The choice of these sectors is motivated both by how significant the sector is to the problem described in the issue statement and the reference modes and by the level of power the Exchange and regulators have in controlling the variables of the sector. In addition, in the interests of focus and bounded complexity we choose to limit the choice of sectors and sector components in the following ways:

• The focus is on the dynamics of a market in a single stock. Although specialists usually have one or two dozens of stocks in their portfolio and the behavior of the specialist in one stock may have an influence on the other, this consideration is left to future research and is not considered in this work.

17

- As explained in the introduction, the system of interest in this work is the system of liquidity provision on a particular exchange, the NYSE. The viewpoint from which sectors are chosen and the model is developed is one of an exchange official or regulator who tries to understand and eventually improve conditions on the NYSE. Such a viewpoint constrains the scope of analysis and the policies that can be recommended. It means that markets and participants outside of the NYSE are considered to be exogenous influences on the system, rather than being part of the system itself. Thus, investor order flow in the models below is generated mostly exogenously, except for the influence that specialists think they have via their control of quotes and spreads. Competitors and foreign markets are exogenous except, similarly, for the influence of spreads and liquidity on which competition is based on.

With these guidelines in mind we have identified four sectors to be included as part of the conceptual dynamic model of the NYSE system of liquidity provision:

(1) The Demand for Immediacy sector

(2) The Informational Content of Order Flow sector

(3) The Competition for Liquidity Services Sector

(4) The Regulation Sector

3.4 Understanding Causal Loop Diagrams

As mentioned, the present causal loop model consists of several separate components organized around the key sectors of the system given above. Each component is a set of causal loops, that we will refer to as a 'loopset' (the term was adopted from conversations with my advisor, Dr. James H. Hines), that we think can cause the problem-related dynamics of the corresponding sector. One often finds components of this sort used to organize the presentation of a loop model, however, most of the time they are eventually put together to form a single causal loop model. In the case of the present model, a sector-by-sector view is sufficient to uncover the relevant dynamics. At the same time, aggregating the sectors might have drowned out the focus on individual forces that drive dynamics and made the model more difficult to grasp. In order to be able to talk about sector loopsets separately we sometimes have to duplicate loops that apply to two or more sectors. We try to only duplicate the loops that are of particular importance to both sectors.

The loopset for each sector is presented in stages. A sector model starts out with the most basic loops connecting the basic variables of that sector. This compact set of loops is the first 'view' of the sector loopset. In the next view, extra variables and loops are added to expand the understanding of the dynamics of the sector. The process continues until the last view contains a complete loopset. We will refer to this method of gradual presentation of model structure as 'layering' (the term was also adopted from conversations with Dr. Hines). The arrows for the loops that have been presented in prior views are dotted and the variables unused by any of the new loops presented in the view are grayed-out. That is done simply to concentrate the attention of the reader on the original loops being presented in the view. Loops presented for the first time are characterized by solid arrows.

Loops can be distinguished from one another by the widths of the arrows that comprise them. These widths do not in any way correspond to the relative weights or strengths of the loops. They simply serve as markers to ease the identification of a particular loop and help set loops apart from one another. New loops are also identified by circular loop symbols inserted in a central location inside the loop. The loop symbol consists of a circular arrow, a picture on the inside of the arrow and a capital letter underneath the arrow. The circular arrow points clockwise or counter-clockwise corresponding to the orientation of the loop it represents. The picture at the center of the symbol illustrates the polarity of the loop. A drawing of scales denotes a balancing loop, a negative feedback cycle. A picture of a snowball rolling down the mountain is indicative of a reinforcing loop, a positive feedback cycle. The letter underneath the shape is a label that is used to name loops of a particular view. Loops are labeled alphabetically, starting with the most basic loops of a view and continuing with more ancillary cycles. We use these labels whenever we refer to loops in the discussion.

Some loops share one or more of the constituent arrows. In that case, the arrow width of the loop with the foremost letter is used. In this way, the secondary loops inherit parts of the structure of a basic loop. In other situations, two or more loops of the same polarity and orientation that share most of their structure and conceptual backbone are grouped together as one. The combined loop is identified by a single loop symbol and contains multiple parallel arrows of the same width. Exogenous inputs to the loopset are marked by dashed lines. We include only those inputs that have a particularly strong influence on the dynamics represented by the loops or provide practical levers for policymakers.

19

3.5 Causal Loops by Sector

3.5.1 The Demand for Immediacy Sector

Introduction

The first sector focuses on the magnitude and composition of order flow and its interplay with the provision of immediacy in the market. It examines the network externalities of trading (Stoll, 1992) or the ways in which order flow today attracts more order flow in the future. It also considers the ability of investors to adjust to various levels of immediacy available in the market by switching between market and limit order strategies and the effects of this adjustment on order flow and immediacy.

Loop Descriptions

A. Number of Spreads Paid (marked by the thickest arrows in the view)
This loop is the first of three reinforcing loops in this sector that can collectively be summarized as the network externalities of order flow. The loop in particular shows that less order flow allows for fewer investor-to-investor matches, increasing the need for services of immediacy providers and thus the average costs of immediacy. The increased costs of obtaining immediate execution induce more people to wait for "better times" rather than trading immediately. One way of doing this is to submit limit orders, and optionally convert them to market orders when conditions become more favorable.

When there is insufficient influx of public order flow into the market, chances that another investor would be willing to trade with you immediately at a price close to the price of the last trade are low. Thus, the investors are not providing immediacy themselves and the need for liquidity providers, such as the specialist or limit order traders, increases. Dealing with providers means having to pay a spread. Since the bid-ask spread usually accounts for a high percentage of order execution costs, having to pay the spread more often can sharply increase the average costs of immediacy. At these levels of immediacy costs more people may decide to wait for better market conditions, further depressing order arrival.

As with any reinforcing loop, "Paying More Spreads" further degrades a situation of low order flow but it can really help to amplify any improvements in order flow, as well. One of the

20

main strengths of the NYSE market structure is that it allows exploiting the favorable operation of this loop, a. k. a. "Avoiding the Spread". When order arrival increases due to exogenous factors, more immediacy can be provided by investors themselves without the need for specialist intervention. This means fewer spreads and lower costs of immediacy that encourage more investors to stop waiting to submit their orders and seek immediate execution. This mechanism is a unique property of auction markets, where investors are able to trade directly with other investors without paying dealer spreads.

B. Competitive Forces and the Size of the Spread (2nd thinnest arrows)

Here we aggregate two loops that describe how declines in order flow cause higher spreads and even lower order flow from spread-conscious investors. The first loop notes that lower order flow means that it is more difficult for the specialist to find a counterparty to adjust her inventory position back to desired levels. That means that the specialist will demand more payment for taking on undesired inventory and will raise the bid-ask spread accordingly. (For background on the inventory management behavior of the specialist please refer to Schwartz, 1991).

The second loop finds that inventory considerations may be even more serious for the specialist's competitors on the regional exchanges, the third market, etc. (Please refer to the Competition for Order Flow sector below for a comprehensive discussion of the different types of competitors to the specialist). These entities are not in any way obligated to make markets in particular securities and are not regulated to keep their spreads down. They also don't have the informational benefits of seeing the depth of the limit order book or the size of the trading crowd in a security. The informational disadvantages mean that competitors face higher inventory risks than the specialist, and will utilize their ability to stay out of the market when times are bad. Lower competitive pressures on the spread mean that the specialist will be able to raise spreads even higher.

C. Specialist Voluntary Participation (the thinnest arrows)

The last of the reinforcement or network externality loops looks at the voluntary participation of the specialist as a speculator or trader. The specialist has an affirmative obligation to maintain a "fair and orderly market" in every one of the stocks assigned to her. To accomplish this task the specialist is often forced to assume losses and take on inventory she would not otherwise take on as an ordinary rational investor. While obligatory trading in itself is usually beneficial to the market quality of the stock, the specialist will participate only as much as she absolutely has to, to

fulfil the obligation. On the other hand, the specialist may be interested in voluntarily trading certain stocks for her personal (or specialist firm) account to obtain extra revenues to supplement revenues from spreads. Since the specialists are largely prohibited from acting in a destabilizing way, the extra speculative trading by the specialist is a beneficial infusion of order flow to the market.

With respect to this voluntary trading behavior, the specialist is likely to behave similarly to her regional / third market competitors, except for certain restrictions imposed by Exchange regulation mentioned above. Thus, lower public order flow and the ensuing inventory holding risk are likely to reduce the desire of the specialist to trade for own account. The immediacy and market depth that would ordinarily be enhanced by specialist voluntary participation decline. Lower immediacy via investor order flow means that more spreads will have to be paid to liquidity providers. Lower market depth, or density of orders in the limit order book, and specialist inventory will mean that larger market orders will have to pay a higher price for their size due to moving the market more in the unfavorable direction. These extra costs will have negative impact on the demand of immediacy and its realization in order flow.

D. Adjustment Via Limit Order Strategies (the 2nd thickest arrows)
One way in which investors can adjust their cost of execution to their current immediacy demands is by choosing between market or limit order strategies. Choosing a limit order strategy allows one to forego paying the current spread and opt to wait for execution in the future. When costs of immediacy are high and the level of immediacy demand at that level of costs is low, more investors may decide to follow the limit order route. Since limit orders essentially compete with the specialist for the provision of liquidity, they increase competitive pressures on the spreads and force them lower. At the same time they enhance market depth by forming barriers against the movement of price due to the execution of market orders (Schwartz, 1991). Both of these effects reduce immediacy costs and allow more order flow, especially in the form of market orders. This balancing loop keeps execution costs from moving too much and allows the market to adjust to different levels of urgency in the order flow.

Summary

The three reinforcing loops in this loopset have undoubtedly played a role in the liquidity dynamics of thin stocks, such as some non-U.S. issues. They show us that risks of not finding counterparties to future trades hamper the level of order flow today. These risks reinforce

downward trends in order flow and give a significant first-mover advantage (Stoll, 1992) to the first liquid market in the stock. The NYSE enjoyed primary market status in almost all of its issues throughout its 200-year history. Thus, it is very used to being on the liquidity-enhancing side of these positive loops. The reality with non-U.S. stocks, especially those from developed countries, is that they already have a primary market in their home country before coming to the NYSE. The Exchange is an underdog in its struggle for liquidity in these markets. The network externalities of order flow, described above, are acting against the NYSE and for the "home market advantage" of its foreign competitors. These are just the first in a series of forces that make the non-U.S. issues fundamentally different from their domestic counterparts on the NYSE. Recognition of these differences by the Exchange in its regulatory and issuer relations policies is essential to the ability of the NYSE to enhance its liquidity and reach primary status in the trading of its non-U.S. shares.

3.5.2 The Informational Content of Order Flow Sector

Introduction

Investors submit order flow for a wide variety of reasons. Some investors follow risk-reducing strategies that call for profit taking after a certain percentage return has been obtained. Others submit sell orders in order to obtain cash to finance buying a house or a car. Yet others periodically rebalance their portfolios to keep constant their exposure to various types of market risks. Investors who are trading for one of the reasons mentioned above are usually called 'liquidity' traders. Their orders are motivated by their personal situations and carry little or no informational content about the future price behavior of the stock. On the other hand, some investors spend a lot of time researching for stock information that has not yet been incorporated into market prices (often referred to as 'private' information) and therefore has non-zero expected influence on the underlying fair price of the stock. If they uncover such information they can expect to profit from it by buying shares of stock if the information implies that shares are undervalued at current prices or selling stock if the shares are overvalued. Orders motivated in this way are called 'informed' and the traders who submit them are referred to as 'information traders'. To make sharper the contrast between information traders and liquidity traders, we sometimes refer to liquidity traders as 'uninformed', as they deliver no new information to the market.

In this sector we look at the influence of the informational content of orders, measured by the ratio of informed to uninformed orders (hereafter referred to as the 'Ratio'), on average transaction costs and order submission patterns. The informational content has strong influence on the willingness of specialists and limit order traders to provide liquidity. We first examine this influence and then tackle the relationship of current informational content and the incentives to engage in more research to uncover private information. Please be aware that informational content refers to the amount of new, private information carried by orders. It does not directly reflect the amount of public information available for the investors in the stock. Ironically, stocks with very high levels of publicly available information will usually have a very low level of informational content, as information is less likely to be held private. We examine the effects of public disclosure by companies on the ratio of privately informed traders and recommend policies that can favorably influence liquidity by reducing this ratio.

Informed traders are very likely to use market orders as the order type of choice, since the expected gains from private information usually far exceed the bid-ask spread and the opportunity costs of non-execution for an informed limit order are too high. Consequently, we can avoid some complexity while remaining true to the dynamically important aspects of informational trading by considering only the market-order portion of informational trades and ignoring the small residual of limit orders.

Loop Descriptions

View #1: Effects on the Liquidity Providers

Here we concentrate on the risks of liquidity providers of losing money through trades with better-informed individuals and how these risks affect the mechanisms for the provision of liquidity.

A. Limit Order Adverse Selection (the thinnest arrows)
The first loop of this view resists changes in the ratio of informed and uninformed market order traders. When the number of informed traders is high, a limit order trader has a very high chance of trading with someone who possesses more private information. They can expect to lose money from this trade, since informed traders will only submit orders if their information is more valuable than the transaction costs of trading (Stoll, 1992). Thus, no matter how high the spreads, limit order traders have an adverse selection risk of trading with a more informed counterparty

24

that increases with the informational content of the order flow. This adverse selection risk will discourage the submission of limit orders by the uninformed and increase the proportion of market orders as part of the uninformed order flow, reducing the proportion of informed market orders.

B. Specialist Adverse Selection (the 2nd thinnest arrows)

Specialists also face adverse selection risks. Since they cannot avoid trading with informed traders when nobody else is willing to take trade, they try to recoup their losses to the informed by setting high bid-ask spreads. This effectively shifts the burden for adverse selection on the shoulders of the uninformed public orders. As the Ratio of informed market orders grows, specialist adverse selection risks grow and spreads are increased, the relative costs of paying the spread become too high and people switch to limit orders to avoid paying the spread. Since it is assumed that informed traders always trade using market orders, they will not try to switch to limit order strategies. The proportion of informed market orders will increase even further (Please refer to View #2 of this sector to see the effects of the Ratio and spreads on the behavior of the informed.)

C. The Effect on Average Transaction Costs (the 2nd thickest arrows)

From the previous two loops we can see that adverse selection risks caused by a rise in the Ratio increase the costs of both the limit and market order strategies. Effectively the average transaction costs in the market for a stock are raised, forcing out orders by cost-sensitive individuals and making some investors wait for better times to trade. Although informed traders will also be affected by the rising costs (see View #2 below), their opportunity costs of waiting will exceed the costs of immediacy. By waiting, the informed would be undertaking a risk that the private information they hold will leak to the market and will not be of profitable use any more. The reduction of order flow will, therefore, be mostly due to uninformed traders and that means that the Ratio will be further increased.

D. Limit Order Competition to the Specialist (the thickest arrows)

The last loop on this diagram reminds us that the costs of the market and limit order options for the investors are linked by the competition that limit orders exert on the specialists. Spreads will not be raised too high due to the direct effect of loop B and the indirect effect of other loops in the diagram, because people will switch to limit orders and avoid paying spreads.

Insights: The Effect of Public Disclosure

Improved public disclosure increases the amount of publicly accessible information about the issuing company in the form of better annual reports, public releases, analyst conference calls, etc. By releasing more information the company puts a higher percentage of the available information in front of the public and prevents a high ratio of privately informed individuals that hurts both the transaction costs and liquidity of the stock through loops B and C above.

The level of public disclosure is one of the levers that the company has in combating liquidity problems that are caused by people's fears to trade with the privately informed. The stock exchange should continue encouraging and facilitating means of public disclosure, since it too can only benefit by a lower ratio of privately informed traders.

View #2: Effects on the Informed

This view explores the effects of informational content and spreads on the behavior of the informed. The effects of the first two balancing loops in this view become perceptible when the Ratio of informed to uninformed traders is high.

A. Adverse Selection Risk for the Informed (the 2nd thickest arrows)

In situations when the reinforcing loops of the previous view are not kept in check by their balancing counterparts and public disclosure is insufficient, the Ratio of informed orders might reach very high levels. At these levels, the limit orders book will be almost empty and the specialist will do everything she can to try to match market orders with each other to avoid trading with them. One of the techniques she may use is 'stopping the stock' in which she guarantees execution at the quote that existed when the market order arrived and then searches for a counterparty in the crowd or the orders arriving through SuperDOT. This technique may achieve a better execution cost (or "price improvement") for the order because there is a chance it will trade against another market order. However, in a market with high concentration of private information it may mean that informed orders may have to trade with one another. This chance creates for informed traders a risk of trading with a counterparty whose private information has a higher expected impact on the stock's fair price (Black, 1986). This adverse selection risk will mean that traders possessing private information with low expected price impact may not chose to capitalize on this information. An increase in the Ratio of informed to uninformed, in a situation where the ratio is already rather high will cause an increase in the adverse selection risk described

above and lower order flow from informed. All else equal that will keep the Ratio from going even higher.

B. Spreads and Informed Orders (the thinnest arrows)

Higher spreads will mean that fewer bits of information will have an expected gain that will exceed roundtrip execution costs (costs of both buying and selling). Some trade ideas will not be acted upon. The order flow from the informed will fall, reducing the Ratio, liquidity risks for the specialist and the spreads charged by the specialist. This effect will be extremely weak in normal circumstances and liquid markets because average spreads as percent of the trade price are negligible compared to typical returns from information trading. However, the effect may start being important when the Ratio is really high because the levels of spreads that specialists will have to resort to may become comparable with the expected values of private information.

C. Incentives to do Research (the thickest arrows)

Private information is a product of careful research and data analysis. The research is a cost to the investor and will only be undertaken if the expected revenues of trading on newly discovered information outweigh these costs. When the private information content of the market and bid-ask spreads are high, efforts to do more research are curtailed because trading is expensive and the chance of trading against someone with better information is high. As spreads and the ratio rise, the incentives to do research decline, fewer trade ideas are generated, less informed order flow can be submitted and the Ratio and the spread are kept from rising higher.

Insights: More on the Effect of Public Disclosure

Public disclosure increases the costs and efforts of researching for private information because less information is left hidden. Yet, since the informational content is reduced (see View #1 above) trading against the uninformed is more commonplace and the utility of extra information grows. Private information is both more expensive and more utilizable.

Heightened disclosure policies are likely to improve overall conditions by lowering the private informational content of orders and eliminating the fears of liquidity providers of being squeezed by informed traders. However, they have to be kept from going overboard and and making it prohibitively expensive to engage in more research.

3.5.3 The Competition for Liquidity Provision Sector

In the Background section we overview the types of competitors the NYSE faces in the race for order flow and the provision of liquidity services. Below we identify how these mechanisms affect liquidity and trading costs on the NYSE itself.

Loop Descriptions

View #1 - U.S. Competition for Liquidity Provision

The U.S. regional exchanges and the third market are providing more and more competition to the Big Board in the last few years. They are able to capitalize on and thus eliminate situations when the NYSE specialist is charging too high of a price for her liquidity provision in the form of the spreads. The competition is still rather limited in scope to the more liquid issues. Also, the regional and third market competitors do not have an affirmative obligation to engage as dealers and do not face the same level of regulation as the NYSE specialist. These freedoms give them competitors a markup advantage with respect to the specialist and, as a result, they are not as driven to compete on the basis of spreads.

A. Competition Drives Down Spread Markups (the loop with the thickest arrows)
When off-Board competitors perceive that the spread has gotten too 'rich' for the level of risks inherent in the stock, they engage in tougher spread competition and drive down the NYSE specialist's markups.

B. Lower Spreads Attract Order Flow from Other Markets (the thinnest arrows)
Lower markups result in fairer spreads that bring the order flow back to the exchange from the competitors, increasing the volume, liquidity and perceived profits. The latter further increase competition and drive the markups even further down.

C. Higher Liquidity Attracts Order Flow from Other Markets (the 2nd thickest arrows)
Higher liquidity exhibits network externalities (see the Demand for Immediacy sector above) and attracts order flow from other markets. Liquidity also reduces inventory holding costs for the specialist and thus the spreads that the specialist charges, attracting investors from other markets based on lower transaction costs.

Insights: Differential Regulation and Off-Board Markup Advantages

28

The NYSE is the primary market for all of the shares that it lists. Being a primary has certain advantages from the point of view of information and control. Thus, regulators are keener to what is happenning on the NYSE than elsewhere in the U.S National Market System. Extra regulation implies that NYSE specialists cannot have high markups on their spreads. Their off-board competitors, however, are not that hard-pressed to keep down their markups. They have to match the spread on the NYSE in order to be able to execute orders, but they compete much less on a spread basis and obtain their order flow through contractual arrangements rather than promising lower execution costs.

View #2 - Arbitrage

We use the example of ADR's to show the effect of intermarket arbitrage on spreads, liquidity and competition. The example can be extended to cover intermarket arbitrage that occurs between the equity markets and its derivatives, options on individual stocks and stock index futures. Index arbitrage is an especially powerful influence on NYSE market. In the ADR example, the foreign market shares represented by the ADR's (referred to as the 'underlying' shares) can be traded contemporaneously with the ADR's during the period of overlap in trading hours between the foreign home market and the NYSE. This fragmentation of markets allows for arbitrage based on temporary misalignments in the quotes posted in the two markets.

A lower spread for an American Depository Receipt means that the price of the underlying home market security is more likely to end up outside of the NYSE quote. If the home market ask price is lower than the NYSE bid, arbitrageurs will buy at the home market ask price and sell at the NYSE bid price ("pick off" buy orders) making a riskless profit. If the home market bid price is higher than the NYSE ask, arbitrageurs will buy at the NYSE and sell in the home mkt. Lower spreads increase the risk exposure to the price process of the underlying and increase the rate of specialist being "picked off" by arbitrageurs.

A. "Watch out for Arbitrageurs!" (the thickest arrows)
In situations where it is possible for specialists to attain more accurate information about the home market, a higher rate of "pick offs" will lead to higher levels of home market monitoring and better alignment of NYSE quotes with the quotes in the home market. Improvements in alignment will act remove the arbitrage opportunities that cause "pick offs".

B. "The Spreads will Keep them from Picking Me Off." (the 2nd thinnest arrows)

In reality, home market information may not be readily available on the trading floor, especially on a quote-by-quote or trade-by-trade frequency level. That hampers the ability to adjust the level of home monitoring based on the level of "pick offs". A more likely outcome of the "pick offs" is that the specialist will preventively raise his quoted bid-ask spread to minimize the chance of an arbitrage opportunity. The spread determines the lower bound on the size of the quote misalignment with the home market that is necessary to have a profitable arbitrage opportunity. Raising the spread raises the minimum necessary misalignment and reduces the chance that profitable "pick-off" opportunities arise. The specialist will often be willing to trade between the bid and ask quotes with counterparties that she suspects are not involved in arbitrage, to keep the effective spread the same as before. However, since investors cannot guarantee themselves a trade within the spread, perceived spreads would still be higher as a result of this behavior.

C. Home Market Advantage (the thinnest arrows)

The inflated spread is an artificial barrier to arbitrage, which may mean that the NYSE price may now deviate from the home market price for prolonged periods of time. This inefficiency in pricing will likely mean that investors may migrate to the home market, where price discovery is happening, in their search for the true price. This migration will dry up volume and liquidity of the ADR on the NYSE, reduce U.S. competitive forces in the stock and increase spreads. Spreads will reduce the pressures to reduce arbitrage exposure and home market monitoring intensity by the specialist, will make the quotes even further misaligned.

D. Monitoring Costs (the 2nd thickest arrows)

Home market monitoring activity of the specialist has a number of costs associated with it, including the costs of special equipment and information sources needed for monitoring and the opportunity cost of the time spent by the specialist on monitoring. These operational costs will be reflected in the specialist's quoted spreads and will automatically contribute to higher barriers for intermarket arbitrage, leading to lower monitoring pressures and fewer monitoring costs.

Insights

The moral of the loop above is that lack of adequate information about the quotes on competing exchanges will have a permanently damaging effect on efficiency. The reason is that, when faced with "pick offs", specialists will try to raise spreads to protect themselves from arbitrageurs. Since raising the spread reduces order flow and increases criticism from exchange

regulators, the spreads will not be able to go up by much. However, they will go up by as much as is necessary to keep the "pick offs" at an acceptable level. The only feasible way to eliminate the inflated spreads is to improve information delivery. Limit order traders will not act to narrow the spread without better information, because if they often submit orders that better the quote, the specialist will let them be "picked off" by the arbitrageurs. Since public (off-the-floor) traders have a much lower ability to adjust their order prices to changing circumstances, the pick offs cannot be easily avoided and the submission of competitive limit orders will be scarce (Stoll, 1992)

This effect is especially strong in stocks in which the NYSE is not the primary market, meaning that most of the trades with informational content and price discovery happen in another market. In that case when something of significant price impact occurs on the home market, the dealers or floor traders in that market are able to exploit their informational advantage and "pick off" NYSE quotes. The reverse actually happens quite a lot on the NYSE, as it happens to be the primary market in the great majorities of its listed securities. Thus, NYSE specialists are able to utilize their primary market advantage and often "pick off" dealers on competing regional exchanges using the Intermarket Trading System that connects all U.S. exchanges.

Policies

The exchange has a policy lever in the face of information delivery and monitoring technology that can help reduce the problems shown in this sector almost completely. Improvements in technology will both empower the "Watch out for Arbitrageurs" loop in fighting quote misalignments and arbitrageurs and reduce marginal monitoring costs making the use of monitoring an economically more feasible option. The NYSE proposal for a separate Non-U.S. trading floor may facilitate to reduction of home monitoring fees and give the specialists more tools to be able to align to the home market. From this perspective, the Non-US floor should be organized by country. Country-specific information and feeds from each particular home market can be delivered exclusively to those specialists who need them. Specialist firms will have to assign individuals who will focus on making markets in Non-U.S. securities on the new designated floor. These individuals will develop country and stock specific expertise that will allow them to grasp home country situations better and compete more squarely with the foreign markets.

3.5.4 The Regulation Sector

The NYSE is a Self Regulatory Organization that reports and gets audited by the Securities and Exchange Commission (SEC) of the United States government. It is largely responsible for day to day monitoring of exchange members and correcting their behavior when it gets out of line. Since Exchange regulators are usually pretty good at what they do, most of the time the SEC has few reasons to intervene. Thus, in our Regulation sector we focus on the effects of self-regulation and assume that any extra intervention by the SEC has largely the same effects on the market making system.

The regulators closely watch the trading of the specialists in particular and make sure that both the affirmative obligation to provide for a "fair and orderly market" and the negative obligations, which restrict the specialists menu of trading options, are fulfilled.

View #1 - Direct Effects on Spreads and Continuity

We first examine the direct effects of regulation on transaction cost measures, such as the spread and continuity of trading.

A. "Lower Markups, Please!" (the thinnest arrows)
Regulators keep a close eye on the spreads charged by specialists to make sure that they do not abuse their unique position of designated market makers on the primary market to attempt to earn monopoly profits via their spreads. To accomplish that, regulators can try to estimate the costs inherent in providing a market in the stock and make sure that the markups over those costs are reasonable. More practically, regulators watch for abnormally high spreads and talk to specialists about the reasons behind them. Thus, high spreads (which also hurt continuity due to the bid-ask bounce, please refer to the Background section and the explanation of continuity) lead to higher pressure by the regulators and adjustments of the spread markups sufficient to keep regulators "off the back".

The regulators have a number of ways of enforcing their policies on the specialist. Obvious ones are disciplinary actions, such as warnings, fines and in extreme cases the reallocation of specialist securities to another specialist unit. More subtly, regulators have input into the decisions of the Allocation Committee regarding the specialist firms that gets assigned new stocks when they first list on the NYSE. Specialists are very eager to obtain 'hot' new listings and try to cooperate with regulators as much as they can.

32

B. "Stabilize, Please!" (the 2nd thickest arrows)

In addition, regulators make sure that specialists "pay their dues" to the Exchange and abide by their affirmative obligation and stabilize the market by putting in more of their own order flow, when other stabilizing factors are weak. For example, if the limit order book is 'shallow' and a large market order comes in it is likely to move the market price a lot. The specialist interferes and takes the other side of a part or whole of the order to maintain price continuity. Higher regulation puts more pressure on the specialists to engage in stabilization, increases continuity and eventually brings regulatory pressures back to normal levels.

C. Specialist Voluntary Participation (the thickest arrows)

As was already mentioned in the Demand for Immediacy sector, specialists can trade for their own accounts as regular investors. In this role they enjoy some informational advantages, such as the knowledge of the order book , and some restrictions in the form of negative obligations. The advantage bestowed by the knowledge of the order book is largely gone since, today, the electronic order book is quite visible to floor brokers on the NYSE floor. When spreads are low, specialists look for other means of obtaining revenues and start trading more for their own accounts. This has a positive effect of further stabilizing the market, reduces regulatory pressures and allows the spreads to pick up a bit.

View #2 – Specialist's Informational Competitiveness

A. Regulation Implies Costs (the thickest arrows)

High regulation increases specialist's costs through putting various restrictions on her behavior that usually turn up as losses on her balance sheet. These costs increase the spreads charged by the specialists and lead to more regulation. The effect of this loop may be rather slight because regulators inquire about the nature of the costs and adjust their own behavior in response. This is an advantage of a self-regulatory system, regulators are fast and they constantly interact with the people they regulate. As a matter of fact, some of the people in the Regulation Department at the Exchange are former floor brokers or specialists themselves.

B. Costs Reduce Informational Competitiveness (the thinnest arrows)

Higher costs from regulation mean lower profits for the specialists. Some of the ways that specialist firms use their profits is by reinvesting them into their own business. In order to compete with third market dealers and upstairs brokers who have very advanced computer

systems and informational sources, specialists need to have money to upgrade their own technology. The NYSE provides a broad range of technological services, but they are rather uniform across specialists. Thus, for more specific needs, such as improved information about foreign home markets for Non-U.S., the specialists are on their own. They also have to provide their own means of execution on foreign markets if they want to maintain positions in the underlying security and hedge their risk in the ADR's. As a matter of fact, the Exchange hampers the ability of the specialists to hedge even if they can provide the technology, by restricting their options trading possibilities and telecommunications access to the outside world necessary to submit orders. The combined effect is that the specialist is short-handed with respect to informational competitiveness with institutional traders. As up-to-date information becomes more and more key to trading and the institutional investors capitalize on their ability to get the latest technologies, specialist's losses from better informed traders will mount and his costs will go up even further.

View #3 - Effects on Liquidity and Competition with Other Markets

A. Regulation Keeps Spreads Down, Helps Compete with Other Markets (the thickest arrows)
The high regulatory environment of the NYSE drives spreads down very close to the level of costs. Essentially the name of the game on the NYSE changes from that of profit-maximization to that of loss reduction and compliance followed by distant rewards in the form new allocations (Abolafia, 1996). This is not so across the board as some firms are able to capitalize on specialist advantages more than others. Yet, the spreads on the NYSE are still very low which keeps the Big Board extremely competitive on the spread basis. This competitiveness does mean higher order flow from spread-conscious consumers, and higher order flow keeps the exchange from wondering if the spreads are low enough and if they should be lowered even further through regulation.

B. Illiquidity Hurts Continuity, Increases Regulation (the 3rd thinnest arrows)
In the situation where order flow is low (please refer to loops C through E below), liquidity suffers and continuity goes down because fewer orders exist to stabilize markets. The effect propagates to the regulatory authorities which increase pressures on spreads attempting to increase competitiveness, if spreads happen to be the reason for low liquidity, competitiveness increases and order flow comes back. The problem is that often high spreads are not the causes of low order flow and further increasing spreads will hurt the system by increasing specialist

costs (Loop B of the last view) and hurting the discovery mechanism (see Loop D of this view below).

C. Order flow attracts Order Flow (the 2nd thinnest arrows)

We include the network externalities of order flow here (also present in the Demand for Immediacy and Competition Sectors) to emphasize that competitiveness with other markets is based not only on low (sometimes artificially low) spreads but also on the level of liquidity of the market. The NYSE regulation concentrates on spreads and often disregards directly attacking liquidity because of the belief that spreads alone will attract investors. In reality in thin NYSE stocks, the bottleneck for order flow may not be transaction costs, which are justifiably high because of costs, but the lack of a sufficient investor base that operates in the market.

D. Informational Disadvantages Hurt Price Discovery and Competitiveness (the thinnest arrows)

Lower specialist profits reduce the investments in information technology and the informational competitiveness of the specialist. This informational disadvantage is likely to result in impaired ability to set proper quotes and may lead to temporary mispricings. These impairments in price discovery are detrimental to investors and will contribute to investors to turning to various competitive mechanisms, not just for liquidity or transaction costs but also price discovery.

E. Lower Profits Hurt Innovation (the 2nd thickest arrows)

A major shortcoming of low specialist profits is the lack of incentives for the specialist to innovate. Innovation and entrepreneurial spirit are absolutely key to the ability of the Exchange to survive in a rapidly changing business, with more and more new potential competitors springing up. These qualities are also absolutely essential in the underdog situation that exists in many non-U.S. securities, where the NYSE has to play a novel role of a secondary market.

3.6 Insights and Policies from Causal Loop Modeling

The main insight of the causal loop model is that illiquidity, as a quality of markets problem, needs to be dealt with on its own and separately from bid-ask spreads. Low spreads are very good at encouraging order flow in liquid markets but may actually hurt market quality in 'thin' markets. One needs to seek the root of the liquidity problem in a particular stock or class of stocks, instead of simply blaming it on high spreads. The problem may be caused by a high degree of private information in the market, in which case higher disclosure and better investor

35

relations programs will help solve the problem, while spreads are likely to only worsen it. Alternatively, it may be that the specialist's meager profits are preventing her from getting the kind of technological and informational sources that she needs to be on the same level with her customers, the institutional investors. Moreover, the level of profits and regulatory environment may be so limiting that the incentives to innovate, looking for new ways of attracting liquidity and offering better services are missing. In that case, regulatory pressures should probably be lifted to generate incentives and appeal to the entrepreneurial spirit of the specialists. More personalized information feeds will help specialists monitor home market risk in non-U.S. securities and avoid high losses from arbitrageurs and having to raise spreads. In this respect, the NYSE idea of adding a non-U.S. floor may help to accomplish this 'personalization' of information feeds on a grand scale.

Last but not least, the Exchange may benefit from further acting on the realization that it is in a very different position with respect to non-U.S. securities than it is for domestic ones. For stocks from emerging market countries, it may be relatively easier to obtain a primary market status due to the world reputation and reliability of the NYSE, the likely superiority of its technology and an ability to offer great enhancements in liquidity. However, for developed country shares and emerging market shares listed on London and Frankfurt exchanges, network externalities (refer to the Demand for Immediacy Sector above) may prove to be very strong forces that will counteract the migration to the NYSE. Thus, the Exchange will need to learn how to operate as a secondary market and that this mode of operation and competition may be very different from what it is used to.

Chapter 4

The Stock and Flow Model

4.1 Motivation

The loops provide a qualitative description of the processes that drive the dynamics of liquidity provision and consumption on the NYSE. In order to understand how the feedback processes represented by the loops interact with each other and influence aggregate dynamic behavior, we have to choose one or more of the loop-sets and simulate them in a way that allows more precise analysis. The analysis of the simulation will tell us which loops dominate given a particular combination of model inputs. It may bring insights about behavior and underlying structures that were not visible through the loops. The simulation may uncover surprises that will lead to deeper insights into the real behavior. Finally, the simulation and analysis results will be good bases for the construction and advocating policies to improve overall exchange efficiency or to correct problems with liquidity in certain 'problematic' classes of stock.

4.2 Selection of Causal Loops to Model

The fundamental relationship being modeled is the positive feedback loop of liquidity in a stock on its risk characteristics, and feedback of the risk through market maker spreads back to liquidity. The model, as mentioned, is an attempt to further the understanding of this dynamic via exploring the underlying market structures and mechanisms involved.

4.3 Model Overview

What are the main concepts of the Stock and Flow model? We present the concepts in bullet point format here and expand on their implementation, justification and limitations in the sections to follow:

- The cornerstone of the model is the realization that the liquidity on the floor of the NYSE is largely driven by the interplay of market and limit orders and the specialist. Limit orders and the specialist represent the suppliers of liquidity, while market orders represent the demand side. The price of liquidity is the bid-ask spread that market orders pay. The lower the spread, the more liquidity market order investors will demand, although diminishing returns are exhibited since the liquidity is a function of percentage spread (as it detracts from the expected percentage returns on investment). The higher the spread, the more limit orders would be attracted to supply liquidity, since the opportunity cost of using the market order strategy becomes prohibitive for more and more investors. The specialist is also willing to supply more liquidity as she is compensated for more expected losses from taking on excessive inventory. Thus the spread dynamics is driven by the interplay of the combined supply curves of the limit order investors and the specialist and the demand curve of market orders.

- There is a fundamental distinction between the liquidity impact of the two main types of orders submitted to the NYSE: market and limit orders. Market orders consume available liquidity / immediacy and pay for this service, either to the specialist or to a limit order trader, in the amount of the bid-ask spread. Limit orders in the specialist's book provide liquidity/immediacy to incoming market orders and avoid paying the spread. In return, market orders enjoy almost instantaneous execution, while limit orders can remain in the book for long periods of time and may expire without ever executing.

- Fraction of market / limit orders is determined by the relative perceived costs of these two order strategies. The focus in this model is on the interaction of market and limit order risks and costs in general and not on the precise formulation of these costs. A comprehensive cost analysis of order strategies a very large number of risks and costs have to be taken into account: brokerage commissions, counterparty risk, information disadvantage, slippage,

market impact, etc. However, attempting to capture all of these costs in the model is likely not to be very rewarding from the point of view of insights about real investor behavior. Individuals often employ only one execution cost measure, the quoted bid-ask spread, into account, when making their order placement decisions. Institutional investors care about two cost variables – the spread and the expected loss due to adverse market impact (Schwartz and Whitcomb, 1988). Thus, the model uses a rather simplified cost structure with only one cost considered per each type of order. We discuss augmenting the number of costs considered with adverse selection and market impact in the Future Enhancement section below.

- Limit order costs are primarily associated with delays in execution and exposure to adversely changing market conditions in the interim. We may assume that limit order traders minimize waiting-related costs, by changing the terms of their order to market orders, after some time. The waiting costs can consequently be subdivided into the opportunity cost of waiting and an expected cost of switching into a market order at an unfavorable set of market conditions. The present model only considers the first component of cost – the interest rate associated with waiting a certain number of hours for execution. We do, however, work out the conceptual model of the expected switching cost in the Future Enhancement section.

- Market order costs are captured by the bid-ask spread, since market orders either trade with limit orders or the specialist and are required to pay a spread in both cases. Spreads are usually set so as to cover the costs incurred by the specialists from processing the orders left with him and taking on undesirable inventory. Most of this cost for order processing is covered by the commissions the specialist charges as a broker, however orders crossed in less than 2 minutes after arrival are not charged commissions by the specialist. The inventory holding costs are exacerbated by the specialist's inability to get rid of inventory quickly, thus they are inversely proportional to the number of limit orders in the book (the specialist uses the limit order book as the primary means of adjusting inventory). In addition, the spreads are bounded from below by the minimum tick, which is the minimum allowable price deviation on the NYSE. Currently the tick is a 'steenth' or one sixteenth of a dollar. It has been an eighth of a dollar for many years and until June of 1997.

- The magnitude of order arrival is primarily exogenous to the market mechanism. However, investors certainly become aware of transaction costs once they move significantly. Higher transaction costs will eventually drive investors to seek alternate trading mechanisms, other

ways of investing in the same underlying risk (e.g. through stock option contracts) or potentially even switching to other types of investments.

4.4 A Guided Tour of the Model

This section attempts to put the pieces of the model together in the reader's mind. The focus is on the interplay of the concepts presented in the last section that forms the carcass of the model. The style of presentation is based on a walk-through of the model, identifying key relationships and ideas and noting how these relationships are represented via equations.

The model, same as the causal loops developed in the earlier sections, had been implemented using a system dynamics package called Vensim from Ventana Systems, Inc (specifically, Vensim DSS32 Version 3.0B has been used). Vensim allows one to view and modify the model both graphically and textually. The graphical mode is most useful for a bird's-eye view of the model structure and an appreciation of the interactions between different variables. The textual mode presents the equations that govern the behavior of each variable, as well as the values of all constants and lookup functions corresponding to a particular simulation run. We provide the printout of both the graphical views and the equations from the textual view. These two are given in the Appendix. The present section refers continuously to specific views and variables in the graphical representation and sometimes to equations in the textual model. Thus, it may be helpful for the reader to follow the models in the above-mentioned Appendix to facilitate the understanding of the text below.

The following conventions are used in clarifying the description below:

- Model variable names are italicized, e.g. the *Spread* variable.

- The names of different graphical (Vensim) views, which hold groups of related variables are underlined, e.g. the Main view contains the variable *Market Orders*.

- The names of molecules, which are common system dynamics building blocks, will be designated by Courier New font as follows: `Split Flow`.

- To avoid confusion, the word 'stock' will be capitalized when used to designate a system dynamics term for an accumulation variable, e.g. Inventory Stock, and left as is when it refers to the financial term for a common share or shares of a public company, e.g. shares of stock.

We start our journey through the model with the variable called *order arrival* that appears to the left and bottom of the <u>Main</u> graphical view of the model. In the model we assume that the order generation process is mainly exogenous to the market mechanism (represented by the constant *normal order arrival rate)*, except for some smoothed effect of abnormally high or low transaction costs (*effect of perceived tcosts on order arr)*. *Order arrival* rate corresponds to the number of orders that arrive to the market every hour. Some of these orders are of market order type and some are limit orders. The proportion of order arrival that goes into *Market Orders* is given by the *fraction of market orders* variable. The rest of the arriving orders make up *limit order arrival* which is the inflow if the *Limit Order* Stock. The formulation of the total order flow splitting into subflows of *Market Orders* and *Limit Arrival Rate* is modeled after a common system dynamics structure or molecule (Hines, 1997) called a `Split Flow`.

The *Limit Order* Stock has a relation to the specialist's limit order book. A very important distinction between the Stock and a real order book is a lack of price structure. In effect, the *Limit Order* Stock is a counter of the number of limit orders in the order book. We assume that limit orders all arrive at the inside quote (specialist's best bid and ask), the price corresponding to the midpoint of the quote never changes (no price dynamics) and the numbers of buy and sell orders always equal (no buy or sell order imbalances). In that case, a simple counter of the number of orders together with the size of the spread (*Spread*) captures all of the relevant information about the order book.

One may ask why is it that *Limit Orders* are given the privilege of being modeled as a Stock, while *Market Orders* are forced to accept the poor lot of an Auxiliary variable. Indeed, in the real world, market orders are also temporarily stored in the limit order book until execution. At the same time, the average speed of execution for a market order is measured in seconds. For the purposes of the simulation, we assume that market orders execute during the same period they arrive and therefore need not be represented with a Stock.

On the floor of the NYSE, public (by public we mean orders that originated off of the trading floor) market orders always trade against public limit orders on the opposite side of the prevailing quote (buy market orders trade with sell limit orders), if such limit orders are available. If they are not available, market orders trade against specialist inventory. (Orders can also trade against interest in the trading crowd of floor brokers. However, we do not explicitly model the trading crowd, as a model simplification – please see the discussion in the Simplification section below). This ordering of liquidity sources is determined by the absolute precedence that public orders

have over specialist's interests on the NYSE. In the model, *Market Orders* are first matched with the available limit orders from the *Limit Order* Stock. The matching is implemented via a specification of the *limit execution rate* (the outflow of the *Limit Order* Stock) to be the minimum of the number of market orders and the size of the *Limit Orders* Stock. Thus, during a certain period, if the number of market orders exceeds the number of limit orders, all of the available limits are executed first (the *Limit Order* Stock becomes 0) and the rest of the market orders trade with the specialist. The simulated specialist always accommodates when her liquidity services are needed. The result is that specialist participation rate is not a value directly controlled by the specialist, but rather is a calculable fraction of current period market orders not matched by limit orders. If all *Market Orders* can be executed against the limit order book, the specialist's services never become needed and the resulting specialist participation rate for the period is 0.

This model assumes a constant trade size (all orders are for the same amount of shares) and focuses on the other determinant of share volume - the number of transactions per hour. We now take a short digression from the Main view to have a quick run through the laconic Volume view. Here we see that the only information-carrying variable that is needed to calculate the *Number of Trades per Hour* is the number of *Market Orders* (here *Market Orders* is shown in gray and in angle brackets to designate that it is a 'shadow' variable, defined in a different view). Indeed, since (1) every incoming market order is executed the same period it arrives, and (2) for every trade there is exactly one market order being executed - either against a limit order or the specialist, it follows that the number of trades per period is exactly equal to the number of market orders that arrive that period. *Trades per Market Order* always equals to unity and serves to convert the units from orders per hour for *Market Orders* to trades per hour for *Number of Trades per Hour*. The last step to calculating the desired *Share Volume per Hour* is to multiply the number of trades just obtained by the constant trade size captured by *Avg. Trade Size*. Note that since the constant number of shares being exchanged has been obtained as an empirical average trade size (= 2,000, Source: NYSE Consolidated Trade database for May, 1997) and the value taken on by Market Orders for most simulation runs falls within the usual range of number of trades per hour for an average, the resulting Share Volume Per Hour corresponds to reasonable NYSE hourly volumes.

The interaction between market orders, limit orders and the specialist described above forms the crux of the order execution mechanics of the model. The rest of the model is concerned, as explained below, with the measurement, perception and incorporation of trading costs into

42

investor decisions about the amount and type of orders to submit. We welcome the reader back to the <u>Main</u> view, to take a look at the *Avg. waiting time* variable, just below the Limit Order level. *Avg. waiting time* is an instance of the `Residence Time` molecule and it measures how long on average a limit order stays in the *Limit Orders* Stock. This quantity corresponds to an average time a limit order investor has to wait for her order to execute.

Execution waiting time is one of the variables limit order traders need to calculate the relative costs of the limit order strategy. However, *Avg. waiting time* is not immediately available to limit order trades. It takes time and perseverance to figure out how long a submitted order will spend in the limit order book. Limit order traders as a population figure out this value based on how long it took them to execute limit orders in the past. We capture this effect by using 'exponential smoothing', a procedure that delays and gradually incorporates the information about the value of a variable. *Perceived Wait Time* is thus a 'Smooth' function (please refer to Ventana Systems (1997a, 1997b)) of *Avg. waiting time*. The time constant or perception lag for smoothing is given by *wait perception time*. It takes about three time constants for a change in *Avg. waiting time* to be fully reflected in *Perceived Wait Time*. Limit order traders may have different costs associated with waiting an hour for an order to execute. Different average levels of urgency of limit orders result in different marginal costs of waiting. The latter is captured by the *cost per hour wait* variable.

One way of motivating the cost of waiting for limit order execution is to think of a limit order as a free option extended to the market by the investor. The option arises because the limit order specifies a firm price at which other investors can trade if they wish. For a buy limit order, for example, if the market price falls below the limit price of the order, other investors will choose to exercise the free option and the limit order will lose its value at maturity (however much the market price is lower than the limit price). Limit order traders want to reduce the value of the option they give away. One of the determinants of this option value is the maturity of the option. The longer the maturity the more likely it is, all else equal, that the option will be exercised with losses to the originator of the order. The cost of waiting variable captures this effect. The less the order has to wait for execution, the less likely it is that it will be providing an "in-the-money" option. (Stoll, 1992, p.85)

Chapter 5

Conclusions

Market making today is a complex and extremely dynamic business, affected by a variety of social and economic forces and requiring a fine balance between the pursuit of profit and the compliance to rules and regulations. At the same time, the performance of market makers and specialists has a profound effect on the quality and efficiency of securities exchange markets, the foundation of the modern financial system. While economic literature has devoted a lot of time to the analysis of market maker behavior, especially in the areas of optimal pricing and inventory control, we are yet to understand the behavior of real-world market makers. Academic studies of this behavior tend to focus on singular aspects of this behavior, and there is more to be done to appreciate the richness of interactions between the various social and economic roles specialists and market makers play.

This study has chosen to employ the system dynamics approach, able to model systems with complex dynamic interactions and "soft" variables required to represent social effects. Modeling proceeded in stages, with the first stage being the conceptualization of the dynamic effects of regulation, competition, the informational content of orders and the investors' demand of immediacy on the NYSE specialist's ability to provide sufficient liquidity. During that stage we extracted a number of objects in the real world and captured them in model variables. We obtained information about the behavior of key variables with respect to time to make sure that the model corresponds to reality and is able to reproduce the current mode of behavior. We then established the causal relationships of the variables based on their dynamic interactions in the real-world system. That gave us an ability to obtain a set of feedback loops that conceptualized the endogenous behavior of the system. The loops were useful in explaining current behavior and also in identifying opportunities to change behavior for the better. Specifically we found that the current policy of increasing regulation to depress spreads, with an intent to evoke higher volumes has a number of adverse side-affects. These side-affects are only visible when one looks at a broader picture of the system and takes into account issues such as informational trading and the arbitrage that is associated with market linkages. The state of affairs in non-U.S. stocks on the

NYSE presents a situation where the "lower spreads" policy is not able to produce the behavior it desires because of an overwhelming number of negative side-affects. The analysis of the causal loop calls for a closer look at the exact causes of the break down of liquidity in these stocks and reveals levers that policymakers can use to fight specific causes.

In the next phase of this project we chose to take a closer look at some of the loops discovered by building a simulation model of a set of loops from the Demand for Immediacy sector. In building that model, we tried to obtain an operational definition of the variables in the loops and to formulate a functional form of the response of some variables to changes in others. An operational definition called for a formalization of the demand and supply sides of immediacy, which we captured by focusing on the dichotomy between market and limit orders. The strengths of the resulting model are its simplicity coupled with a fair ability to capture real-world phenomena. The key to obtaining these strengths was a careful choice of assumptions that limit the dynamic complexity, while maintaining real-world basis. Another key to success in modeling is an iterative approach, where elements of structure are added gradually, and an extra element is not added until a very good understanding of the behavior of the existing model is attained.

There is a number of possible ways to take the present work into the future. First of all, the simulation model of the demand for immediacy sector has a lot of potential enhancements that may bring further understanding. One area that is particularly interesting to explore is the addition of trade size dynamics. Our model presently does not consider variations in the mix of institutional and individual investors. In the market, this mix determines the level of "demand for depth", since institutional traders tend to trade in larger quantities and care a lot more about market impact costs.

Value can also be obtained by simulating the other sectors of the model. For example, the current simulation model is particularly well suited to the analysis of adverse selection risks that are the cornerstone of the Informational Content of Order Flow Sector. Adding the Ratio of informed to uninformed market orders as a variable in the model, allows us to quantify the probability of a liquidity provider dealing with a better informed counterparty. There is a certain monetary loss for either a limit order trader or a specialist from trading with the informed. The expected value of the loss from adverse selection, is simply the probability of dealing with informed times the loss. The expected adverse selection loss is simply another one of the costs associated with both the limit order strategy for investors and the specialist. The specialist will

45

use the cost in calculating the minimum spread to be charged. This example shows the extensibility of the present model to enhancements that at first seem to be well outside its scope.

Last, but not least, yet another set of enhancements can be obtained from looking at issues outside the causal loop model. One of these issues is the consideration of the specialist's total portfolio of stocks as opposed to the current focus on individual stocks. A rational specialist uses his portfolio of securities to try to diversify as many unsystematic risks as possible. Thus, the relative costs and risks of stocks in the portfolio as well as the covariance of the risks across stocks will affect the amount of commitment the specialist will try to obtain in each individual stock. A model of cross-stock decisions by the specialist may reveal further insights about the process of liquidity provision. It may also have important policy implications about how to take into account the existing specialist portfolios in deciding how to allocate new securities to specialists.

Bibliography

Abolafia, M. Y., 1996, *Making Markets: Opportunism and Restraint on Wall Street* (Harvard University, Cambridge, MA).

Adler, P. and P. Adler, eds., 1984, *The Social Dynamics of Financial Markets* (JAI Press, Greenwich, CT).

Black, F., 1971, "Toward a Fully Automated Stock Exchange", Parts I and II, *Financial Analysts Journal*, July-August, 29-35, 44, November-December, 24-28, 86-87.

Black, F, 1986, "Noise", *The Journal of Finance* 51, 529-543.

Cohen, K., Maier S., Schwartz R., and D. Whitcomb, 1983, "A Simulation Model of Stock Exchange Trading", *Simulation* (November), 181-191.

Cohen, K., Maier. S., Schwartz, R., and D. Whitcomb, 1986, *The Microstructure of Securities Markets* (Prentice Hall, Englewood Cliffs, NJ)

Copeland, T. and D. Galai, 1983, "Information Effects and the Bid-Ask Spread", *Journal of Finance* 38, 1457-1469.

Demsetz, H., 1968, "The Cost of Transacting", *Quarterly Journal of Economics* 82, 33-53.

Easley, D., and M. O'hara, 1987, "Price, Trade Size, and Information in Securities Markets", *Journal of Financial Economics* 19, 69-90.

Garman, M., 1976, "Market Microstructure", *Journal of Financial Economics* 3, 257-275.

Glosten, L. and Milgrom, P., 1985, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders", *Journal of Financial Economics* 14 (March) 71-100.

Hakansson, N., A. Beja and J. Kale, 1985, "On the Feasibility of Automated Market Making by a Programmed Specialist", *The Journal of Finance* 40, 1-20.

Hasbrouck, J. and G. Sofianos, 1993, "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists", *Journal of Finance* 48, 1565-1594.

Hasbrouck, J., Sofianos, G., and D. Sosebee, 1993, "New York Stock Exchange Systems and Trading Procedures", Working Paper 93-01 (New York Stock Exchange, New York, NY).
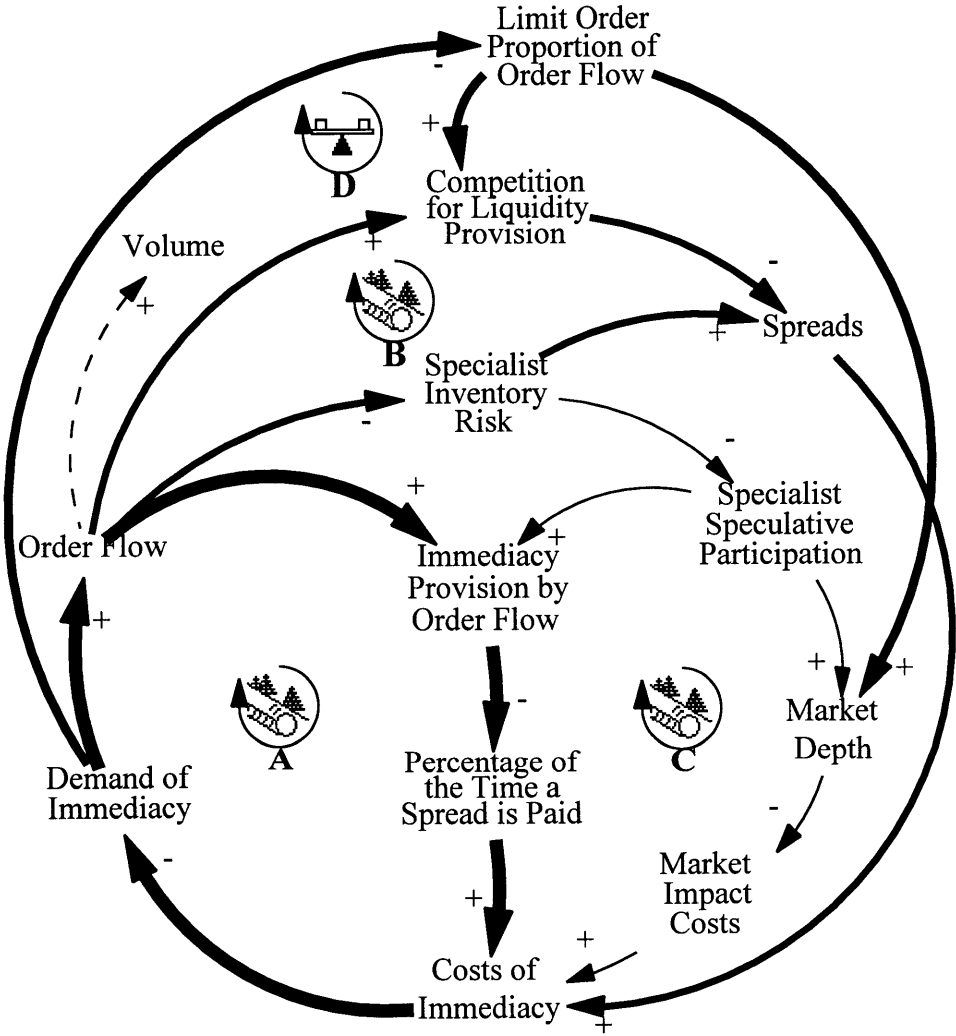
Hines, J, 1987, "A Behavioral Theory of Interest Rate Behavior", Working Paper WP-1951-87 (Sloan School of Management, MIT, Cambridge, MA).

Hines, J., 1997, "Molecules of Structure: Building Blocks for System Dynamics Models", Version 1.3.

Ho, T. and H. R. Stoll, 1981, "Optimal Dealer Pricing Under Transactions and Return Uncertainty", *Journal of Financial Economics* 9 (March).

Ho, T. and H. R. Stoll, 1983, "The Dynamics of Dealer Markets Under Competition", *Journal of Finance* 38, 1053-1074.

Huang, R. G. and H. R. Stoll, 1996, "Dealer versus Auction Markets: A Paired Comparison of Execution Costs on NASDAQ and the NYSE", *Journal of Financial Economics* 41, 313-357.

Macey, J. R. and M. O'Hara, 1997, "The Law and Economics of Best Execution", Working Paper 97-05 (New York Stock Exchange, New York, NY).

Madhavan, A. and G. Sofianos, 1994, "Auction and Dealer Markets: An Empirical Analysis of NYSE Specialist Trading", Working Paper 94-01 (New York Stock Exchange, New York, NY).

O'Hara, M., 1995, *Market Microstructure Theory* (Blackwell, Cambridge, MA).

Peterson, C., 1996, "Risk and Return Characteristics of American Depository Receipts", *The Journal of Investing* (Fall), 74-85.

Schwartz, R., 1991, *Reshaping the Equity Markets: A guide for the 1990s* (HarperBusiness, New York, NY).

Smidt, S., 1971, "Which Road to an Efficient Stock Market", *Financial Analysts Journal* 27 (September-October), 18-20, 64-69.

Sofianos, G., 1995, "Specialist Gross Trading Revenues at the New York Stock Exchange", Working Paper 95-01 (New York Stock Exchange, New York, NY).

Sofianos, G. and I. Werner, 1997, "The Trades of NYSE Floor Brokers", Working Paper 97-04 (New York Stock Exchange, New York, NY).

Stoll, H. R., 1985, "The Stock Exchange Specialist System: An Economic Analysis", *Monograph Series In Finance And Economics*, 1985-2 (New York University, Salomon Brothers Center, New York, NY).

Stoll, H. R., 1992, "Principles of Trading Market Structure", *Journal of Financial Services Research* 6, 75-107.

Stoll, H. R., 1997, "Affirmative Obligation of Market Makers: An Idea Whose Time Has Passed?", Working Paper 97-14 (Financial Markets Research Center, Owen Graduate School of Management, Nashville, TN).

U.S. Securities and Exchange Commission, 1971, *Institutional Investor Study Report* (U.S. Government Printing Office, Washington, D.C.).

Velli, J., 1994, "American Depository Receipts: An Overview", *Fordham International Law Journal* 17, S38-S57.

Ventana Systems, 1997a, "Vensim Modeling Guide", Version 3.0B.

Ventana Systems, 1997b, "Vensim Reference Manual", Version 3.0B.

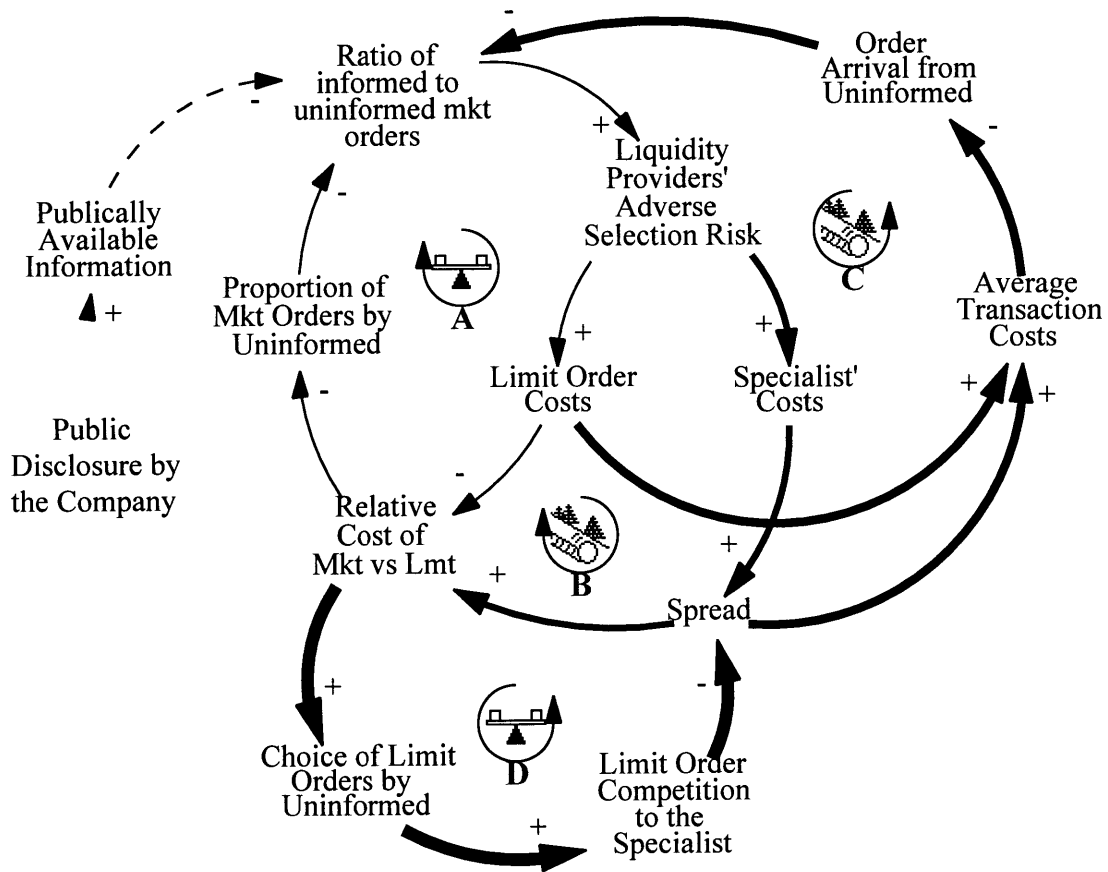Ventana Systems, 1997c, "Vensim User's Guide", Version 3.0B.

Appendix

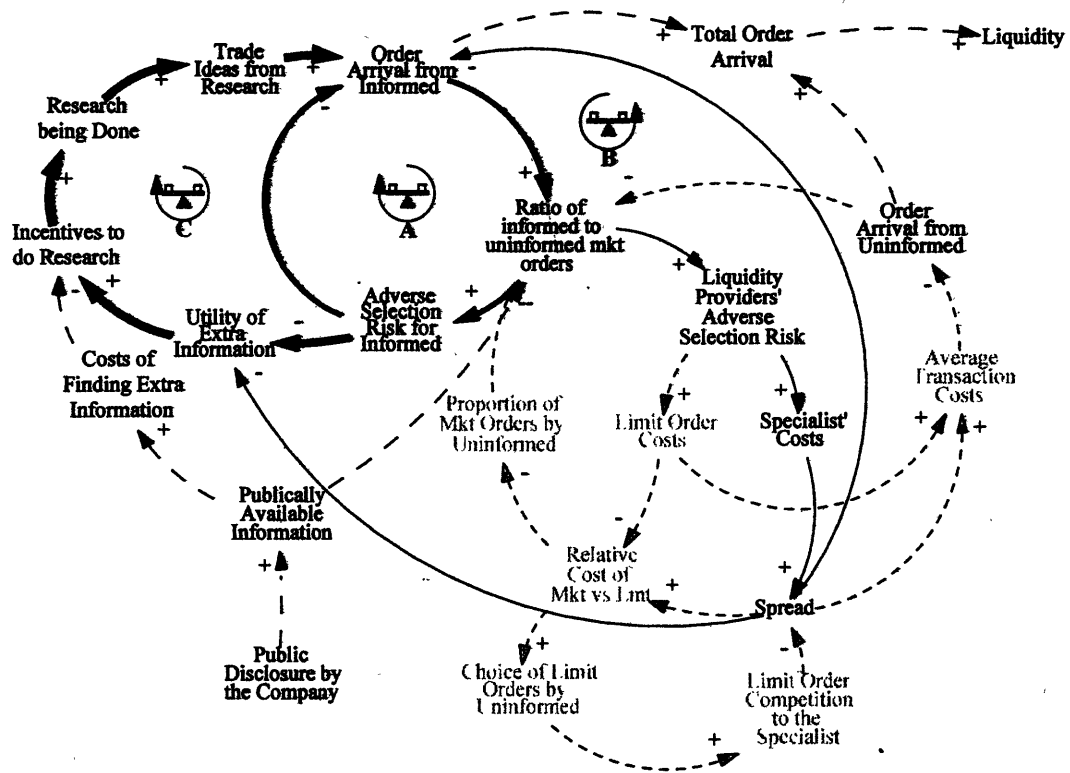## *A.1 Causal loops*

The Demand for Immediacy Sector

Limit Order
Proportion of
Order Flow

Competition
for Liquidity
Provision

D

Volume

B   Specialist
Inventory
Risk

Spreads

Order Flow

Specialist
Speculative
Participation

Immediacy
Provision by
Order Flow

Market
Depth

Demand of
Immediacy

A

Percentage of
the Time a
Spread is Paid

C

Market
Impact
Costs

Costs of
Immediacy

The Informational Content of Order Flow Sector

View #1: Effects on the Liquidity Providers

The Competition for Liquidity Provision Sector
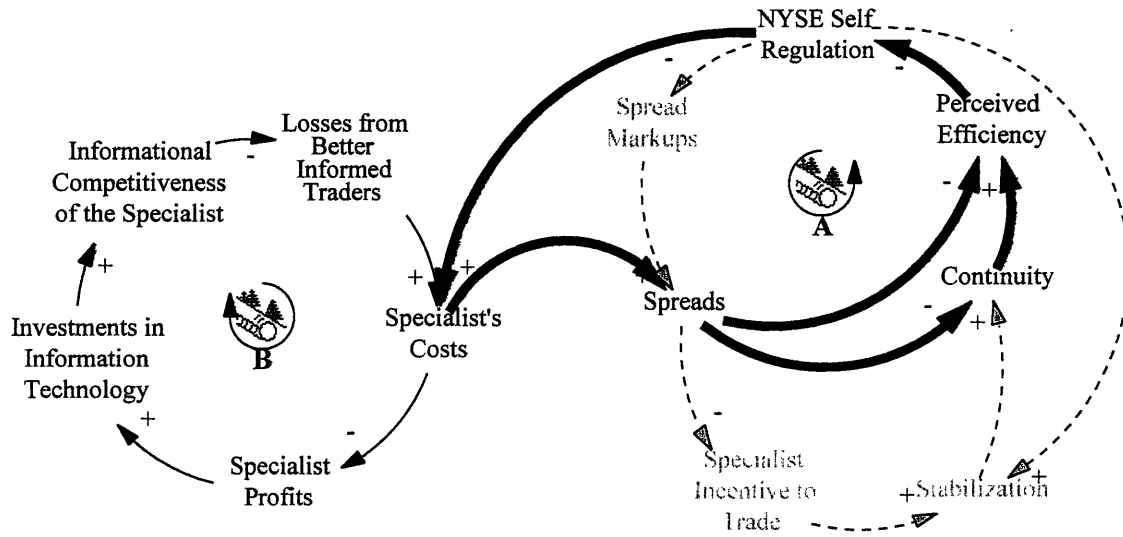
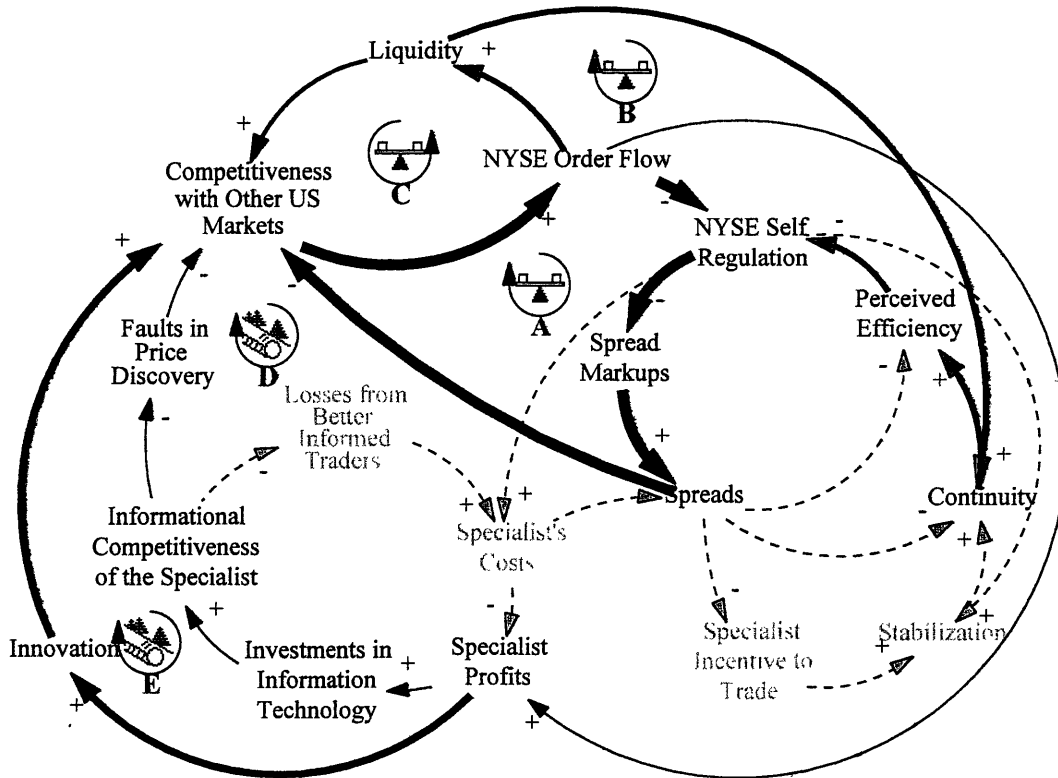*View #1: U.S. Competition for Liquidity Provision*
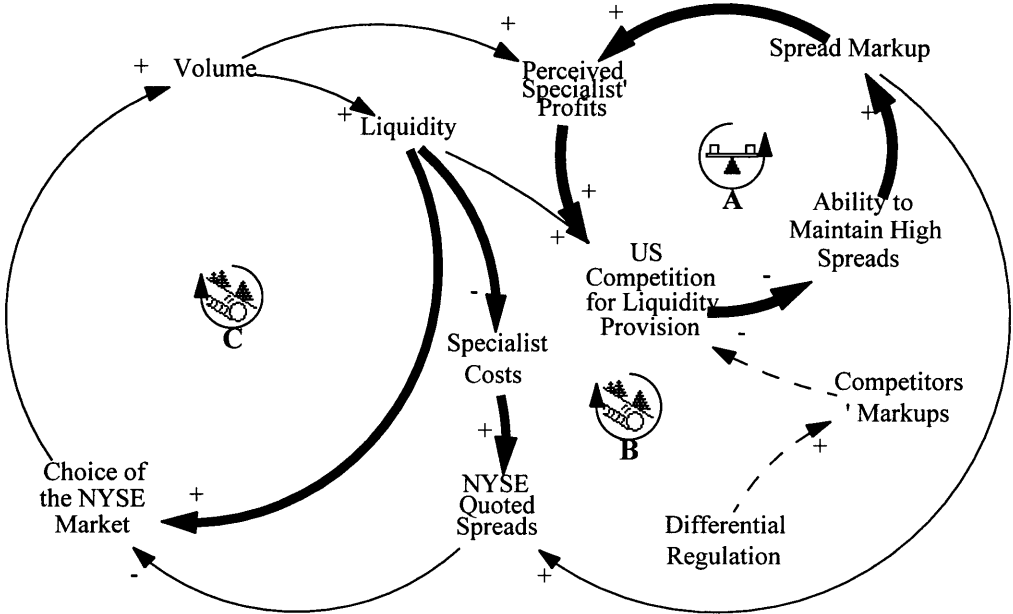
The Regulation Sector

View #1: Direct Effects on Spreads and Continuity
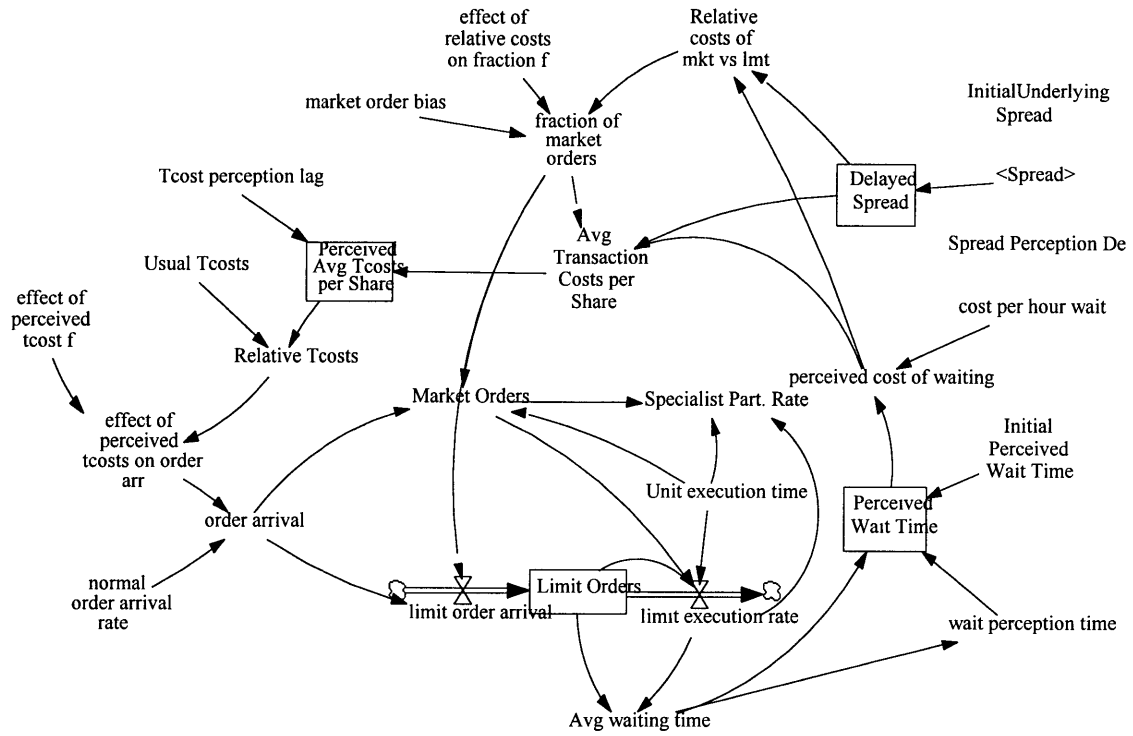
View #2: Specialist's Informational Competitiveness

Losses from
Better
Informational — Informed
Competitiveness Traders
of the Specialist

NYSE Self
Regulation

Spread
Markups

Perceived
Efficiency

Investments in
Information
Technology

Specialist's
Costs

Spreads

Continuity

Specialist
Profits

Specialist
Incentive to
Trade

+Stabilization

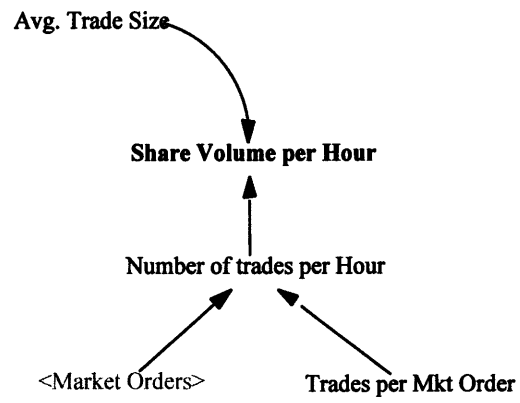# A.2 Stock and Flow Model

View #1: Main

View #2: Volume

Avg. Trade Size

Share Volume per Hour

Number of trades per Hour

<Market Orders>   Trades per Mkt Order

59