Systematic Liquidity and Learning about the Risk Premium

Gideon Saar\textsuperscript{1}

This version: August 2006

\textsuperscript{1}Johnson Graduate School of Management, 455 Sage Hall, Cornell University, Ithaca, NY 14853, e-mail: gs25@cornell.edu. Tel: 607-255-7484. Fax: 607-255-5993. I wish to thank Yakov Amihud, Shmuel Baruch, Henry Cao, David Easley, Joel Hasbrouck, Robert Jarrow, Charles Lee, Richard Lyons, Roni Michaely, Maureen O’Hara, Bhaskaran Swaminathan, Ingrid Werner, Lei Yu and seminar participants at Cornell University, Dartmouth College, New York University, the University of California at Berkeley, the University of Texas at Austin, and the University of Wisconsin at Madison for helpful comments.
Abstract

This paper presents a rationale for systematic liquidity and links time variation in the risk premium to liquidity. The driving force behind endogenous liquidity in the model is uncertainty about the preferences and endowments of investors. Learning about the risk premium gives rise to the price impact of trades because the order flow is used to make an inference about the preferences and endowments of the investors who submit orders. The model suggests that shocks that increase uncertainty about the risk premium would harm liquidity. Both the risk premium and liquidity are time-varying even between such shocks because transaction prices incorporate current beliefs about the investor population, and these beliefs are changing as the market learns from the order flow. Liquidity is therefore related to the risk premium in this framework not because it is a cost that needs to be priced but because it is how the market learns about the risk premium.
1 Introduction

Recent empirical evidence on systematic liquidity and the way in which time variation in liquidity relates to expected returns raises important conceptual questions. It is unclear what economic forces cause these patterns, as determinants of liquidity have been discussed most often in the context of a specific asset rather than the entire market. This paper contributes to the literature by showing how uncertainty about investors’ preferences and endowments creates uncertainty about the risk premium in the economy, which in turns generates endogenous illiquidity. Price impact or illiquidity is related to the risk premium in this model not because it is a cost (or risk) that needs to be priced but because illiquidity is the result of the market’s attempt to learn about the true state of the risk premium.

Extant literature suggests that illiquidity is caused by order processing costs, inventory costs, and adverse selection costs. Several recent empirical papers note that existing paradigms could potentially account for commonality in liquidity (e.g., Chordia, Roll, and Subrahmanyam, 2000; Huberman and Halka, 2001; Hasbrouck and Seppi, 2001). While inventory costs depend on the holdings of a specific asset, the inventory risk of dealers in individual assets could be affected by systematic variations in volume or volatility as well as changes in interest rates. Adverse selection costs are usually discussed in the context of private information about the future cash flows of a single asset, but investors could have private information about common factors (e.g., Subrahmanyam, 1991). As for the dynamic relation between market liquidity and returns, Amihud (2002) and Jones (2002) suggest an explanation in the spirit of Amihud and Mendelson (1986). The relation they posit is that lower liquidity should be associated with a lower price level and higher expected returns to compensate for the higher costs of trading.

This paper provides a rationale for systematic liquidity and links time variation in the risk premium to liquidity. The driving force behind endogenous liquidity in this paper is

\[1\] Market-wide variations in the level of noise or liquidity trading could also affect the adverse selection costs of all stocks.
uncertainty about the preferences and endowments of investors. This assumption seems fairly intuitive since these attributes of investors are inherently unobservable.\textsuperscript{2} In addition, different investors arrive to trade in financial markets at different times, further complicating the task of learning about the overall distribution of preferences and endowments of the investor population. This uncertainty about the investor population creates a problem with respect to pricing assets because it implies uncertainty about the risk premium. The order flow communicates the trading desires of investors and can be used to extract information about their preferences and endowments and hence about the risk premium.

I develop a simple single-asset model with two classes of investors who differ with respect to their risk aversion and endowments. The model resembles a traditional sequential trade model but has some unique features.\textsuperscript{3} Unlike the typical setup in a traditional sequential trade model, there is no information asymmetry about future cash flows in the economy. Rather, the distribution of investors’ preferences and endowments is uncertain. While the price of an asset in an information-asymmetry-driven sequential trade model is the conditional expected value of the asset’s future cash flows, prices here are set by an expected market-clearing condition that can be viewed as the essence of “reduced-form” market makers who care about supply and demand during the trading period (see also Garman, 1976; Amihud and Mendelson, 1986; Brock and Kleidon, 1992).

The model maintains the rational expectations requirement of the traditional information-asymmetry-driven sequential trade models that each order is executed at a price that reflects its information content. Learning about the risk premium gives rise to the price impact of trades because the order flow can be used to make an inference about the preferences and endowments of investors who submit orders. For example, an investor who submits a buy

\textsuperscript{2}Rubinstein (2001) writes “Perhaps the most important missing generalization in almost all work on asset prices thus far is uncertainty about the demand curves (via uncertainty about endowments or preferences) of other investors. This uncertainty injects a form of “endogenous uncertainty” into the economy that may be on a par with exogenous uncertainty about fundamentals” (p.23).

\textsuperscript{3}I am using the term “traditional sequential trade models” to denote the models of Glosten and Milgrom (1985), Easley and O’Hara (1987, 1991, 1992), Diamond and Verrecchia (1987), and others who have employed a similar framework.
order reveals that he is less risk averse or has a smaller endowment (or both) than an investor who submits a sell order. Market prices are then raised to reflect the new information about the investor population that there may be more investors who are less risk averse or have small endowments.4

The price impact of trades (or the spread) in the model is therefore a consequence of the uncertainty about the preferences and endowments of investors who submit market orders. This creates a direct link between the asset’s risk premium and liquidity. One can interpret the risky asset in the model as representing the market portfolio. This interpretation seems natural because uncertainty about the investor population should be common to all assets. The model demonstrates that any shocks that would increase uncertainty about the risk premium of the market portfolio would result in deterioration of liquidity for all risky assets. In addition, the risk premium is time-varying even between such shocks because transaction prices incorporate current beliefs about the investor population and those beliefs are changing as the market learns from the order flow. The extent of illiquidity is therefore also time-varying as it reflects the dynamics of changed beliefs about the risk premium.

The model also establishes a link between liquidity and welfare. When there is greater uncertainty about the risk premium, a buyer is forced to transact at a higher price and a seller at a lower price. Faced with less favorable prices, investors want to buy or sell smaller quantities and therefore there is less risk sharing in the economy, lowering the welfare of all investors.

Most theoretical papers that look at the relation between liquidity and returns on risky assets exogenously assume the existence of illiquidity (or transaction costs) in the market (e.g., Amihud and Mendelson, 1986; Constantinides, 1986; Vayanos, 1998; Acharya and Pedersen, 2005; Lo, Mamaysky, and Wang, 2004). In contrast, illiquidity in this paper is endogenous because it is created from basic uncertainty about investors’ preferences and

---

4 An alternative view of uncertainty about investors in this paper is to see it in terms of information about future order flow. The arrival of an investor is used to make an inference about the entire population of investors like the inference from a sample about a population. Hence, the arrival of an order provides information about the nature of future order flow and therefore causes a change in the price of the asset.
endowments, which gives rise to the price impact of trades as the market learns about the risk premium. This paper therefore is related to some recent attempts to link endogenous liquidity and returns of risky assets. Eisfeldt (2004) studies a dynamic economy where productivity level affects both returns and the amount of rebalancing trades (hence liquidity). She does not, however, look at the price impact dimension of liquidity. Baker and Stein (2004) use information asymmetry about future cash flows, short sale constraints, and irrational investors who underreact to information in the order flow to generate a link between returns and liquidity.

In the context of risk-free bonds, Gallmeyer, Hollifield, and Seppi (2005) look at liquidity and prices using a three-period model where there are investors with different holding periods. Short-horizon investors do not know the time preference of long-horizon investors and therefore are uncertain about the future price of the bond, which creates a link between liquidity and returns. Their model shares with the information asymmetry paradigm the approach whereby there are two types of investors, informed and uninformed. Instead of information asymmetry about the future cash flows of the asset, long-horizon (informed) investors know their own time-preference parameter, while short-horizon (uninformed) investors do not know the long-horizon investors’ time-preference parameter. Gallmeyer, Hollifield, and Seppi also discuss an extension of their model with multiple risky assets where this information asymmetry about time preferences is associated with commonality in return movements across stocks.

Since the driving force in the model presented here is uncertainty about investors in the market, this paper is also related to the literature that investigates the effect of market participation on the prices of assets. In particular, Kraus and Smith (1989) stress that uncertainty about future prices can reflect the beliefs, preferences, and endowments of the participants in the economy. They refer to this uncertainty as “market created risk” to emphasize that its source is the investors themselves rather than the future cash flows of a

---

The approach taken here differs from theirs along several dimensions, among which is the sequential arrival of investors and, most importantly, the recognition that information about the investor population is carried by the order flow and affects prices in the market. The model in this paper suggests an important link between market microstructure and asset pricing. A standard result in asset pricing is that if there are more (less) risk-averse investors in the economy, the prices of risky assets will be lower (higher). Here I show that if there is more (less) uncertainty about risk aversion of investors, the liquidity of risky assets would be lower (higher). Therefore, the model offers an explanation for common factors and intertemporal patterns in liquidity that have been documented in the literature (e.g., Chordia, Roll and Subrahmanyam, 2000, 2001; Hasbrouck and Seppi, 2001; Huberman and Halka, 2001; Fujimoto, 2004). The model also links time-varying risk premia to time variation in liquidity, a relation that was empirically studied by Amihud (2002) and Jones (2002). Finally, the model provides implications about the relation between uncertainty about the risk premium and liquidity that could be tested in future empirical work.

The remainder of the paper is organized as follows. Section 2 describes the economy and establishes the existence of an equilibrium. Section 3 investigates the implications of uncertainty about the preferences and endowments of investors for liquidity and welfare. Section 4 concludes with a discussion of the approach pursued in the paper.

2 The Economy

There are two assets in the economy: a risky asset that pays $\tilde{u}$ dollars at time $T'$, where $\tilde{u}$ is normally distributed with mean $\theta$ and variance $\sigma^2$, and a riskless bond that pays $R$ dollars.

---

6A few other recent papers also emphasize the existence of information in financial markets other than information about future cash flows. See Madrigal (1996), Lyons (1997), and Cao, Evans, and Lyons (2006).

7There is also a literature that considers the effects of random preferences and endowments in Walrasian exchange economies (e.g., Hildenbrand, 1971; Bhattacharya and Majumdar, 1973; Mendelson, 1985). The implication that demand variability relates to trading costs is also discussed in Spiegel and Subrahmanyam (1995). They use simulations to investigate the effects of exogenous supply shocks on the intraday risk premium.

8There are also papers that construct measures of market liquidity (or liquidity innovations) and relate them to the cross-section of expected returns (see Pastor and Stambaugh, 2003; Acharya and Pedersen, 2005).
at time $T'$. Trading in the risky asset takes place in discrete intervals of time, called trading periods, denoted $t = 1, 2, \ldots, T$, where trading ends before the liquidating dividend of the risky asset is realized ($T < T'$). As in traditional sequential trade models (e.g., Glosten and Milgrom, 1985; Easley and O’Hara, 1987, 1992), each trading period is long enough to accommodate at most one trade and is comprised of the following sequence of events. Market makers post the “quote” that constitutes a firm commitment to buy or sell the specified quantities at the specified prices. An investor who arrives at the market decides optimally how many shares to buy or sell at the quoted prices. He then submits an order, the order gets executed, and the investor leaves the market. The market makers have the opportunity to revise their quote before the arrival of the next investor, and the sequence of events starts again.

2.1 Investors

There are two classes of investors in the population indexed by $i \in \{1, 2\}$. All investors maximize Constant Absolute Risk Aversion (CARA) expected utility of their wealth at time $T'$ (when the liquidating dividends of the assets are realized).\footnote{Negative exponential utility functions and normally distributed liquidation payoffs are used here for mathematical tractability. This framework has known drawbacks (e.g., unlimited liability) and results in prices that can be negative for some parametric specifications. In what follows, I assume that the parameters of the economy are such that prices are non-negative.} Investor classes differ with respect to their endowments ($\bar{X}_i$ of the risky asset and $\bar{Z}_i$ of the riskless bond) and their coefficient of absolute risk aversion, $\alpha_i$. The fraction of class 1 (class 2) investors in the population is $q$ ($1 - q$), which in the model is the same as the fraction of trades coming from class 1 (class 2) investors. An investor who arrives at the market trades to optimally rebalance a portfolio that consists of his initial endowments. Investors behave competitively in that they take market prices as given and decide how much to invest in the risky asset and the riskless bond.

Because the optimization problem of an arriving investor is the same in each trading period, I drop the subscript $t$ to simplify the exposition of the model in this section. Under
these assumptions, a class $i$ investor who arrives at the market solves the following problem:

$$\max_{Y_i} \ E \left[ -e^{\alpha_i W_i,T'} \right]$$

s.t. $RZ_i + \bar{u}Y_i = W_i,T'$

$$Z_i + P_iY_i = \bar{Z}_i + P_i\bar{X}_i$$

where $Y_i$ is the investor’s demand for the risky asset, $P_i$ is the price at which the investor can transact in the risky asset, and the price of the riskless bond is set to unity. The solution to this problem is well known and the optimal demand is: \(^{10}\)

$$Y_i^* = \frac{\theta - RP_i}{\alpha_i \sigma^2}$$

When an investor arrives at the market, he submits an order,

$$X_i = Y_i^* - \bar{X}_i$$

where $X_i > 0$ ($X_i < 0$) is interpreted as a buy (sell) order.

As in traditional sequential trade models, each period an investor is randomly and anonymously selected to trade from among the pool of investors.\(^{11}\) The probability that an investor who arrives to trade belongs to class 1 is $q$ and the probability that he belongs to class 2 is $1 - q$. In this paper I focus on the effects of uncertainty about the distribution of investors’ preferences and endowments. This uncertainty is modeled by assuming that no one knows the value of the relative population weight $q$.

**2.2 Market Clearing**

Asset pricing models usually generate prices using a market-clearing condition. Such a condition requires that all investors submit orders at the same time, and hence it is difficult to apply in sequential markets where orders arrive one at a time. However, the intuition

\(^{10}\)See, for example, Grossman and Stiglitz (1980).

\(^{11}\)Nothing changes in the results if the model is extended to allow periods without trading. As I will show later, prices are adjusting only when an investor arrives and his order reveals to which class he belongs (in the absence of public information arrival). Unlike in Easley and O’Hara (1992), prices here do not change in periods without trading.
behind market clearing as the determinant of prices remains in that prices in a secondary market must adjust over sufficiently long horizons such that the number of shares bought and sold by investors is the same (without the creation of additional shares of the risky asset). This intuition is implemented here by requiring that prices are set such that the expected excess demand in each period conditional on all available information is equal to zero. The expected market-clearing condition can be viewed as the essence of “reduced-form” market makers who help the market overcome the problem of intertemporal disaggregation. Market makers are needed because investors arrive one at a time, but market makers simply balance the expected flow of shares bought and sold.12

One way to motivate such an expected market-clearing condition is to adopt a formulation of market makers similar to that used by Garman (1976), Amihud and Mendelson (1986), and Brock and Kleidon (1992), where market makers maximize expected profit per period subject to the constraint that the expected number of shares bought and sold in a period is equal to zero. As in Amihud and Mendelson (1986), “The market makers’ inventories fluctuate over time to accommodate transitory excess demand or supply disturbances, but their expected inventory positions are zero” (p. 225). If market makers earn zero expected profit, it can be shown that solving their profit maximization problem subject to a zero inventory-drift constraint yields the same price as the one that comes out of the expected market-clearing condition.

There are other potential ways in which the intuition of long-run market clearing can be implemented in a sequential market. For example, the market makers’ objective could be that they end trading at time $T$ with expected inventory of zero. Market makers would

---

12The notion of prices that are determined by equating the flow of shares demanded and supplied rather than by forecasting the future cash flows of a firm seems to correspond rather well to the activity of market makers. Bagehot (1971) writes that “it is well known that market makers of all kinds make surprisingly little use of fundamental information. Instead, they observe the relative pressure of buy and sell orders and attempt to find a price that equilibrates these pressures” (p. 14). Mayer (1988) also notes that market makers are not interested in taking a position in the stock based on long-term forecasts. Rather, they are constantly searching for the prices at which the flows of shares bought and sold are approximately equal. Their constant search for a market-clearing price keeps them in business since their inventories do not drift without bounds. This behavior sets them apart from investors who trade infrequently to re-balance a portfolio of investments they hold for prolonged periods of time.
then adjust prices to affect the order sizes of arriving investors so that their inventories are brought back to zero. Uncertainty about investors’ preferences and endowments in this alternative framework produces implications that are similar to those presented here.\textsuperscript{13} My focus in this paper is on the information effects created by learning about the risk premium, and therefore I abstract from specific characteristics of market makers (such as their degree of risk aversion or wealth constraints) that would cause prices to be affected by the level of inventory. Market makers are not real agents in the model; rather, the expected market-clearing condition simply formalizes the idea that market makers’ inventory should not have a drift ex-ante.

The use of the expected market-clearing condition to generate prices is probably the least familiar and more objectionable element in the model. Its main advantage is a simple correspondence from demand to prices that enables us to study in a very transparent framework how information about the risk premium is carried by the order flow and how it relates to liquidity.\textsuperscript{14}

\subsection*{2.3 Rational Expectations}

The equilibrium price of the risky asset in traditional sequential trade models is the conditional expected value of the liquidating dividend.\textsuperscript{15} In other words, the preferences and endowments of investors do not matter for pricing. In contrast, these attributes of investors affect prices in this paper since prices are determined by the expected market-clearing condition and hence reflect the supply and demand of shares by investors. Uncertainty about preferences and endowments creates a problem, however, with respect to assessing demand and supply in the market. If no one in the economy actually knows the distribution of classes in the investor population, the prices set by market makers cannot reflect that information.

\textsuperscript{13}The results with alternative formulations of market makers are available from the author upon request.

\textsuperscript{14}Gallmeyer, Hollifield, and Seppi (2005) use a different approach, a batch-trading market where a class of short-horizon investors provides liquidity, but some of the insights on the effects of uncertainty about preferences are similar in the two models. Hence, I do not believe that the specific formulation of the expected market-clearing condition is crucial for the insights that come out of the results in section 3.

\textsuperscript{15}This is usually motivated by saying that prices are set by competitive, risk-neutral market makers who care only about their wealth at the end of the economy, when the liquidating dividend of the asset is realized.
While the market makers' information set does not include \( q \), I assume that they have a prior on \( q \) at time zero denoted by \( f^0(q) \). The prior distribution can be rather general but its support should be in \([0, 1]\). The prior can be interpreted as the experience market makers develop by regularly observing the investor clientele who trade the risky asset or it can simply be the public information set. Each time an investor arrives and submits an order, market makers (and all other market observers) use Bayes’ rule to update their beliefs about the distribution of investors’ preferences and endowments.

At the beginning of each trading period, market makers are required to post a binding quote consisting of prices and associated depths. Then, an investor arrives, maximizes expected utility taking the quoted prices as given, and submits his order (which executes at the quoted price).\(^{16}\) Trading is anonymous in the sense that the only information market makers have about an arriving investor is the order the investor submits.\(^{17}\) Since orders convey information but the quote needs to be posted before investors arrive, the pricing strategy of the market makers should reflect their rational expectations about the incoming orders. Market makers post “regret-free” prices in that the information inferred from different order sizes is used to calculate the prices at which the orders will be executed.

This rational expectations requirement is similar to the one used in traditional sequential trade models. Glosten and Milgrom (1985) discuss price setting in their model as follows: “In this model, prices exhibit a semi-strong form of efficiency: Indeed, they may reflect slightly more information than was available to the specialist at the time he set the bid and ask prices. The explanation for this seemingly strange conclusion lies in the observation that the specialist does not set a single price. The ask price, for example, specifies what

\(^{16}\)The timing of events, where the binding quote is posted prior to the arrival of the investor, is similar to that in traditional sequential trade models and is very important for the results in the paper. Still, this is probably the most natural sequence of events that represents the interaction in actual markets. For example, dealer markets feature binding quotes at which customers can execute orders, where the quotes are set prior to the orders’ arrival. Similarly, limit orders in electronic limit order books must be posted to the book before an arriving order can execute.

\(^{17}\)The requirement that information can be extracted only from the order size seems quite realistic. Modifying the problem by giving market makers additional information about the identity of arriving investors does not materially affect the analysis.
the price will be if the next customer is a buyer. Consequently, the ask price can be (and
at equilibrium is) set using both the current information and the information that will be
inferred if the next customer turns out to be a buyer” (p. 73). More specifically, the price
that market makers set in traditional sequential trade models to execute a quantity of shares
(positive for buys or negative for sells), \( Q \), is just the expectation of the asset’s liquidation
payoff conditional on the information set that includes the arrival of the order, \( P_Q = \frac{E[u|Q]}{R} \).
Since a buy order (\( B \)) contains different information than a sell order (\( S \)), prices that are
conditional on these orders will be different, \( P_B = \frac{E[u|B]}{R} \neq \frac{E[u|S]}{R} = P_S \). The market makers’
quote therefore consists of two prices, each conditional on a different information set.

A similar situation exists in this paper, where each price in the quote is the solution to
the expected market-clearing condition using a different information set. Formally, I define
the quote posted by the market makers at the beginning of each period as follows:

**Definition (Quote):** The market makers’ *Quote* is a pair of prices and depths \( \{P(\Phi_i), D_i\}_{i \in \{1, 2\}} \)
such that (i) \( D_i = X_i \), and (ii) \( P(\Phi_i) \) is the solution to the expected market-clearing condition
\[
E[qX_1 + (1-q)X_2 | \Phi_i] = 0
\]  
(6)
where \( X_1 \) (\( X_2 \)) is the order size of class 1 (class 2) investors, and the conditioning information
set \( \Phi_i \) includes all information up to the trading period and the next arrival of an order for
\( X_i \) shares.

The quote specifies a price at which a buy order will be executed and a (potentially
different) price at which a sell order will be executed. Since there are two classes of investors
in the economy, it is natural to restrict attention to the order sizes that the two classes of
investors will find optimal.\(^{18}\) An arriving investor does not submit a demand schedule but
rather uses the quoted prices to maximize his expected utility and decide whether he wants to
buy or to sell and what size order to submit. Let \( \Phi_1 \) (\( \Phi_2 \)) be the information set that includes
the information in the next arrival of an order of a class 1 (class 2) investor for \( X_1 \) (\( X_2 \))

\(^{18}\)As Easley and O’Hara (1987) note, though there is a large number of potential price-quantity pairs,
market makers need only determine prices for quantities that potential traders desire to trade.
shares in addition to all public information and past order flow. Since market makers post
the binding quote at the beginning of a trading period before an investor arrives, they would
like to execute an arriving order for $X_1$ (or $X_2$) shares at a price that reflects the information
that can be inferred from such an order. In other words, calculating an execution price for
an order of size $X_1$ conditioning on $\Phi_1$ (or an order of size $X_2$ conditioning on $\Phi_2$) would
result in a “regret-free” (or rational expectations) transaction price.

2.4 Equilibrium

The definition of the quote in the previous section already contains two important require-
ments for an equilibrium: (i) that prices are computed from the expected market-clearing
condition, and (ii) that prices reflect rational expectations (i.e., the price at which an arriving
order executes is computed using an information set that includes the arrival of that order).

Using this definition of the quote, we can define a separating equilibrium in the market.

**Definition** (Separating Equilibrium): A *Separating Equilibrium* in the market consists of
(i) a Quote $\{P(\Phi_i), D_i\}_{i \in \{1,2\}}$, and (ii) an order size for each investor class $\{X_i\}_{i \in \{1,2\}}$, such
that:

1. At the beginning of each trading period, market makers commit to the Quote assuming
   that they can identify the class affiliation of an arriving investor from his order.

2. An optimizing class $i$ investor who arrives at the market chooses the pair $(P(\Phi_i), D_i)$
   by submitting an order for $X_i$ shares, which is his optimal order size given $P(\Phi_i)$, such
   that $X_i = D_i$, and $X_1 \neq X_2$.

To simplify the exposition, assume that the parameters of the economy are such that
$\Delta \alpha \bar{X} = \alpha_1 \bar{X}_1 - \alpha_2 \bar{X}_2 > 0$. This will result in an equilibrium where class 1 investors are
sellers and class 2 investors are buyers (since class 1 investors are more risk averse and/or
have a larger endowment of the risky asset than class 2 investors). The case of $\Delta \alpha \bar{X} < 0$ is
symmetric, and in equilibrium class 1 are buyers and class 2 are sellers.\textsuperscript{19} This assumption is without loss of generality since we could always rename the two classes of investors.\textsuperscript{20} The following proposition establishes the existence of an equilibrium:

**Proposition 1** There exists a Separating Equilibrium where (i) the market makers’ Quote is comprised of the following prices and depths:

\[
\text{Ask Price: } P(\Phi_2) = \frac{\theta}{R} - \frac{\alpha_1\alpha_2\sigma^2 \bar{X}(\Phi_2)}{R\alpha(\Phi_2)}, \quad \text{Ask Depth: } D_2 = X_2
\]  

\[
\text{Bid Price: } P(\Phi_1) = \frac{\theta}{R} - \frac{\alpha_1\alpha_2\sigma^2 \bar{X}(\Phi_1)}{R\alpha(\Phi_1)}, \quad \text{Bid Depth: } D_1 = X_1
\]  

(ii) An arriving class 1 investor submits the order:

\[
X_1 = - \frac{(1 - q_1)\Delta \alpha \bar{X}}{\alpha(\Phi_1)} < 0
\]  

and an arriving class 2 investor submits the order:

\[
X_2 = \frac{q_2\Delta \alpha \bar{X}}{\alpha(\Phi_2)} > 0
\]  

where \( q_1 = E[q | \Phi_1], q_2 = E[q | \Phi_2], \bar{X}(\Phi_1) = q_1 \bar{X}_1 + (1 - q_1) \bar{X}_2, \bar{X}(\Phi_2) = q_2 \bar{X}_1 + (1 - q_2) \bar{X}_2, \alpha(\Phi_1) = q_1 \alpha_2 + (1 - q_1) \alpha_1, \alpha(\Phi_2) = q_2 \alpha_2 + (1 - q_2) \alpha_1, \) and \( \Delta \alpha \bar{X} = \alpha_1 \bar{X}_1 - \alpha_2 \bar{X}_2. \)

All proofs are provided in the Appendix. The proof of this proposition follows the standard structure. I start by assuming that market makers can identify the class affiliation (preferences and endowments) of investors who arrive at the market from their order sizes, and are thus able to calculate the quote that specifies the prices at which investors can transact. Then, the investors’ participation and incentive compatibility conditions are analyzed to show that an investor will in fact self-select to trade using the price and depth that the

\textsuperscript{19}When \( \Delta \alpha \bar{X} = 0, \) investors choose not to trade for risk sharing. This is also the case in which there will be no risk sharing among investors in an equivalent market cleared by a Walrasian auctioneer. This case is less interesting and will not be pursued further in the paper.

\textsuperscript{20}The proofs of all results in the paper are done for the general case of \( \Delta \alpha \bar{X} \neq 0. \)
market makers had set for investors who belong to his class, completing the requirements of
the separating equilibrium.\footnote{The proof shows that an investor would choose the price-depth pair that was set for his class over the price-depth pair that was set for the other class of investors. Investors cannot choose order sizes that are not equal to the quoted depths. In the separating equilibrium, however, the depth set by the market makers is equal to the optimal order size (given the price) of the investor class that selects that price-depth pair.}

Proposition 1 shows that, similarly to the situation in traditional sequential trade models, market makers post two prices: one at which investors can buy the risky asset (an “ask”) and another at which investors can sell the risky asset (a “bid”). These prices depend on the information set of the market makers and are “regret-free” in that they incorporate the information that the market could infer from the executed order. Prices have the usual structure from asset pricing models: a risk-neutral component, the mean of the asset’s payoffs divided by the risk-free rate, and a risk premium that depends on the relative population weight $q$.\footnote{The price that comes out of the expected market-clearing condition also preserve an important notion of optimality in that if market makers have full information (they know $q$), the price will be equal to the competitive equilibrium price in the economy.}

Note, however, that since nobody knows $q$, the aggregate demand and the harmonic mean of the risk aversion coefficients are calculated using the conditional expectation of $q$ rather than the parameter itself. This is a key element of the model as uncertainty about preferences and endowments inevitably results in uncertainty about the risk premium. Market makers are able to post prices in advance of investor arrival because the expected market-clearing condition translates that uncertainty into an expression for the risk premium that uses the expectation of the relative population weight. Since the order flow conveys information about investors, market makers learn about the risk premium by updating their expectation of the population parameter $q$ as investors arrive at the market and trade.\footnote{One subtlety in the modeling framework used here concerns the interaction between uncertainty about $q$ and the correspondence between the expected market-clearing condition and the optimization problem of competitive market makers subject to a no-drift constraint. The inventory positions of the competitive market makers would not drift when the bid and ask prices converge to a single price. However, as long as market makers learn about the investor population, their inventories can drift (ex-post) in one direction or the other depending on the relation between their prior expected value of $q$ and the true population parameter. This situation is similar to that of market makers in traditional sequential trade models. As Glosten and Milgrom (1985) note, the inventory of the specialist in their model would not drift only in the limit. As long as some investors have private information and they trade on it, inventory may drift. A positive or a negative inventory drift is due to the assumption of risk neutrality on the part of market makers and the assumption of no inventory carrying costs. Inventory models such as Ho and Stoll (1981) and O’Hara and}
3 Results

This section presents some of the model’s implications for the relation between uncertainty about the preferences and endowments of investors and liquidity of the risky asset. The results demonstrate how shocks to uncertainty about the risk premium translate into systematic liquidity shocks, and how learning about the risk premium is associated with time-varying liquidity. The model is also used to demonstrate a link between liquidity and the welfare of investors.

Before turning to the results, it will be useful to establish how beliefs about the distribution of preferences and endowments in the population are updated as investors arrive at the market and trade.

Proposition 2

\[ q_{1,t} \equiv E[q | \Phi_{1,t}] = q_t + \frac{V_t[q]}{q_t} > q_t \]  

\[ q_{2,t} \equiv E[q | \Phi_{2,t}] = q_t - \frac{V_t[q]}{1 - q_t} < q_t \]

where \( \Phi_t \) is the information set of market makers at the beginning of period \( t \), \( \Phi_{1,t} \) (\( \Phi_{2,t} \)) is the information set of market makers conditional on the arrival in period \( t \) of a class 1 (class 2) investor, \( q_t = E[q | \Phi_t] \), and \( V_t[q] = E[q^2 | \Phi_t] - q_t^2 \).

Oldfield (1986) show how risk averse market makers adjust prices to avoid inventory accumulation. When I replace the expected market-clearing condition with an objective function for market makers that requires an expected end-of-trading inventory of zero, I can show how learning about the risk premium occurs in a more complicated model where market makers adjust prices to eliminate the ex-post inventory drift. The implications of this model for the relation between liquidity and uncertainty about the risk premium are similar to the implications of the simpler model presented here.
a random draw from the true distribution. Bayes’ Rule then dictates that the market makers update their beliefs about the population of investors by giving more weight to class $i$. The inference is not about the class of the arriving investor, which is fully known to the market makers in equilibrium when the investor submits his order. Rather, learning occurs about the true distribution of investors in the population.

While market makers may have a rather general prior on the distribution of the population parameter $q$, (11) and (12) show that only the first two moments of that distribution matter for pricing the risky asset. The extent of uncertainty about investors in the market can therefore be represented by the variance of beliefs about the population parameter $q$. The smaller is $V_t[q]$, the tighter is the distribution around the mean, which implies less uncertainty about the distribution of preferences and endowments in the market.

The next result, which follows directly from Proposition 1, establishes that the economy is characterized by a positive spread between the bid and ask prices, and hence that the risky asset is not perfectly liquid.

**Proposition 3** Whenever there is uncertainty about the distribution of preferences and endowments of investors ($V_t[q] > 0$), there exists a strictly positive spread between the bid and ask prices:

$$S_t = \frac{\alpha_1 \sigma^2 | \Delta \alpha X | V_t[q]}{R \alpha(\Phi_{1,t}) \alpha(\Phi_{2,t}) q_t (1 - q_t)}$$

What is the intuition behind the spread? Uncertainty about the risk premium (or the preferences and endowments of investors) means that prices reflect the beliefs about $q$. Market makers learn about the population of investors from the arriving orders, and their updating of beliefs causes prices to move with each order. When market makers observe in period $t$ a sell order for $X_{1,t}$ shares, they learn that a class 1 investor arrived at the market. Since $\alpha_1 X_1 > \alpha_2 X_2$, a class 1 investor is more risk averse and/or has a larger endowment of the risky asset than a class 2 investor. Market makers then update their beliefs about the distribution of investors so that their expectation of $q$ increases. Hence, prices must decrease to reflect the information that there are more investors in the market with large endowments...
and/or higher risk aversion. The arriving sell order suffers from a price impact as prices adjust downward. Similarly, an arriving buy order will identify the investor as belonging to class 2 and will cause an increase in prices to reflect the belief that there are more investors who are less risk averse or who have smaller endowments. As in traditional sequential trade models, the spread is the sum of the price impacts of a buy order and a sell order.24

The next two results examine how changes to uncertainty about the risk premium affect liquidity.

**Proposition 4** Shocks that increase (decrease) the variance of the market makers’ beliefs about \( q \) cause liquidity to worsen (improve).

Higher uncertainty about the risk premium in the market is costly to investors because it causes the ask to be higher and the bid to be lower.25 This result is driven by the learning process of the market about the distribution of investors in the economy. Proposition 2 shows that a larger \( V_t[q] \) implies a greater “distance” between the prior and posterior expectations of \( q \). Greater revisions in the estimate of \( q \) cause larger price impacts and therefore the ask is higher and the bid is lower. This result provides a key link between the asset’s risk premium and liquidity. The risky asset in the model can be viewed as the market portfolio in the economy because it seems intuitive that uncertainty about investors’ preferences and endowments affects all risky assets. Therefore, shocks that cause more uncertainty about the risk premium would result in deterioration of liquidity for all risky assets.

The risk premium, however, is time-varying in this model even between shocks to uncertainty about the distribution of preferences and endowments. The reason for this is that each transaction price reflects current beliefs about the expected value of \( q \) and these beliefs

---

24While the result that prices move on trades may cause us to question the price-taking assumption, there is a sense in which investors do take into account their impact on the price. The price charged of an investor is adjusted for the information market makers learn from his order. When an investor arrives, he calculates his optimal demand using the price that already reflects his order. Hence, this rational expectations feature of the prices set by the market makers creates a situation in which price taking is consistent with a quasi-strategic behavior in which an investor’s demand reflects his impact on the price.

25The proposition considers the effect of changing the variance of \( q \) holding its expected value constant.
are changing as market makers learn from the order flow. As a consequence, illiquidity also varies through time as it reflects the dynamics of the risk premium.

**Proposition 5** $V_t[q]$ decreases on average when orders arrive. If we let $t$ (and therefore $T$ as well) go to infinity, $V_t[q]$ would approach zero, the market makers would learn the true risk premium almost surely, and the risky asset would become perfectly liquid as the bid and ask prices would converge to a single price.

The decrease in uncertainty about the risk premium as orders arrive is a consequence of the Bayesian updating. A similar learning process also takes place in traditional sequential trade models that utilize information asymmetry about future cash flows to generate the spread. In these models, it is learning about the informed investors’ signal that drives down the spread as more orders arrive. The price impact of trades should disappear without new private information about future cash flows (in traditional sequential trade models) or renewed uncertainty about the risk premium (here). If we continue to observe the effects of risk premium uncertainty in the market, it must be the case that the economic environment is constantly changing.

For example, one way to think about the manner in which markets operate is that every day there is a different subset of investors who trade: some suddenly need money, others have found the time to go over their finances, and so on. This subset of investors determines the price path on that day as market makers try to learn the distribution of investor classes. Of course, prices that go up or down by a large amount will attract the attention of other subsets of investors who did not plan on arriving at the market on that day. But within some bounds on the price movement, prices are determined at every point in time by this search for the preferences and endowments of a subset of the investor population. Uncertainty about the risk premium also arises when the market learns about changes in the tastes and income of the population of investors as a whole. This hierarchy of inferences, both on a subset and on the population, creates a situation in which market makers can never stop
learning from the order flow and the illiquidity associated with uncertainty about the risk premium is always present.\textsuperscript{26}

A major difference between the traditional sequential trade models that generate illiquidity by hypothesizing information asymmetry about future cash flows and the approach pursued in this paper lies in the ability to examine welfare implications. The traditional models are unable to provide a welfare analysis since informed investors profit at the expense of uninformed investors (and the uninformed investors in most models do not have an explicit utility function that can be evaluated). Here, the model spells out a clear welfare result:

**Proposition 6** The welfare of all investors in the economy is decreasing in \( V_t[q] \).

Since prices and hence optimal demands are functions of \( V_t[q] \), uncertainty about the risk premium enters the indirect utility functions of investors and is shown to hurt all investors in the market. It therefore follows that design and regulation of markets aimed at reducing uncertainty about the risk premium will make all investors in the market better off. This result also suggests that market makers who are experts in assessing the nature of the investor population (i.e., have tighter priors on \( q \)) can offer better prices to investors (and also increase volume, as it is straightforward to show that volume is lower when uncertainty about the risk premium is higher). Hence, the expertise of market makers constitutes a positive externality that benefits all investors in the market.

This result also illustrates a link between liquidity and welfare: Investors are better off when the risky asset is more liquid. The link is created because when uncertainty about the risk premium is higher, the ask is higher and the bid is lower. Faced with less favorable prices,\textsuperscript{26}Note that uncertainty about preferences and endowments can exist in markets even when all investors are present. It is intuitively clear, however, that such uncertainty would be greater in sequential markets where not all investors are in the market at any given time. While a limit order book is not explicitly modeled in this paper, it also seems reasonable to conjecture that arrival of limit orders can provide information about the preferences and endowments of investors. However, as the discussion in the text emphasizes, uncertainty about the risk premium is constantly created in the market and therefore a limit order book could not eliminate it.

\textsuperscript{19}
investors want to buy or sell smaller quantities and therefore the welfare of all investors is lower (due to less risk sharing).

The causal relationship demonstrated by the model is from uncertainty about the risk premium to illiquidity. The implications of the model are consistent with the empirical findings of systematic liquidity and the observed relation between time-varying expected returns and time-varying liquidity. A different causal relationship that has been suggested in the literature goes from liquidity to future expected returns. For example, Amihud (2002) and Jones (2002) discuss the idea that lower liquidity should be associated with a lower price and hence higher expected returns to compensate for the higher costs of trading. This is a dynamic equivalent of the cross-sectional ideas in Amihud and Mendelson (1986).

The primary relation demonstrated by the results in this section is between liquidity and uncertainty about the risk premium rather than between liquidity and the risk premium itself (or the level of prices). There is however a secondary effect that relates liquidity to the price level under certain conditions. For example, if the two investor classes have the same endowments but different risk aversions, then it can be shown that an increase in the variance of $q$ lowers the expected transaction price of the next trader (or the quote midpoint) at the same time as it causes worsened liquidity. In other words, greater uncertainty about investors’ preferences and endowments lowers the price of the risky asset and therefore increases its expected return. The reason for this result is that the price is a concave function of the population parameter $q$ (because the risk premium is proportional to the harmonic mean of investors’ risk aversion coefficients). Therefore, the magnitude of the change in the ask will be smaller than the magnitude of the change in the bid. This will hold more generally with heterogeneous endowments whenever the more (less) risk-averse investor sells (buys).\footnote{There can be a case where the less risk-averse investor has a very large endowment and in the resulting equilibrium he will be the seller. In this case, the property that increased uncertainty about investors’ preferences and endowments lowers the price level will not hold. The exact condition for this property to hold is: $\Delta\alpha X(\alpha_2 - \alpha_1) < 0$.}
4 Discussion and Concluding Remarks

This paper advances the idea that there is a relation between the risk premium and the liquidity of risky assets created by the process of learning about the preferences and endowments of investors. At the very basic level both the risk premium and liquidity are related because they have to do with determining an asset’s price.\textsuperscript{28} The illiquidity of an asset is simply an indication that trades convey information to the market that causes a reevaluation of its price. The information featured here is about the preferences and endowments of investors, which is crucial in determining the risk premium. Hence, the price impact of trades is a manifestation of learning about the risk premium and this relationship evolves over time as uncertainty is gradually resolved through trading and new uncertainty is introduced by shocks to the economic environment. The model therefore motivates empirical findings of systematic liquidity, a link between the risk premium and liquidity, and time variation in both.

One contribution of the model is to demonstrate another underlying reason for the price impact of trades—uncertainty about the preferences and endowments of investors. The driving force behind this friction is therefore very different from the one advanced by traditional sequential trade models, where the price impact is due to the trading of a subset of investors with private information about the future cash flows of the asset. Much empirical work in the market microstructure literature has been devoted to identifying and investigating informational effects in prices. Econometric spread decomposition procedures were developed to measure the “adverse selection” component of the spread that is attributed to information asymmetry about the firm.\textsuperscript{29} In general, these methodologies identify the permanent component of the price impact of trades and attribute it to information, while the temporary component is attributed to order-processing costs and inventory costs. Since uncertainty

\textsuperscript{28}A similar idea is discussed in Gallmeyer, Hollifield, and Seppi (2005).

about the risk premium generates a price impact for trades similar to that described by the 
traditional sequential trade models but for a completely different reason, a question arises 
as to what these methodologies exactly capture.

Using data generated by simulating the model, I find that the adverse selection com-
ponent of the spread produced by such econometric procedures picks up the effects of un-
certainty about the risk premium. This suggests caution in interpreting the output of 
these econometric techniques, as they seem to bundle information about the risk premium 
with private information about the firm. Attributing the estimated informational effects in 
prices to either uncertainty about the risk premium or information asymmetry about future 
cash flows therefore requires looking at the economic context. When the assets analyzed are 
stocks, a plausible interpretation could be that the systematic component in liquidity is due 
to uncertainty about the risk premium while the firm-specific component of the permanent 
price impact is due to private information about the future cash flows of the firm. How-
ever, the possibility that investors have private information about common factors (as in 
Subrahmanyam, 1991) means that both explanations for the price impact could be driving 
systematic liquidity. An interesting direction for future empirical work would be to come up 
with empirical proxies that separate these two effects.

The liquidity frictions generated by uncertainty about investors’ preferences and endow-
ments also provide an alternative way of interpreting empirical evidence of permanent price 
impact of trades in assets that do not fit easily into the asymmetric information paradigm. 
One such example is the foreign exchange market. Since “inside information” in the usual 
sense is less relevant in the foreign exchange market, much of the day-to-day pricing reflects 
the demand of different users. A foreign exchange dealer receives orders, infers the demand, 
and sets quotes much like the market makers in this paper (see Lyons, 1995). Hence, we can

---

30The simulations are available from the author upon request.
31For example, Edelen and Warner (2001) show a relation between aggregate flows into U.S. equity funds 
and market returns. These flows could serve as a proxy for changes in endowments and preferences of 
investors, and therefore the model presented here motivates looking for a link between these flows and 
market liquidity.
expect that uncertainty about investors’ preferences and endowments would play an important role in this market.\(^{32}\) Similarly, evidence of informational effects in prices of closed-end mutual funds whose net asset values are known (Neal and Wheatley, 1998) or Treasury securities (Green, 2004) may predominantly reflect uncertainty about the risk premium in these markets.

On the theoretical side, the approach pursued in this paper is subject to several limitations. The first limitation, which is also shared by the traditional sequential trade models, is the probabilistic selection of investors that precludes dynamic trading strategies. The second limitation is that market makers are not modeled as agents and thus I abstract from possible intertemporal strategies that utility-maximizing market makers can adopt. In some sense, ignoring the possibility of strategic behavior on the part of market makers is also a characteristic of most traditional sequential trade models, where the requirement (formalized in Glosten and Milgrom, 1985) that the price of each trade reflects a zero expected profit condition results in a convenient pricing rule but does not lend itself easily to dynamic considerations. Leach and Madhavan (1993), for example, show how more complex patterns of price experimentation can arise when market makers solve dynamic problems.

It is also interesting to note that transaction prices in the model need not be a martingale. While martingale prices are an implication of traditional sequential trade models (because prices are simply conditional expected values of an exogenous random variable), other market microstructure models demonstrate how market frictions can result in non-martingale prices. For example, Amihud and Mendelson (1980) show how inventory control considerations give rise to non-martingale prices. Leach and Madhavan (1993) show that prices do not follow a martingale when market makers implement price experimentation to try to profit from discovering the private information of traders.\(^{33}\)

While the simplified assumptions of the model are helpful in the investigation of uncer-

\(^{32}\)For an argument in support of the traditional form of adverse selection in the foreign exchange market see Naranjo and Nimalendran (2000).

\(^{33}\)Empirically, Hasbrouck and Ho (1987) show that neither quote midpoints nor transaction prices (controlling for the bid-ask bounce) follow a martingale.
tainty about the risk premium and its effects on liquidity, I believe that the intuition behind the relation between uncertainty about investors’ preferences and endowments and the price impact of trades is more general than the model itself and would hold in more complex settings. Extending the model to accommodate \( n > 2 \) classes of investors who differ with respect to their preferences and endowments is conceptually straightforward. While the proof of the equilibrium would be much more complex due to the greater number of incentive compatibility conditions, the result would likely be a schedule of prices and associated depths at which investors can trade. A large buy order, for example, would imply that more investors are characterized with much lower risk aversion or much smaller endowments and therefore would result in a larger price impact than that created by a small buy order.

The use of a single risky asset in the model lends itself easily to discussions in terms of systematic liquidity because the risky asset can be taken to represent the market portfolio and it seems plausible that uncertainty about the risk premium would affect all assets in the market. The investor recognition hypothesis (e.g., Merton, 1987; Shapiro, 2002) suggests, however, that information costs could create a situation where different investors focus on different subsets of the assets in the economy. If clienteles with different distributions of preferences and endowments are formed for different groups of assets, the cross-section of risky assets could feature several systematic liquidity components. While it remains to be seen whether there is empirical support for a more complex structure for systematic liquidity, the implications of the model in this paper can be used to motivate new directions in the endeavor to characterize the relationship between investor trading, liquidity, and returns.
Appendix

Proof of Proposition 1:
Assume that market makers can identify the class affiliation of an investor from his order size, and that class 1 (class 2) investors submit orders for \(X_1\) \((X_2)\) shares. First I will establish the expected clearing price conditional on \(\Phi_1\) (the information set that includes the arrival of an order for \(X_1\) shares). \(P(\Phi_1)\) can be found by plugging the optimal orders into the expected market-clearing condition in (6) as follows:

\[
q_1 \left( \frac{\theta - RP(\Phi_1)}{\alpha_1 \sigma^2} - \bar{X}_1 \right) + (1 - q_1) \left( \frac{\theta - RP(\Phi_1)}{\alpha_2 \sigma^2} - \bar{X}_2 \right) = 0
\]  
(14)

where \(q_1 = E[q | \Phi_1]\). Solving for \(P(\Phi_1)\) we get,

\[
P(\Phi_1) = \frac{\theta}{R} - \frac{\bar{X}(\Phi_1) \alpha_1 \alpha_2 \sigma^2}{R \alpha(\Phi_1)}
\]  
(15)

where \(\bar{X}(\Phi_1) = q_1 \bar{X}_1 + (1 - q_1) \bar{X}_2\) and \(\alpha(\Phi_1) = q_1 \alpha_2 + (1 - q_1) \alpha_1\). The expected clearing price conditional on \(\Phi_2\) can be found in a similar fashion. Let \(q_2 = E[q | \Phi_2]\). Then,

\[
P(\Phi_2) = \frac{\theta}{R} - \frac{\bar{X}(\Phi_2) \alpha_1 \alpha_2 \sigma^2}{R \alpha(\Phi_2)}
\]  
(16)

where \(\bar{X}(\Phi_2) = q_2 \bar{X}_1 + (1 - q_2) \bar{X}_2\) and \(\alpha(\Phi_2) = q_2 \alpha_2 + (1 - q_2) \alpha_1\). Since in equilibrium the order of a class 1 investor will be executed at the price \(P(\Phi_1)\), his optimal order size is:

\[
X_1 = \frac{\theta - RP(\Phi_1)}{\alpha_1 \sigma^2} - \bar{X}_1 = -\frac{(1 - q_1) \Delta \alpha \bar{X}}{\alpha(\Phi_1)} < 0
\]  
(17)

where \(\Delta \alpha \bar{X} = \alpha_1 \bar{X}_1 - \alpha_2 \bar{X}_2 > 0\) by assumption. This assumption is without loss of generality since we could always rename the two classes of investors (in which case class 1 investors will be the buyers and class 2 investors the sellers). Similarly, in equilibrium the order of a class 2 investor will be executed at the price \(P(\Phi_2)\), and so his optimal order size is:

\[
X_2 = \frac{\theta - RP(\Phi_2)}{\alpha_2 \sigma^2} - \bar{X}_2 = \frac{q_2 \Delta \alpha \bar{X}}{\alpha(\Phi_2)} > 0
\]  
(18)

Market makers specify \(P(\Phi_1)\) as the bid price with a bid depth that is equal to the optimal order size of class 1 investors and \(P(\Phi_2)\) as the ask price with an ask depth that is equal to the optimal order size of class 2 investors. Note that \(X_1 \neq X_2\) (one is a buy order and the other a sell order), and hence market makers can differentiate between them. This completes the first requirement of the equilibrium.

Since investors face two prices with associated depths and are free to choose between them, we need to check that class 1 investors will choose \((P(\Phi_1), X_1)\) while class 2 investors...
will choose \((P(\Phi_2), X_2)\). For that consider the following participation and incentive compatibility conditions:

\[
P.1. \quad U_1(P(\Phi_1), X_1) \geq U_1(P, 0) \\
IC.1. \quad U_1(P(\Phi_1), X_1) \geq U_1(P(\Phi_2), X_2) \\
P.2. \quad U_2(P(\Phi_2), X_2) \geq U_2(P, 0) \\
IC.2. \quad U_2(P(\Phi_2), X_2) \geq U_2(P(\Phi_1), X_1)
\]

where \(P\) in the participation conditions denotes any arbitrary price. The participation condition P.1 can be written as follows:

\[
R\tilde{Z}_1 - RP(\Phi_1)X_1 + (\tilde{X}_1 + X_1)\theta - \frac{\alpha_1}{2}(\tilde{X}_1 + X_1)^2\sigma^2 \geq 0 
\]

which simplifies to:

\[
R\tilde{Z}_1 - RP0 + (\tilde{X}_1 + 0)\theta - \frac{\alpha_1}{2}(\tilde{X}_1 + 0)^2\sigma^2 
\]

Denote the second term by \(B\). The first term in the left-hand side of the above expression is \(A\). Using (16) and (18):

\[
A = X_1 \left[ -R \left( \frac{\theta}{R} - \frac{\alpha_1\alpha_2\bar{X}(\Phi_1)\sigma^2}{R\alpha(\Phi_1)} \right) + \theta - \frac{\alpha_1\sigma^2 X_1^2}{2} - \frac{\alpha_1\sigma^2 X_1^2}{2} \right] 
\]

which simplifies to:

\[
-A(\bar{\Phi}_2) = X_1 \left[ (1 - q_1)(-\Delta \alpha \bar{X}) \right] = \frac{(1 - q_1)\alpha_1\sigma^2 (\Delta \alpha \bar{X})^2}{2\alpha(\Phi_1)^2} > 0 
\]

Hence, P.1. always holds. The incentive compatibility condition I.C.1. can be written as follows:

\[
R\tilde{Z}_1 - RP(\Phi_1)X_1 + (\tilde{X}_1 + X_1)\theta - \frac{\alpha_1}{2}(\tilde{X}_1 + X_1)^2\sigma^2 \geq 0 
\]

which simplifies to:

\[
\left[ -RP(\Phi_1)X_1 + X_1(\theta - \alpha_1\sigma^2 \tilde{X}_1) - \frac{\alpha_1\sigma^2 X_1^2}{2} \right] + \left[ RP(\Phi_2)X_2 - X_2(\theta - \alpha_1\sigma^2 \tilde{X}_1) + \frac{\alpha_1\sigma^2 X_2^2}{2} \right] \geq 0 
\]

The first term in the left-hand side of the above expression is \(A\), which was shown to be positive. Denote the second term by \(B\). It can be rewritten using (16) and (18) as follows:

\[
B = X_2 \left[ R \left( \frac{\theta}{R} - \frac{\alpha_1\alpha_2\bar{X}(\Phi_2)\sigma^2}{R\alpha(\Phi_2)} \right) - \theta + \frac{\alpha_1\sigma^2 \tilde{X}_1}{2} \right] 
\]

which simplifies to:

\[
\frac{X_2\alpha_1\sigma^2}{2\alpha(\Phi_2)^2} \left[ (2 - q_2)\Delta \alpha \bar{X} \right] = \frac{q_2(2 - q_2)\alpha_1\sigma^2 (\Delta \alpha \bar{X})^2}{2\alpha(\Phi_2)^2} > 0 
\]
Hence, I.C.1. always holds. The participation condition P.2. can be written as follows:

\[
R\bar{Z}_2 - RP(\Phi_2)X_2,t + (\bar{X}_2 + X_2)\theta - \frac{\alpha_2}{2}(\bar{X}_2 + X_2)^2 \sigma^2 \geq 0
\]

(32)

\[
R\bar{Z}_2 - RP0 + (\bar{X}_2 + 0)\theta - \frac{\alpha_2}{2}(\bar{X}_2 + 0)^2 \sigma^2
\]

(33)

which simplifies to:

\[-RP(\Phi_2)X_2 + X_2(\theta - \alpha_2\sigma^2\bar{X}_2) - \frac{\alpha_2\sigma^2X_2^2}{2} \geq 0
\]

(34)

Denote the left-hand side of the above expression by \(C\). Using (16) and (18):

\[
C = \frac{X_2\alpha_2\sigma^2}{2} \left[ \frac{2\alpha_1\bar{X}(\Phi_2)}{\alpha(\Phi_2)} - 2\bar{X}_2 - X_2 \right]
\]

(35)

\[
= \frac{X_2\alpha_2\sigma^2}{2\alpha(\Phi_2)} [q_2\Delta\alpha\bar{X}] = \frac{q_2^2\alpha_2\sigma^2(\Delta\alpha\bar{X})^2}{2\alpha(\Phi_2)^2} > 0
\]

(36)

Hence, P.2. always holds. The incentive compatibility condition I.C.2. can be written as follows:

\[
R\bar{Z}_2 - RP(\Phi_2)X_2 + (\bar{X}_2 + X_2)\theta - \frac{\alpha_2}{2}(\bar{X}_2 + X_2)^2 \sigma^2 \geq 0
\]

(37)

\[
R\bar{Z}_2 - RP(\Phi_1)X_1 + (\bar{X}_2 + X_1)\theta - \frac{\alpha_2}{2}(\bar{X}_2 + X_1)^2 \sigma^2
\]

(38)

which simplifies to:

\[
\left[ -RP(\Phi_2)X_2 + X_2(\theta - \alpha_2\sigma^2\bar{X}_2) - \frac{\alpha_2\sigma^2X_2^2}{2} \right]
\]

(39)

\[
+ \left[ RP(\Phi_1)X_1 + X_1(\theta - \alpha_2\sigma^2\bar{X}_2) - \frac{\alpha_2\sigma^2X_1^2}{2} \right] \geq 0
\]

(40)

The first term in the left-hand side of the above expression is \(C\), which was shown to be positive. Denote the second term by \(D\). It can be rewritten using (15) and (17) as follows:

\[
D = X_1 \left[ R \left( \frac{\theta}{R} - \frac{\alpha_1\alpha_2\bar{X}(\Phi_1)^2}{R\alpha(\Phi_1)} \right) - \theta + \alpha_2\sigma^2\bar{X}_2 + \frac{\alpha_2\sigma^2X_1}{2} \right]
\]

(41)

\[
= \frac{X_1\alpha_2\sigma^2}{2\alpha(\Phi_1)} \left[ -(1 + q_1)\Delta\alpha\bar{X} \right] = \frac{(1 - q_1)(1 + q_1)\alpha_2\sigma^2(\Delta\alpha\bar{X})^2}{2\alpha(\Phi_1)^2} > 0
\]

(42)

Hence, I.C.2. always holds. These four conditions show that investors self-select to the pairs of prices and depths that reveal their class affiliations to the market makers, and hence the second requirement of the equilibrium is satisfied.

Q.E.D.
Proof of Proposition 2:
Let \( f^{\Phi_t}(q) \) be the prior distribution of \( q \) given all public information including the order flow up to time \( t \), and \( \Phi_{t,t} = \{ \Phi_t, X_{t,t} \} \) be the information set that also includes an incoming order of a class \( i \) investor in period \( t \). By Bayes' Law,

\[
f^{\Phi_{1,t}}(q) = \frac{q f^{\Phi_t}(q)}{\int_0^1 q f^{\Phi_t}(q) dq} \tag{43}
\]

\[
E[q \mid \Phi_{1,t}] = \int_0^1 q f^{\Phi_{1,t}}(q) dq = \frac{\int_0^1 q^2 f^{\Phi_t}(q) dq}{\int_0^1 q f^{\Phi_t}(q) dq} = \frac{V_t[q] + q_t^2}{q_t} = q_t + \frac{V_t[q]}{q_t} \tag{44}
\]

where \( E_t[\cdot] = E[\cdot \mid \Phi_t] \), \( q_t = E_t[q] \), and \( V_t[q] = E_t[q^2] - q_t^2 \). Similarly,

\[
f^{\Phi_{2,t}}(q) = \frac{(1-q) f^{\Phi_t}(q)}{\int_0^1 (1-q) f^{\Phi_t}(q) dq} \tag{45}
\]

\[
E[q \mid \Phi_{2,t}] = \int_0^1 q f^{\Phi_{2,t}}(q) dq = \frac{\int_0^1 q(1-q) f^{\Phi_t}(q) dq}{\int_0^1 (1-q) f^{\Phi_t}(q) dq} = \frac{q_t - E_t[q^2]}{1-q_t} = \frac{q_t - V_t[q] - q_t^2}{1-q_t} \tag{46}
\]

Q.E.D

Proof of Proposition 3:
From (17) and (18), class 2 investors are buyers (sellers) and class 1 investors are sellers (buyers) if and only if \( \Delta \alpha \bar{X} = \alpha_1 \bar{X}_1 - \alpha_2 \bar{X}_2 > 0 (< 0) \). Hence, the spread \( S_t \) can be defined as:

\[
S_t = (P(\Phi_{2,t}) - P(\Phi_{1,t})) \text{sign}(\Delta \alpha \bar{X}) \tag{47}
\]

Using (15), (16) and Proposition 2,

\[
S_t = \frac{\alpha_1 \alpha_2 \sigma^2}{R} \left[ \frac{\bar{X}(\Phi_{1,t})}{\alpha(\Phi_{1,t})} - \frac{\bar{X}(\Phi_{2,t})}{\alpha(\Phi_{2,t})} \right] \text{sign}(\Delta \alpha \bar{X})
\]

\[
= \frac{\alpha_1 \alpha_2 \sigma^2}{R \alpha(\Phi_{1,t}) \alpha(\Phi_{2,t})} |\Delta \alpha \bar{X}| \left( \text{sign}(\Delta \alpha \bar{X}) \right)^2 (q_{1,t} - q_{2,t})
\]

\[
= \frac{\alpha_1 \alpha_2 \sigma^2}{R \alpha(\Phi_{1,t}) \alpha(\Phi_{2,t}) q_t} V_t[q] (1-q_t) > 0 \tag{48}
\]

Q.E.D
Proof of Proposition 4:
Using (15) and Proposition 2,
\[
\frac{\partial P(\Phi_{1,t})}{\partial V_t[q]} = -\frac{\alpha_1\alpha_2\sigma^2\Delta \alpha \bar{X}}{R\alpha(\Phi_{1,t})^2 q_t} \tag{49}
\]
Hence, if \(\Delta \alpha \bar{X} < 0\) and class 1 investors are buyers, the ask increases with \(V_t[q]\). If \(\Delta \alpha \bar{X} > 0\) and class 1 investors are sellers, the bid decreases with \(V_t[q]\). Using (16) and Proposition 2,
\[
\frac{\partial P(\Phi_{2,t})}{\partial V_t[q]} = \frac{\alpha_1\alpha_2\sigma^2\Delta \alpha \bar{X}}{R\alpha(\Phi_{2,t})^2 (1 - q_t)} \tag{50}
\]
Hence, if \(\Delta \alpha \bar{X} < 0\) and class 2 investors are sellers, the bid decreases with \(V_t[q]\). If \(\Delta \alpha \bar{X} > 0\) and class 2 investors are buyers, the ask increases with \(V_t[q]\). Using the definition of the spread from (47),
\[
\frac{\partial S_t}{\partial V_t[q]} = \left[\frac{\alpha_1\alpha_2\sigma^2}{R\alpha(\Phi_{2,t})^2 (1 - q_t)} + \frac{\alpha_1\alpha_2\sigma^2}{R\alpha(\Phi_{1,t})^2 q_t}\right] |\Delta \alpha \bar{X}| > 0 \tag{51}
\]
Q.E.D

Proof of Proposition 5:
These are implications of standard Bayesian results. Using the Law of Iterated Expectations (where \(E[\cdot]\) denotes the expectation and \(V(\cdot)\) denotes the variance), \(E_{X_t} [V(q;X_t)] = V(q) - V_{X_t} [E(q;X_t)] \leq V(q)\). Also, since \(q_t\) converges almost surely to \(q\), \(V_t[q]\) goes in the limit to zero. Using Proposition 2 and the definitions of \(\bar{X}(\Phi_{1,t})\), \(\alpha(\Phi_{1,t})\), \(\bar{X}(\Phi_{2,t})\), and \(\alpha(\Phi_{2,t})\):
\[
\lim_{t \to \infty} P(\Phi_{1,t}) = \frac{\theta}{R} - \frac{\lim_{t \to \infty} \bar{X}(\Phi_{1,t})\alpha_1\alpha_2\sigma^2}{\lim_{t \to \infty} R\alpha(\Phi_{1,t})} = \frac{\theta}{R} - \frac{\bar{X}^*\alpha_1\alpha_2\sigma^2}{R\alpha^*} \tag{52}
\]
\[
\lim_{t \to \infty} P(\Phi_{2,t}) = \frac{\theta}{R} - \frac{\lim_{t \to \infty} \bar{X}(\Phi_{2,t})\alpha_1\alpha_2\sigma^2}{\lim_{t \to \infty} R\alpha(\Phi_{2,t})} = \frac{\theta}{R} - \frac{\bar{X}^*\alpha_1\alpha_2\sigma^2}{R\alpha^*} \tag{53}
\]
where \(\bar{X}^* = q\bar{X}_1 + (1 - q)\bar{X}_2\) and \(\alpha^* = qa_2 + (1 - q)a_1\).
Q.E.D

Proof of Proposition 6:
The indirect utility function of a class 1 investor is:
\[
U_{1,t} = -\exp\left\{-\alpha_1 \left[ R\bar{Z}_t + R\bar{X}_t P(\Phi_{1,t}) + \frac{(\theta - R P(\Phi_{1,t}))^2}{2\alpha_1\sigma^2}\right]\right\} \tag{52}
\]
Let \(\bar{X}(\Phi_t) = q_t\bar{X}_1 + (1 - q_t)\bar{X}_2\) and \(\alpha(\Phi_t) = q_t a_2 + (1 - q_t) a_1\). Using (15) and Proposition 2, \(P(\Phi_{1,t})\) can be written as:
\[
P(\Phi_{1,t}) = \frac{\theta}{R} - \frac{(\bar{X}(\Phi_t) - \frac{V_t[q](\bar{X}_2 - \bar{X}_1)}{q_t})\alpha_1\alpha_2\sigma^2}{R (\alpha(\Phi_t) + \frac{V_t[q](\alpha_2 - \alpha_1)}{q_t})} \tag{53}
\]
Plugging the price into the indirect utility function and differentiating with respect to $V_t[q]$,

$$\frac{\partial U_{1,t}}{\partial V_t[q]} = -\exp \left\{ -\alpha_1 \left[ R\bar{Z}_1 + R\bar{X}_1 P(\Phi_{1,t}) + \frac{(\theta - RP(\Phi_{1,t}))^2}{2\alpha_1\sigma^2} \right] \right\}$$

$$\frac{\alpha_1^2 \alpha_2 \sigma^2 q_t (\Delta \alpha \bar{X})^2 (1 - q_{1,t})}{\alpha(\Phi_{1,t}) [q_t \alpha(\Phi_t) + V_t[q] (\alpha_2 - \alpha_1)]^2} < 0 \quad (54)$$

Similarly for a class 2 investor,

$$P(\Phi_{2,t}) = \frac{\theta}{R} - \frac{\left( \bar{X}(\Phi_t) + \frac{V_t[q](\bar{x}_2 - \bar{x}_1)}{1 - q_t} \right) \alpha_1 \alpha_2 \sigma^2}{R \left( \alpha(\Phi_t) - \frac{V_t[q](\alpha_2 - \alpha_1)}{1 - q_t} \right)} \quad (55)$$

$$\frac{\partial U_{2,t}}{\partial V_t[q]} = -\exp \left\{ -\alpha_2 \left[ R\bar{Z}_2 + R\bar{X}_2 P(\Phi_{2,t}) + \frac{(\theta - RP(\Phi_{2,t}))^2}{2\alpha_2\sigma^2} \right] \right\}$$

$$\frac{\alpha_1 \alpha_2 \sigma^2 (1 - q_t) (\Delta \alpha \bar{X})^2 q_{2,t}}{\alpha(\Phi_{2,t}) [(1 - q_t) \alpha(\Phi_t) - V_t[q] (\alpha_2 - \alpha_1)]^2} < 0 \quad (56)$$

Q.E.D


References


