Quote Setting and Price Formation in an Order Driven Market

Puneet Handa *

Robert Schwartz **

Ashish Tiwari ***

Current Draft: January 2002

Journal of Financial Markets, forthcoming

* Robert L. Jensen Fellow in Finance, Henry B. Tippie College of Business, University of Iowa, Iowa City 52242, Ph.: 319-335-2731, Fax: 319-335-3690, E-mail: puneet-handa@uiowa.edu.

** Marvin M. Speiser Professor of Finance, Zicklin School of Business, Baruch College, CUNY, One Bernard Baruch Way, Box 10-225 New York, NY 10010-5585, Ph.: 646-312-3467, Fax: 646-312-3530, E-mail: robert_schwartz@baruch.cuny.edu.

*** Assistant Professor, Henry B. Tippie College of Business, University of Iowa, Iowa City 52242, Ph.: 319-353-2185, Fax: 319-335-3690, E-mail: ashish-tiwari@uiowa.edu.

We would like to thank Yakov Amihud, Utpal Bhattacharya, Bruno Biais, Didier Davidoff, Tom George, Larry Glosten, Bertrand Jacquillat, Larry Harris, Tom Rietz, Avanidhar Subrahmaniam and the seminar participants at NYU, University of Iowa, SUNY-Binghamton, Dartmouth College, LSE Market Microstructure Conference (London, 1997), SBF Bourse Equity Markets Conference (Paris, 1996), and The Pacific Basin Conference (New Brunswick, 1996) for their comments and suggestions. We are also most grateful to the referee for his or her helpful insights and suggestions. All errors remain ours.

Abstract

This paper models quote setting and price formation in a non-intermediated, order driven market where trading occurs because investors differ in their share valuations and the advent of news that is not common knowledge, and tests the model using transaction data on individual stocks in the ParisBourse CAC40 index. As an extension of Foucault (1999), we show that the size of the spread is a function of the differences in valuation among investors and of adverse selection. Both GMM estimation of the model parameters and empirical evidence on spread behavior as the relative proportion of buyers and sellers in the market changes, provide strong support for the model. Our analysis yields further insight into the dynamic process of price formation and into the market clearing process in an order driven market.

Quote Setting and Price Formation in an Order Driven Market

1. Introduction

This paper models quote setting and price formation in a non-intermediated, order driven market, and tests the model using transaction data on individual stocks in the ParisBourse CAC40 index. In an extension of Foucault (1999) where investor share valuations for a security differ, we examine the impact of asymmetric information on posted bid and ask prices. Our primary objective is to analyze the determinants of the bid-ask spread in an order driven market and our main contribution is to test the empirical implications of the extended Foucault model.

The theoretical model we present in this paper shows, and our empirical tests confirm, that the size of the inside spread in an order driven market is a function of differences in valuation among investors. As we will explain, our analysis in this regard may be viewed as an extension of Cohen, Maier, Schwartz and Whitcomb (1981) who demonstrate that a natural bid-ask spread exists in an order driven market because of the "gravitational pull" an already posted order has on a new, incoming order.[1] The current paper also extends Handa and Schwartz (1996) who analyze the rationale and profitability of limit order trading, but who do not explicitly model a trader's decision as to whether to place a limit order or a market order. Our analysis is also closely related to Glosten (1994) who examines the inevitability of an electronic limit order book when there is an exogenous supply of limit orders.

Increased attention has been focused on order driven markets in recent years. This is in large part attributable to the U.S. Securities and Exchange Commission's new order handling rules;[2] Nasdaq's development of SuperMontage; the growth of new Electronic Communications

---

[1] Specifically, the gravitational pull argument refers to the attractiveness to a participant of trading with certainty at a posted quote, rather than placing a limit order to buy (or sell) within the immediate neighborhood of an already posted offer (or bid) and risking non-execution of the order.

[2] The new order handling rules, which were phased in in 1997, require that market makers holding customer limit orders display those limit orders in their quotes. The new rules also require market makers who post more aggressively priced quotes in an ECN to update their own quotes in Nasdaq to match their ECN quotes if the ECN does not itself display the top of its book in the Nasdaq quote montage.

Networks (ECNs); and the development of electronic limit order book trading platforms in virtually all of the market centers in Europe.[3]

Concurrently, more attention has been given in the microstructure literature to limit order trading. The significance of the order driven environment has been underscored by theoretical research such as Glosten (1994). Parlour (1998) presents a dynamic model of a limit order market in which the decision to submit a market order or a limit order depends critically on the state of the limit order book and the trader's place in the limit order queue. Models of liquidity provision in a market that includes a specialist along with limit order traders, have been examined by Rock (1996) and Seppi (1997), among others.[4] Rock (1996) solves for the equilibrium limit order book when a competitive risk-averse specialist competes with off-exchange investors. Seppi (1997) analyzes the limit order book in an environment where a specialist with market power competes with other potential liquidity providers, namely, limit order traders and the trading crowd. However, neither Rock nor Seppi analyze the choice between limit orders and market orders faced by traders.

As noted, trading occurs in our model because investors differ in their share valuations. We assume all investors fit into one of two groups, one of which places a higher value on the security than the other.[5] In addition, some traders receive private information that becomes public knowledge at the end of one trading period, at which time all participants revise their valuations by the amount of the innovation. Investors arrive in the marketplace sequentially and choose to trade via a market order or a limit order.

Quote setting and price formation are not trivial when investors have differing share valuations. The high valuation investors would like to buy shares at the price assessed by the low valuation shareholders, but if they were to do so the low valuation investors would receive no gains from trade. Similarly, the low valuation investors would like to sell shares at the price assessed by the high valuation shareholders but, if they were to do so, the high valuation investors would receive

---

[3] Many exchanges around the globe, such as the Toronto Stock Exchange, the ParisBourse, the Australian Stock Exchange and the Tokyo Stock Exchange operate without dealers or market makers and offer viable bid-ask spreads and prices for their stocks.

[4] Other related papers include Angel (1995), Chakravarty and Holden (1995), Harris (1994) and more recently Kavajecz (1999).

[5] Difference in valuation may simply arise from taxes, liquidity shocks or other portfolio considerations, or may be an outcome of a more complex belief system where some traders are over-confident. See, for example, Daniel, Hirshleifer and Subrahmanyam (1998) and Odean (1998). Also see Harris and Raviv (1993) for an alternative interpretation based upon differences in opinion.

no gains from trade. We show that quotes are optimally set between the two valuations according to the risks of non-execution faced by investors in each of the two groups. In equilibrium, the spread increases monotonically with the difference in valuation between low and high valuation investors. Moreover, for a given difference, the spread is smallest when investors are predominantly in one or the other of the two groups.

Our model is an extension of Foucault (1999). Foucault models the choice between limit orders and market orders. Trading occurs in Foucault's paper because of differences in valuation, but there is no private information in his model. Instead, asset value evolves as public information arrives. Generally, Foucault assumes an equal proportion of buyers and sellers in the marketplace. As a special case he allows this proportion to be unrestricted but only when the asset value does not change over time.

In contrast, our model focuses on the bid-ask spread where the proportion of buyers and sellers is free to vary without restriction. Additionally, our limit order traders face adverse selection due to the presence of privately informed traders. As a result, in our model, the bid-ask spread is a function of both the adverse selection cost and differences in valuation.

We test the model using transaction data on individual stocks in the ParisBourse CAC40 index. The evidence on the size of the spread as the proportion of traders becomes concentrated on one side of the market or the other is consistent with our model's predictions. We perform a GMM estimation of the underlying model parameters for the CAC40 stocks on a stock by stock basis for each of three intra-day trading intervals (morning, midday and afternoon). We fail to reject our model for 94 of the 120 stock-intervals examined (i.e., 77.78% of the cases).

The paper is organized as follows. In Section 2, we describe a model of a pure order driven market. We analyze the equilibrium in Section 3. In Section 4, we present an extension of our model when there are large information shocks. In Section 5 we derive a key empirical implication of our model. Section 6 presents an empirical test and our findings using data from the ParisBourse. Our conclusions appear in Section 7.

2. Model Assumptions

Assumption 1(Trading Mechanism): There is a single risky asset that trades in a continuous market environment. Investors arrive sequentially to trade one share of the risky asset via either a market order or a limit order. A market order can be placed only if a counterpart limit order exists, and in such a case the market order executes with certainty at the counterpart limit order's price. If a limit order is placed, its execution is not certain; the probability of execution depends on the price at which it is placed, on the proportion of buyers and sellers in the market, and on the private information which exists. A limit order is held until after the next trader arrives, at which point it either executes or expires.

Assumption 2 (Asset Valuation): There are two groups of investors, one which attaches a high value V_h to the asset and the other which attaches a low value V_l, where $V_h > V_l$.[6] Vis-à-vis each other, the first group is potential *buyers* and the second group is potential *sellers*. For the buyers, V_h can be interpreted as the reservation price they are willing to pay for one share of the risky asset. Similarly, for the sellers, V_l can be interpreted as the reservation price they are willing to receive for one share of the risky asset. The exogenously determined proportion of investors of type V_h is given by k and of type V_l by $1-k$. The risky asset has an expected liquidation value equal to $\overline{V} + \widetilde{\varepsilon}$ where $\varepsilon = \pm H$ with equal probability and $\overline{V} = kV_h + (1-k)V_l$. All parameters of the model, including V_h, V_l, k and H are known to the investors.

Assumption 3 (Preferences): Traders are risk-neutral expected utility maximizers. The expected utility from a *buy* order at a specified price P is $E(U) = \phi(V_i - P)$, $i = l, h$ where ϕ is the probability of execution of the order. Similarly, the expected utility from a *sell* order at a specified price P is $E(U) = \phi(P - V_i)$, $i = l, h$. In the absence of a trade, the utility is normalized to zero.

Assumption 4 (Asymmetric Information): Before an order is placed, a proportion of investors denoted by δ, privately observe the value of the innovation ε in the asset value. The private information has a value $+H$ with probability one-half and a value $-H$ with probability one-half. The information remains private for one trade, after which it is publicly revealed and traders of both types revise their subjective valuations of the asset upward or downward by an amount H. The short-lived nature of the private information implies that informed traders only trade by means of market orders. We assume that the proportion of traders who become informed is identical across trader types V_h and V_l. We also initially assume that the magnitude of the innovation in the asset value is bounded such that a type V_h (V_l) trader does not switch to become a type V_l (V_h)

---

[6] These differences in valuation are an outcome of taxes, liquidity shocks or other portfolio considerations.

trader after observing an innovation with a value equal to $-H$ ($+H$).[7] We show later that our model and its predictions are qualitatively unchanged if we relax the assumption that the magnitude of the innovation is bounded from above.

3. Equilibrium Bid-Ask Prices and Spread in an Order Driven Market

Equilibrium in our model

Equilibrium is defined as a set of mutual strategies such that each trader's strategy is optimal given the strategies of other traders. A strategy involves the choice of (a) the type of order (market order versus limit order), and (b) if a limit order is chosen, the bid or ask price at which the order is placed. The equilibrium is characterized by optimal bid and offer prices, B^* and A^*, that just induce a counterpart trader arriving next period to trade via a market order.

The Informed Trader's Problem

An informed trader who attaches a high value V_h to the asset and observes a private signal $+H$ will buy if and only if $V_h + H > A^m$, where A^m is the market ask. In our initial parameterization, this trader does not trade if he or she observes a private signal $-H$.

Similarly, an informed trader who attaches a low value V_l to the asset and observes a private signal $-H$ will sell if and only if $V_l - H < B^m$, where B^m is the market bid. Again, in our initial parameterization, this trader does not trade if he or she observes a private signal $+H$.

The Uninformed Trader's Problem

The uninformed buyer (a trader with a high valuation V_h for the asset), faces a decision tree shown in the upper panel of Figure 1. There are two cases when the limit order placed by an uninformed buyer does not execute: (1) the arriving trader is a buyer, (2) the arriving trader is a seller with good news. Normalizing the trader's utility to zero if the limit order expires without executing and simplifying, we can write the expected utility of a limit buy order placed at a bid

[7] We formally establish this result and the corresponding constraint on the value of the parameter H in Proposition 1. Intuitively, exogenous shocks to the system in the form of private signals must be small compared to the difference in valuation, and not large enough to make a seller (buyer) out of a potential buyer (seller).

price B as:

$$E(U) = (1-k)\cdot\{(1-p)\cdot(V_h - B) - pH\}$$

where $p = \delta/2$. An uninformed buyer is indifferent between trading via limit order or trading via market order if the utility derived from trading via a market order equals that from trading via a limit order, i.e.,

$$V_h - A^m = (1-k)\cdot\{(1-p)\cdot(V_h - B) - pH\}$$

Hence, the optimal ask price A^* that will induce an incoming uninformed buyer to trade via a market order is given by:

$$A^* = V_h - (1-k)\left[(1-p)\cdot(V_h - B) - pH\right]$$

Similarly, the optimal bid price B^* that will induce an incoming uninformed seller to trade via a market order is given by:

$$B^* = V_l + k\left[(1-p)\cdot(A - V_l) - pH\right]$$

Equilibrium Bid and Ask Prices

The equilibrium is characterized by the optimal bid and ask prices $\{B^*, A^*\}$ which we establish in Proposition 1. As we will see, the equilibrium bid and ask prices depend upon (1) the difference in valuation in the marketplace, and (2) the adverse selection problem faced by an investor who chooses to place a limit order. In Proposition 1, for tractability, we restrict the value of the information shock H so that an informed trader who observes adverse news chooses not to trade. Later on, in proposition 5 we explore the situation when informed traders choose to trade after receiving adverse news.

Proposition 1

There exist parameter values for V_h, V_l, H, δ and k for which equilibrium bid and ask prices are given by:

$$B^* = \lambda\cdot V_\ell + (1-\lambda)\cdot(V_h - qH)$$
$$A^* = \mu\cdot V_h + (1-\mu)\cdot(V_\ell + qH)$$

$$\text{where } q = \frac{p}{1-p}$$

$$\lambda = \frac{1-k(1-p)}{1-k(1-k)(1-p)^2}$$

$$\mu = \frac{1-(1-k)(1-p)}{1-k(1-k)(1-p)^2}$$

and H satisfies the following bounds:

$$Max\left[\frac{1-\lambda}{1+(1-\lambda)q}(V_h - V_l), \frac{1-\mu}{1+(1-\mu)q}(V_h - V_l)\right]$$
$$< H <$$
$$Min\left[\frac{\mu}{1-(1-\mu)q}(V_h - V_l), \frac{\lambda}{1-(1-\lambda)q}(V_h - V_l)\right]$$

Proof: See Appendix.

Let us first consider the optimal bid price, B^*. It is a weighted combination of V_l (the reservation value of an uninformed market order seller) and $V_h - qH$ (the reservation value of an uninformed limit order buyer adjusted for the expected loss to an informed trader), with λ and $(1-\lambda)$ being the weights placed on these two terms, respectively. As λ approaches its maximum value of unity, B^* approaches V_l and an uninformed market order seller's gain from trade goes to zero. On the other hand, as λ approaches its minimum value of zero, B^* approaches $V_h - qH$ and the uninformed limit order buyer's expected gain from trade goes to zero. More generally, λ determines the location of B^* between V_l and $V_h - qH$, and thus how the benefits of trading are shared between an uninformed limit order buyer and an uninformed market order seller in the case when the buyer arrives first.

Additional insight can be gained by closely examining the term λ. The unconditional probability of an uninformed buyer arriving in the marketplace is given by $k(1-p)$. For a trade to occur between an uninformed buyer and an uninformed seller, the two participants must arrive sequentially and the joint probability is given by $k(1-k)(1-p)^2$. Thus λ, the ratio of the complement of the two probabilities, reflects the relative risk of non-execution to an uninformed buyer vis-à-vis the risk of non-execution to both parties. The uninformed limit order buyer optimally places the bid to apportion the gain from trade in accordance with this ratio. Similar interpretations can be given to the optimal ask price, A^*, and the parameter μ.

Note that the quotes, A^* and B^*, that we solve for are "shadow quotes". In a one period model, only one of the two can be observed at any time.

Proposition 2

A natural positive spread exists in the market that is given by:

$$\pi = \varphi \cdot (V_h - V_\ell) + (1-\varphi) \cdot qH$$
$$\text{where } \varphi = [1-(1-k)(1-p)]\lambda = [1-k(1-p)]\mu$$

Proof: See Appendix.

In the order driven market we study, the spread depends on (1) the difference in valuation and (2) adverse selection. The term $(V_h - V_l)$ captures the effect of difference in valuation and the term qH captures the effect of adverse selection. This is in contrast to a dealer market where the spread is a function of inventory costs and adverse selection. Notice that for our initial parameterization (where $V_h - H > B$) if the difference in valuation $(V_h - V_l)$ gets small then so must the information shock H.

With a difference in valuations, a spread exists in our order driven market even in the absence of inventory costs and asymmetric information. Interestingly, it is the size and location of the spread that allows the market to clear for any value of k. Viewed through the eyes of a buyer (the seller's view is symmetrical), the expected utility from placing a bid at B^* and accepting non-execution risk, equals the expected utility from transacting at the offer with certainty by market order when the offer is at A^*. Thus, a willingness exists on the part of participants to take risk, and it is this willingness that enables the market to clear for all k. For $k = ½$, the number of buyers equals the number of sellers at any price in the range V_h and V_l and there is no rationing problem. For $k > ½$, buyers outnumber sellers and for $k < ½$, sellers outnumber buyers, and no single value of price exists that just clears the market. Consequently, some participants must "go away empty handed" in our model, these are the limit order traders on the heavy side of the market who have been selected by chance rather than by an explicit rationing device. In our market, the spread covers the cost of non-execution that is borne by limit order traders who supply liquidity to liquidity demanding market order traders. The effect of k on the location of the optimal bid and offer (B^* and A^*) between V_h and V_l reflects the balance of relative execution probabilities for buyers and sellers, as discussed above in relation to λ and μ.

An alternative interpretation of the spread in our market is possible. At any moment, we know the bid price that would be posted (if no ask price exists) and the ask price that would be posted (if no bid price exists). Both these quotes are set so that the next counterparty is just induced to trade with certainty by market order.[8] Thus, in our market the bid-ask spread is the smallest distance between two counterpart quotes that can be sustained without the two orders being drawn to each other because of the desirability of transacting with certainty via a market order [this is equivalent to Cohen, Maier, Schwartz and Whitcomb (1981)'s interpretation of spread and their notion of "gravitational pull"].

The effect on the spread of the distribution of participants by V_h and V_l is of further interest. We now show that the spread in an order driven market attains it maximum value when there is an equal proportion of traders on the two sides of the market, i.e., when k is ½:

Proposition 3

The spread is a concave function of k, and is maximized when $k = ½$, where its value is:

$$\pi = \varphi \cdot (V_h - V_\ell) + (1-\varphi) \cdot qH$$
$$\text{where } \varphi = \frac{(1+p)}{(3-p)}.$$

Proof: See Appendix.

Corollary

The spread is minimized when $k = 0$ and when $k = 1$, where its value is

$$\pi = \varphi \cdot (V_h - V_\ell) + (1-\varphi) \cdot qH$$
$$\text{where } \varphi = p.$$

Proof: See Appendix.

We have shown that the spread is maximized when the traders are on both sides of the market in equal proportion, and that it is minimized when k is close to zero or one. This result generalizes the result in Foucault (1999) to include adverse selection. Specifically, when the proportion of sellers is large ($k \approx 0$) the bid and ask prices are both close to V_l and when the proportion of

[8] In the spirit of marginal analysis, the next counterparty is indifferent between trading via a market order and trading via a limit order.

buyers is large ($k \approx 1$) the bid and ask prices are both close to V_h. This result can be understood intuitively as follows. Suppose buyers substantially outnumber sellers in our market, which means that more buy than sell orders will not execute.[9] Our optimality conditions are satisfied by both the bid and offer being close to V_h and the spread being relatively tight. Because the attractiveness to buyers of trading via market orders has been increased by the relative dearth of sellers and the inter-temporal competition that exists between a buyer today and future buyers, sellers post their asks aggressively close to V_h. Concurrently, the attractiveness to sellers of trading via market orders is relatively low, and thus buyers also place their bids close to V_h in order to induce the seller to trade via a market order. We see that buyers' willingness to place limit orders acts as an implicit rationing mechanism as several of them eventually walk away empty-handed. This feature of the market allows it to clear even in the face of an apparent buy-sell imbalance. A similar result holds when sellers outnumber buyers in our market. Hence, all traders are satisfied ex-ante and have no ex-post regret.

In our framework the unconditional value of the asset V, can be one of two values V_h or V_l. The unconditional mean value of the asset may be expressed as

$$\overline{V} = kV_h + (1-k)V_l .$$

We establish below that the bid-ask spread is related to the variance of the unconditional value of the asset V, which captures the difference in valuation.

Proposition 4

The bid-ask spread is proportional to the variance of the unconditional value of the asset V.

Proof: See Appendix.

The result in Proposition 4 is consistent with theoretical models of the bid-ask spread in dealer markets [see, for example, Copeland and Galai (1983) and Easley and O'Hara (1987)] and with the interpretation of a limit buy (sell) order as a free put (call) option offered by the limit order trader.

---

[9] This is the analog of the result in a dealer market when a dealer acquires a large unbalanced inventory position (in the present case, a large short position).

4. Model Extensions

As a first step toward understanding the price formation process in a non-intermediated environment, we have attempted to focus our attention on the investor's order placement decision in a fairly structured environment. Additional important and interesting extensions would include (i) long-lived private information (to allow the book to be informative), (ii) multiple-share orders (to study order-size and volume effects), and (iii) simultaneous placement of multiple limit orders, including both buy and sell limit orders from the same participant, among others.

We now examine the equilibrium when the magnitude of the information shock H is not restricted by the upper and lower bounds given in Proposition 1. First consider the case when the economy is characterized by information shocks that are relatively small and violate the lower bound in Proposition 1. In this setting, the informed buyer (seller) even after observing a negative (positive) information shock, chooses to buy at the posted ask (bid) price. Intuitively, the informed investor behaves similar to the uninformed investor. The net effect is as if there was no information in the model. Hence, the resulting equilibrium is a special case of Proposition 1 with $H = 0$.[10]

We next turn to the case when the economy is characterized by information shocks that are relatively large.

4.1 Equilibrium With Large Information Shocks

When the information shock is positive and exceeds the upper bound placed on H in Proposition 1, informed sellers switch to become potential buyers. Conversely, if the information shock is negative and exceeds the bounds, informed buyers switch to become potential sellers. In this case, the equilibrium is qualitatively similar to that in Proposition 1 and is presented below in Proposition 5.

Proposition 5

There exist parameter values for V_h, V_l, H, p and k for which equilibrium bid and ask prices are given by:

$$B^* = \lambda V_l + (1-\lambda)V_h + pH\left\{\frac{[k-(2k-1)p]-1}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}\right\}$$

where $\lambda = \dfrac{(1-k)+(2k-1)p}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}$ and

$$A^* = \mu V_h + (1-\mu)V_l + pH\left\{\frac{1-[(1-k)-(1-2k)p]}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}\right\}$$

where $\mu = \dfrac{k-(1-2k)p}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}$.

The spread is given by:

$$A^* - B^* = \phi(V_h - V_l) + (1+\phi)H$$

where $\phi = \dfrac{[(1-k)-(1-2k)p][k-(2k-1)p]}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}$.

Proof: See Appendix.

5. Deriving a Testable Implication of the Model

To derive a testable implication of our model, we first consider the implications of introducing a risk-neutral competitive dealer in our framework. The dealer posts a bid price and an ask price to trade one share of the security. Following Glosten and Milgrom (1985), the bid price set by the dealer is the expected value of the security conditional upon a sell order. Similarly, the ask price is the expected value of the security conditional upon a buy order. Proposition 6 shows the bid and ask that the dealer would set in our environment.

Proposition 6

The optimal bid and ask prices set by a risk-neutral dealer are given by:

$$Bid = \overline{V} - qH$$

$$Ask = \overline{V} + qH\ .$$

Thus, the dealer bid-ask spread is given by:

$$Spread = 2qH$$

---

[10] We thank the referee for drawing our attention to this case.

Proof: See Appendix.

In the context of our model, the dealer spread consists of an adverse selection component that is free of k, the proportion of traders of type V_h. This is in contrast with the results in Proposition 3 where we showed that in an order driven market, the spread is maximized when $k = ½$. In fact, risk aversion on the part of the dealer would imply a spread that is minimized at $k = ½$ as the dealer would tend to increase the spread in one-sided markets to avoid going excessively short or long on the risky asset. Hence, the functional form of the spread in Proposition 6 when contrasted with that in Proposition 3 provides a testable implication of our model.

As an example, suppose the traders with a high valuation assess the stock is at V_h = \$105.00 and the traders with a low valuation assess it at V_l = \$95.00. We will consider the equilibrium bid and ask prices as we allow k (the proportion of traders who believe that the value is V_h) to vary. Assume that the unconditional probability p of a trade with an informed investor is given by 0.375 and that the information shock (H) is \$5.00. With these parameter values, the conditions presented in *Proposition 1* are satisfied. We plot the limit order spreads and the dealer spreads in *Figure 2*. The limit order spread attains a maximum value at $k = ½$ and minimum values at $k = 0$ and $k = 1$. The dealer spread, on the other hand, is constant and free of k. Notice that the order driven market spread exceeds the dealer spread for intermediate values of k. However, as k tends to zero or one, resulting in a one-sided market, the order driven market spread declines.

Intuitively, in a one-sided market, a limit order trader on the thicker side of the market is willing to concede a large portion of the trading surplus in order to induce an opposite side participant to trade. A natural question that arises is what is the source of this trading surplus for a limit order trader? To understand this recall that , unlike the dealer, a limit order trader starts out with a potential gain from trade measured from their high or low valuation of the asset. This allows the trader more room to bargain as compared to a dealer who is concerned with the adverse selection problem and controlling inventory by the location of the spread relative to the average value of the asset. This phenomenon may also help explain the claim often made by the New York Stock Exchange that by allowing investors to trade directly with each other through the limit order book, it offers opportunities of price improvement vis-à-vis a pure dealer quote. It also provides justification for why the SEC special order handling rules forcing dealers to allow investors to submit limit orders, might lead to price improvement.

6. Empirical Analysis

Our model implies that in an order driven market, the bid-ask spread is related to the proportion of buyers (k) in the market. This proportion of buyers (relative to all traders) can be viewed as a measure of imbalance in the market place during a period. More specifically, the spread is a concave function of k and is at a maximum at $k = 0.5$. Therefore, in the region where k is less than 0.5, the spread is positively related to the imbalance measure and in the region where k is greater than 0.5, the spread is negatively related to it. As we will discuss, our evidence supports the hypothesized relationship between the spread π and the variable k.

6.1. Data

We use the ParisBourse BDM (*Base de Donnees de Marche*) data base which consists of intra-day data on trades and quotes for all equities traded on the Bourse. The ParisBourse is a centralized order-driven market which provides an ideal testing ground for the implications of our model. Securities listed on the ParisBourse are placed under one of two categories: the official list or the second market. The official list includes large French and foreign companies and is divided further into two segments: the *Reglement Mensuel* (RM) where the most active stocks are traded and the *Comptant* where less active stocks of the official list are traded. The second market caters primarily to medium-sized companies and has less strict inclusion criteria. Trading takes place in a computerized limit order environment (the market mechanism is known as the CAC) through members acting as brokers or principals. The market opens at 10:00 A.M. with a batch auction and trading takes place on a continuous basis until 5:00 P.M. There are no market makers in this market; hence it is particularly suited to our purposes.[11]

We examine the stocks that comprise the CAC40 index. The CAC40 index, launched in 1988, is computed every 30 seconds during the trading day. It is based on a sample of 40 French equities that are selected from the top 100 market capitalization firms on the *Reglement Mensuel* (RM) section of the market.

The database consists of an order file containing information on both limit and market orders placed on the ParisBourse for each stock of the CAC40 for the twelve month period, January

[11] However, member firms are allowed to act as market makers for certain medium-sized stocks and in the case of block trades to enhance liquidity. See Biais, Hillion and Spatt (1995) for a detailed discussion of the ParisBourse.

1995 through December 1995. The file has one time-stamped record for every order entered for each stock throughout the trading day. Each record contains the order price and quantity, type of order (market or limit), classification (good until cancelled, good until end of day, etc.) and other details. This allows us to construct a proxy for the measure of imbalance k. For each stock we divide the trading period into 5-minute intervals and calculate the ratio of the number of trades at the ask plus the number of limit buy orders submitted relative to the total number of limit order submissions (i.e., limit buys and sells) and trades (at the ask and bid). The choice of the interval length is a tradeoff between too much aggregation on the one hand and too noisy a dataset on the other. [12] Over a 5-minute interval, our measure of k is given by:

$$k = \frac{Number\ of\ Trades\ at\ the\ Ask + Limit\ Buy\ Orders\ Submitted}{Number\ of\ Trades\ at\ the\ Ask\ and\ Bid + Limit\ Buy\ and\ Sell\ Orders\ Submitted}$$

To measure the spread, we use the Best-Limit file. The spread is given by:

$$Spread = Best\ Ask\ Price - Best\ Bid\ Price$$

The relative spread is:

$$Rel.Spread = \frac{100 \cdot \left(Ask\ Price - Bid\ Price\right)}{\left(Ask\ Price + Bid\ Price\right)/2}$$

Given the above definitions, we have a measure of spread and relative spread for each observed quote and a measure of k based upon the 5-minute interval a quote falls in. In our tests we use observations on the spread and *contemporaneous* observations of k as this choice is most consistent with the rational expectations setting of our model. Table 1 presents the mean values of these measures for all CAC40 stocks during the six bi-monthly subperiods in 1995. The number of observed quotes, also given in the Table, varied between 775,611 in the sub-period Jan-Feb 1995 to 1,460,430 in the sub-period Nov-Dec 1995. The sub-period average of the imbalance parameter k varied from 0.445 for Nov-Dec 1995 to 0.505 from Mar-Apr 1995. Mean spread varied from a low of 1.168 French Francs in July-August 1995 to 1.693 FFs in Jan-Feb 1995. Finally, the relative spread varied from 0.260% in May-June 1995 to 0.325% in Sep-Oct 1995.

[12] As a robustness check we also measured k over a 15-minute interval. Our results were qualitatively similar.

6.2. Relation between spread, relative spread and the imbalance measure

We examine the relation between spread and the imbalance measure by first classifying spread by quintiles of k. We do this for each spread observation for each CAC40 index stock. The results appear in Table 2 with quintile 1 consisting of observations with the lowest value of k. Our model's prediction is that spread will be lower for the extreme quintiles and higher for the middle quintile, in contrast to the prediction of dealer-based models with inventory costs and asymmetric information. The results support our model. Overall, the spread is low at 1.340 FFs and 1.336 FFs for quintiles 1 and 5 respectively and attains its highest value for quintile 3 (1.456 FFs). The results are robust and fairly consistent across the six subperiods. For each sub-period, we test the null hypothesis that the quintile mean spreads are equal using an F-test. The purpose of the test is to confirm if the inter-quintile differences in spreads are statistically significant. As the low p-values of the F-statistic indicate, the quintile mean spreads are statistically not equal. Figure 3 graphs the difference between the quintile spread and the overall spread across quintiles. A clear inverted U-shaped spread behavior is evident as the imbalance measure is varied.

Table 3 presents results on relative spread classified by quintiles of the imbalance parameter k. Overall, quintiles 1 and 5 have a low relative spread of 0.269% and 0.274% respectively. Relative spread is highest for quintile 3 (0.292%). The pattern of results is fairly consistent across all sub-periods. For each sub-period, we again test the null hypothesis that the quintile mean relative spreads are equal using an F-test. As the low p-values of the F-statistic indicate, the quintile mean relative spreads are statistically different. Figure 4 is a graph that plots the difference between the quintile relative spread and the overall relative spread, and it clearly has an inverted U-shape as the imbalance parameter varies.

Overall, both the behavior of the spread and of the relative spread over values of the imbalance parameter k is consistent with our model. Both variables follow an inverted U pattern, being low when the market is concentrated on one side (proportionately lower order flow on the buy or the sell side) and high when the order flow is relatively well-balanced.

6.3. Correlation tests

We examine the correlation between the spread and the imbalance measure, and also between the relative spread and the imbalance measure. Our prediction is that when $k \leq 0.5$, an increase in k results in an increase in both the spread and relative spread, and conversely, that when $k > 0.5$, an increase in k results in a decrease in both the spread and relative spread.

Table 4 presents the average correlation observed between the spread and the imbalance parameter for the CAC40 firms over the six bi-monthly subperiods. The average correlation is, indeed, positive when $k \leq 0.5$. It varies between 0.040653 and 0.082339, is positive for 82.5% of the firms in January-February 1995, and for 100% of the firms in two of the other sub-periods. Similarly, it is negative when $k > 0.5$, varying between -0.072456 and -0.108300. We test whether each of the average correlations are significantly different from zero using a t-test. The results indicate that in each case when $k \leq 0.5$, the observed correlation is significantly positive and when $k > 0.5$, it is significantly negative.

6.4. A GMM Test of the Model

Our model of the bid-ask spread suggests that in an order driven market, the bid-ask spread (π) for a stock is a function of four underlying parameters: the imbalance parameter, proxied by k defined in Section 6.1; a parameter governing the probability of execution against an informed trader, p; the difference in valuation, $\left(V_h - V_l\right)$; and the information shock, H. More formally,

$$\pi = f\left(k, p, V_h - V_l, H\right)$$

Depending on the magnitude of the information shock H, for small shocks the spread is given by Proposition 1 and for large shocks it is given by proposition 5. We would like our analysis to help us decide which of the two versions of the equilibrium is empirically valid. To do this we turn to the general expression of the spread that encompasses the two versions (Equation 1.9 in the Appendix). This equation allows us to write the spread as:

$$\pi = \phi(V_h - V_l) + (1-\phi)H + [(1-k)(2p-1)\lambda + k(2p-1)\mu]H$$

where $\phi = \dfrac{[1-k'(\alpha - p)][1-k''(\beta - p)]}{1-k'k''(\alpha - p)(\beta - p)}$, and

$k' = (1-k) - kr$, $\alpha. = \dfrac{(1-k)}{k'}$, $k'' = k - (1-k)r$ and $\beta = \dfrac{k}{k''}$.

Note that $r = 0$ corresponds to Proposition 1 and $r = 1$ corresponds to Proposition 5. In order to rule out potential identification problems, we transform the above expression by dividing both sides of the equation by φ:

$$\frac{\pi}{\varphi} = (V_h - V_\ell) + \frac{(1-\varphi)}{\varphi} \cdot H + \frac{[(1-k)(2p-1)\lambda + k(2p-1)\mu]}{\varphi} \cdot H$$

The unknown parameters of the model are p, r, $\left(V_h - V_l\right)$ and H. We carry out a test of our model based upon the Generalized Method of Moments (GMM) proposed by Hansen (1982).

The ***GMM*** error terms are of the form:

$$\varepsilon_t\left(k, p, r, V_h - V_l, H\right) = \frac{\pi}{\varphi} - \left(V_h - V_\ell\right) - \frac{(1-\varphi)}{\varphi} \cdot H - \frac{\left[(1-k)(2p-1)\lambda + k(2p-1)\mu\right]}{\varphi} \cdot H$$

where subscripts are omitted for notational simplicity. Notice that our model implies that $E_{t-1}(\varepsilon_t) = 0$ where ε is a vector of error terms for a particular stock i derived from the time series of spread observations π_i and the corresponding value of the imbalance parameter k_i for a particular stock.

Let z_{t-1} denote a $L \times 1$ vector of variables in the information set of the agents at time $t-1$. Further, define:

$$g_t(\theta) = \varepsilon_t \otimes z_{t-1}$$

where θ is the vector of parameters to be estimated. Our model implies that $E(g_t(\theta)) = 0$. This represents a set of L orthogonality conditions. The GMM estimator of the unknown model parameters, θ, is obtained by minimizing the following criterion function:

$$J_T = G_T' W_T G_T$$

where G_T represents the sample counterpart to the population moment, $E(g_t(\theta))$, and W_T is the inverse of a consistent estimate of the covariance matrix for the orthogonality conditions.

Conditions under which consistent parameter estimates may be obtained and the minimized value of the quadratic form is asymptotically chi-square under the null hypothesis, are discussed by Hansen (1982) as also is the appropriate choice of the weighting matrix, W. The minimised quadratic form is used to test for the goodness-of-fit of the model. The model is over-identified provided the number of orthogonality conditions (L) exceeds the number of parameters being estimated.

We test the model using data on each CAC40 stock for the period Jan-Feb, 1995. To control for volume and volatility effects, we divide each trading day into three sub-periods: morning (10:00 A.M. to 12:00 noon), midday (12:01 P.M. to 3:00 P.M.) and afternoon (3:01 P.M. to 5:00

P.M.).[13] For each stock, we estimate the model over each of the three sub-periods. We employ five instrumental variables for estimation and testing. These include a constant term and four lagged values of the imbalance parameter, k. We estimate the four unknown parameters, H, p, and $(V_h - V_l)$ and r simultaneously. Since we employ five instrumental variables and estimate four parameters, the test statistic is asymptotically distributed as a chi-square variate with one degree of freedom.

6.5. GMM Results

We present a summary of the GMM results in Table 5. The table shows the estimated values of $\left(V_h - V_l\right)$, p and H, averaged across all CAC40 stocks, for the each of the three sub-periods examined by us. Overall, the estimates are fairly stable over the three intraday sub-periods. The "difference in valuation" estimate $\left(V_h - V_l\right)$, measured in Francs, is highest during midday, a period characterized by lower volume. It is the least during the afternoon sub-period. The estimate of p, the unconditional probability of trading with an informed investor, increases during the day from 31.8% to 35.7%.[14] Corresponding to this, the estimate of the magnitude of the information shock H, is minimum during the midday period, and highest in the afternoon. The model rejection rate at the 5% level of significance varies from 17.949% in the midday to 25.641% during the morning. Overall, the GMM results are encouraging.

Results on the parameter r are not reported in the table. The value of r was estimated as zero in all but one case; for the firm **Elf Aquitaine** during the midday sub-period r was estimated as one. As stated in Section 6.4, $r = 0$ corresponds to the version of the equilibrium presented in Proposition 1 where H is bounded from above. The case $r = 1$ corresponds to the equilibrium presented in Proposition 5.

The results for the individual CAC40 stocks, aggregated across the three intraday sub-periods, are presented in Table 6. The "difference in valuations" parameter is estimated between 0 to 4.5 Francs for most of the stocks (37 out of 40). It is the highest (48.422 Francs) for **Legrand Ord.** that has an average stock price of 6453.15 Francs. The parameter $\left(V_h - V_l\right)$ has an average

---

[13] We thank the referee for suggesting that we control for volume and volatility effects in our tests.

[14] Previous studies on intra-day return, volume and spread behavior [for example, Harris (1986), Jain and Joh (1988), McInish and Wood (1992)] suggest that there may be information spikes at the open of a trading day. We do not find such as spike, possibly because we aggregate data over the first two hours of the day.

value of zero in just one case (**Oreal**). The unconditional probability of trading against an informed trader, given by parameter *p*, is between 0 and 50% for 34 out of 40 firms. The magnitude of the information shock, given by parameter *H* is estimated to be between 0 and 4.5 Francs for all but 3 firms. The parameter *H* has an average value of zero Francs in just one case (**LaFarge**). Hence, our results indicate that in the vast majority of cases, the spread has both the difference in valuation component and the adverse selection component.

Overall, detailed results for the forty CAC40 stocks and for three sub-periods, not presented here, indicate that the estimation algorithm converged for 117 cases out of 120 total cases (97.5% cases). The model was rejected 26 times at the 5% level of significance, out of the 117 cases (22.22% rejection rate).[15]

7. Conclusion

We have modeled the process of quote setting and price formation in a non-intermediated, order driven market. Our major finding is that in an order driven market the location of the bid and offer quotes and the size of the bid-ask spread depend on three things: the difference in the valuation among groups of investors, the proportions of investors in each of the groups and adverse selection. In our model, the spread widens as the difference in valuation between the two groups increases and, ceteris paribus, it is at a maximum when the proportion of investors in the two groups is equal. Empirical tests using quote data for the CAC40 index stocks from the ParisBourse genarally validate these implications: the model was rejected at the 5% significance level for only 22% of the firms in our sample. Given the complexity of spread determination and all the factors which affect spreads, we find the GMM results encouraging.

The analysis generates an important insight into the market clearing process itself. One might expect that if natural buyers outnumber natural sellers or vise versa (a condition commonly referred to as a *one way market*), the market would not be able to clear without a dealer intermediating as a seller or buyer of shares. For instance, if the proportion of potential buyers to sellers equals 2:1 up to the highest valuation placed on shares, then apparently at least half of the potential buyers would not be able to purchase shares in the absence of dealer provided immediacy. Nevertheless, we find that the order driven market rations itself as participants resolve the tradeoff between accepting non-execution risk and paying the bid-ask spread. The bid

[15] We conducted Monte Carlo simulations to study the power of the GMM test to reject our model when it is false. The rejection rate was 96% at a 5% level of significance. Details are available from the authors.

and offer quotes are set at a level that just induces just enough participants to accept the risk of non-execution and to save paying the spread by placing limit orders. Ex-post, a requisite proportion of the participants' orders simply expire unfilled. For instance, if the proportion of potential buyers to sellers is high, the quotes will be close to the price assessed by the high valuation group. Some potential sellers will place limit orders because, for them, the risk of non-execution is low, and some potential buyers will place limit orders because, for them, the benefits of trade are low. In equilibrium, sufficiently more buy orders than sell orders fail to execute, some buyers go home empty handed but no participant has ex-post regret.

References

Angel, J., 1995, Limit versus Market Orders, working paper, Goergetown University.

Amihud, Yakov and Haim Mendelson, 1980, Dealership market: market making with inventory, Journal of Financial Economics 8, 31-53.

Bagehot, Walter (pseudonym), 1971, The only game in town, The Financial Analysts Journal 8, 31-53.

Biais, Bruno, Pierre Hillion and Chester Spatt, 1995, An empirical analysis of the limit order book and the order flow in the Paris Bourse, Journal of Finance 50, 1655-89.

Chakravarty, S. and C. Holden, 1995, An integrated model of market and limit orders, Journal of Financial Intermediation 4, 213-241.

Cohen, Kalman, Steven Maier, Robert Schwartz and David Whitcomb, 1981, Transaction costs, order placement strategy, and existence of the bid-ask spread, The Journal of Political Economy April, 287-305.

Copeland, Thomas and Dan Galai, 1983, Information effects on the bid-ask spreads, Journal of Finance 38, 1457-69.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, Journal of Finance 53, 1839-1885.

Easley, David and Maureen O'Hara, 1987, Price, trade size, and information in securities markets, Journal of Financial Economics 19, 69-90.

Foucault, Thierry, 1999, Order flow composition and trading costs in a dynamic order driven market, Journal of Financial Markets 2, 99-134.

Glosten, Lawrence and Paul Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 14, 71-100.

Glosten, Lawrence, 1994, Is the electronic open limit order book inevitable? Journal of Finance 49, 1127-1161.

Handa, Puneet and Robert Schwartz, 1996, Limit order trading, Journal of Finance 51, 1835-1861.

Hansen, Lars P., 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029-1054.

Harris, Milton, and Artur Raviv, 1993, Differences of opinion make a horse race, Review of Financial Studies 6, 473-506.

Harris, L., 1994, Optimal dynamic order submission trading strategies in some stylized trading problems, working paper, University of Southern California.

Harris, L., 1986, A transaction data study of weekly and intradaily patterns in stock returns, Journal of Financial Economics, 16, 99-118.

Ho, Thomas and Hans R. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, Journal of Financial Economics 9, 47-73.

Jain, Prem and Gun-Ho Joh, 1988, The dependence between hourly prices and trading volume, Journal of Financial and Quantitative Analysis 23, 269-284.

Kavajecz, Kenneth A., 1999, A specialist's quoted depth and the limit order book, Journal of Finance 54, 747-771.

McIninsh, Thomas H. and Robert A. Wood, 1992, An analysis of intraday patterns in bid/ask spreads for NYSE stocks, Journal of Finance, 47, 753-764.

Odean, Terrance, 1998, Volume, volatility, price, and profit when all traders are above average, Journal of Finance 53, 1887-1934.

Parlour, Christine A., 1998, Price dynamics in limit order markets, Review of Financial Studies, 11, 789-816.

Rock, K., 1996, The specialist's order book and price anomalies, working paper.

Seppi, D., 1997, Liquidity provision with limit orders and a strategic specialist, Review of Financial Studies 10, 103-150.

Stoll, Hans R., 1978, The supply of dealer services in securities markets, Journal of Finance 33, 1133-51.

Appendix

Proof of Proposition 1

By Assumption 4, the magnitude of the information shock, H, is bounded such that a type V_h (V_l) trader does not switch to become a type V_l (V_h) trader after observing an innovation with a value equal to $-H$ ($+H$). Hence the probability of a trader switching is zero. More generally, it can assumed that this probability has a value r where $0 \leq r \leq 1$. Our proof uses this more general setting.

Given our framework, the expected utility of a limit buy order is:

$$EU_{Limit\ Buy} = (1-k)\left\{(1-\delta)\left[\frac{1}{2}(V_h - B + H) + \frac{1}{2}(V_h - B - H)\right] + \frac{\delta}{2}(V_h - B - H)\right\} + kr\frac{\delta}{2}(V_h - B - H) \qquad (1.1)$$

i.e.,

$$EU_{Limit\ Buy} = (V_h - B)\{(1-k) - p[(1-k) - kr]\} - pH\{(1-k) + kr\} \qquad (1.2)$$

where $p = \dfrac{\delta}{2}$.

Let $(1-k) - kr = k'$ and $\dfrac{(1-k)}{k'} = \alpha$.

The we have

$$EU_{Limit\ Buy} = (V_h - B)(\alpha - p)k' - pH\{(1-k) + kr\} \qquad (1.3)$$

In order to be indifferent between a market buy order and a limit buy order, it must be the case that:

$$V_h - A = (V_h - B)(\alpha - p)k' - pH\{(1-k) + kr\} \qquad (1.4)$$

This gives:

$$A = V_h - (V_h - B)(\alpha - p)k' + pH\{(1-k) + kr\} \qquad (1.5)$$

Following a line of reasoning similar to buy orders, we obtain an expression for the bid price that would make the seller indifferent between a market sell order and a limit sell order:

$$B = V_l - (A - V_l)(\beta - p)k'' - pH\{k + (1-k)r\} \tag{1.6}$$

where $k - (1-k)r = k''$ and $\dfrac{k}{k''} = \beta$.

Solving equations (1.5) and (1.6) simultaneously, we get the optimal bid and ask prices:

$$B^* = \lambda V_l + (1-\lambda)V_h + pH\left\{\frac{k''(\beta - p)[(1-k) + kr] - [k + (1-k)r]}{1 - k'k''(\alpha - p)(\beta - p)}\right\} \tag{1.7}$$

where $\lambda = \dfrac{1 - k''(\beta - p)}{1 - k'k''(\alpha - p)(\beta - p)}$ and

$$A^* = \mu V_h + (1-\mu)V_l + pH\left\{\frac{[(1-k) + kr] - k'(\alpha - p)[k + (1-k)r]}{1 - k'k''(\alpha - p)(\beta - p)}\right\} \tag{1.8}$$

where $\mu = \dfrac{1 - k'(\alpha - p)}{1 - k'k''(\alpha - p)(\beta - p)}$.

The spread is given by:

$$A^* - B^* = \phi(V_h - V_l) + (1-\phi)H + [(1-k)(2p-1)\lambda + k(2p-1)\mu]H \tag{1.9}$$

where $\phi = \dfrac{[1 - k'(\alpha - p)][1 - k''(\beta - p)]}{1 - k'k''(\alpha - p)(\beta - p)}$.

Equations (1.7), (1.8) and (1.9) hold when r, the probability of traders switching from buying to selling or vice versa, is $0 \le r \le 1$. For the special case when $r = 0$, we get:

$$B^* = \lambda \cdot V_\ell + (1-\lambda)\cdot(V_h - qH) \tag{1.10}$$

$$A^* = \mu \cdot V_h + (1-\mu)\cdot(V_\ell + qH). \tag{1.11}$$

where $q = \dfrac{p}{1-p}$

$$\lambda = \frac{1 - k(1-p)}{1 - k(1-k)(1-p)^2}$$

$$\mu = \frac{1 - (1-k)(1-p)}{1 - k(1-k)(1-p)^2}$$

In the proof above, we have assumed that in equilibrium (1) a seller who privately observes a signal -H will sell at the posted bid price B^*, (2) a seller who privately observes a signal +H will not sell at the posted bid price B^*, and (3) a seller who observes $+H$ will not buy at the ask price A^*. These conditions hold provided we have:

$$B^* < \left(V_l + H\right) < A^* \tag{1.12}$$

This condition translates to the following relation:

$$\frac{1-\lambda}{1+(1-\lambda)q}\left(V_h - V_l\right) < H < \frac{\mu}{1-\left(1-\mu\right)q}\left(V_h - V_l\right) \tag{1.13}$$

It can be verified that this region is feasible.

We have also assumed in the proof above that (1) a buyer who privately observes a signal $+H$ will buy at the posted ask price A^*, (2) a buyer who privately observes a signal $-H$ will not buy at the posted ask price A^*, and (3) a buyer who observes $-H$ will not sell at the bid price B^*. These additional conditions hold provided we have:

$$B^* < \left(V_h - H\right) < A^* \tag{1.14}$$

This condition translates to the following relation:

$$\frac{1-\mu}{1+\left(1-\mu\right)q}\left(V_h - V_l\right) < H < \frac{\lambda}{1-(1-\lambda)q}\left(V_h - V_l\right) \tag{1.15}$$

It can be verified that this region is also feasible.

Hence, a sufficient condition for the equilibrium bid and ask prices to hold is that the parameter H is constrained to the region given by:

$$\begin{aligned} &Max\left[\frac{1-\lambda}{1+(1-\lambda)q}\left(V_h - V_l\right), \frac{1-\mu}{1+\left(1-\mu\right)q}\left(V_h - V_l\right)\right] \\ &< H < \\ &Min\left[\frac{\mu}{1-\left(1-\mu\right)q}\left(V_h - V_l\right), \frac{\lambda}{1-(1-\lambda)q}\left(V_h - V_l\right)\right] \end{aligned} \tag{1.16}$$

It can be verified that the region of feasible values for H is a non-empty set.

Proof of Proposition 2

The spread π is given by:

$$\pi = A^* - B^* = \left[\mu \cdot V_h + (1-\mu) \cdot (V_\ell + qH)\right] - \left[\lambda \cdot V_\ell + (1-\lambda) \cdot (V_h - qH)\right]$$

$$= \varphi \cdot (V_h - V_\ell) + (1-\varphi) \cdot qH$$

where $\varphi = [1-(1-k)(1-p)]\lambda$

Proof of Proposition 3

As $V_h - V_\ell > qH$, it suffices to show that φ is maximized at $k = ½$.

$$\text{Now } \varphi = [1-(1-k)(1-p)]\lambda$$
$$= \frac{1+p}{1-(1-k)(1-p)^2} - 1$$

Taking the first derivative, we get

$$\frac{d\varphi}{dk} = -\frac{(1+p)}{y^2} \cdot \frac{dy}{dk}$$

where $y = 1-k(1-k)(1-p)^2$

and $\dfrac{dy}{dk} = (2k-1)(1-p)^2$

Therefore, $\dfrac{d\varphi}{dk} = 0$ when $k = ½$ as $p \in [0, ½]$.

To confirm that $k = ½$ is indeed the maximum, we take the second derivative w.r.t. k and get

$$\frac{d^2\varphi}{dk^2} = -\frac{2(1+p)}{y^3} \cdot \frac{dy}{dk} \cdot (2k-1)(1-p)^2 - \frac{2(1+p)}{y^2}(1-p)^2$$

$$< 0 \text{ when } k = ½.$$

Note that at k = ½ we have $\varphi = \dfrac{(1+p)}{(3-p)}$.

Proof of Corollary

As φ attains a maximum at $k = ½$ and has no other extreme value, it must attain minimum value at the upper and lower bounds of k, i.e. at $k = 0$ and at $k = 1$. Also, we have $\varphi = p$ when $k = 0$ and when $k = 1$.

Proof of Proposition 4

The value of the asset is $V+\varepsilon$ where $\varepsilon = \pm H$. The variance of $V+\varepsilon$ is:

$$\sigma^2_{v+\varepsilon} = \frac{k}{2}\left[\left(V_h + H - \overline{V}\right)^2 + \left(V_h - H - \overline{V}\right)^2\right] + \frac{(1-k)}{2}\left[\left(V_l + H - \overline{V}\right)^2 + \left(V_l - H - \overline{V}\right)^2\right]$$

$$= k(1-k)(V_h - V_l)^2 + H^2$$

Hence, the variance of the unconditional value of the asset V is given by $k(1-k)(V_h - V_l)^2$. Solving for $(V_h - V_l)$ and substituting it in the expression for spread in *Proposition 2* gives us the result.

Proof of Proposition 5

We use the results presented in Equations (1.7), (1.8) and (1.9) under the proof for Proposition 1. When the magnitude of the information shock, H, is large, a type V_h (V_l) trader switches to become a type V_l (V_h) trader after observing an innovation with a value equal to $-H$ ($+H$). Hence the probability of a trader switching is one. For the special case when $r = 1$, Equations (1.7), (1.8) and (1.9) yield:

$$B^* = \lambda V_l + (1-\lambda)V_h + pH\left\{\frac{[k-(2k-1)p]-1}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}\right\} \tag{5.1}$$

where $\lambda = \dfrac{(1-k)+(2k-1)p}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}$ and

$$A^* = \mu V_h + (1-\mu)V_l + pH\left\{\frac{1-[(1-k)-(1-2k)p]}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}\right\} \tag{5.2}$$

where $\mu = \dfrac{k-(1-2k)p}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}$.

The spread is given by:

$$A^* - B^* = \phi(V_h - V_l) + (1+\phi)H \tag{5.3}$$

where $\phi = \dfrac{[(1-k)-(1-2k)p][k-(2k-1)p]}{1-[(1-k)-(1-2k)p][k-(2k-1)p]}$.

Proof of Proposition 6

The bid price set by the dealer is the expected value of the security conditional upon a sell order:

$$Bid = \frac{(1-k)(1-\delta)}{(1-k)(1-\delta/2)}\overline{V} + \frac{(1-k)\delta/2}{(1-k)(1-\delta/2)}\left(\overline{V} - H\right)$$

$$= \overline{V} - qH$$

Similarly, the ask price is the expected value of the security conditional upon a buy order:

$$Ask = \frac{k(1-\delta)}{k(1-\delta/2)}\overline{V} + \frac{k \cdot \delta/2}{k(1-\delta/2)}\left(\overline{V} + H\right)$$

$$= \overline{V} + qH$$

Table 1

Descriptive Statistics: CAC40 at the ParisBourse in 1995

The number of observed quotes for the CAC40 firms, the average imbalance measure (k), the mean spread (π) in French Francs and the mean relative spread ($r\pi$) in % are reported for the six-subperiods in 1995.

We compute k over 5-minute intervals as $k = \dfrac{Number\ of\ Trades\ at\ the\ Ask + Limit\ Buy\ Orders\ Submitted}{Number\ of\ Trades\ at\ the\ Ask\ and\ Bid + Limit\ Buy\ and\ Sell\ Orders\ Submitted}$.

For each quote, $Spread = Best\ Ask\ Price - Best\ Bid\ Price$ and $Rel.Spread = \dfrac{100 \cdot (Ask\ Price - Bid\ Price)}{(Ask\ Price + Bid\ Price)/2}$

	Sub-periods in 1995					
	Jan-Feb	**Mar-Apr**	**May-Jun**	**Jul-Aug**	**Sep-Oct**	**Nov-Dec**
Quotes observed	775,611	1,310,856	1,365,738	1,102,412	1,303,518	1,460,430
Mean k	0.496	0.505	0.459	0.471	0.499	0.445
Mean Spread (FF)	1.693	1.601	1.431	1.168	1.244	1.275
Mean Rel. spread (%)	0.264	0.296	0.260	0.263	0.325	0.287

Table 2

Mean Spread (π) classified by quintiles of the imbalance measure in the market (*k*) for the CAC40 at the ParisBourse in 1995

Mean spreads, classified by quintiles of the imbalance measure (*k*), are shown for each of the six two-month sub-periods during 1995. The measure of market imbalance (*k*), varies between 0 and 1, and is defined over a 5-minute interval as:

$$k = \frac{Number\ of\ Trades\ at\ the\ Ask + Limit\ Buy\ Orders\ Submitted}{Number\ of\ Trades\ at\ the\ Ask\ and\ Bid + Limit\ Buy\ and\ Sell\ Orders\ Submitted}.$$

The last row presents p-values corresponding to a F-test that tests the null hypothesis that the quintile means are equal.

Spread (FF)	Sub-periods in 1995						
	Jan-Feb	Mar-Apr	May-Jun	Jul-Aug	Sep-Oct	Nov-Dec	Overall
Quintile 1 *(low k)*	1.663	1.521	1.364	1.122	1.141	1.226	1.340
Quintile 2	1.720	1.663	1.479	1.189	1.263	1.294	1.435
Quintile 3	1.739	1.684	1.493	1.223	1.290	1.309	1.456
Quintile 4	1.748	1.652	1.465	1.190	1.288	1.320	1.444
Quintile 5 *(high k)*	1.596	1.487	1.356	1.114	1.237	1.228	1.336
Average	1.374	1.304	1.160	0.945	0.996	1.030	1.135
p-value for F-test of equality of quintile means	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Table 3

Mean relative spread ($r\pi$) classified by quintiles of the imbalance measure in the market (k) for the CAC40 at the ParisBourse in 1995

Mean relative spreads, classified by quintiles of the imbalance measure (k), are shown for each of the six two-month sub-periods during 1995. The measure of market imbalance (k), varies between 0 and 1, and is defined over a 5-minute interval as:

$$k = \frac{Number\ of\ Trades\ at\ the\ Ask + Limit\ Buy\ Orders\ Submitted}{Number\ of\ Trades\ at\ the\ Ask\ and\ Bid + Limit\ Buy\ and\ Sell\ Orders\ Submitted}.$$

The last row presents p-values corresponding to a F-test that tests the null hypothesis that the quintile means are equal.

Rel. spread (%)	Sub-periods in 1995						
	Jan-Feb	Mar-Apr	May-Jun	Jul-Aug	Sep-Oct	Nov-Dec	Overall
Quintile 1 *(low k)*	0.257	0.287	0.246	0.254	0.302	0.269	0.269
Quintile 2	0.267	0.302	0.271	0.267	0.329	0.287	0.287
Quintile 3	0.272	0.308	0.269	0.274	0.335	0.295	0.292
Quintile 4	0.275	0.300	0.265	0.270	0.337	0.298	0.291
Quintile 5 *(high k)*	0.252	0.285	0.252	0.251	0.319	0.286	0.274
Average	0.264	0.296	0.260	0.263	0.325	0.287	0.283
p-value for F-test of equality of quintile means	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Table 4

Average correlation between the bid-ask spread (π) and the imbalance measure in the market (k) when k ≤ 0.5 and when k > 0.5 for the CAC40 stocks for 1995.

For a 5-minute interval, $k = \frac{Number\ of\ Trades\ at\ the\ Ask + Limit\ Buy\ Orders\ Submitted}{Number\ of\ Trades\ at\ the\ Ask\ and\ Bid + Limit\ Buy\ and\ Sell\ Orders\ Submitted}$.

For each quote, $Spread = Best\ Ask\ Price - Best\ Bid\ Price$.

k ≤ 0.5	**Jan-Feb**	**Mar-Apr**	**May-Jun**	**Jul-Aug**	**Sep-Oct**	**Nov-Dec**
Average correlation	**0.040653****	**0.057797****	**0.063115****	**0.056398****	**0.082339****	**0.059689****
% of firms with positive correlation	82.5	90.0	92.5	90.0	100.0	100.0
% of firms with positive correlation significant at 5% or better	80.0	82.5	90.0	70.0	100.0	90.0
% of firms with positive correlation significant at 1% or better	70.0	82.5	85.0	67.5	97.4	82.5
k > 0.5	**Jan-Feb**	**Mar-Apr**	**May-Jun**	**Jul-Aug**	**Sep-Oct**	**Nov-Dec**
Average correlation	**-0.108300****	**-0.072456****	**-0.087438****	**-0.089998****	**-0.076476****	**-0.076061****
% of firms with negative correlation	100.0	92.5	92.5	97.5	89.7	100.0
% of firms with negative correlation significant at 5% or better	97.5	97.5	90.0	95.0	87.2	95.0
% of firms with negative correlation significant at 1% or better	97.5	95.0	90.0	92.5	87.2	95.0

** Significantly different from zero at the 1% level of significance.

Table 5

Summary of GMM results for the CAC40 stocks (Jan-Feb, 1995)

The table below presents summary results from estimating the model of bid-ask spread given by the equation:

$$\frac{\pi}{\varphi} = \left(V_h - V_\ell\right) + \frac{(1-\varphi)}{\varphi} \cdot H + \frac{[(1-k)(2p-1)\lambda + k(2p-1)\mu]}{\varphi} \cdot H$$

using the Generalized Method of Moments. Results are presented for all CAC40 stocks and for the period January-February, 1995. During the sample period, each trading day is divided into three sub-periods: morning (10:00 AM to 12 noon), midday (12:01 PM to 3:00 PM) and afternoon (3:01 PM to 5:00 PM) to control for volume and volatility effects. The estimation is carried out for individual stocks for each sub-period separately. The results reported below are parameter estimates averaged across all stocks for the three sub-periods. Also reported are the proportion of times the model is rejected at a 5% significance level.

Period of Day	Average estimate of $(V_h - V_l)$ (Francs)	Average estimate of p	Average estimate of H (Francs)	Average p-value for $\chi^2(1)$	Rejection rate (percent) at the 5% level of significance (%)
Morning (10:00 AM to 12 noon)	3.995	0.318	1.402	0.223	25.641
Midday (12:01 PM to 3:00 PM)	4.075	0.339	1.211	0.239	17.949
Afternoon (3:01 PM to 5:00 PM)	2.362	0.357	2.374	0.295	23.077
Full Sample	3.477	0.338	1.662	0.252	22.222

Table 6

GMM results for individual firms for the CAC40 stocks (Jan-Feb, 1995)

The table below presents results from estimating the model of bid-ask spread given by the equation:

$$\frac{\pi}{\varphi} = (V_h - V_\ell) + \frac{(1-\varphi)}{\varphi}\cdot H + \frac{[(1-k)(2p-1)\lambda + k(2p-1)\mu]}{\varphi}\cdot H$$

using the Generalized Method of Moments. Results are presented for all CAC40 stocks and for the period January-February, 1995. During the sample period, each trading day is divided into three sub-periods: morning (10:00 AM to 12 noon), midday (12:01 PM to 3:00 PM) and afternoon (3:01 PM to 5:00 PM) to control for volume and volatility effects. The estimation is carried out for individual stocks for each sub-period separately. The results reported below are parameter estimates averaged across the three sub-periods for the CAC40 stocks. Also reported are the average p-values associated with the chi-square goodness-of-fit statistic.

CAC40 Firm	Average mid-quote (Francs)	Average estimate of $(V_h - V_l)$ (Francs)	Average estimate of p	Average estimate of H (Francs)	Average p-value for $\chi^2(1)$
SCHNEIDER	363.68	2.603	0.272	1.394	0.485
AIR LIQUIDE	732.06	1.523	0.221	2.297	0.217
CARREFOUR	2133.56	2.676	0.335	4.745	0.094
SANOFI	261.05	2.220	0.345	0.984	0.238
TOTAL	299.04	1.109	0.232	1.335	0.216
OREAL	1120.92	0.000	0.103	4.534	0.137
ACCOR	551.27	2.228	0.333	1.492	0.308
ELF AQUITAINE	376.76	1.740	0.259	0.376	0.102
BOUYGUES	524.61	0.498	0.267	2.992	0.494
LYONNAISE DES EAUX	431.68	0.567	0.527	1.341	0.212
LAFARGE	357.76	3.679	0.187	0.000	0.044
LEGRAND ORD.	6453.15	48.422	0.386	12.259	0.189
AXA	228.41	1.001	0.471	0.040	0.086
GROUPE DANONE	741.98	2.022	0.409	1.542	0.079
PERNOD-RICARD	314.38	4.286	0.157	0.509	0.564

CFF	714.32	4.064	0.561	1.991	0.190
LVMH MOET VUITTON	829.18	2.150	0.492	0.753	0.361
CGIP	1050.71	7.515	0.223	1.249	0.231
PROMODES	951.78	1.269	0.397	5.623	0.138
MICHELIN CAT.B	208.66	1.877	0.377	0.006	0.168
THOMSON-CSF	136.6	2.049	0.179	0.943	0.270
BANCAIRE(CIE)	485.66	1.953	0.557	2.233	0.264
PEUGEOT	473.57	3.229	0.381	1.060	0.215
SAINT-GOBAIN	713.22	3.664	0.162	1.272	0.360
PARIBAS(FIN.)	622.19	0.762	0.280	0.946	0.130
CANAL	320.23	2.788	0.509	2.417	0.704
CASINO GUICHARD	787.15	1.443	0.235	0.660	0.347
CCF	135.96	3.324	0.332	0.357	0.426
EURO DISNEY SCA	210.51	0.180	0.365	0.375	0.052
ALCATEL ALSTHOM	11.85	0.999	0.570	0.461	0.318
HAVAS	438.49	2.615	0.383	1.116	0.188
LAGARDERE GROUPE	372.33	2.014	0.375	0.192	0.235
SAINT LOUIS	117.38	14.051	0.315	3.281	0.108
SOCIETE GENERALE A	1362.72	1.200	0.566	1.292	0.341
SUEZ COMPAGNIE	519.37	0.954	0.381	0.141	0.338
UAP	230.03	0.706	0.339	0.934	0.415
CREDIT LCL FRANCE	123.5	1.587	0.427	0.226	0.234
RHONE-POULENC A	398.35	0.622	0.276	0.726	0.061
BNP	129.18	0.566	0.280	1.135	0.416
SCHNEIDER	242.18	2.603	0.272	1.394	0.485

Figure 1
The decision tree faced by uninformed buyers and sellers.

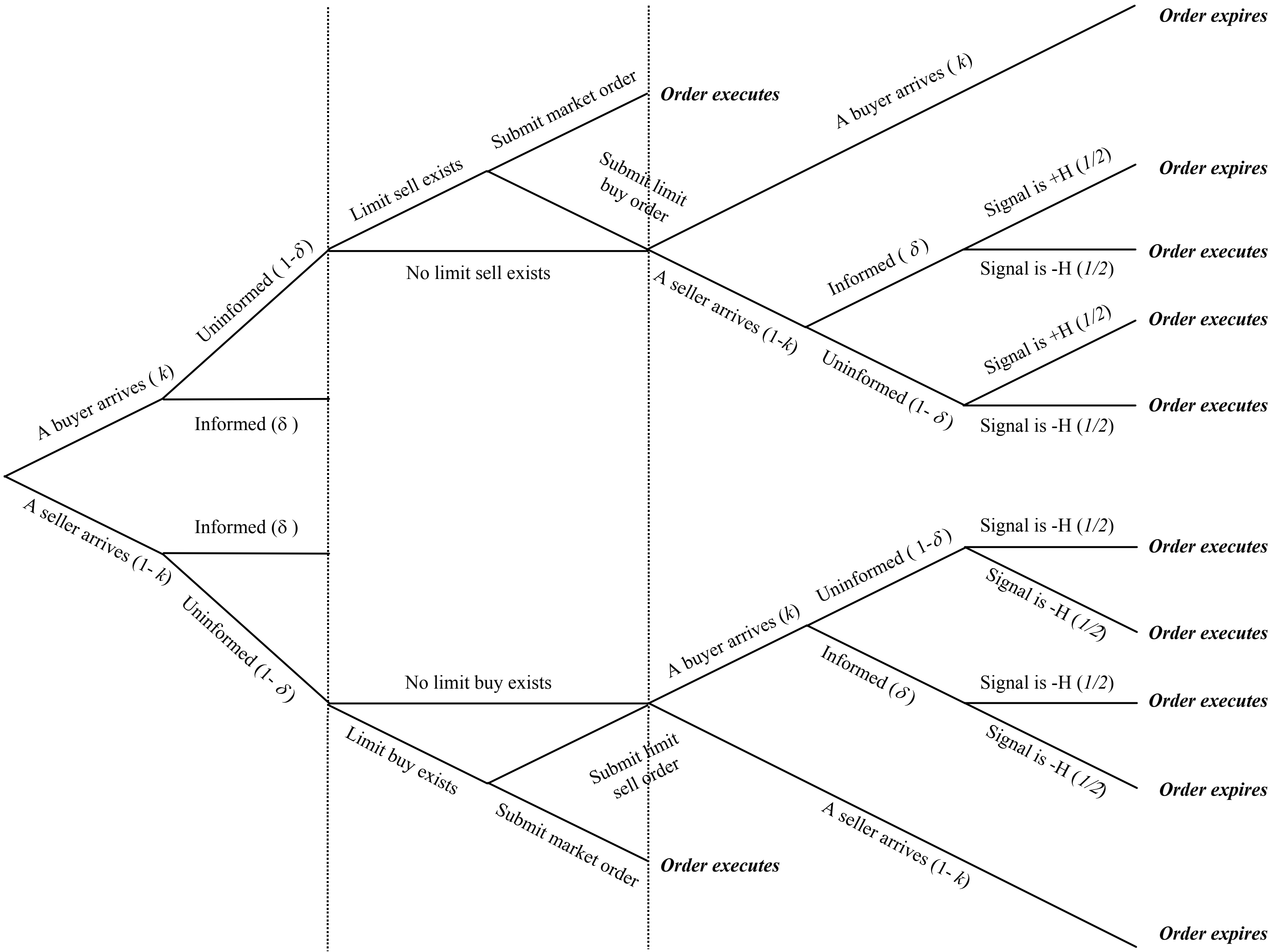

Figure 2

Spread behavior in an order driven market as proportion of buyers is varied

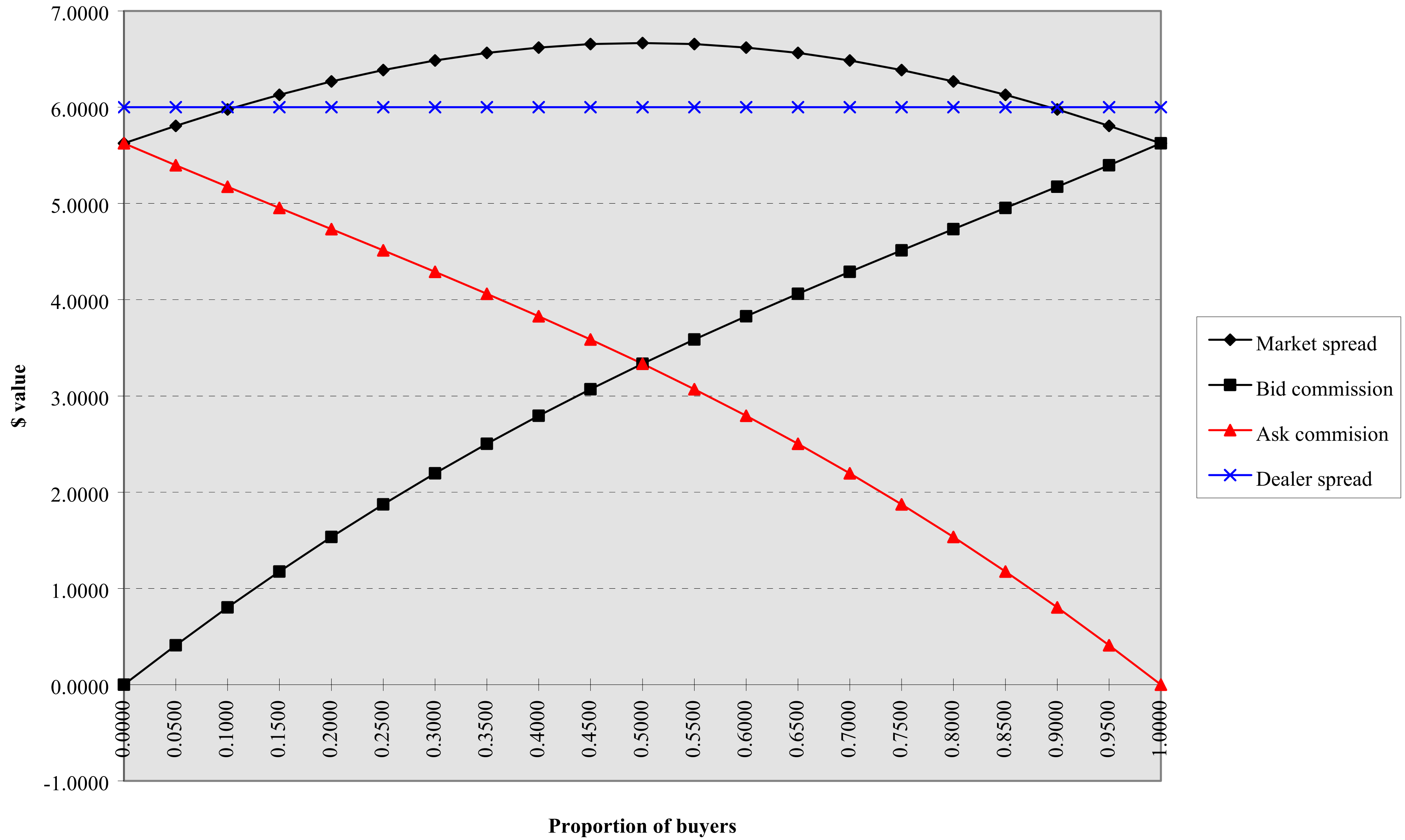

Figure 3

Plot of difference between quintile mean spread and grand mean spread when quotes are classified into quintiles using the proportion of buyers in the market for the CAC40 at the Paris Bourse for six sub-periods in 1995

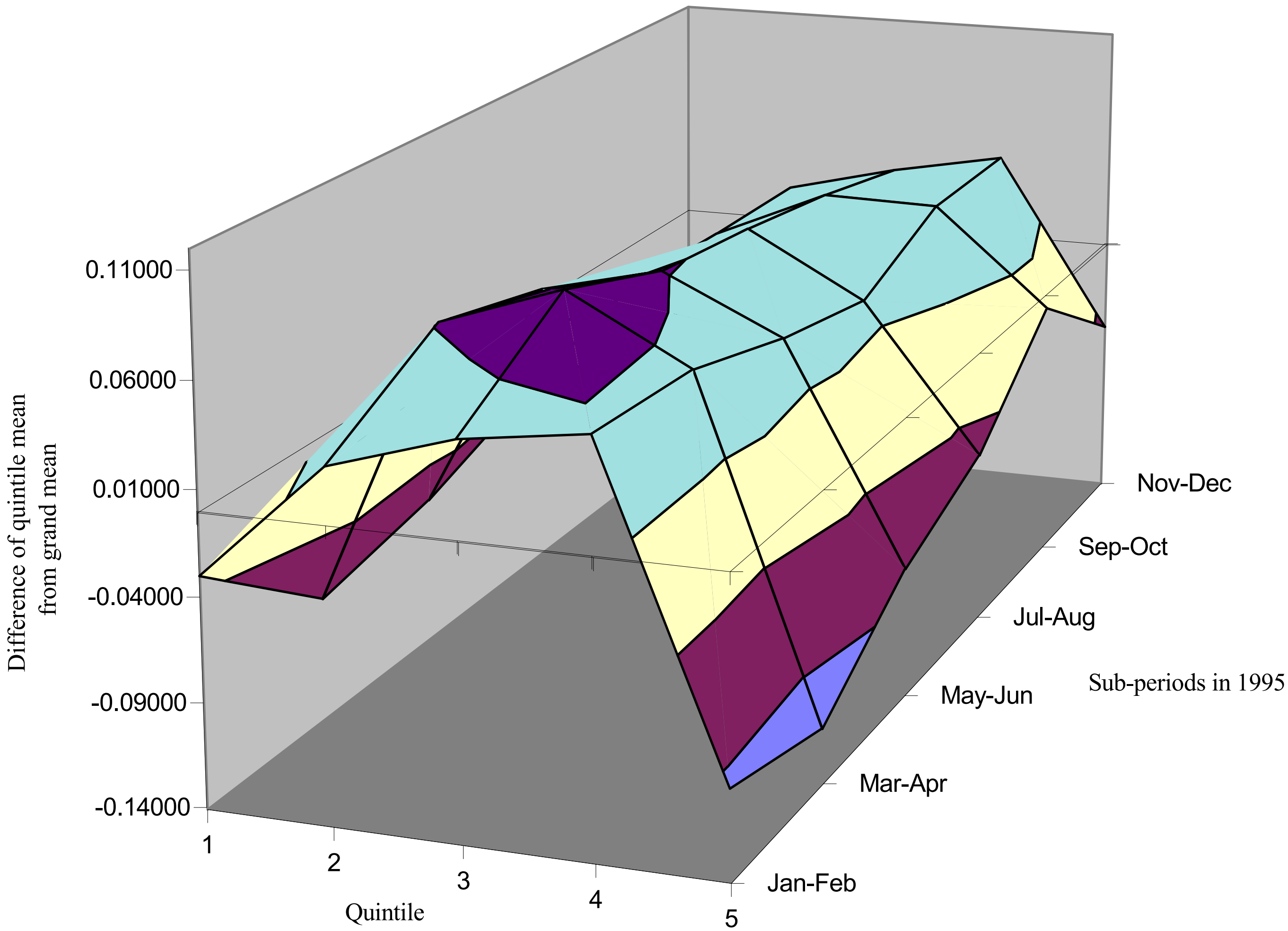

Figure 4

Plot of difference between quintile mean relative spread and grand mean relative spread when quotes are classified into quintiles using the proportion of buyers in the market for the CAC40 at the Paris Bourse for six sub-periods in 1995