BRUSSELS ECONOMIC REVIEW - CAHIERS ECONOMIQUES DE BRUXELLES
VOL. 46 - N°4 WINTER 2003

A Comparison of Alternative Spread Decomposition Models on Euronext Brussels

Rudy De Winne and Christophe Majois*

Abstract:

This paper checks the relevance of alternative spread decomposition models in an order-driven environment. Using intraday data from Euronext Brussels, we compute estimates of the bid-ask spread components provided by eight models. Our results support the hypothesis of no inventory holding costs in order-driven markets. Focusing on adverse selection component, we find high correlation across five models assuming no inventory holding cost. In order to assess the reliability of the "best" models, i.e., Huang & Stoll's (1997) 2-way decomposition, Madhavan et al.'s (1997) method and Lin et al.'s (1995) procedure, we compare their adverse selection cost estimates with five information asymmetry proxies. However, results on that point do not allow us to draw definitive conclusions.

Résumé:

Cet article analyse la pertinence de différents modèles de décomposition de la fourchette de prix dans un marché dirigé par les ordres. Au moyen de données intra-journalières sur Euronext Bruxelles, nous calculons les estimations des composantes de la fourchette de prix fournies par huit modèles. Nos résultats supportent l'hypothèse d'absence de coûts d'inventaire sur les marchés dirigés par les ordres. En nous focalisant sur la composante de sélection adverse, nous mettons en évidence une corrélation élevée entre cinq modèles basés sur l'absence de coûts d'inventaire. Afin d'évaluer la validité des "meilleurs" modèles, à savoir la décomposition en deux parties de Huang & Stoll (1997), la méthode de Madhavan et al. (1997) et la procédure de Lin et al. (1995), nous comparons leur estimation du coût de sélection adverse avec cinq mesures d'asymétrie d'information. Les résultats de cette analyse ne nous permettent cependant pas de tirer des conclusions définitives.

JEL Classification: G14, G10.

Keywords: Market Microstructure, Bid-Ask Spread Components, Order-Driven Markets, Information Asymmetry, Inventory Holding Costs.

Mots-clés: Microstructure des marchés, Composantes de la fourchette de prix, Marchés dirigés par les ordres, Asymétrie d'information, Coûts d'inventaire.

---

* Department of Finance, Facultés Universitaires Catholiques de Mons (FUCaM - Catholic University of Mons), Chaussée de Binche 151, 7000 Mons, Belgium. Phone: +32 65 32 33 34 and +32 65 32 33 36. Fax: +32 65 32 32 23. E-Mail: rudy.dewinne@fucam.ac.be and christophe.majois@fucam.ac.be. We wish to thank Alexis Cellier, Catherine D'Hondt, participants at the 8th Belgian Financial Research Forum, participants at the 20th AFFI Annual International Conference and two anonymous referees for their helpful comments and suggestions. All remaining errors are entirely ours.

INTRODUCTION

The bid-ask spread, i.e., the difference between the best quotes available at the buy and sell sides of the market, has been the center of numerous studies in recent years. This spread represents the cost that an investor would incur if he bought and sold at the same time one unit of a financial asset, what is generally known as a "round-trip trade". Alternatively, the spread represents the gain that a market maker – the liquidity supplier in a quote-driven market – would realize following a round-trip trade.

It is usually considered that the spread helps to cover three different costs:

- The first one, and historically the oldest identified in the literature, is the *order processing cost*; one can think for instance about different kinds of administrative costs linked to order execution.

- Some authors then argued that market makers use the spread to compensate for some unwanted inventory positions. Because of their role of liquidity suppliers, market makers are obliged to constantly post quotes and must be ready to play as counterparts for each trade; as a result, they bear an inventory risk due to positions away from their desired target level. The second spread component is thus the *inventory holding cost.* The relationship between the spread and inventory costs has been studied, among others, by Stoll (1978), Ho & Stoll (1981) and Amihud & Mendelson (1980).

- The third component is known as the *information asymmetry cost* or *adverse selection cost.* The basic idea, which was first developed by Bagehot (1971), is that a market maker always loses to informed traders, but recovers his losses with gains he earns from transactions with uninformed traders. The relationship between information asymmetry and the bid-ask spread has then been the object of numerous theoretical studies (see for instance Copeland & Galai (1983), Glosten & Milgrom (1985), Easley & O'Hara (1987)).

Since the mid 80's, several authors have attempted to empirically evaluate these various components of the bid-ask spread. The methods essentially vary in (*i*) the number of components they take into account (broadly speaking, some methods consider the existence of the inventory holding cost while the others do not) and (*ii*) the way they try to identify the considered components (a first class of methods is based on serial covariance properties of price changes, while a second class estimates the spread components thanks to regressions on a trade indicator variable).

These spread decomposition models have been largely used in empirical studies for various purposes. The examples we give hereunder are far from being exhaustive. By comparing the spread components between open outcry auction and automated order execution on the Sydney Futures Exchange, Wang (1999) showed that floor traders are better able to assess the presence of adverse information than screen traders. Saporta et al. (1999) proved that a reduction in the delay of reporting on the London Stock Exchange in January 1996 did not have any impact on the relative components of the spread, thereby suggesting that this change did not affect market liquidity. Menyah & Paudyal (2000) studied how the spread components are affected by stock liquidity on the London Stock Exchange. Dennis & Weston (2001) analyzed the relationship between the adverse selection component of the bid-ask spread and the ownership structure of firms and suggested that institutions and insiders were better informed than individual investors. Elder et al. (2004) consider the impact of tracking stocks on the adverse selection component of spreads posted by market makers.

All the papers we have cited here above focus on quote-driven markets. Indeed, the intuition that lies behind the existence of the spread components, as well as the models proposed to estimate them, have all been developed within the context of quote-driven markets, where a market maker is obliged to post quotes at which he must be prepared to trade. However, the existence of the bid-ask spread in an order-driven environment has been theoretically showed (Foucault (1999), Glosten (1994)), and most of the spread decomposition models have been applied in order-driven markets (e.g., Silva & Chavez (2002), Vandelanoite (2002), Declerck (2002)). The choice of the model is an important question, but we do not know any article that compares the results provided by various models within an order-driven environment. The idea that often guides the choice of the model to apply in an order-driven market is that the inventory holding component of the spread is not relevant because no market maker is committed to supply liquidity. While intuitively appealing, this hypothesis has nevertheless never been checked in the literature. Some features of order-driven markets, such as the presence of implicit market makers or what is known as "liquidity provider agreements", prevent us from definitely eliminating models that estimate an inventory holding component. Our objective in this paper is to compare the results provided by a wide range of spread decomposition models, including those that contain an inventory holding component, when they are applied in order-driven markets. Our will to achieve exhaustiveness in this particular environment is our main contribution to the literature.

A first step can consist in checking whether the estimates provided by the alternative models are plausible and consistent with each other. A second step can analyze if the components measure what they pretend to measure. Indeed, while the spread decomposition models are often used in the empirical literature, there are still some doubts about whether the adverse selection component of the spread is really a proxy for information asymmetry. Neal & Wheatley (1998), for instance, do not find any significant difference between adverse selection components estimated for closed-end funds and for a matched sample of common stocks. This result seems puzzling for the authors who predicted lower adverse selection for closed-end funds, because they report their net asset values weekly which thus eliminates uncertainty about their current liquidation value. According to Neal & Wheatley, this suggests either that adverse selection arises primarily from factors other than current liquidation value or that the empirical models are misspecified.

Van Ness et al. (2001) examine the performance of five spread decomposition models by comparing the adverse selection estimates to other measures of information asymmetry – volatility, analysts forecast errors, dispersion of analysts earnings forecasts, etc. – and informed trading – the number of analysts that follow the stock and the percentage of shares owned by institutions. As the adverse selection components appear unrelated to measures of uncertainty, the authors suggest that "*the adverse selection models measure adverse selection weakly at best*". Another explanation, that they consider as less likely, is the weak ability of their corporate finance variables to measure information asymmetry. Clarke & Shastri (2000) obtain more optimistic results using a similar approach. They find that adverse selection estimates tend to be related to firm characteristics that should be ex-ante associated with information asymmetry, and that monthly changes in these estimates are significantly correlated with annual changes in the corporate finance proxy variables.

In the same spirit as Van Ness et al. (2001) and Clarke & Shastri (2000), we want to assess the validity of spread decomposition models, but once again within the context of an order-driven market. We will compare the adverse selection estimates with other information asymmetry proxies derived from the limit order book.

To fulfill our objectives, we use intraday data for 19 stocks belonging to the Belgian BEL20 index traded on Euronext Brussels, which is a pure order-driven market that has rarely been analyzed until now.

We encounter some difficulties in applying both models that take the inventory holding cost into account. While Stoll's (1989) model produces estimates of the probability of trade reversal that seem inconsistent with our data set, Huang & Stoll's (1997) 3-way decomposition procedure provides values that are either insignificant or irrelevant. But these results are consistent with the hypothesis that liquidity providers in order-driven markets do not have to really manage their inventory as they are not obliged to trade.

The spread decomposition models that consider only adverse selection and order processing costs produce more reliable results. At the exception of George et al.'s (1991) method, the adverse selection components provided by these models are highly correlated with each other. In particular, we suggest that Huang & Stoll's (1997) 2-way decomposition, Lin et al.'s (1995) method and Madhavan et al.'s (1997) procedure – the latter to a lesser extent – are the "best" models to use in an order-driven market.

We then concentrate on adverse selection estimates provided by these three models and compare them with five information asymmetry proxies derived from market microstructure literature. We find that these estimates are positively correlated with only two of our proxies. This can indicate either that some of our information asymmetry proxies are not appropriate, or that adverse selection components do not perfectly measure information asymmetry. Results on this subject are thus inconclusive.

This paper is organized as follows. In Section 1, we provide a clear typology of the existing spread decomposition models. Section 2 discusses spread decomposition in order-driven environments, describes the microstructure of Euronext Brussels and presents some descriptive statistics of our sample. Spread components estimates given by the alternative models are provided in Section 3. Section 4 focuses on a comparison between adverse selection components and other information asymmetry proxies. Section 5 concludes.

1. Two Decades of Spread Decomposition Models

This section is devoted to a quick survey of the main spread decomposition models developed in the literature. For each model, we identify the components that are estimated as well as other more specific characteristics. As our goal here is to provide a clear typology of these models, we will stay at a general level, leaving the presentation of more detailed equations for Section 3 where we apply some of those methods.

Broadly speaking, there are two general classes of spread decomposition models: (*i*) the *covariance-based models* and (*ii*) the *trade indicator models.*

The covariance-based models find their origin in Roll's (1984) spread estimator. Under the assumption that the market maker faces only order processing cost, Roll proposes to estimate the quoted spread with the following formula :

$S = 2\sqrt{-COV(\Delta P_t, \Delta P_{t-1})}$, where P_t is the transaction price at time t.[1]

In Stoll's (1989) model, following the introduction of the inventory holding and adverse selection components, the *realized spread*, defined as the difference between the price at which a dealer sells at one point in time and the price at which he buys at an earlier point in time, can be inferior to the *quoted spread.* On the basis of two covariances, the covariance of price change and the covariance of quote change, his model provides an estimate of the realized spread expressed as a fraction of the quoted spread. The difference between the realized spread and the quoted spread is the adverse selection component. The realized spread itself can be decomposed into the order processing and the inventory holding costs. The spread components are in fact combinations of two other estimated parameters: π , the probability of trade reversal, and ∂, the magnitude of a price change expressed as a fraction of the spread.

The π parameter is particularly interesting. If a market maker is risk-averse, he will try to reduce his inventory risk: after a buy (sell) trade from an investor, the market maker, who does not want to depart too far from a pre-specified target level of inventory in the stock, will revise his bid and ask quotes upward (downward), in order to induce sell

[1] Alternatively, you can work with the natural log of the transaction price. In that case, ΔP_t is the continuously compounded rate of return and S is an estimate of the relative spread, i.e., the spread expressed as a fraction of the price.

(buy) transactions from the next investor. So the conditional probability that a particular trade is buyer- (seller-) initiated given that the previous one was seller- (buyer-) initiated is *higher* than 0.5. The presence of an inventory holding component is thus associated with a π value superior to 1/2.[2]

George et al. (1991) show that the spread estimates defined by Roll (1984) and Stoll (1989) are downward biased, because they do not take into account the time variation in expected returns. They propose a new approach that provides unbiased and efficient estimators of the spread and its components. Their approach uses a spread measure based on the serial covariance of the *difference* between transaction returns and returns using bid prices (which helps to eliminate the time-varying expected return). George et al. decompose the bid-ask spread in only two components: the order processing cost and the adverse selection cost.

In the covariance-based models of Stoll (1989) and George et al. (1991), two steps are necessary to obtain the spread components: (*i*) various serial covariances first need to be computed; (*ii*) these serial covariances are then regressed on the quoted spread.

In the trade indicator models, the spread components are estimated through a regression of price changes on a trade indicator variable, i.e., a variable that equals +1 if a trade is buyer-initiated and -1 if it is seller-initiated. George et al.'s (1991) model has been adapted by some authors (Neal & Wheatley (1998), Van Ness et al. (2001)) in order to obtain the spread components by directly regressing the difference between transaction returns and returns using bid prices on the trade indicator variable. This version of George et al. (1991) enters into the second category of spread decomposition models.

Glosten & Harris (1988) propose a model where the bid-ask spread is divided into order processing and adverse selection costs. The main originality of their approach is to consider these two components as linear combinations of trade size so that the spread itself can vary with trade size. After several trials, they argue that the best model specification is characterized by a constant order processing cost and an adverse selection cost that increases with trade size. This view is consistent with informed traders trading larger quantities in order to maximize the return of their private information (Easley & O'Hara (1987)).

Lin et al. (1995) also provide a 2-way decomposition model, with order processing and adverse selection costs. They have a slightly different approach, since their model is essentially based on the *effective spread*, i.e., the difference between transaction price and quote midpoint. Their model also provides a third parameter that

---

[2] Note that the unconditional probability that a given trade is buyer- or seller-initiated remains unchanged at 0.5.

represents the extent of order persistence. The various parameters are obtained through the non-simultaneous estimation of three different equations.

As in Glosten & Harris (1988) and Lin et al. (1995), the model of Madhavan et al. (1997) produces order processing and adverse selection spread components. In addition, their method provides an estimate for three other parameters: the probability a transaction takes place inside the spread, the probability of trade reversal and the first-order autocorrelation of the trade initiation variable.

Finally, Huang & Stoll (1997) propose a model which can encompass most of the previous ones: Roll (1984), Stoll (1989), George et al. (1991), Glosten & Harris (1988) and Madhavan et al. (1997). In the *2-way decomposition*, the first parameter is the sum of the adverse selection and inventory holding costs, while the second parameter represents the order processing component. In the *3-way decomposition*, Huang & Stoll further decompose the first parameter from the 2-way decomposition into its two components, by the use of the probability of trade reversal, once again supposed to be superior to 0.5.

The typology given in Table 1 provides a helpful summary of the models described in this section. It is interesting to notice that only two models estimate the three classical components of the bid-ask spread: Stoll (1989) and Huang & Stoll's (1997) 3-way decomposition. All the other models do not take the inventory holding cost into account.

We must highlight here that other spread components have been proposed in the literature next to the three traditional ones. One of them is the market maker *rent* or *profit markup*, resulting from the noncompetitive behavior of market makers (see for instance Levin & Wright (2004)). Another component is the *search cost* identified in an experimental inter-dealer market in Flood et al. (1998). We mention these components here, but we will not consider them in the rest of the paper. Both components are indeed marginally treated in the literature and they are clearly related to dealer markets, while we will focus on an order-driven market.

TABLE 1. TYPOLOGY OF THE MOST IMPORTANT SPREAD DECOMPOSITION MODELS

Model	*Model category*	*Order processing*	*Inventory holding*	*Adverse selection*	*Other estimated parameters*
Stoll (1989)	Covariance based	✓	✓	✓	π: probability of trade reversal ∂: magnitude of a price change (fraction of the spread)
George et al. (1991)	Covariance based	✓		✓	
George et al. (1991) adapted (e.g. Van Ness et al. (2001))	Trade indicator	✓		✓	
Glosten & Harris (1988)	Trade indicator	✓		✓	
Lin et al. (1995)	Trade indicator	✓		✓	Θ: reflects the extent of order persistence δ: (1+Θ)/2: probability of order persistence
Madhavan et al. (1997)	Trade indicator	✓		✓	λ: probability a transaction takes place inside the spread ρ: first-order autocorrelation of the trade indicator variable π: probability of trade reversal
Huang & Stoll (1997) 2-way decomposition	Trade indicator	✓	✓		
Huang & Stoll (1997) 3-way decomposition	Trade indicator	✓	✓	✓	π: probability of trade reversal

This table presents the main characteristics of 8 usual spread decomposition models. In the columns *Order processing*, *Inventory holding* and *Adverse selection*, a ✓ indicates that the model provides an estimate for this component of the bid-ask spread.

2. Euronext Specific Features for Spread Decomposition

2.1. Spread components estimation in an order-driven environment

Although Glosten (1994) suggests certain advantages for the order-driven environment, the superiority of the limit order book remains an open empirical issue. Indeed, while several papers evidence lower trading costs on order-driven markets (see for example Swan & Westerholm (2004)), other papers suggest that fully automated trading systems may not be able to replicate the benefits of human intermediation on a trading floor (Venkataraman (2001)).[3] Anyway, the growth of new Alternative Trading Systems, Nasdaq development of SuperMontage, and the choice of an order-driven environment by most European exchanges[4] make it more relevant to understand the behavior of bid-ask spreads observed on order-driven markets.

Although more and more papers[5] now focus on limit order trading, there is a need for additional research dealing with the components of the bid-ask spread. At this point in time, no spread decomposition model has been developed in the context of purely order-driven markets.

In an order-driven market, order processing and adverse selection costs[6] can affect order placement of limit order traders and it is thus natural to see these costs as bid-ask spread components. However inventory holding cost does not seem that relevant since no market maker is committed to supply liquidity. Let us notice that, in most order-driven markets, we observe the existence of implicit market makers who are simultaneously submitting buy and sell orders for their own account. Even if they are not committed to supply liquidity, these traders can face a deviation in their inventory position and adjust their order placement in such a way that the bid-ask spread can be enlarged. In addition, liquidity provider agreements sometimes exist in order-driven markets in order to enhance liquidity for low volume stocks.[7]

The features and behaviors observed in order-driven markets prevent us to automatically reject spread decomposition models including an inventory cost component, even if

---

[3] Let us notice that the author thinks that future developments could improve trading rules in order to better meet the requirements of a variety of market participants.

[4] Other exchanges around the world, such as the Australian Stock Exchange, the Stock Exchange of Hong Kong or the Toronto Stock Exchange, also work with a limit order book and without market makers.

[5] For example, Parlour (1998) and Foucault (1999) analyze the choice between limit order and market order. Handa et al. (2003) model quote setting in an order-driven environment and evidence several factors affecting the size of the spread.

[6] As shown by Glosten (1994), the existence of adverse selection costs generates positive bid-ask spreads in an order-driven market.

[7] A liquidity provider agreement is a contract between the stock exchange and a dealer. The latter commits himself to continuously quote a bid price and an ask price during the trading session. Actually, he has to submit buy and sell orders in order to guarantee that the spread does not exceed a given level. An example of these agreements is analyzed by Declerck & Hazard (2002). Analyzing 38 representative stock exchanges in the world, Swan & Westerholm (2004) show that European hybrid markets, where continuous trading occurs in a limit order book and dealers are present in medium to low liquidity stocks, have the lowest effective spreads.

intuition leads to think that this component is less important in an order-driven environment. So, in order to be exhaustive, all the models presented in Section 1 will be estimated in this paper.

Even if spread components estimation in order-driven markets did not lead to any dedicated theoretical proposition, several empirical investigations were conducted on these markets using models developed in the context of quote-driven markets.

Using Stoll's (1989) model, Silva & Chavez (2002) estimate the spread components on a sample including stocks from the NYSE and from the Mexican Stock Exchange (MSE). As the MSE is an order-driven market where liquidity suppliers have no affirmative obligation to provide liquidity, their expectation was that the inventory cost component should be lower on this market than on the NYSE. Their results show that order processing cost and inventory holding cost are not significantly different across exchanges but that adverse selection cost is much more important on the Mexican Stock Exchange. Ahn et al. (2002) estimate the spread components in the limit order book of the Tokyo Stock Exchange using Madhavan et al.'s (1997) model. According to the authors, this model has an attractive feature since it considers that inventory costs are of a less important concern for limit order traders. The same reasoning drove the choice made by Vandelanoite (2002) and Brockman & Chung (1999) to use Lin et al.'s (1995) model. While Vandelanoite (2002) examines the adverse selection cost around takeover announcements in the French market, Brockman & Chung (1999) analyze the bid-ask spread components on a purely order-driven market : the Hong Kong Stock Exchange. Another analysis of bid-ask spread components has been proposed by Declerck (2002) who used Huang & Stoll's (1997) model to estimate the three cost components.

2.2. Euronext Brussels microstructure

Euronext is the result of the merger of the Paris, Brussels and Amsterdam Stock Exchanges in September 2000.[8] These stock exchanges now operate through an homogeneous trading system, which is called NSC (Nouveau Système de cotation). This system is a continuous electronic order-driven system, where transactions result from the crossing of buy and sell orders placed by investors. Two kinds of orders are often used on Euronext: *(i) limit orders*, that specify a limit price above (under) which the investor is not ready to buy (sell); *(ii) market orders*, which do not specify any limit price, but are treated as limit orders presenting a limit price equal to the best price available at the opposite side of the market at the time they are submitted. Whenever an order is placed at a price which is equal to or better than the best price proposed at the opposite side of the market, a transaction takes place. Orders that are not immediately fully executed are recorded in an electronic limit order book; the system enforces price and time priority.

[8] Three markets have recently joined the three initial partners into Euronext: the Lisbon and Porto Stock Exchanges and the LIFFE.

Market participants can see at any time the five best limits at each side of the market, with the associated displayed depth.[9] Even if Euronext can be considered as a rather transparent market, participants can use hidden orders, i.e., orders whose some part of the quantity is not disclosed to other participants. This hidden quantity keeps price priority but loses time priority. Except for the most liquid stocks, Euronext authorities have sometimes an agreement with a liquidity provider who is then committed to supply liquidity to the market during the whole trading session. Actually, this liquidity provider has to submit buy and sell orders in order to guarantee that the spread does not exceed a given level.

The market opens at 9:00 AM with a call auction, which helps to determine an efficient opening price after the night interruption (see Biais et al. (1995)). This auction is preceded by a pre-opening period (from 7:15 AM) where all investors can place, modify or cancel orders without any transaction taking place. The market closes at 5:30 PM with another call auction preceded by a 5-minute pre-closing period in the same spirit as the pre-opening one. A "trading at last" period was recently introduced on Euronext Brussels and gives investors the opportunity to trade until 5:40 PM at the closing price, i.e., at the price determined by the auction mechanism at 5:30 PM.

2.3. Data construction and descriptive statistics

As Euronext Brussels is not very liquid compared with the French and Dutch market segments, we focus on the most traded Belgian stocks. The data used in this paper were given by Euronext authorities. They include all trades, quotes and orders for the companies belonging to the Belgian Bel20 index for the months of October, November and December 2002 (64 trading days).[10] The ID codes of Euronext members who placed the orders were also provided.[11]

Our initial data set contains a total of 524 198 transactions. This represents an average of approximately 431 transactions a day for each stock. We decide to exclude from our sample the so-called "applications", since they are in fact prearranged trades that do not result from the classical crossing of buy and sell orders in the order book (there were 2 853 applications in our sample). We also decide to drop transactions that emerged from the auction mechanisms at the beginning and at the end of the trading day, as well as "trading at last" transactions. Applying these rules lets us with a subset of 491 249 trades.

---

[9] Since December 1, 2003, participants have access to the whole limit order book.

[10] As GIB was replaced by Mobistar in the index composition during this period, we excluded both stocks from our sample. There are thus 19 stocks in our final sample.

[11] Actually, these ID codes are not sufficient to know the market members' identities but allow us to isolate the whole set of orders or trades associated to a given member from the other orders and trades of the sample. They also allow for an identification of the order generating a trade.

The next step in our data set construction is to sign transactions, i.e., to determine whether a transaction was buyer- or seller-initiated. This is necessary for trade indicator spread decomposition models. Most of the studies that need to sign transactions use the well-known Lee & Ready's (1991) algorithm based on the position of the transaction price relative to the prevailing bid-ask spread, but we did not follow this approach. Thanks to the ID codes, we can identify for each transaction both orders generating the trade. By comparing the time submission of each order, we can find the trade sign: a transaction is buyer- (seller-) initiated if the buy (sell) order was submitted after the sell (buy) order. In cases where both orders were submitted exactly at the same time, we arbitrarily assigned a trade indicator value. These cases represent only 0.7% of our sample of trades and should not bias our results.

Some studies aggregate all trades that are recorded at the same time, based on the hypothesis that these transactions result from a single order (see for instance Vandelanoite (2002)). Our data allow us to see whether these transactions actually emerge from a single order, so that our aggregation procedure is more accurate. For each of the aggregated transactions, we compute a new single price, which is the mean of all prices weighted by the number of shares traded at each price. This aggregation leaves us with a total of 332 121 transactions.

The last step in our data set construction is to associate to each transaction the prevailing quoted bid-ask spread. We therefore rebuild the limit order book, following the methodology developed by De Winne & D'Hondt (2004). This provides us with a comprehensive data set, giving for each state of the book detailed information such as the five best limits with the associated displayed and hidden depths. If different events affect the order book within the same second, we only keep the last state in our rebuilt order book file.

Table 2 provides some descriptive statistics based on our final sample including 332 121 transactions. In the upper panel, we find 12 stocks for which liquidity is exclusively supplied by limit order traders. For the other 7 stocks in the lower panel, a liquidity provider is committed to supply additional liquidity. Globally, there exists an important disparity in terms of liquidity across the 19 stocks of our sample, as the two most traded companies (Fortis and Dexia) represent 35.72% of the total number of trades, while the three less traded companies (IBA, D'Ieteren and Tessenderlo) each account for only 1 % of that amount. The mean daily number of transactions ranges from 55.09 (IBA) to 996.89 (Fortis), with a sample average of 273.13. For some companies like Umicore, D'Ieteren and IBA, the minimum number of transactions per day can be very low. As expected, stocks with a liquidity provider are less traded than the other ones, but there are three exceptions. IBA and Barco, despite having no designated liquidity provider, are not heavily traded. On the other hand, we may wonder why a liquidity provider has been affected to Omega Pharma.

On average, the quoted spread is equal to 0.1666 euros, with values ranging from 0.0187 to 0.807 euros. This quoted spread represents on average 0.42% of the mid-quote. We can highlight a negative relationship between spread and trading volume, as the companies presenting the highest relative spread also have the lowest mean number of transactions (e.g., Tessenderlo, D'Ieteren, IBA, Barco). For all stocks, the effective spread[12] is higher than the quoted spread. This result is a direct and logical consequence of the aggregation procedure: lots of the aggregated transactions occur at a (new) price which is outside the quoted spread

3. THE COMPONENTS OF THE BID-ASK SPREAD ON EURONEXT BRUSSELS

We now turn to the spread components estimates. In addition to the set of cost components estimated through a model, another important feature of the model is the assumption about the probability of trade reversal. While Lin et al.'s (1995) model is based on the idea that buyer- (seller-) initiated trades tend to follow buyer- (seller-) initiated trades, Huang & Stoll (1997) assume a probability of trade reversal higher than 0.5. For each stock, we provide in Table 2 an estimate for the trade reversal probability, simply computed as the proportion of transactions that represent a reversal compared to the previous one (a buy that follows a sell or a sell that follows a buy). For all stocks in our sample, this probability is *lower* than 0.5, showing a clear tendency toward *trade continuation*. No difference appears between the stocks with and without a liquidity provider. These results tend to support the hypothesis that the inventory holding component is not relevant in an order-driven market. However, as our goal is to achieve exhaustiveness, we will use all the models presented in Section 1.

We will begin with models that estimate all three parameters. We will then present the results for the 2-component models, that we will group according to their similarities in the estimation procedures.

[12] The effective spread is measured as $2\ |P_t - M_t|$ where M_t is the mid-quote prevailing before a transaction occurring at a price P_t.

TABLE 2. DESCRIPTIVE STATISTICS

Stock	*Total # trades*	*Mean # trades*	*Min # trades*	*Mean volume (1000 euros)*	*Mean # shares (1000 shares)*	*Price (euros)*	*Quoted spread (euros)*	*Relative spread (%)*	*Effective spread (euros)*	*Probability of trade reversal*
SOLVAY	14168	221.38	24	4223.01	67.88	62.49	0.1355	0.22	0.1455	0.3551
DELHAIZE GROUP	26349	411.70	37	4558.75	257.83	17.72	0.0596	0.34	0.0644	0.3723
KBC	21619	337.80	71	6133.27	193.25	31.76	0.0698	0.22	0.0746	0.3498
ELECTRABEL	18959	296.23	72	9397.28	39.50	237.81	0.2726	0.11	0.2936	0.3844
DEXIA	54818	856.53	131	14772.82	1363.37	10.91	0.0187	0.17	0.0206	0.3725
UCB	19028	297.31	68	4916.22	185.19	26.58	0.0830	0.31	0.0947	0.3312
FORTIS	63801	996.89	145	22169.16	1296.96	16.98	0.0236	0.14	0.0261	0.3469
AGFA-GEVAERT	17190	268.59	55	2799.50	143.36	19.63	0.0454	0.23	0.0482	0.3172
IBA	3526	55.09	9	99.83	18.31	5.36	0.0714	1.39	0.0808	0.3798
BARCO (NEW)	4331	67.67	17	538.69	11.71	45.86	0.2985	0.66	0.3162	0.3349
INTERBREW	27587	431.05	73	10614.58	483.57	21.98	0.0499	0.23	0.0540	0.3574
GBL	11859	185.30	34	3714.41	96.63	38.80	0.1610	0.42	0.1681	0.3276
TESSENDERLO	3648	57.00	24	447.55	16.43	27.18	0.2020	0.76	0.2126	0.3418
UMICORE	5492	85.81	6	824.52	21.01	39.43	0.1581	0.41	0.1646	0.3261
D'IETEREN	3606	56.34	8	348.32	2.65	131.32	0.8070	0.62	0.8347	0.3137
ALMANIJ	9870	154.22	24	1958.15	58.42	33.50	0.1333	0.40	0.1408	0.3227
COLRUYT	7129	111.39	16	1064.15	20.97	50.62	0.2206	0.44	0.2373	0.3128
BEKAERT	4133	64.58	21	643.58	15.97	40.37	0.2224	0.56	0.2296	0.3559
OMEGA PHARMA	15008	234.50	43	2836.93	88.11	32.92	0.1337	0.42	0.1743	0.3400
ALL	332121	273.13	6	4845.30	230.59	46.91	0.1666	0.42	0.1779	0.3443

This table provides statistics based on our final sample of 332 121 transactions for the 19 stocks belonging to the Belgian BEL20 index and for the months of October, November and December 2002. Stocks in italic in the lower panel are those for which there exists a liquidity provider. *Total # trades* is the total number of transactions in the final sample. The columns *Mean # trades, Mean volume* and *Mean # shares* are daily averages. *Min # trades* is the daily minimum number of transactions in our period. *Price, Quoted spread, Relative spread* and *Effective spread* are averages over the sample period. *Probability of trade reversal* is the ratio of trades that represent a reversal compared to the previous trade (a buy following a sell and a sell following a buy) to the total number of trades.

3.1. Stoll's (1989) model

Defining P_t and B_t as the transaction price and the quoted bid price at time t, Stoll (1989) first defines the following serial covariances[13]:

$$COV_T = COV(\Delta P_t, \Delta P_{t-1}) \quad (1)$$

$$COV_Q = COV(\Delta B_t, \Delta B_{t-1}) \quad (2)$$

Both covariances are then regressed against the square of the quoted spread S :

$$COV_T = a_o + a_1 S^2 + u \quad (3)$$

$$COV_Q = b_o + b_1 S^2 + v \quad (4)$$

where u ad v are random errors. Under the assumption of market efficiency, parameters a_0 and b_0 should equal 0. The estimated parameters a_1 and b_1 are then related to two other essential parameters in the model, π (probability of price reversal) and ∂ (magnitude of a price change expressed as a fraction of the spread), by the equations hereunder:

$$a_1 = \partial^2(1-2\pi)-\pi^2(1-2\partial) \quad (5)$$

$$b_1 = \partial^2(1-2\pi) \quad (6)$$

The realized spread (expressed as a proportion of the quoted spread) is then measured as $2(\pi-\partial)$. It is the sum of the order processing (γ) and inventory holding (β) costs. The difference between the quoted spread and the realized spread is the adverse selection cost α[14]:

$$\alpha = 1-2(\pi - \partial) \quad (7)$$

$$\beta = 2(\pi - 0.5) \quad (8)$$

$$\gamma = 2(0.5-\partial) \quad (9)$$

---

[13] COV_Q can alternatively be calculated with ask prices A_t instead of bid prices. Results should normally be equivalent under the assumption of a constant spread.

[14] In the rest of this paper, we will keep letters α , β and γ to designate respectively the adverse selection, inventory holding and order processing components expressed as a proportion of the spread. In the same aim of coherence, we will always use the same notation for the same concepts. The logical consequence is that our notation can differ from the one presented in the original papers.

Stoll (1989) uses data from Nasdaq and for the months of October, November and December 1984. He provides estimates of the spread components globally for all stocks in his sample. With $\pi = 0.550$ and $\partial = 0.265$, the realized spread represents 57% of the quoted spread and the spread components are: $\alpha = 0.43$, $\beta = 0.10$, and $\gamma = 0.47$.

In contrast to Stoll (1989), we want to provide estimates of the spread components for each of the 19 stocks of our sample. We first calculate COV_T , COV_Q and S for each stock and for each day (this needs working with intraday prices and quotes). For each stock, we then estimate Equations (3) and (4), using for each regression 64 observations corresponding to the 64 trading days in our sample.[15] This provides for each stock the parameters a_1 and b_1 necessary to compute the spread components (Equations (5) to (9)).

Let's first concentrate on results presented in Table 3, where COV_Q is calculated with bid quotes. We note that, consistent with the hypothesis of efficient markets, most of a_0 and b_0 estimates are not significantly different from 0. The parameter a_1 is almost always significant at the 1% level, while lots of b_1 values are significant either at the 1 or 5% level (these results seem consistent with Stoll's (1989) ones). We think the most important result lies in values taken by the probability of trade reversal, π , which is higher than 0.5 for all stocks. This is inconsistent with the results we have presented in Table 2 where we highlighted a clear tendency of trade continuation, but this is coherent with the intuition of the model. Of course, the higher the π parameter, the higher the inventory holding component β . Stocks with a liquidity provider do not seem to exhibit an inventory holding part superior than for other stocks. Some parameters present inconsistent values: α is negative for Electrabel and Dexia, which consequently present a realized spread superior to the quoted spread; γ is also negative for GBL. The adverse selection component α lies in a wide range, between 10.3% (Delhaize Group) and 92.14% (D'Ieteren).

---

[15] 63 observations are used in the case of Umicore, as it is impossible to calculate COV_T and COV_Q for one day in our sample due to a too small number of trades – see Table 2.

TABLE 3. ESTIMATES FROM STOLL'S (1989) MODEL - USING BID QUOTES

Stock	a_0	a_1	b_0	b_1	∂	*Probability of trade reversal* π	*Adverse selection* α	*Inventory holding* β	*Order processing* γ	*Realized spread*
SOLVAY	0.00011	-0.2106***	0.00011	-0.0682***	0.3721	0.7462	0.2519	0.4923	0.2558	0.7481
DELHAIZE GROUP	0.00011	-0.2599***	0.00011	-0.0936***	0.3784	0.8269	0.1030	0.6538	0.2432	0.8970
KBC	0.00019	-0.2567***	-0.00022	-0.0555***	0.3262	0.7608	0.1308	0.5216	0.3475	0.8692
ELECTRABEL	0.00670***	-0.3458***	0.00069	-0.0820***	0.3291	0.8785	-0.0988	0.7570	0.3417	1.0988
DEXIA	0.00004***	-0.3144***	-0.00001**	-0.0352***	0.2594	0.7616	-0.0045	0.5232	0.4813	1.0045
UCB	-0.00030	-0.1274***	0.00018	-0.1114***	0.4852	0.7365	0.4974	0.4731	0.0295	0.5026
FORTIS	-0.00002*	-0.1459***	0.00002***	-0.0863***	0.4426	0.7203	0.4445	0.4407	0.1148	0.5555
AGFA-GEVAERT	-0.00002	-0.1490***	-0.00012	-0.0327**	0.3539	0.6307	0.4463	0.2614	0.2923	0.5537
IBA	-0.00011	-0.0620***	-0.00019**	-0.0055	0.3946	0.5176	0.7539	0.0353	0.2108	0.2461
BARCO (NEW)	-0.00266	-0.0780***	-0.00047	-0.0281***	0.4252	0.5777	0.6950	0.1554	0.1496	0.3050
INTERBREW	-0.00004	-0.2057***	-0.00008	-0.0489***	0.3433	0.7074	0.2718	0.4148	0.3134	0.7282
GBL	-0.00140*	-0.0774***	0.00211*	-0.1591***	0.5744	0.7412	0.6664	0.4824	-0.1487	0.3336
TESSENDERLO	-0.00147	-0.0976***	0.00180*	-0.0914***	0.4934	0.6877	0.6114	0.3754	0.0132	0.3886
UMICORE	0.00121	-0.1648***	-0.00182*	-0.0147	0.2851	0.5907	0.3886	0.1815	0.4299	0.6114
D'IETEREN	-0.05735***	-0.0197	-0.01722**	-0.0017	0.4646	0.5040	0.9214	0.0079	0.0707	0.0786
ALMANIJ	-0.00122*	-0.1050***	-0.00036	-0.0258***	0.3850	0.5870	0.5960	0.1740	0.2301	0.4040
COLRUYT	-0.00371**	-0.0948***	-0.00098	-0.0355*	0.4180	0.6015	0.6331	0.2030	0.1640	0.3669
BEKAERT	-0.00257	-0.1050***	-0.00128	-0.0328**	0.4005	0.6022	0.5965	0.2045	0.1990	0.4035
OMEGA PHARMA	0.00085	-0.1883***	-0.00051	-0.0268***	0.3052	0.6438	0.3228	0.2875	0.3897	0.6772

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from Stoll's (1989) model. Stocks in italic in the lower panel are those for which there exists a liquidity provider.

Parameters a_0, a_1, b_0 and b_1 are estimated from: $COV_T = a_0 + a_1 S^2 + u$ and $COV_Q = b_0 + b_1 S^2 + v$, where COV_T and COV_Q are the serial covariances in transaction price changes and bid quotes changes. S is the quoted spread and u and v are random errors.

Parameters ∂ and π are the solutions of: $a_1 = \partial^2(1-2\pi) - \pi^2(1-2\partial)$ and $b_1 = \partial^2(1-2\pi)$

α, β and γ are then computed as: $\alpha = 1-2(\pi-\partial)$, $\beta = 2(\pi-0.5)$, $\gamma = 2(0.5-\partial)$

Realized spread is the sum of the β and γ columns.

The marks ***, ** and * represent significance at respectively the 1, 5 and 10% level.

Table 4 provides the results where we have used ask quotes to compute COV_Q . We once again observe some inconsistent estimates: α is still negative for Electrabel and Dexia, and β is negative for Dexia - because the π parameter is inferior to 0.5. If the realized spread, and thus the parameter α , is of the same magnitude as in Table 3, there are some important differences in the inventory holding and order processing components. Let's take the example of Delhaize Group, that presents in both tables a realized spread of approximately 0.9, but where the parameter β (γ) shows a value of 0.65 (0.24) with bid quotes and of 0.36 (0.52) with ask quotes.

TABLE 4. ESTIMATES FROM STOLL'S (1989) MODEL - USING ASK QUOTES

Stock	a_0	a_1	b_0	b_1	∂	*Probability of trade reversal* π	*Adverse selection* α	*Inventory holding* β	*Order processing* γ	*Realized spread*
SOLVAY	0.00011	-0.2106***	-0.00057*	-0.0344***	0.3092	0.6796	0.2592	0.3593	0.3815	0.7408
DELHAIZE GROUP	0.00011	-0.2599***	-0.00017***	-0.0210*	0.2417	0.6800	0.1233	0.3600	0.5167	0.8767
KBC	0.00019	-0.2567***	0.00004	-0.0759***	0.3576	0.7968	0.1217	0.5936	0.2847	0.8783
ELECTRABEL	0.00670***	-0.3458***	0.00056	-0.0675***	0.3090	0.8534	-0.0889	0.7068	0.3821	1.0889
DEXIA	0.00004***	-0.3144***	-0.00003***	0.0036	-0.2319	0.4661	-0.3960	-0.0678	1.4638	1.3960
UCB	-0.00030	-0.1274***	-0.00011	-0.0642***	0.4303	0.6733	0.5138	0.3467	0.1395	0.4862
FORTIS	-0.00002*	-0.1459***	-0.00001	-0.0326***	0.3566	0.6283	0.4565	0.2567	0.2868	0.5435
AGFA-GEVAERT	-0.00002	-0.1490***	-0.00003	-0.0408***	0.3712	0.6481	0.4463	0.2961	0.2576	0.5537
IBA	-0.00011	-0.0620***	-0.00017*	-0.0227**	0.4372	0.5594	0.7557	0.1187	0.1255	0.2443
BARCO (NEW)	-0.00266	-0.0780***	-0.00336**	0.0000	0.3439	0.5000	0.6879	-0.0000	0.3122	0.3121
INTERBREW	-0.00004	-0.2057***	0.00009	-0.1110***	0.4268	0.8046	0.2444	0.6092	0.1463	0.7556
GBL	-0.00140*	-0.0774***	-0.00067	-0.0629**	0.4820	0.6353	0.6933	0.2707	0.0360	0.3067
TESSENDERLO	-0.00147	-0.0976***	0.00052	-0.0529***	0.4443	0.6339	0.6208	0.2679	0.1113	0.3792
UMICORE	0.00121	-0.1648***	0.00039	-0.1043***	0.4476	0.7601	0.3750	0.5203	0.1047	0.6250
D'IETEREN	-0.05735***	-0.0197	-0.01999*	-0.0078	0.4779	0.5172	0.9214	0.0343	0.0442	0.0786
ALMANIJ	-0.00122*	-0.1050***	-0.00051*	-0.0280***	0.3901	0.5920	0.5961	0.1841	0.2198	0.4039
COLRUYT	-0.00371**	-0.0948***	0.00077	-0.0706***	0.4721	0.6584	0.6275	0.3168	0.0558	0.3725
BEKAERT	-0.00257	-0.1050***	-0.00087	-0.0585***	0.4447	0.6480	0.5934	0.2960	0.1107	0.4066
OMEGA PHARMA	0.00085	-0.1883***	0.00035	-0.0820***	0.4054	0.7494	0.3119	0.4989	0.1892	0.6881

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from Stoll's (1989) model. Stocks in italic in the lower panel are those for which there exists a liquidity provider.
Parameters a_0, a_1, b_0 and b_1 are estimated from: $COV_T = a_0+a_1S^2+u$ and $COV_Q = b_0+b_1S^2+v$, where COV_T and COV_Q are the serial covariances in transaction price changes and ask quotes changes, S is the quoted spread and u and v are random errors.
Parameters ∂ and π are the solutions of: $a_1 = \partial^2(1-2\pi)-\pi^2(1-2\partial)$ and $b_1 = \partial^2(1-2\pi)$
α, β and γ are then computed as: $\alpha = 1-2(\pi-\partial)$, $\beta = 2(\pi-0.5)$, $\gamma = 2(0.5-\partial)$
Realized spread is the sum of the β and γ columns.
The marks ***, ** and * represent significance at respectively the 1, 5 and 10% level.

So we do not think that the results of Stoll's (1989) procedure are very relevant, as they can vary depending on the side of the quote used, and because they show a tendency of trade reversal where there is in fact a trend toward trade continuation. We however do not rule out the possibility that part of the inconsistency in our results is driven by the relatively low number of observations used in the estimation of regressions (3) and (4).

3.2. Huang & Stoll's (1997) 3-way decomposition model

We now turn to the second "complete" spread decomposition model. Defining M_t as the quote midpoint at time t ($M_t = (B_t+A_t)/2$), S_t as the posted spread at time t and Q_t as the trade indicator variable (equal to +1 if a transaction is buyer-initiated and -1 if it is seller-initiated), Huang & Stoll (1997) measure the three components of the spread α, β and $\gamma = 1-\alpha-\beta$ thanks to the simultaneous estimation of the two equations hereunder:

$$E(Q_{t-1}|Q_{t-2}) = (1-2\pi)Q_{t-2} \tag{10}$$

$$\Delta M_t = (\alpha+\beta)\frac{S_{t-1}}{2}Q_{t-1} - \alpha(1-2\pi)\frac{S_{t-2}}{2}Q_{t-2} + e_t \tag{11}$$

where e_t is an error term.

We use a GMM procedure in order to provide for each stock an estimate for all parameters present in the above model. The results are given in Table 5. They once again do not seem very consistent, as 15 out of the 19 α estimates are negative. Moreover, six α values, including three positive ones, are not significantly different from 0. It is striking to note that these six values concern stocks for which there exists a liquidity provider, but we cannot find any plausible reason to explain this finding. The inventory holding component of the spread is however not systematically higher for these stocks.

TABLE 5. ESTIMATES FROM HUANG & STOLL'S (1997) 3-WAY DECOMPOSITION MODEL

Stock	*Probability of trade reversal* π	*Adverse selection* α	*Inventory holding* β	*Order processing* γ
SOLVAY	0.3569***	-0.6463***	0.2210***	1.4253
DELHAIZE GROUP	0.3728***	-1.0394***	0.2879***	1.7515
KBC	0.3502***	-0.9204***	0.2576***	1.6628
ELECTRABEL	0.3840***	-0.9528***	0.2717***	1.6811
DEXIA	0.3728***	-0.7256***	0.3614***	1.3642
UCB	0.3326***	-0.8144***	0.2309***	1.5835
FORTIS	0.3466***	-0.7614***	0.5731***	1.1884
AGFA-GEVAERT	0.3172***	-2.6937***	0.5440***	3.1498
IBA	0.3826***	-1.1439***	0.2714***	1.8724
BARCO (NEW)	0.3364***	-1.0334***	0.2955***	1.7379
INTERBREW	0.3581***	-1.3797***	0.4171***	1.9627
GBL	0.3278***	0.5509**	0.1826***	0.2665
TESSENDERLO	0.3459***	-6.8909	1.1636	6.7273
UMICORE	0.3263***	0.4350	0.1493**	0.4157
D'IETEREN	0.3185***	0.0469	0.2886***	0.6645
ALMANIJ	0.3237***	-0.0473	0.2398***	0.8075
COLRUYT	0.3141***	0.0671	0.1530***	0.7799
BEKAERT	0.3547***	-0.1126	0.2670***	0.8456
OMEGA PHARMA	0.3393***	-1.9049***	0.0856*	2.8193

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from Huang & Stoll's (1997) 3-way decomposition model. Stocks in italic in the lower panel are those for which there exists a liquidity provider.
Parameters π, α and β are obtained through the simultaneous estimation of the following two equations:

$$E(Q_{t-1} \mid Q_{t-2}) = (1-2\pi)Q_{t-2}$$

$$\Delta M_t = (\alpha+\beta)\frac{S_{t-1}}{2}Q_{t-1} - \alpha(1-2\pi)\frac{S_{t-2}}{2}Q_{t-2} + e_t$$

where Q_t is a trade indicator variable (+1 if a trade is buyer-initiated and -1 if it is seller-initiated), M_t is the quote midpoint, S_t is the quoted spread and e_t is an error term.
γ is calculated as 1-α-β.
The marks ***, ** and * represent significance at respectively the 1, 5 and 10% level

Estimation problems were also encountered by Huang & Stoll (1997), as well as by other authors who applied this model (e.g., Van Ness et al. (2001), De Winne & Platten (2003)) and are thus not specific to order-driven markets, as all these authors worked on dealer markets.[16] The origin of these bad results lies in the low value of π, which is lower than 0.5 for all stocks and is thus not coherent with the intuition of the

[16] Declerck (2002) does not highlight the problem when she applies Huang & Stoll's (1997) 3-way decomposition to 37 stocks from the CAC40 index traded at the Paris Bourse. We only know that, in her results, the adverse selection component ranges from -0.24 to 0.43 with an average of 0.10. We do not have any information about the number of irrelevant values she obtained.

model (as explained in Section 1). However, the estimates of π provided by Huang & Stoll's (1997) 3-way decomposition are both highly significant and highly correlated with our own estimates of the probability of trade reversal given in Table 2 (the correlation coefficient is almost 1). These estimates are thus consistent with our data set.

Huang & Stoll (1997) argue that a possible source of the positive serial covariance in trade flows is that orders are broken up as they are executed. Orders could be negatively serially correlated as their theory suggests, but the trades are positively serially correlated. Their approach to deal with that problem is to bunch related data, i.e., to treat a cluster of trades at the same price and unchanged quotes as a single order. This procedure does not have any sense in our case, as we already aggregated trades that happen at the same time and result from a single order.

In a recent paper, Henker & Wang (2004) showed that there is a misspecification in Huang & Stoll's (1997) 3-way decomposition, which is the origin of the low or even negative values often found for α in many studies. They propose a correction for the model, and show that this correction indeed better performs than the original model.[17] However, most of the stocks in their sample exhibit a π value higher than 0.5, and the authors do not present results for the stocks where π is inferior to 0.515. So their correction is not expected to provide any improvement with our data. We applied the corrected model and indeed found results qualitatively similar to those provided in Table 5.[18]

3.3. George et al.'s (1991) model

This model is based on the difference between transaction price change and quoted bid change: $RD_t = \Delta P_t - \Delta B_T$ (where the subscript T indicates that the quoted bid change is measured on the basis of the bid price immediately following the transaction at time t). Their estimator of the realized spread, that we denote S_R, is very similar in appearance to Roll's (1984) estimator:

$$S_R = 2\sqrt{-COV(RD_t, RD_{t-1})} \tag{12}$$

The following regression provides an estimate for the order processing cost γ :

$$S_R = c + \gamma S + \epsilon \tag{13}$$

where S is the quoted spread, c must normally be equal to 0 and ϵ is the error term. The adverse selection cost is $\alpha = 1-\gamma$.

We first compute S_R and S for each stock and for each day. We then estimate Equation (13) for each stock. Results are provided in Table 6. We see that most of c values are not

[17] For instance, for S&P500 stocks listed on the NYSE during the year 1998, only 28.86% of adverse selection parameter estimates are larger than zero with the original model, while 65.67% are positive with the corrected specification.

[18] Results are not presented here but are available from the authors upon request.

significantly different from 0 while all γ estimates are highly significant (but six of them present incoherent values, i.e., values superior to 1). The values of γ range from 0.4197 to 1.5043, and most of them are relatively high, 13 being superior to 0.7; it thus seems that the order processing cost is the most important part of the bid-ask spread. The results provided by that procedure are globally not very consistent. However, we must be cautious in the interpretation, as the maximum number of observations in estimating Equation (13) for each stock is equal to 64.[19]

TABLE 6. ESTIMATES FROM GEORGE ET AL.'S (1991) MODEL

Stock	*c*	*Order processing γ*	*Adverse selection α*
SOLVAY	0.0141	1.0220***	-0.0220
DELHAIZE GROUP	-0.0021	1.1493***	-0.1493
KBC	0.0024	1.1338***	-0.1338
ELECTRABEL	-0.0808***	1.5043***	-0.5043
DEXIA	-0.0022	1.2203***	-0.2203
UCB	0.0161*	0.8636***	0.1364
FORTIS	0.0026	1.1082***	-0.1082
AGFA-GEVAERT	0.0089**	0.8305***	0.1695
IBA	0.0264***	0.4197***	0.5803
BARCO (NEW)	0.0501*	0.6565***	0.3435
INTERBREW	0.0158**	0.8473***	0.1527
GBL	0.0445*	0.5845***	0.4155
TESSENDERLO	0.0333	0.6932***	0.3068
UMICORE	0.0059	0.8910***	0.1090
D'IETEREN	0.2152***	0.5187***	0.4813
ALMANIJ	0.0156	0.8228***	0.1772
COLRUYT	0.0435	0.7569***	0.2431
BEKAERT	0.0577**	0.6447***	0.3553
OMEGA PHARMA	0.0078	0.8740***	0.1260

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from George et al.'s (1991) model. Stocks in italic in the lower panel are those for which there exists a liquidity provider.

Parameters c and γ are obtained through the estimation of the following equation: $S_R = c+\gamma S+\epsilon$ where $SR=2\sqrt{-COV(RD_t, RD_{t-1})}$, S is the quoted spread and ϵ is the error term. $RD_t = \Delta P_t - \Delta B_T$ (where the subscript T indicates that the quoted bid change is measured on the basis of the bid price immediately following the transaction at time t).
α is calculated as $1-\gamma$.
The marks ***, ** and * represent significance at respectively the 1, 5 and 10% level.

[19] It is sometimes impossible to compute $COV(RD_t, RD_{t-1})$ due to a lack of transactions. It can also be impossible to compute S_R when $COV(RD_t, RD_{t-1})$ is positive. The lowest number of observations used for an estimation equals 54.

3.4. Glosten & Harris' (1988) model

The main assumption of Glosten & Harris' (1988) model is that the spread components are linear functions of the transaction size V_t.[20] As a consequence, the spread can vary through time. Define $Z = z_0 + z_1 V_t$ as one-half of the adverse selection component and $C = c_0 + c_1 V_t$ as one-half of the order processing component, both expressed in monetary terms. The four parameters z_0, z_1, c_0 and c_1 are estimated on the basis of the following regression:

$$\Delta P_t = c_0 \Delta Q_t + c_1 \Delta(Q_t V_t) + z_0 Q_t + z_1 Q_t V_t + \epsilon_t \tag{14}$$

where ϵ_t is an error term.

Defining $\overline{V}$ as the average transaction size, the spread is estimated by:

$$S = 2(c_0 + c_1 \overline{V}) + 2(z_0 + z_1 \overline{V}) \tag{15}$$

while the adverse selection and order processing components expressed as a fraction of the spread are given by:

$$\alpha = \frac{2(z_0 + z_1\overline{V})}{2(c_0 + c_1\overline{V}) + 2(z_0 + z_1\overline{V})} \tag{16}$$

$$\gamma = \frac{2(c_0 + c_1\overline{V})}{2(c_0 + c_1\overline{V}) + 2(z_0 + z_1\overline{V})} = 1 - \alpha \tag{17}$$

Glosten & Harris (1988) find that the best model specification is to consider that the order processing cost is constant ($C = c_0$) and that the adverse selection cost varies with transaction size but without constant term ($Z = z_1 V_t$). In this paper, we decide to estimate two model specifications: (1) the general model with all parameters and (2) the constrained model as advised by Glosten & Harris (1988). Results are provided in Tables 7 and 8 respectively.

[20] Transaction size is the number of shares traded per transaction.

TABLE 7. ESTIMATES FROM GLOSTEN & HARRIS' (1988) MODEL (1)

Stock	c_0	c_1	z_0	z_1	*Average transaction size* $\overline{V}$	*Adverse selection* α	*Order processing* γ	*Estimated spread* S	*Quoted spread*
SOLVAY	-0.0096	0.000180	-0.0272**	0.000147***	306	0.2818	0.7182	0.1268	0.1355
DELHAIZE GROUP	0.0800	-0.000102	-0.0042	0.000020***	626	0.3401	0.6599	0.0487	0.0596
KBC	0.0229***	-0.000002	0.0077***	0.000004	572	0.3094	0.6906	0.0627	0.0698
ELECTRABEL	0.1197	-0.000195	0.0504	-0.000080	133	0.2979	0.7021	0.2669	0.2726
DEXIA	0.0134***	-0.000005**	0.0002	0.000002***	1591	0.4547	0.5453	0.0179	0.0187
UCB	0.0212***	-0.000000*	0.0118***	-0.000000**	622	0.3558	0.6442	0.0653	0.0830
FORTIS	-0.1072	0.000090	-0.0006	0.000004	1301	0.2955	0.7045	0.0281	0.0236
AGFA-GEVAERT	0.0144***	-0.000001	0.0031*	0.000003	533	0.2525	0.7475	0.0375	0.0454
IBA	0.0179***	-0.000003	0.0040***	0.000005	332	0.2523	0.7477	0.0449	0.0714
BARCO (NEW)	0.1186***	-0.000183***	0.0070	0.000144**	173	0.2685	0.7315	0.2378	0.2985
INTERBREW	0.0283***	-0.000012**	-0.0008	0.000008*	1121	0.3680	0.6320	0.0454	0.0499
GBL	0.0305*	0.000021	0.0210	-0.000008	521	0.2897	0.7103	0.1162	0.1610
TESSENDERLO	0.0685***	-0.000032	0.0093	0.000044	288	0.2699	0.7301	0.1621	0.2020
UMICORE	0.0409***	0.000014	0.0203***	-0.000012	244	0.2825	0.7175	0.1234	0.1581
D'IETEREN	0.2052***	0.000148	0.1280***	-0.000044	47	0.3724	0.6276	0.6760	0.8070
ALMANIJ	0.0401***	0.000003	0.0184**	-0.000010	378	0.2615	0.7385	0.1114	0.1333
COLRUYT	0.0743***	0.000023*	0.0223***	0.000012	188	0.2376	0.7624	0.2066	0.2206
BEKAERT	0.0819*	-0.000076	0.0069	0.000082	247	0.3015	0.6985	0.1807	0.2224
OMEGA PHARMA	0.0610***	-0.000044***	-0.0056**	0.000047***	375	0.2123	0.7877	0.1131	0.1337

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from Glosten & Harris' (1988) model – version (1). Stocks in italic in the lower panel are those for which there exists a liquidity provider.

Parameters c_0, c_1, z_0 and z_1 are obtained through the estimation of equation:

$\Delta P_t = c_0\Delta Q_t + c_1\Delta(Q_t V_t) + z_0 Q_t + z_1 Q_t V_t + \epsilon_t$

where P_t is the transaction price, Q_t the trade indicator variable (+1 if a trade is buyer-initiated and -1 if it is seller-initiated), V_t the transaction size in number of shares and ϵ_t is an error term.

$\overline{V}$ is the average transaction size. α and γ are computed as follows:

$$\alpha = \frac{2(z_0+z_1\overline{V})}{2(c_0+c_1\overline{V})+2(z_0+z_1\overline{V})} \quad \text{and} \quad \gamma = \frac{2(c_0+c_1\overline{V})}{2(c_0+c_1\overline{V})+2(z_0+z_1\overline{V})} = 1-\alpha$$

The estimated spread S is given by: $S = 2(c_0+c_1\overline{V})+2(z_0+z_1\overline{V})$.

Quoted spread is the average quoted spread over the period. The marks ***, ** and * represent significance at respectively the 1, 5 and 10% level.

TABLE 8. ESTIMATES FROM GLOSTEN & HARRIS' (1988) MODEL (2)

Stock	c_0	z_1	*Average transaction size* $\bar{V}$	*Adverse selection* α	*Order processing* γ	*Estimated spread* S	*Quoted spread*
SOLVAY	0.0435***	0.000063***	306	0.3067	0.6933	0.1254	0.1355
DELHAIZE GROUP	0.0194***	0.000013***	626	0.2972	0.7028	0.0553	0.0596
KBC	0.0222***	0.000016***	572	0.2977	0.7023	0.0631	0.0698
ELECTRABEL	0.0998***	0.000278***	133	0.2712	0.7288	0.2737	0.2726
DEXIA	0.0051***	0.000002***	1591	0.4387	0.5613	0.0181	0.0187
UCB	0.0257***	0.000000***	622	0.0118	0.9882	0.0521	0.0830
FORTIS	0.0056***	0.000005***	1301	0.5374	0.4626	0.0243	0.0236
AGFA-GEVAERT	0.0141***	0.000009***	533	0.2486	0.7514	0.0375	0.0454
IBA	0.0176***	0.000014***	332	0.2059	0.7941	0.0443	0.0714
BARCO (NEW)	0.0890***	0.000184***	173	0.2633	0.7367	0.2417	0.2985
INTERBREW	0.0149***	0.000007***	1121	0.3582	0.6418	0.0463	0.0499
GBL	0.0412***	0.000033***	521	0.2915	0.7085	0.1164	0.1610
TESSENDERLO	0.0602***	0.000074***	288	0.2614	0.7386	0.1629	0.2020
UMICORE	0.0450***	0.000068***	244	0.2712	0.7288	0.1235	0.1581
D'IETEREN	0.2038***	0.002797***	47	0.3922	0.6078	0.6707	0.8070
ALMANIJ	0.0408***	0.000039***	378	0.2667	0.7333	0.1112	0.1333
COLRUYT	0.0792***	0.000128***	188	0.2326	0.7674	0.2064	0.2206
BEKAERT	0.0627***	0.000111***	247	0.3036	0.6964	0.1802	0.2224
OMEGA PHARMA	0.0451***	0.000032***	375	0.2101	0.7899	0.1141	0.1337

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from Glosten & Harris' (1988) model – version (2). Stocks in italic in the lower panel are those for which there exists a liquidity provider.

Parameters c_0 and z_1 are obtained through the estimation of equation:

$\Delta P_t = c_0 \Delta Q_t + z_1 Q_t V_t + \epsilon_t$

where P_t is the transaction price, Q_t the trade indicator variable (+1 if a trade is buyer-initiated and -1 if it is seller-initiated), the transaction size in number of shares and V_t is an error term.

$\bar{V}$ is the average transaction size. α and γ are computed as follows:

$$\alpha = \frac{2(z_1\bar{V})}{2(c_0)+2(z_1\bar{V})} \quad \text{and} \quad \gamma = \frac{2(c_0)}{2(c_0)+2(z_1\bar{V})} = 1-\alpha$$

The estimated spread S is given by: $S = 2(c_0)+2(z_1\bar{V})$.

Quoted spread is the average quoted spread over the period. The marks ***,** and * represent significance at respectively the 1, 5 and 10% level.

In the general version of the model, the parameter c_0 is the only one that is significant for a great majority of stocks. Parameter z_0 presents significant values for half of the stocks. Estimates of c_1 and z_1 are only marginally significant and sometimes – even often for c_1 – present negative values, meaning that the corresponding cost would *decrease* with transaction size. There are two stocks – Electrabel and Fortis – for which none of the estimated parameters is significant. This general model seems to be too exhaustive, but highlights the importance of the constant order processing component in the spread. The constrained version of the model provides estimates of c_0 and z_1 that are significantly different from 0 at the 1% level for all stocks, giving support to Glosten & Harris' hypothesis of a constant order processing cost and an adverse selection cost linearly related to transaction size.

Both versions of Glosten & Harris' (1988) model provide estimates of α and γ that are of the same magnitude, with the order processing component being the most important part of the spread. There are two exceptions. The first one is UCB, which presents an adverse selection component of 35.58% in model (1) but only 1.18% in model (2). The difference is due to the *constant* adverse selection component z_0, that represents 36% of the spread in model (1)[21], but disappears in model (2). The second exception is Fortis, with adverse selection components of 29.55% and 53.74% in models (1) and (2) respectively. The origin of the difference lies in the *variable* order processing cost, that represents more than 8 times the spread in model (1), but disappears in model (2).

The consistency of the model can be appreciated through the estimated spread S, which is very close to the average posted spread for each stock, whatever the considered version. But even if, as noted earlier, the effective spread is higher than the quoted spread, the estimate given by S is slightly lower than the quoted spread for all stocks but Fortis.

3.5. Huang & Stoll's (1997) 2-way decomposition model

The basic regression model estimated in Huang & Stoll (1997) is the following:

$$\Delta P_t = \frac{S}{2}\Delta Q_t + (\alpha + \beta)\frac{S}{2}Q_{t-1} + e_t \tag{18}$$

where e_t is the error term. This regression thus provides estimates for the spread S, considered as being constant, and for (α+β), the sum of the adverse selection and the inventory holding components. The order processing cost γ is equal to 1-α-β. The 3-way decomposition (or how to distinguish α and β) is proposed in Section 3.2.

If we consider that the inventory holding cost β is equal to 0, we can rewrite Equation (18) as follows:

$$\Delta P_t = \frac{S}{2}\Delta Q_t + \alpha\frac{S}{2}Q_{t-1} + e_t \tag{19}$$

The order processing cost γ is then equal to 1-α. In their paper, Huang & Stoll prove that the model described by Equation (19) is equivalent to Glosten & Harris' (1988) model where we make the assumption that the spread is constant, i.e., where $C = c_0$ and $Z = z_0$. To show this, let's define $z_0 = \alpha\frac{S}{2}$ and $c_0 = (1-\alpha)\frac{S}{2}$; Equation (19) can then be rewritten as:

$$\Delta P_t = c_0\Delta Q_t + z_0 Q_t + e_t \tag{20}$$

We estimated Equations (19) and (20) separately and indeed found the same results. Estimates given by Equation (19) are presented in Table 9. All parameter estimates are

[21] It is even superior to the global adverse selection component of the spread because the variable part z_1 is *negative*.

statistically significant at the 1% level. The adverse selection component lies in a relatively small range, between 19.5% and 55.3%. The order processing component is once again the main part of the bid-ask spread. The spread estimated by the model is very close to the average posted spread. These results are logically very similar to those we have obtained in previous section for Glosten & Harris' (1988) model. The presence of a constant adverse selection component (z_0) in Huang & Stoll's (1997) 2-way decomposition explains why the estimates for UCB correspond to those given by the version (1) of Glosten & Harris' (1988) model. The absence of a variable order processing component can explain why the estimates for Fortis are closer to those given by the second version of Glosten & Harris' (1988) model.

TABLE 9. ESTIMATES FROM HUANG & STOLL'S (1997) 2-WAY DECOMPOSITION MODEL

Stock	*Adverse selection α*	*Order processing γ*	*Estimated spread S*	*Quoted spread*
SOLVAY	0.3106***	0.6894	0.1249***	0.1355
DELHAIZE GROUP	0.3059***	0.6941	0.0550***	0.0596
KBC	0.3106***	0.6894	0.0628***	0.0698
ELECTRABEL	0.2878***	0.7122	0.2726***	0.2726
DEXIA	0.4452***	0.5548	0.0180***	0.0187
UCB	0.3558***	0.6442	0.0654***	0.0830
FORTIS	0.5530***	0.4470	0.0242***	0.0236
AGFA-GEVAERT	0.2531***	0.7469	0.0374***	0.0454
IBA	0.2567***	0.7433	0.0450***	0.0714
BARCO (NEW)	0.2665***	0.7335	0.2392***	0.2985
INTERBREW	0.3654***	0.6346	0.0461***	0.0499
GBL	0.2899***	0.7101	0.1166***	0.1610
TESSENDERLO	0.2730***	0.7270	0.1621***	0.2020
UMICORE	0.2829***	0.7171	0.1234***	0.1581
D'IETEREN	0.3715***	0.6285	0.6760***	0.8070
ALMANIJ	0.2624***	0.7376	0.1113***	0.1333
COLRUYT	0.2378***	0.7622	0.2067***	0.2206
BEKAERT	0.3020***	0.6980	0.1807***	0.2224
OMEGA PHARMA	0.1950***	0.8050	0.1142***	0.1337

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from Huang & Stoll's (1997) 2-way decomposition model. Stocks in italic in the lower panel are those for which there exists a liquidity provider.
Parameters α and *S* are obtained through the estimation of equation:

$$\Delta P_t = \frac{S}{2}\Delta Q_t + \alpha\frac{S}{2}Q_{t-1} + e_t$$

where P_t is the transaction price, Q_t the trade indicator variable (+1 if a trade is buyer-initiated and -1 if it is seller-initiated) and e_t the error term.
γ is calculated as 1-α and *Quoted spread* is the average quoted spread over the period.
The marks ***, ** and * represent significance at respectively the 1, 5 and 10% level.

3.6. Madhavan et al.'s (1997) model

The equation used by Madhavan et al. (1997) to estimate the spread components is the following:

$$\Delta P_t = \phi \Delta Q_t + \zeta (Q_t - \rho Q_{t-1}) + e_t \tag{21}$$

where ρ is the first-order autocorrelation in the trade-initiation variable, and e_t is the error term.

Let us notice the similarity with Equation (20) above. In fact, in Madhavan et al. (1997), the revision in beliefs, and hence the asymmetric information parameter, is not measured by the observed order flow Q_t as it is the case in Huang & Stoll (1997), but rather by the innovation or "surprise" in order flow, i.e., the difference between the observed and the expected order flow $(Q_t - \rho Q_{t-1})$.

In the original model developed by Madhavan et al. (1997), a parameter λ is estimated that represents the probability that a trade takes place inside the spread.[22] While this view is consistent with quote-driven markets, where market makers can improve their quotes depending on the investor, so that some transactions occur at a price inside the quoted spread, it makes no sense in an order-driven market.[23] We thus decide to constrain the λ parameter to be equal to zero in our procedure.[24]

Equation (21) is estimated using a GMM procedure. The spread is then given by:

$$S = 2(\zeta + \phi) \tag{22}$$

The adverse selection and order processing components expressed as spread proportions are computed as follows:

$$\alpha = \frac{2\zeta}{2(\zeta + \phi)} \tag{23}$$

$$\gamma = \frac{2\phi}{2(\zeta + \phi)} \tag{24}$$

Finally, the probability of trade reversal π is estimated by $\frac{1-\rho}{2}$.

The estimation results are provided in Table 10. All estimated parameters are significant at the 1% level, at the exception of ϕ for D'Ieteren. The estimated spread is very close

[22] This parameter does not appear in the basic equation, but is present in the population moments used to estimate the model with a GMM procedure.

[23] On the contrary, we already showed that our aggregation procedure led to some transactions outside the spread.

[24] A similar approach has been followed in Ahn et al. (2002).

to the spread observed in the data for all stocks. The π estimates are lower than our own estimates in Table 2, but there is a 0.97 correlation between both measures. These results seem to suggest that the model well captures the price behavior in our data.

The proportion of adverse selection in the spread lies from 38.4% to 90.7%, which is a relatively wide range. For all stocks in the sample, this component is higher than found with previous models of Glosten & Harris (1988) and Huang & Stoll (1997) – it is sometimes even the double. For 14 stocks out of 19, the adverse selection component represents more than half of the spread. These results are consistent with Van Ness et al. (2001) and Clarke & Shastri (2000).

3.7. Lin et al.'s (1995) model

In addition to the classical adverse selection and order processing costs, Lin et al. (1995) propose to estimate a third component that represents the extent of order persistence. Defining x_t as one-half the signed effective spread ($x_t = P_t - M_t$), the three equations below are used to estimate the relevant parameters:

$$\Delta M_t = \alpha x_{t-1} + e_t \quad (25)$$

$$x_t = \theta x_{t-1} + \eta_t \quad (26)$$

$$\Delta P_t = -\gamma x_{t-1} + u_t \quad (27)$$

where e_t and η_t are the disturbance terms assumed to be uncorrelated, $u_t = e_t + \eta_t$ and $\gamma = 1-\alpha-\theta$. The parameter θ, which is the one that represents the extent of order persistence, can be used to measure the probability of trade continuation $\delta = \frac{1+\theta}{2}$ or if we prefer the probability of trade reversal $\pi = 1-\delta$.[25]

Results of that procedure are given in Table 11.[26] Most parameter estimates are highly significant, the only exception being the α value for Omega Pharma, which is not significant at all. It is easy to check that the sum of the three estimated parameters is approximately equal to 1 for all stocks. The adverse selection component lies within a relatively small range, between 2.13% and 39.92%. The range is even smaller if we do not take into account UCB and Omega Pharma that both present particularly low α and π values. The order processing component is smaller than in previous models, but this is due to the presence of the θ parameter. Parameter π is lower than 0.5 for all stocks, which is consistent with our data set. However, the correlation coefficient between our own estimates of the probability of trade reversal given in Table 2 and π values provided by Lin et al.'s (1995) procedure is only 0.32 and is not statistically significant. If we do not take into account UCB and Omega Pharma, for which Lin et al.'s (1995) estimated π values are very low, the correlation coefficient still remains relatively low at 0.51 but becomes significant at the 5% level.

[25] This θ parameter is very similar in meaning to the parameter ρ estimated in Madhavan et al. (1997).

[26] The empirical procedure actually uses the natural logarithm of the transaction price and the quote midpoint in each of the equations.

TABLE 10. ESTIMATES FROM MADHAVAN ET AL.'S (1997) MODEL

Stock	ϕ	ζ	ρ	Probability of trade reversal π	*Adverse selection* α	*Order processing* γ	*Estimated spread* S	*Quoted spread*
SOLVAY	0.0310***	0.0315***	0.4340***	0.2830	0.5040	0.4960	0.1249	0.1355
DELHAIZE GROUP	0.0147***	0.0129***	0.3666***	0.3167	0.4659	0.5341	0.0552	0.0596
KBC	0.0154***	0.0162***	0.4404***	0.2798	0.5119	0.4881	0.0632	0.0698
ELECTRABEL	0.0788***	0.0577***	0.3402***	0.3299	0.4225	0.5775	0.2730	0.2726
DEXIA	0.0030***	0.0060***	0.3562***	0.3219	0.6637	0.3363	0.0180	0.0187
UCB	0.0126***	0.0204***	0.5040***	0.2480	0.6171	0.3829	0.0660	0.0830
FORTIS	0.0011***	0.0108***	0.4236***	0.2882	0.9072	0.0928	0.0238	0.0236
AGFA-GEVAERT	0.0100***	0.0091***	0.5442***	0.2279	0.4756	0.5244	0.0381	0.0454
IBA	0.0136***	0.0093***	0.3879***	0.3060	0.4071	0.5929	0.0458	0.0714
BARCO (NEW)	0.0563***	0.0644***	0.5075***	0.2462	0.5337	0.4663	0.2414	0.2985
INTERBREW	0.0097***	0.0136***	0.4117***	0.2941	0.5824	0.4176	0.0466	0.0499
GBL	0.0253***	0.0325***	0.5167***	0.2417	0.5620	0.4380	0.1155	0.1610
TESSENDERLO	0.0388***	0.0421***	0.4935***	0.2532	0.5203	0.4797	0.1617	0.2020
UMICORE	0.0264***	0.0353***	0.5422***	0.2289	0.5722	0.4278	0.1233	0.1581
D'IETEREN	0.0331	0.3029***	0.5861***	0.2070	0.9014	0.0986	0.6722	0.8070
ALMANIJ	0.0255***	0.0307***	0.5385***	0.2307	0.5463	0.4537	0.1122	0.1333
COLRUYT	0.0477***	0.0541***	0.5746***	0.2127	0.5315	0.4685	0.2035	0.2206
BEKAERT	0.0415***	0.0488***	0.4523***	0.2738	0.5407	0.4593	0.1806	0.2224
OMEGA PHARMA	0.0346***	0.0216***	0.4839***	0.2581	0.3839	0.6161	0.1125	0.1337

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from Madhavan et al.'s (1997) model. Stocks in italic in the lower panel are those for which there exists a liquidity provider.

Parameters ϕ, ζ and ρ are obtained through the GMM estimation of equation:

$\Delta P_t = \phi \Delta Q_t + \zeta(Q_t - \rho Q_{t-1}) + e_t$

where P_t is the transaction price, Q_t the trade indicator variable (+1 if a trade is buyer-initiated and -1 if it is seller-initiated) and e_t the error term.

π, α and γ are computed as follows:

$\pi = \frac{1-\rho}{2}$, $\alpha = \frac{2\zeta}{2(\zeta+\phi)}$ and $\gamma = \frac{2\phi}{2(\zeta+\phi)}$.

The estimated spread is given by: $S = 2(\zeta+\phi)$

Quoted spread is the average quoted spread over the period.

The marks ***, ** and * represent significance at respectively the 1, 5 and 10% level.

TABLE 11. ESTIMATES FROM LIN ET AL.'S (1995) MODEL

Stock	*Adverse selection α*	*Order persistence θ*	*Order processing γ*	*Probability of trade reversal π*
SOLVAY	0.2609***	0.3177***	0.4098***	0.3411
DELHAIZE GROUP	0.2543***	0.3140***	0.4323***	0.3430
KBC	0.2649***	0.3143***	0.4182***	0.3429
ELECTRABEL	0.2610***	0.2381***	0.4945***	0.3810
DEXIA	0.3735***	0.2467***	0.3831***	0.3767
UCB	0.0556**	0.9885***	0.1043***	0.0057
FORTIS	0.3992***	0.2429***	0.3390***	0.3785
AGFA-GEVAERT	0.1996***	0.4511***	0.3502***	0.2745
IBA	0.2339***	0.4783***	0.2899***	0.2608
BARCO (NEW)	0.1994***	0.4639***	0.3371***	0.2680
INTERBREW	0.2962***	0.3169***	0.3883***	0.3416
GBL	0.2780***	0.4096***	0.3118***	0.2952
TESSENDERLO	0.2379***	0.4301***	0.3324***	0.2850
UMICORE	0.3056***	0.3887***	0.3055***	0.3057
D'IETEREN	0.3217***	0.4251***	0.2586***	0.2874
ALMANIJ	0.2250***	0.4234***	0.3382***	0.2883
COLRUYT	0.2579***	0.3646***	0.3788***	0.3177
BEKAERT	0.2634***	0.3923***	0.3412***	0.3038
OMEGA PHARMA	0.0213	0.8885***	0.0640***	0.0558

This table provides for 19 stocks belonging to the Belgian BEL20 index the results from Lin et al.'s (1995) model. Stocks in italic in the lower panel are those for which there exists a liquidity provider. Parameters α ,θ and γ are obtained respectively through the estimation of equations:

$\Delta M_t = \alpha x_{t-1} + e_t$

$x_t = \theta x_{t-1} + \eta_t$

$\Delta P_t = -\gamma x_{t-1} + u_t$

where P_t is the transaction price, M_t the quote midpoint and x_t the signed effective spread ($x_t = P_t - M_t$). e_t and η_t are uncorrelated error terms, and $u_t = e_t - \eta_t$.

π is calculated as $\frac{1-\theta}{2}$.

The marks ***, ** and * represent significance at respectively the 1, 5 and 10% level.

3.8. GLOBAL COMMENTS

If we want to summarize results from this section, it first seems that "complete" spread decomposition models, i.e., those that provide an estimate for all three classical spread components, are not relevant in an order-driven market. Indeed, results obtained with Stoll (1989) and Huang & Stoll's (1997) 3-way decomposition are inconsistent, even for the stocks of our sample for which a liquidity provider is committed to supply liquidity. These results thus give support to empirical papers that choose not to apply these models in order-driven markets based on the hypothesis of no inventory holding cost.

Then we find that the decomposition proposed in George et al. (1991) does not seem to be effective, as one third of the α estimates it provides are negative. These poor results may be due to the covariance-based type of the model (similar to Stoll (1989)).

As for the five other 2-component spread decomposition models, they seem to globally produce more consistent results. Parameters are usually highly significant. The spread components are all positive. The estimated spread provided by both versions of Glosten & Harris (1988) (respectively GH1 and GH2 in the following), Huang & Stoll's (1997) 2-way decomposition (HS2) and by Madhavan et al.'s (1997) model (MRR) are always very close to the quoted spread found in the data.

In order to analyze the coherence between those five models, we will now focus on the adverse selection component. This component is particularly interesting because it is the focus of most empirical studies in the literature. In GH1, GH2 and HS2, which can be considered as three variations around a same model, adverse selection is the smallest component. This seems true also for Lin et al. (1995) (LSB), while MRR is the only case where adverse selection represents the major part of the spread. Table 12 provides the Spearman rank correlation coefficients for the adverse selection components measured by the five considered models.

TABLE 12. CORRELATION COEFFICIENTS FOR THE ADVERSE SELECTION ESTIMATES

	GH1	*GH2*	*HS2*	*MRR*	*LSB*
GH1	1.0000				
GH2	0.6632	1.0000			
HS2	0.8860	0.7912	1.0000		
MRR	0.5772	0.5333	0.6737	1.0000	
LSB	0.6018	0.8526	0.6930	0.6035	1.0000

This table presents Spearman rank correlation coefficients for the adverse selection components provided by the five following models: Glosten & Harris (1988) – versions (1) and (2) [respectively GH1 and GH2], Huang & Stoll's (1997) 2-way decomposition [HS2], Madhavan et al. (1997) [MRR] and Lin et al. (1995) [LSB].

All coefficients are positive and significant at the 1% level, except the correlation between MRR and GH2 that is significant at the 5% level. Correlations among GH1, GH2 and HS2 were expected to be high, as these are actually three different versions of a same model. This seems to be the case, even if the correlation between GH1 and GH2 can be considered as rather low (0.66). This may be driven by the relatively unstable estimates given for UCB and Fortis.[27] While lower in absolute value, correlations between MRR, LSB and the other models tend to show that there is some consistency among the models.

[27] Indeed, when we do not take Fortis and UCB into account, the correlation coefficients among GH1, GH2 and HS2 all increase to 0.9 or more.

In summary, the five spread decomposition models used in Table 12 could be considered as alternative or substitute methods to analyze spread components in an order-driven market. But as GH1 and GH2 sometimes provide different estimates – remember our results for Fortis and UCB – we would recommend to limit the choice by considering only HS2, MRR and LSB methods.

4. Adverse Selection Components and Information Asymmetry

As already mentioned, there are some doubts about whether adverse selection estimates from spread decomposition models really measure information asymmetry. In the same spirit as Van Ness et al. (2001) and Clarke & Shastri (2000), we will now compare the results provided by the spread decomposition models to other information asymmetry metrics. Building on the conclusions drawn from previous section, we will concentrate on HS2, MRR and LSB models.

Let us look more carefully at Tables 9 to 11. The cross-sectional mean for the α parameter equals 0.3119, 0.5605 and 0.2479 respectively for HS2, MRR and LSB. According to HS2 and LSB, information asymmetry does not represent on average the major part of the bid-ask spread for BEL20 index stocks, while the opposite result holds for MRR. HS2 and LSB both classify Fortis and Dexia as having the largest adverse selection spread component. MRR provides the same conclusion if we do not take into account D'Ieteren for which one parameter is not significant at all. All three models attribute the smallest α parameter to Omega Pharma, but LSB estimate is not significant.

While those results are consistent with each other, the fact that the two most liquid stocks on Euronext Brussels are associated with the highest level of information asymmetry is puzzling, as these stocks are generally more carefully followed by financial analysts, thereby reducing the risk of private information. It is usually considered that small and illiquid stocks are more subject to information asymmetry risk (see for instance Easley et al. (1996)), so we would have expected to obtain highest α values for stocks such as IBA, D'Ieteren or Tessenderlo.[28]

This observation further encourages us to assess the validity of adverse selection components. However, contrary to Van Ness et al. (2001) and Clarke & Shastri (2000), we will stay in the field of market microstructure, which also allows us to exploit information available from the rebuilt order book.

[28] On the other hand, we cannot forget that the 19 stocks from our sample are part of Belgian main stock index, so even the most illiquid ones must be under the attention of financial analysts.

4.1. Our information asymmetry proxies

Ideally, our information asymmetry proxies should be based on some theoretical results found in the literature. However, this is not an easy task. Indeed, many theoretical frameworks dealing with limit order book, such as Foucault (1999) or Parlour (1998), do not consider asymmetric information. As explained in Bloomfield et al. (2004), *"adding in asymmetric information generally renders the problem unsolvable."*[29] To incorporate adverse selection, papers must introduce some restrictive assumptions about the behavior of informed traders. The most popular one, that we can found for instance in Glosten (1994), is that informed traders are impatient and use only market orders. But the experimental setting in Bloomfield et al. (2004) tells a completely different story, as informed traders introduce more limit orders than market orders compared to the uninformed investors.

Given these difficulties, our information asymmetry proxies will be based on empirical and experimental results provided for limit order books, as well as on other measures which were developed in quote-driven markets. We consider the following proxies:

- **Relative hidden depth** As explained in our description of Euronext microstructure (see Section 2.2), investors can use hidden orders, i.e., orders with some part of the quantity not being displayed to other market participants. Intuitively, we may think that hidden orders are used by informed traders who want to trade upon their information without paying the costs of demanding liquidity. An alternative hypothesis is that hidden orders are used by uninformed traders in order to mitigate adverse selection costs. Working on Nasdaq stocks, Tuttle (2003) shows that there is a positive relationship between hidden order use and the probability of an informational event proxied by market-model residual. She also shows that hidden depth, especially when it results from investment banks and wirehouses, is predictive of future market price movements, which could support the first hypothesis. Results obtained by Pardo & Pascual (2003) on the Spanish Stock Exchange are in favor of the second explanation, as they conclude that traders submit iceberg orders to manage the information asymmetry risk. Choosing between both explanations is difficult, but whatever the reason, hidden order use is in both cases positively linked to information asymmetry.

Denoting $Total\ depth_t$ the total number of shares (displayed and undisclosed) available at the five best limits of the order book at time t , and $Hidden\ depth_t$ the undisclosed number of shares available at these five best limits, we define relative hidden depth at time t as follows:

[29] Section 2 of Bloomfield et al. (2004) provides an interesting summary of theoretical models dedicated to limit order markets.

$$Relative\ hidden\ depth_t = \frac{Hidden\ depth_t}{Total\ depth_t} \tag{28}$$

Our first information asymmetry proxy is then a time-weighted average of *Relative hidden depth*$_t$ over the sample period.

- **Trade Imbalance** One of the most popular measures of information asymmetry is the *probability of informed trading* (PIN) developed in the context of a quote-driven market by Easley et al. (1996). One particular thing that contributes to a higher PIN is the imbalance between buy and sell trades. We therefore use this imbalance as a second proxy. Defining *Buy* as the number of buyer-initiated trades during the sample period and *Sell* as the number of seller-initiated trades, we derive our second proxy as follows:

$$Trade\ imbalance = \frac{|Buy - Sell|}{Buy + Sell} \tag{29}$$

- **Limit order proportion** The most striking result from Bloomfield et al.'s (2004) experiment is that informed traders use a lot more limit orders than do uninformed traders. If we denote *All orders* the total number of orders introduced during the sample period, and *Limit orders* the number of limit orders introduced during this same period, then our third proxy is:

$$Limit\ order\ proportion = \frac{Limit\ orders}{All\ orders} \tag{30}$$

- **Depth imbalance** Given the result found in Bloomfield et al. (2004) that informed traders act as liquidity providers, the expected imbalance due to private information would translate into the depth available in the order book. Denoting *Depth at bid*$_t$ (respectively *Depth at ask*$_t$) the total number of shares, both displayed and hidden, available at the five best bid (respectively ask) limits of the order book at time t, we define total depth imbalance at time t:

$$Depth\ imbalance_t = \frac{|Depth\ at\ bid_t - Depth\ at\ ask_t|}{Depth\ at\ bid_t + Depth\ at\ ask_t} \tag{31}$$

Our fourth proxy is thus the time-weighted average of *Depth imbalance*$_t$

- **Volatility** It is often argued that volatility is the consequence of trades generated by informed traders and is thus positively correlated with information asymmetry.[30] To compute volatility, we decided to divide each trading day in seventeen 30-minute intervals and to keep only the last transaction in each of these intervals.[31]

[30] See for instance French & Roll (1986).

[31] The last day interval lasts only 25 minutes.

If for some stocks and some days, there is no transaction in a given interval, we replicate the transaction price of the last available interval. We then calculate the return between two consecutive intervals. This provides us with 16 returns for each trading day. *Volatility* is measured as the standard deviation of these returns over the period under study.

All our information asymmetry proxies are constructed in such a way that we expect a positive relationship between them and the adverse selection spread components.

4.2. Data and results

We propose to analyze the relationship between the adverse selection components and our information asymmetry proxies through correlation coefficients. We acknowledge that using 19 observations might not be sufficient to draw solid conclusions. In order to increase the sample size, we will use subsamples of one week for each stock.[32] As we only want "complete" weeks, i.e., weeks that contain five trading days, we need to drop some trading days from the data set we used until here. We finally keep 11 complete weeks, corresponding to 55 trading days and 296 952 transactions.

We first compute α estimates for each stock and for each week using HS2, MRR and LSB. Summarized results are given in Table 13, where we present for each stock the minimum, maximum and mean values based on the 11 estimates.[33]

Results for HS2 and LSB confirm those presented in Section 3. The mean values are very close to the estimates obtained in Tables 9 and 11, at the exception of LSB estimates for UCB and Omega Pharma – but remember that α values in Table 11 were particularly low for both stocks. Fortis and Dexia once again have largest adverse selection components with both models, whereas Omega Pharma has the smallest α value with LSB, and the second smallest value with HS2. It also seems that adverse selection components for less liquid stocks tend to lie in a wider range than for other stocks. It is worth noticing that almost all estimates are significant at the 1% level.[34]

---

[32] Ideally, we would like to work with daily estimates, but the small number of trades for some stocks and some days prevents us from doing so.

[33] Comprehensive results are available from the authors upon request. Note that as HS2 procedure cannot provide estimates for IBA for 2 out of the 11 weeks, we only have 9 values for that stock.

[34] This information is not present in Table 13.

Results for MRR procedure are more contrasted. Firstly, there is a non negligible number of cases where the ϕ parameter used to compute the α component is not significant.[35] Secondly, we have some inconsistent α values, either negative or higher than 1, and not only for small stocks. Finally, the mean values are generally farther from the estimates given in Table 10. Globally, results from Table 13 suggest that our weekly HS2 and LSB estimates can be considered as consistent, while we have more concern with MRR weekly results.

We now turn to correlations of these estimates with our information asymmetry proxies. The coefficients provided in Table 14 are obtained using 207 observations, corresponding to 18 stocks with 11 observations and one stock – UCB – with 9 observations. All coefficients which are higher than 0.15 in absolute value are significantly different from zero at the 1% or 5% level, the other ones are not significant. Correlations among our HS2, MRR and LSB α estimates are of the same magnitude as in Table 12. Our information asymmetry proxies are on general not correlated to each other, except the positive link between *Relative hidden depth* and *Depth imbalance*, and a negative one between *Relative hidden depth* and *Volatility*.

The other results are not very enthusiastic. Only two out of the five proxies present the expected positive relationship with spread adverse selection components: *Limit order proportion* and *Volatility*. The magnitude of the coefficients is however rather low, the higher being equal to 0.47.[36]

---

[35] There are 11 stocks for which 4 or more of the 11 weekly ϕ estimates are not significant.

[36] If we consider only the best limit, and not the five best limits, to compute *Relative hidden depth* and *Depth imbalance*, results are qualitatively the same. This is also the case if we replace *Trade imbalance* by an *Order imbalance*. Results are available from the authors upon request.

TABLE 13. SUMMARY OF ADVERSE SELECTION COMPONENTS MEASURED BY WEEK

	HS2			*MRR*			*LSB*		
Stock	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*
SOLVAY	0.2469	0.3960	0.3083	0.4332	0.7635	0.5667	0.2228	0.3488	0.2697
DELHAIZE GROUP	0.2491	0.3597	0.3122	0.3670	0.6145	0.5288	0.2169	0.3040	0.2622
KBC	0.1603	0.4178	0.3154	0.3710	0.7191	0.5894	0.1511	0.3471	0.2552
ELECTRABEL	0.2001	0.3587	0.2655	0.3057	0.6139	0.4391	0.1821	0.2944	0.2288
DEXIA	0.3307	0.5033	0.4441	0.5705	0.9415	0.7330	0.2926	0.4402	0.3831
UCB	0.2356	0.4408	0.3418	0.4338	1.0247	0.7419	0.1198	0.4140	0.2630
FORTIS	0.4784	0.6612	0.5565	0.8235	1.2299	1.0059	0.3187	0.5487	0.4208
AGFA-GEVAERT	0.2047	0.3370	0.2563	0.4190	0.8244	0.6048	0.1401	0.2805	0.1944
IBA	0.0707	0.4768	0.2579	-0.3571	0.8589	0.3774	0.0566	0.3397	0.2256
BARCO (NEW)	0.1291	0.4752	0.2614	0.2582	1.2833	0.6276	0.1174	0.3503	0.2009
INTERBREW	0.2799	0.4406	0.3709	0.4760	0.7801	0.6561	0.2335	0.3645	0.2969
GBL	0.1896	0.4559	0.2944	0.3846	1.0146	0.6728	0.1281	0.3608	0.2411
TESSENDERLO	0.1173	0.4381	0.2622	0.2651	0.9773	0.6278	0.1362	0.2856	0.2123
UMICORE	0.0847	0.5330	0.2671	0.1929	1.0984	0.6908	0.0867	0.4538	0.2651
D'IETEREN	0.2071	0.6069	0.3899	0.7109	1.5396	1.1587	0.2001	0.4996	0.3150
ALMANIJ	0.1505	0.3661	0.2585	0.4195	0.9374	0.6307	0.1510	0.3142	0.2126
COLRUYT	0.1151	0.3446	0.2217	0.3422	1.0039	0.6193	0.1112	0.3241	0.2445
BEKAERT	0.1522	0.4657	0.2951	0.2911	1.2450	0.6409	0.1033	0.3522	0.2479
OMEGA PHARMA	0.1108	0.3201	0.2313	0.2982	0.6466	0.4951	0.0212	0.3075	0.1831

This table provides for 19 stocks belonging to the Belgian BEL20 index the minimum, maximum and mean value of the weekly α estimates for Huang & Stoll's (1997) 2-way decomposition [*HS2*], Madhavan et al. (1997) [*MRR*] and Lin et al. (1995) [*LSB*]. Weekly results are obtained using 296 952 transactions corresponding to 11 "complete" weeks of October, November and December 2002. Stocks in italic in the lower panel are those for which there exists a liquidity provider.

As explained above, there is no well accepted information asymmetry proxy in order-driven markets. If we accept the premise that the spread decomposition models are valid, then we can look critically at our "bad" proxies. *Trade imbalance* was chosen because it is at the heart of the PIN measure developed in Easley et al. (1996). This measure was developed in the context of a quote-driven market and may not be appropriate. The results for *Depth imbalance*, while counter-intuitive at first sight, may be explained by the results in Bloomfield et al. (2004). This experiment highlights that informed traders are the best liquidity providers in an order-driven market, which is supported by the positive relationship between our *Limit order proportion* proxy and the adverse selection components. But this is especially true at the end of the experimental market, when the true value of the stock is bracketed by the best bid and offer in the order book. Thus, if the market is efficient, meaning if the midquote is a good estimate of the fundamental value, then the informed traders will be present both at the bid and at the ask. In this case, a *lower* imbalance in depth can be associated to a higher degree of information asymmetry, which is what we find in Table 14. The negative relationship between *Relative hidden depth* and the adverse selection spread components is difficult to justify. Indeed, this would suggest that if all depth available in the book was displayed, information asymmetry would be at its highest degree.

We cannot forget of course that the correlation coefficients are relatively low in absolute value, and that they measure only linear relations, while the real relations may be more complex. We also recognize that our study can be thought of as a joint test of the spread component models and the information asymmetry proxies we have chosen. We can only conclude that some variables move in the same direction, while other do not. But determining which of these variables proxy information asymmetry is a difficult task, mainly because it is the very nature of private information that it is invisible to other market participants.

TABLE 14. CORRELATION BETWEEN ADVERSE SELECTION COMPONENTS AND INFORMATION ASYMMETRY PROXIES

	HS2	*MRR*	*LSB*	*Relative hidden depth*	*Trade imbalance*	*Limit order proportion*	*Depth imbalance*	*Volatility*
HS2	1.0000							
MRR	0.7059	1.0000						
LSB	0.8075	0.5911	1.0000					
Relative hidden depth	-0.2867	-0.2290	-0.3052	1.0000				
Trade imbalance	-0.3336	-0.0470	-0.3564	-0.0190	1.0000			
Limit order proportion	0.3750	0.4654	0.3193	-0.0972	-0.1028	1.0000		
Depth imbalance	-0.1722	-0.0533	-0.1822	0.5500	0.0714	-0.0570	1.0000	
Volatility	0.2934	0.2869	0.2613	-0.3585	0.1276	-0.0145	-0.0481	1.0000

This table presents Spearman rank correlation coefficients between α estimates provided by Huang & Stoll's (1997) 2-way decomposition [*HS2*], Madhavan et al. (1997) [*MRR*] and Lin et al. (1995) [*LSB*], and five other information asymmetry proxies.

Relative hidden depth is the time-weighted average of

$$Relative\ hidden\ depth_t = \frac{Hidden\ depth_t}{Total\ depth_t}$$

where *Total depth*$_t$ is the total number of shares (displayed and undisclosed) available at the five best limits of the order book at time t and *Hidden depth*$_t$ is the undisclosed number of shares available at these five best limits.

Trade imbalance is defined as: $\frac{|Buy - Sell|}{Buy + Sell}$

where *Buy (Sell)* is the number of buyer- (seller-) initiated trades during the sample period.

Limit order proportion is measured as: $\frac{Limit\ orders}{All\ orders}$

where *All orders* is the total number of orders introduced during the sample period, and *Limit orders* is the number of limit orders introduced during this same period.

Depth imbalance is the time-weighted average of *Depth imbalance*$_t$ computed as follows:

$$Depth\ imbalance_t = \frac{|Depth\ at\ bid_t - Depth\ at\ ask_t|}{Depth\ at\ bid_t + Depth\ at\ ask_t}$$

where *Depth at bid*$_t$ (respectively *Depth at ask*$_t$) is the total number of shares, both displayed and hidden, available at the five best bid (respectively ask) limits of the order book at time t.
Volatility is measured by the standard deviation of 30-minute returns (computed using trade prices). Correlation coefficients are computed using for each stock 11 observations corresponding to 11 "complete" weeks during the months of October, November and December 2002.

Conclusion

In this paper, we used several spread decomposition models to provide estimates of the spread components – inventory holding, adverse selection and order processing – for 19 stocks belonging to the Belgian BEL20 index traded on Euronext Brussels. As these models were originally developed in the context of quote-driven markets, we checked whether they provide consistent results when applied to order-driven markets.

We find that models which estimate an inventory holding component (Stoll (1989) and Huang & Stoll's (1997) 3-way decomposition) do not produce consistent estimates. This result can be viewed as a support for the often used hypothesis relative to the absence of real inventory management in order-driven markets: as limit order traders – who provide liquidity to the market – are not obliged to trade, they do not bear the risk of moving away from an "optimal" inventory level. This seems to be the case also for smaller stocks which are characterized by a so-called liquidity provider in charge of improving the supply of liquidity.

We also found that the covariance-based model developed by George et al. (1991) does not provide consistent estimates of the adverse selection costs.

We showed that the other models focusing exclusively on adverse selection and order processing components, namely the 2 implemented versions of Glosten & Harris (1988), Huang & Stoll's (1997) 2-way decomposition, Madhavan et al. (1997) and Lin et al. (1995) globally provide consistent results, with relatively highly correlated adverse selection components. If researchers have to choose among them in the context of an order-driven market, we suggest that they apply the Huang & Stoll's (1997) 2-way decomposition or Lin et al.'s (1995) procedure.

In order to check how well adverse selection components provided by Huang & Stoll (1997), Madhavan et al. (1997) and Lin et al. (1995) measure adverse selection, we compared them with five information asymmetry proxies: relative hidden depth, trade imbalance, proportion of limit orders, total imbalance in the order book, and volatility. While the adverse selection components are as expected positively correlated with volatility and the proportion of limit orders in the order flow, the correlation coefficients with our three other proxies do not present the hypothesized positive sign. Even if we admit that our information asymmetry proxies can be cast in doubt, we think that whether classical spread decomposition models really measure information asymmetry remains an open question. Finding specific information asymmetry measures in the context of an order-driven market is a challenging research question left for the future.

The main limit of this paper lies in the relatively small number of stocks we used to produce our various estimates. As a consequence, it may be possible that some of our results primarily reflect liquidity effects – opposition between liquid and less liquid stocks. On the other hand, this is also the originality of the paper that drove this small sample size, as we have dealt with a small and relatively illiquid market, Euronext Brussels, for which methods using intraday data have rarely been applied.

REFERENCES

Ahn H.-J., J. Cai , Y. Hamao and R. Y. Ho, 2002. "The components of the bid-ask spread in a limit-order market: Evidence from the Tokyo Stock Exchange", *Journal of Empirical Finance* 9(4), 399-430.

Amihud Y. and H. Mendelson, 1980. "Dealership market: Market-making with inventory", *Journal of Financial Economics* 8(1), 31-53.

Bagehot W., 1971. "The only game in town", Financial Analysts Journal 27, 12-14.

Biais B., P. Hillion and C. Spatt, 1995. "An empirical analysis of the limit order book and the order flow in the Paris Bourse", *The Journal of Finance* 50(5), 1655-1689.

Bloomfield R., M. O'Hara and G. Saar, 2004. "The "make or take" decision in an electronic market: Evidence on the evolution of liquidity", *Journal of Financial Economics*. Forthcoming.

Brockman P. and D. Y. Chung, 1999. "Bid-ask spread components in an order- driven environment", *The Journal of Financial Research* 22(2), 227-246.

Clarke J. and K. Shastri, 2000. "On information asymmetry metrics", *Working Paper.*

Copeland T. E. and D. Galai, 1983. "Information effects on the bid-ask spread", *The Journal of Finance* 38(5), 1457-1469.

De Winne R. and C. D'Hondt, 2004. "Market transparency and traders' behavior: An analysis on Euronext with full order book data", *Working Paper*. Department of Finance, FUCaM - Catholic University of Mons.

De Winne R. and I. Platten, 2003. "An analysis of market makers' behavior on Nasdaq Europe", *Working Paper*. Department of Finance, FUCaM - Catholic University of Mons.

Declerck F., 2002. "Trading costs on a limit order book market: Evidence from the Paris Bourse", *Banque & Marchés* 57, 31-45.

Declerck F. and P. Hazard, 2002. "Impacts de l'animation sur la qualité du second marché", *Banque & Marchés* 60, 5-18.

Dennis P. J. and J. P. Weston, 2001. "Who's informed? An analysis of stock ownership and informed trading", *Working Paper*.

Easley D., N.M. Kiefer, M. O'Hara and J. B. Paperman, 1996. "Liquidity, information, and infrequently traded stocks", *The Journal of Finance* 51(4), 1405-1436.

Easley D. and M. O'Hara, 1987. "Price, trade size, and information in securities markets", *Journal of Financial Economics* 19(1), 69-90.

Elder J., P. K. Jain and J.-C. Kim, 2004. "Do tracking stocks reduce informational asymmetries? An analysis of liquidity and adverse selection", *The Journal of Financial Research* . Forthcoming.

Flood M. D., R. Huisman, K. G. Koedijk and R. K. Lyons, 1998. "Search costs: The neglected spread component", *Working Paper*.

Foucault T., 1999. "Order flow composition and trading costs in a dynamic limit order market", *Journal of Financial Markets* 2(2), 99-134.

French K. R. and R. Roll, 1986. "Stock return variances: The arrival of information and the reaction of traders", *Journal of Financial Economics* 17(1), 5-26.

George T. J., G. Kaul and M. Nimalendran, 1991. "Estimation of the bid-ask spread and its components: A new approach", *The Review of Financial Studies* 4(4), 623-656.

Glosten L. R., 1994. "Is the electronic open limit order book inevitable?", *The Journal of Finance* 49(6), 1127-1161.
Glosten L. R. and L. E. Harris, 1988. "Estimating the components of the bid / ask spread", *Journal of Financial Economics* 21(1), 123-142.
Glosten L. R. and P. R. Milgrom, 1985."Bid, ask and transaction prices in a specialist market with heterogeneously informed traders", *Journal of Financial Economics* 14(1), 71-100.
Handa P., R. Schwartz and A. Tiwari, 2003. "Quote setting and price formation in an order driven market", *Journal of Financial Markets* 6(4), 461-489.
Henker T. and J.-X. Wang, 2004. "On the importance of timing conventions in market microstructure research", *Working Paper.*
Ho T. and H. R. Stoll, 1981. "Optimal dealer pricing under transactions and return uncertainty", *Journal of Financial Economics* 9(1), 47-73.
Huang R. D. and H. R. Stoll, 1997. "The components of the bid-ask spread: A general approach", *The Review of Financial Studies* 10(4), 995-1034.
Lee C. M. C. and M. J. Ready, 1991. "Inferring trade direction from intraday data", *The Journal of Finance* 46(2), 733-746.
Levin E. J. and R. E. Wright, 2004. "Estimating the profit markup component of the bid-ask spread: Evidence from the London Stock Exchange", *Quaterly Review of Economics and Finance* 44(1), 1-19.
Lin J.-C., G. C. Sanger and G. G. Booth, 1995. "Trade size and components of the bid-ask spread", *The Review of Financial Studies* 8(4), 1153-1183.
Madhavan A., M. Richardson and M. Roomans, 1997. "Why do security prices change? A transaction-level analysis of NYSE stocks", *The Review of Financial Studies* 10(4), 1035-1064.
Menyah K. and K. Paudyal, 2000. "The components of bid-ask spreads on the London Stock Exchange", *Journal of Banking & Finance* 24(11), 1767-1785.
Neal R. and S. M. Wheatley, 1998. "Adverse selection and bid-ask spreads: Evidence from closed-end funds", *Journal of Financial Markets* 1(1), 121-149.
Pardo A. and R. Pascual, 2003. "On the hidden side of liquidity", *Working Paper.*
Parlour C. A. (1998), "Price dynamics in limit order markets", *The Review of Financial Studies* 11(4), 789-816.
Roll R., 1984. "A simple implicit measure of the effective bid-ask spread in an efficient market", *The Journal of Finance* 39(4), 1127-1139.
Saporta V., G. Trebeschi and A.Vila, 1999. "Price formation and transparency on the London Stock Exchange", *Working Paper.* Bank of England.
Silva A. C. and G. Chavez, 2002. "Components of execution costs: Evidence of asymmetric information at the Mexican Stock Exchange", *Journal of International Financial Markets, Institutions & Money* 12(3), 253-278.
Stoll H. R., 1978. "The supply of dealer services in securities markets", *The Journal of Finance* 33(4), 1133-1151.
Stoll H. R., 1989. "Inferring the components of the bid-ask spread: Theory and empirical tests", *The Journal of Finance* 44(1), 115-134.
Swan P. L. and J. Westerholm, 2004. "The impact of market architectural features on world equity market performance: A structural equation approach", *Working Paper.*

Tuttle L., 2003. “Hidden orders, trading costs and information”, *Working Paper*.
Van Ness B. F., R. A. Van Ness and R. S. Warr, 2001. “How well do adverse selection components measure adverse selection?”, *Financial Management* 30(3), 77-98.
Vandelanoite S., 2002. “Takeover announcements and the components of the bid-ask spread”, *Working Paper*. University of Paris 1.
Venkataraman K., 2001. “Automated versus floor trading: An analysis of execution costs on the Paris and New York exchanges”, *The Journal of Finance* 56(4), 1445-1485.
Wang J., 1999. “Asymmetric information and the bid-ask spread: An empirical comparison between automated order execution and open outcry auction”, *Journal of International Financial Markets*, Institutions & Money 9(2), 115-128.