

# How Important Is Informed Trading for the Bid-Ask Spread? Evidence from an Emerging Market

Jan Hanousek and Richard Podpiera<sup>1</sup>

*December 2000*

## **Abstract:**

The link between informed trading and the bid-ask spread has been the focus of abundant literature and some authors feared that a large amount of informed trading might lead to shutdown of markets. We explore this issue using data from the Czech Republic. Our estimates confirm that the share of informed trading and its variability is indeed high relative to developed markets, however, share of the adverse selection component is only 14% of the spread. Since the Czech Republic has been known in the financial community as being plagued by informed trading, our findings suggest that the relative importance of adverse selection as a determinant of the spread is generally low across markets.

## **Abstrakt**

V tomto článku odhadujeme rozsah obchodování na základě neveřejných informací v České republice. Rozsah takového obchodování je důležitý neboť podle většiny novějších modelů chování investorů určuje likviditu trhu i náklady jeho účastníků a tím také efektivnost jeho fungování. Naše odhady ukazují, že pravděpodobnost, že daný obchod je motivován znalostí neveřejných informací dosahuje u akcií v systému SPAD hodnoty 37%, což je téměř dvakrát více ve srovnání i s méně likvidními akciemi na NYSE. Navíc také variabilita toku obchodních příkazů je v České republice vyšší, což dále zvyšuje riziko nesené tvůrci trhu a zřejmě i požadovanou kompenzaci za jejich služby – vyšší transakční náklady pak snižují atraktivitu českého trhu.

Keywords: market microstructure, informed trading, bid-ask spread, adverse selection

JEL Classification: G14, G15

---

<sup>1</sup> Both authors are at CERGE-EI, a joint workplace of Charles University and the Academy of Sciences of the Czech Republic, Prague. Jan Hanousek currently fills the Citibank Chair in Financial Markets at CERGE-EI. This research was supported by the National Science Foundation of the United States (grant number SPR-9712336) and by the PHARE/ACE Research Program of the European Union (grant number P97-8118-R). We would like to thank Randall K. Filer and Jan Kmenta for their suggestions and comments. Correspondence to: Richard Podpiera, CERGE-EI, P.O. Box 882, Politických vězňů 7, 111 21, Prague, Czech Republic. E-mail: Richard.Podpiera@cerge.cuni.cz.

## **1. Introduction**

The link between informed trading (adverse selection of traders) and the bid-ask spread has been the focus of plentiful research since the bid-ask spread is an important determinant of trading costs.<sup>2</sup> Some authors have suggested that if informed trading is common, it might lead to market shutdown as the spread would be simply too wide for investors to participate in trading.

Historically, the bid-ask spread has been regarded as a function of order processing and inventory costs. More recent research brought adverse selection and thus informed trading into the picture. In this view, the market maker (as a provider of liquidity who always quotes prices for buying and selling) faces the possibility of trading with agents who have superior information. The market maker will lose money when trading with such individuals, and he or she thus sets a spread between the bid and ask price in order to compensate for this adverse selection problem. Glosten and Milgrom (1985) were the first to show formally that, with informed trading, a bid-ask spread would exist even if there were no order processing and inventory costs. The current literature thus distinguishes three components of the bid-ask spread, that is, inventory, order processing and adverse selection components.

Despite the fact that a number of empirical studies attempted to estimate the components of the bid-ask spread, the problem of the size of the adverse selection component does not have a final solution. Earlier studies suggested that the share of the adverse selection component is large. For instance, Stoll (1989), who used a sample of NASDAQ stocks,

reported that 43% of the quoted spread is due to adverse selection. On the other hand, George, Kaul and Nimalendran (1991) estimate much smaller share of adverse selection costs. For a sample of NYSE, AMEX and NASDAQ stocks, they put this share in the range of 8–13%. Huang and Stoll (1997) estimate a general model and offer two ways of decomposing the spread. They thus also estimate two values of this share—9.6% and 21.5%. It is difficult to explain the differences of the estimated shares due to adverse selection since the cited studies differ both in terms of the data and the estimation methodology they use. Generally, recent studies suggest that the share of the adverse selection component is rather low—at least for developed markets. This induces the question to what extent is the adverse selection component important across markets, especially those plagued by informed trading. The current paper addresses precisely this issue.

The current paper investigates the bid-ask spread components on the Czech stock market, with a focus on the adverse selection component. Besides the fact that it provides evidence for the composition of the bid-ask spread from an emerging market and high frequency data (previous studies dealt almost exclusively with U.S. markets and, except for the most recent studies, had only daily and weekly data), we believe our analysis is interesting from two additional points of view. First, when estimating the adverse selection component, we combine the recent findings of the inventory literature with older methods of estimation. In previous literature, the structure of the market—mainly the distinction between competitive market makers and a single specialist—was to a large extent ignored, possibly due to the fact that less detailed data were used.

---

<sup>2</sup> “Bid-ask spread” is defined as the difference between the price at which the market maker is willing to

Second, the present paper offers an insight into the magnitude of the adverse selection component not only in an emerging market, but in a market which has been known to the international investment community as one being seriously plagued by insider trading. The development of the Czech equity market was very rapid at the outset, since a large number of companies were floated as a result of coupon privatization. The regulation of the market, however, lagged significantly behind. Insider trading, price manipulation, fraud in the investment funds industry, and abuses of minority shareholder rights eroded investor confidence to a large extent. In recent years, regulation has improved somewhat, but enforcement appears to still be rather weak.<sup>3</sup> Especially in terms of the quality and timeliness of information disclosure, the Czech market still offers a lot of space for informed, and more often than not insider, trading.

Should the adverse selection component be small on this market, then it is safe to assume that its practical importance is also low on most other markets. In this sense, we estimate an upper bound of the importance of the adverse selection component.

Besides being known as an insider market, there is another reason why the Czech equity market offers a good opportunity for this kind of inquiry. The setup of its key component, the SPAD trading system, which is based on market makers that quote prices in selected

---

sell and the price at which he is willing to buy.

<sup>3</sup> The reflection of the Czech equity market in the eyes of the international investor community can be illustrated by articles in the press and the evaluation by international organizations. *The Economist* (April 1996) and the *Wall Street Journal* (May 1996), among others, reported on “dealing in Prague as a losers’ guide to investment,” and characterized the Czech capital market as “a muddy market” and as “anarchy to the outsider, sweet profit to those in the know”. More recently, *The Economist* (March 1997) quoted an investor as saying “...[the government should] fight the perception that the Prague stock exchange is just a vehicle for select insiders to enrich themselves at the expense of the ordinary shareholder”. In its 1999 Country Study, the World Bank argues that “The capital market needs to be further strengthened to recover credibility and to be a real source of corporate financing” (Summary Report, page 17). It is also illustrative that the Prague Stock Exchange has been unable to become a member of the Federation of European Stock

stocks, is very suitable for investigating the components of the bid-ask spread and inquiring into the effect of informed trading. It corresponds to the basic setup of many theoretical models, for instance, the trading lots are rather large, so in most trades only one lot changes hands, which allows us to abstract from the trade size. Also, there are several market makers for each stock that compete for order flow. Only the most liquid Czech stocks are included in the system, but these are nevertheless only medium-sized in international terms.<sup>4</sup>

There was one earlier attempt to estimate the extent of insider trading at the Prague Stock Exchange. Němeček (1998) attempted to estimate an Easley et al. (1996b)-type model, but at that time the SPAD system was only at the beginning of its operation thus detailed data were not available. Němeček concluded that the insider trading or adverse selection component of the bid-ask spread was relatively insignificant. However, due to insufficient data (Němeček had to use data from the auction system as no data on individual trades were available), his conclusions were not very robust. Moreover, at that time, the liquidity at the PSE's automated price-setting system was minuscule since a majority of trades were conducted outside of the exchange and it was actually advantageous for informed investors to trade off the exchange. In this paper, we use completely different data, from a new trading system, and from a time period when the vast majority of the order flow is publicly observable.

---

Exchanges, even though both Budapest and Warsaw Stock Exchanges are associate members of this federation.

<sup>4</sup> We thus do not face the problem that our sample would be formed by large companies which tend to have a lower spread in absolute terms and, as shown by Easley et al. (1996b), also tend to have lower adverse selection component of the spread. For a more detailed discussion of the sample selection and comparison with developed markets, see the next section (Trading System and Data Description).

In the next section, we describe the trading system and the data. In Section 3, then, we estimate the extent of informed trading by using the approach of Easley et al. (1996b), in order to ascertain whether informed trading indeed plays an important role in the Czech market. In the following section, we describe the models we use to investigate the components of the spread. At first, we examine the basic characteristics of market maker behavior based on findings of inventory models. Then, we present our strategy for estimating the adverse selection component of the spread and review the Huang and Stoll (1997) approach, which we later use for comparative purposes. Section 5 depicts the estimation results. Section 6 summarizes our findings and concludes.

## **2. Trading System and Data Description**

### *2.1. Trading System*

The SPAD segment of the Prague Stock Exchange is formed by market makers who are obliged to quote prices for sale and purchase. The whole system is screen-based with all the market makers, and, in fact, all the members of the PSE, able to see all the quotes and trades. Members of the PSE who apply and are approved serve as market makers in the SPAD system. This system was launched in May 1998 with the aim of improving the liquidity in the market. It was successful in attracting order flow from the OTC market and, currently, the vast majority of trades in the securities that are listed in SPAD, are channeled through this system.

The most liquid Czech stocks are traded in the system. The number of stocks in SPAD grew from a single stock at the outset to the current number of eight stocks. There are two telecommunications companies (Český telecom and České radiokomunikace), two banks (Česká spořitelna and Komerční banka), a petrochemical company (Unipetrol), an electricity generator (ČEZ), an investment fund (RIF) and a construction company (IPS).<sup>5</sup> Daily trading volume in SPAD amounted to approximately 500 million CZK (some 17 million USD) for most of its existence and only at the beginning of 2000 grew sharply to almost 2 billion (bn) CZK (50 million USD). In the period under study (late 1999) daily trading volume hovered around the (mn) 500 million CZK level.<sup>6</sup>

In late 1999 there were 16 market makers who quoted at least one stock in the system. On average, each market maker quoted approximately 6 stocks. Viewed differently, each of the eight stocks had some 12 market makers who were quoting it. The SPAD rules stipulate that there must be at least 3 market makers for each stock for it to be traded in the system. In reality, the number of market makers for each stock is considerably higher than the required minimum.<sup>7</sup>

Individual market makers are allowed to quote different ask and bid prices, but the maximum spread for each of them is limited. A committee of the PSE sets the exact limits, but we can say that it corresponds roughly to 2.5% of the stock's price (depending

---

<sup>5</sup> One more bank, IPB, had its stock traded in the SPAD system. In June 2000, however, IPB was put under forced administration and its stock was suspended from trading. Because it was introduced into SPAD only in November 1999 and thus was not traded in the system during the time period we focus on, we leave it out altogether.

<sup>6</sup> At the same time, the share of the SPAD system of the whole PSE trading volume was rather high in 1999, at 82%, as it was successful in attracting order flow from the direct trades segment.

<sup>7</sup> Earlier, there was an informal limit on the maximum number of market makers for one stock (10). It was abolished before the time period under study began.

on the development of the stock's price the maximum spread is irregularly changed). The system operates in two phases, closed and open. The closed phase can be viewed as a technical device that allows market makers to clear the trades that they did not manage to conduct during the open phase and is not important for our purposes. The actual trading occurs during the open phase of the system, which lasts from 9.30 a.m. to 4.00 p.m. each trading day and during which the market makers quote firm prices for a fixed amount of shares of each stock. The size of trading lots varies from 1,000 shares for České radiokomunikace to 20,000 shares for Unipetrol or ČEZ. The size of the lots is occasionally changed, depending on price development of the stock. The trading lots are rather large, both as compared with the overall trading volume and with the capital base of some of the market makers.<sup>8</sup> The quotes are firm in the sense that if the quote is the best available on the market and if a counterparty reacts to it in the system by entering an instruction for a trade, the market maker is obliged to enter his instruction so that the trade can be executed. Blocks of shares that differ in size from the trading lot can be settled through the system as well, but these trades are negotiated in advance over the phone and, in fact, are not very frequent.

In order to limit the risk of default, there exists a standard settlement procedure and a guarantee fund, into which market makers must insert a deposit, and there are procedures that come in place if one side of the trade defaulted. Overall, trading in SPAD appears to be safe, since no serious problems of default were reported since its inception.

---

<sup>8</sup> The market value of one trading lot ranged from 0.7 mn CZK to 5.3 mn CZK in late 1999 and averaged slightly below 2.0 mn CZK.



## *2.2. Data Description*

Data on individual trades from SPAD are publicly available from early 1999 to July 2000. However, we were able to obtain more detailed data for two months, September and October 1999, and thus concentrate on this time period. We add data from other months as necessary for estimating some of the models we use below.

For the time period from early 1999 to July 2000, we have the basic data on each trade conducted in the SPAD system. For each trade, our database carries an identification of the stock, transaction price, number of shares, time when the trade was concluded, and the best bid and ask quotes at the time the transaction took place. Also, we are able to identify so called cross trades, that is, we are able to distinguish trades that are conducted between the inventory of a market maker and the market maker's clients, since these must be reported in the system as well. For the September and October 1999 we also observe the identity of the buyer and the seller and an indication of whether there was an obligation to react on the instruction.

Table 1 depicts the basic characteristics of the stocks in our sample—we investigate all eight stocks that have been traded in the system. Even though they can be generally described as the most liquid equities on the Czech market, there are substantial differences among them. Their market capitalization ranges between a mere 2 bn CZK for IPS to 124 bn CZK for SPT Telecom. This difference in size is also reflected in trading activity. We thus see only three trades per day for IPS, on average, and 43 daily trades for SPT Telecom. These two stocks are indeed extremes, the other stocks are concentrated in the range of 10–51 bn CZK of market capitalization and, with the exception of RIF, exhibit decent trading activity.

**Table 1: Basic Characteristics of the Stocks in Our Sample (Sept. and Oct. 1999)**

Stock	Market Cap. (mn CZK)	Average Price (CZK)	Daily Turnover (mn CZK)	Daily Number of Trades	Trading Lot (number of shares)	Avg. Spread* (CZK)	Avg. Spread* (%)	Number of market makers
ČESKÁ SPORITELNA	11,873	183	33	15	10,000	4.1	2.2%	11
Č. RADIOKOMUNIKACE	36,996	1 203	39	26	1,000	16.8	1.4%	13
ČEZ	51,258	87	77	34	20,000	0.9	1.0%	15
IPS	1,831	132	4	3	5,000	3.8	2.9%	11
KOMERČNÍ BANKA	16,865	887	79	25	3,000	11.9	1.3%	12
RIF	12,557	1 309	9	5	1,000	11.5	0.9%	10
SPT TELECOM	124,394	529	243	43	10,000	5.2	1.0%	14
UNIPETROL	10,573	58	50	19	20,000	1.2	2.1%	13
<b>Average</b>	<b>33,293</b>	<b>210</b>	<b>67</b>	<b>170</b>	<b>n.a.</b>	<b>6.7</b>	<b>1.6%</b>	<b>12.4</b>

Source: Prague Stock Exchange and authors' calculation. The average exchange rate during the two months amounted to 34.4 CZK/USD.

Note: There were 42 trading days in September and October 1999. Thus, the total number of trades in these two months ranges between 122 for IPS and 1,823 for SPT Telecom.

\*This is the posted spread, that is, the difference between best bid and best ask prices.

As already mentioned above, the total daily turnover slightly exceeded 500 mn CZK in September and October 1999 and the size of trading lots varied considerably in terms of the number of shares and somewhat less in terms of the market value of trading lots. The average spread varied considerably in CZK terms, but its relative value, in percent of average price, is more important. Here, the variation is smaller, and, again with the exception of RIF, the percentage spread to a large extent depends on the trading activity. It is largest for IPS, at 2.9% and lowest for SPT Telecom, at 1%. For the other stocks, the relation between the percentage spread and the trading activity is rather monotonic.

A comparison of the stocks in our sample with those used in previous studies of informed trading and the components of the bid-ask spread appears to be in order. Naturally, the volumes and market capitalization of SPAD stocks is dramatically smaller than those of

the most liquid U.S. stocks. The average trading volume of the stocks in our sample approximately equals the average trading volume of the fifth decile (thus approximately average) stocks from the NYSE included in Easley et al. (1996b). Huang and Stoll (1997) examined 1992 data for the 20 stocks in the Major Market Index. For these stocks, the mean percentage posted spread has been hovering between 20 and 70 basis points. The NASDAQ stocks used by Stoll (1989) are comparable in terms of both volume and posted percentage spread, and even the reported average number of market makers—13 to 14—corresponds to the number of market makers in SPAD.

Overall, our stocks are less frequently traded and have larger posted spreads than the blue chips from the U.S.; the differences, however, are not so dramatic to prevent us from comparing our results with those of previous studies. In fact, the opposite is true, since our stocks are comparable to an average U.S. stock. If anything, the differences suggest that the adverse selection component should be larger in the Czech equity market, since lower trading volume was shown by Easley et al. (1996b) to be associated with a larger probability of informed trading.

### 3. Estimating the Extent of Informed Trading

#### 3.1. Model and Estimation Method

In order to be able to compare our results to those from developed markets, we use the model developed by Easley et al. (1996b), who, as a by-product of their effort to relate differences in information-based trading and the variation in bid-ask spread, estimated the extent of informed trading for the first, fifth, and eighth deciles of NYSE stocks ranked by trading volume.

The basic setup of their model is as follows.<sup>9</sup> Potentially informed and uninformed traders trade a risky asset with a competitive risk neutral market maker. Time is split into separate trading days and is continuous within each trading day. Before each trading day begins, nature determines whether an information event that influences the value of the asset occurs. These independently distributed information events occur with probability  $\alpha$  and are good news with probability  $1 - \delta$  and bad news with probability  $\delta$ . Naturally, the asset has a higher value when good news is coming to the market and a lower value for bad news. The news is fully realized by the traders by the end of each day—in this sense, the days are separated.

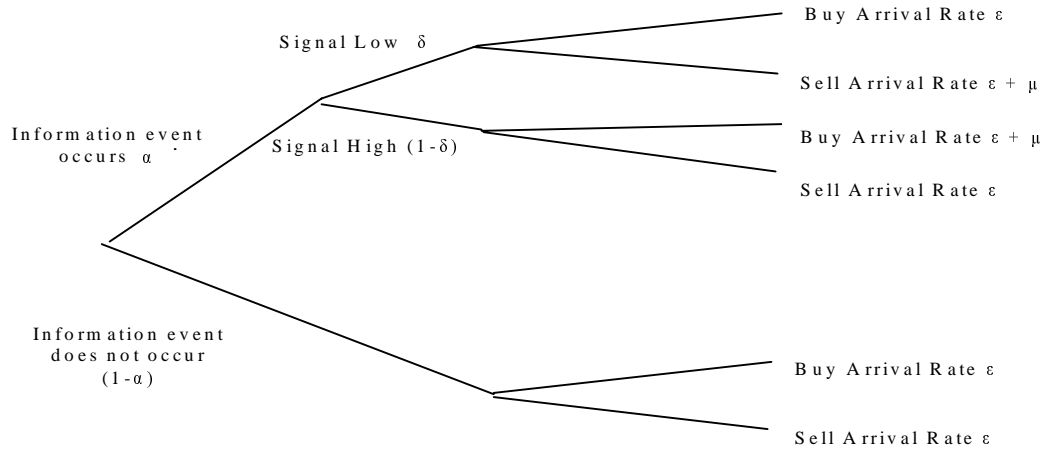
Both uninformed (did not see any signal) and informed (saw some signal) traders arrive to the market. The arrival rates are independent Poisson processes. The arrival rate of uninformed buyers and uninformed sellers is denoted by  $\varepsilon$ . When an information event occurs, informed traders also arrive, at a rate  $\mu$ , which is again the same for informed

---

<sup>9</sup> The description here is very brief due to space constraints. Please check the original Easley et al. (1996b) paper for details.

sellers, who arrive in the case of bad news, and informed buyers, who are attracted by good news.

**Figure 1: Trading Process**



The market maker knows the initial probabilities attached to each branch of the tree and updates his beliefs in a Bayesian manner based on the observed order flow. For instance, a sell order might come from both an informed and an uninformed trader, but its appearance changes the belief of the market maker about bad news.

Thus, there are three possible events—no news, good news and bad news. The arrival rates of traders are represented by the number of buy and sell transactions that occur during the day. The likelihood of observing a given number of buys (B) and sells (S) given a vector of parameters  $\theta = (\alpha \delta \epsilon \mu)$  can be expressed as

$$\begin{aligned}
L((B, S) | \theta) = & (1 - \alpha) * e^{-\epsilon T} \frac{(\epsilon T)^B}{B!} e^{-\epsilon T} \frac{(\epsilon T)^S}{S!} + \alpha \delta * e^{-\epsilon T} \frac{(\epsilon T)^B}{B!} e^{-(\mu + \epsilon) T} \frac{[(\epsilon + \mu) T]^S}{S!} + \\
& + \alpha (1 - \delta) * e^{-(\mu + \epsilon) T} \frac{[(\epsilon + \mu) T]^B}{B!} e^{-\epsilon T} \frac{(\epsilon T)^S}{S!} . \quad (1)
\end{aligned}$$

The first term of the likelihood denotes the “no event” day, the second the “bad event” day and the last the “good event day”. The likelihood contributions of these day-types take this form due to the assumption of (independent) Poisson processes driving the arrival of traders. What makes the difference between the three expressions are the arrival rates of buyers and sellers—that is what we observe, but we cannot say whether some of the traders saw a signal or not. In fact, only the order flow, that is, the number of buys and sells, is used to make an inference about the extent of informed trading in this model. Since days are independent, the likelihood of observing a series of buys and sells is the product of the likelihoods for individual days:

$$L(M | \theta) = \prod_{i=1}^I L(\theta | B_i, S_i) \quad \text{for } M = (B_i, S_i)_{i=1}^I. \quad (2)$$

Given data on the frequency of buy and sell orders, maximizing the likelihood expression above will yield the estimates of the four key parameters ( $\alpha$   $\delta$   $\epsilon$   $\mu$ ) of the model. Even though we do not observe the order flow directly, we can infer from the transaction data which trades were buyer-initiated and which were seller-initiated. We use the simple rule that trades occurring above the quote midpoint, that is, above the average of best bid and ask quotes, are considered to be buys and those below the midpoint are coded as sells.

Since the parameters of the model describe the arrival rates of informed and uninformed traders and the probability that an information event will occur, their estimates can be

used to assess the probability that a transaction will be information-based. The probability of an informed trade is expressed as

$$\Pr(\text{inf}) = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} \quad (3)$$

This expression can be intuitively described as a comparison of the expected arrival of informed traders and the expected arrival of all traders. Thus, the higher the probability that information event will occur and the larger the arrival rate of the informed traders relative to the arrival rate of the uninformed, the higher the probability that a trade is motivated by knowledge of information that is not known to all market participants.

### *3.2. Estimation Results*

We have estimated the model described in Section 3.1 by maximizing the likelihood (2). We have used the ML procedure in TSP 4.4 package. Since the model uses only the number of sells and buys every day to estimate the parameters, we would have a relatively low number of observations if we used only September and October 1999 data. We have thus extended the sample by one month on each side, that is, we have estimated the model on data from the beginning of August 1999 until the end of November 1999. The number of observations stood at 85 (with minor differences among the individual stocks due to several missing data points). The estimates are depicted in Table 2.

**Table 2: Estimates of the Extent of Informed Trading (Aug. through Nov. 1999)**

Company	$\alpha$	$\delta$	$\varepsilon$	$\mu$	Prob(inf)	$\mu / \varepsilon$
ČESKÁ SPOŘITELNA	0.71 (0.13)	0.63 (0.07)	9.03 (0.26)	16.11 (0.77)	<b>0.388</b>	1.8
ČESKÉ RADIOKOMUNIKACE	0.88 (0.08)	0.47 (0.06)	12.93 (0.34)	18.88 (0.81)	<b>0.392</b>	1.5
ČEZ	0.91 (0.05)	0.57 (0.06)	15.72 (0.27)	22.58 (0.62)	<b>0.396</b>	1.4
IPS	0.20 (0.12)	0.36 (0.13)	2.19 (0.11)	11.04 (0.92)	<b>0.333</b>	5.0
KOMERČNÍ BANKA	0.85 (0.10)	0.43 (0.06)	11.47 (0.35)	17.00 (0.80)	<b>0.386</b>	1.5
RIF	0.13 (0.10)	0.23 (0.14)	2.20 (0.11)	13.88 (1.15)	<b>0.291</b>	6.3
SPT TELECOM	0.95 (0.04)	0.61 (0.06)	16.43 (0.35)	22.10 (0.77)	<b>0.390</b>	1.3
UNIPETROL	0.85 (0.10)	0.43 (0.06)	11.48 (0.35)	16.99 (0.79)	<b>0.386</b>	1.5
<b>Average</b>	<b>0.69</b>	<b>0.47</b>	<b>10.2</b>	<b>17.3</b>	<b>0.370</b>	<b>2.5</b>

*Note: Standard errors in parentheses. Prob(inf) denotes the probability of informed trading.*

The estimates of  $\alpha$  are relatively high—on average almost 0.7—suggesting that information events are frequent. There is also tendency of more frequently traded stocks to have higher  $\alpha$ —frequently traded SPT Telecom’s  $\alpha$  is close to 0.95, while the least frequently traded IPS and RIF have  $\alpha$  of 0.20 and 0.13, respectively. This may have a connection with less intensive analyst coverage of the less frequently traded stocks.<sup>10</sup> On average, the estimate of  $\delta$ , which is close to 0.5, confirms the expectations that the good and bad news should be equally likely overall. There is some variation in these estimates among individual stocks, but with the exception of RIF, all are reasonably close to 0.5.

<sup>10</sup> In a relatively illiquid environment, it might be that information events are also of a technical short term nature, for instance the information that there is a large London buyer. That could help explain the high estimates of  $\alpha$ .



The resulting probability of informed trading, denoted as  $\text{Prob}(\text{inf})$  and computed according to equation (3), varied from 0.29 for RIF to 0.40 for ČEZ, with most of the estimates being clustered just below 0.40.<sup>11</sup> On average, it stood at 0.37 and was thus considerably higher than the probability of informed trading on NYSE as estimated by Easley et al. (1996b), who estimated this probability to vary from 0.16 for the most liquid (first decile) stocks to 0.22 for the eighth decile stocks. It is worth noting that the probability of informed trading for Czech stocks is higher not only as compared to results for the fifth decile at NYSE (0.21), which is comparable also in terms of trading volume, but also as compared to the eighth decile (the lowest volume Easley et al. (1996b) used in estimation). In fact, it is virtually double. It should be noted that the average trading volume of the stocks in our sample is considerably higher than that of the eighth decile stocks used by Easley et al.

It appears that informed trading is considerably more important in the Czech Republic than at NYSE, even if we take into account the differences in trading volume. Moreover, the difference between  $\varepsilon$  and  $\mu$  (order flow of uninformed and informed traders, respectively) is greater on the Czech market—on average the ratio of  $\mu$  to  $\varepsilon$  is 2.5 in the case of Czech stocks, while it varied from 0.7 to 1.6 in the Easley et al. study. This is a feature that increases rather than decreases the variability of order flow. If the market maker was not risk neutral (as in the Easley et al. (1996b) model), this would increase the demanded compensation for market making services (and thus also the adverse selection component of the bid-ask spread).

---

<sup>11</sup> Except for the two least liquid stocks (IPS and RIF) all the estimates of informed trading probability are very close to 0.39.

A quick comparison of our results with the previous study that dealt with the extent of informed trading on the Czech market is in order. Němeček (1998) modified the Easley et al. (1996b) model to use data from the auction mechanism (at the time he wrote his study, there was no system of market makers). He found, using data from the beginning of the trading of each stock until late 1996, that the probability of informed trading varied between a negligible 0.002 and very small 0.02. We see two problems that make his estimates weak. First, it is very difficult to estimate the probability of informed trading from auction results (there are problems with identification of the model). A second, and more fundamental, problem is that at that time only a very low fraction of trades were channeled through the auction system. The actual order flow (and therefore the information it contains) was simply not observable in the auction data. Overall, we conclude that informed trading played an important role on the Czech market in the second half of 1999.

#### **4. Estimating the Size of Adverse Selection Component—Models and Estimation Method**

Several different approaches to estimation of the components of the bid-ask spread exist. In our view, the approach pioneered by Roll (1984) and developed by others, namely Stoll (1989) and George et al. (1991) is not suitable in our emerging market setting. This approach, in essence, uses the covariance of transaction prices or quotes to infer the effective spread and its components. In order to do that, market efficiency must be assumed, so that the only source of serial covariance of prices and quotes is the bid-ask

friction and the factors behind it, like adverse selection. This assumption might be warranted in the case of developed markets, but we view it as hardly justifiable in an emerging market setting. We thus prefer another class of models—those that use the trade indicator to analyze the bid-ask spread.<sup>12</sup>

#### *4.1. The Adverse Selection Component of the Bid-Ask Spread—Our Model*

We have already mentioned that the basic aim of this paper is to decompose the bid-ask spread into three components. The first is called the order processing component and is caused simply by the costs market makers incur when providing market making services—these costs include various items starting with renting office space and ending with the salaries of traders. This component does not change with short-term fluctuations of the order flow and is assumed to be constant. This appears to be a safe assumption as our sample period includes only a couple of months. The second, inventory component, is caused by the fact that market makers are exposed to the risk of taking undesired inventory when quoting firm prices. Order flow, which changes the market maker's inventory position, results in quote and price revisions, since the market maker, by changing the quotes, increases the probability of reversal in his inventory position. The third adverse selection component of the spread is induced by the fact that the market maker faces the risk of trading with an agent who has private information about the value of the asset. The market makers will inevitably lose when trading with such informed

---

<sup>12</sup> The trade indicator is usually defined as equal to 1 if the trade is considered to be a buy, -1 for a sell and 0 if the type of the trade is unclear—for instance if the trade was executed exactly at the midpoint between bid and ask quotes. Some studies do not use 0 and classify all the trades as buys or sells by using additional characteristics of the order flow, for instance the direction of previous trades.

agents and they must set a spread between their bid and ask quotes that compensates for this expected loss. We should also point out that there is a difference between the posted (quoted) spread, which is simply the difference between the bid and ask quotes, and the traded (effective) spread. The traded spread takes into account the fact that some trades are conducted within the inside quote, that is, at better prices than best bid or best ask.<sup>13</sup>

As mentioned above, we do not consider the approach started by Roll (1984) and which uses the covariance of prices or quotes as suitable for our emerging market setting. This is due to the fact that it hinges critically on the assumption of market efficiency and makes rather restrictive assumptions in regard to the probabilities of reversal (that is, the probability that a buy will be followed by a sell and vice versa). We start from the general model developed in Huang and Stoll (1997), which uses the trade indicator variable and which, if restricted in an appropriate way, can encompass most of the previous models, including the original Roll (1984) one. It should be noted that many papers before Huang and Stoll (1997) ignored one of the components of the spread.

The most serious problem of estimating all the three components of the bid-ask spread appears to be the distinction between the adverse selection and inventory components. They both lead to the same revision after a trade—an incoming buy order lowers the market maker’s inventory so he increases the quotes in order to raise the probability of a sell order which would balance his inventory position. Similarly, a buy order in an environment with nonzero probability of informed trading increases the market maker’s beliefs about the value of the asset and leads to an upward revision of quotes. Thus, buy orders lead to quote increases and sell quotes to quote decreases due to both adverse

---

<sup>13</sup> The inside quote is formed by the best bid and best ask quotes.

selection and inventory effects and it is difficult to distinguish these two effects empirically.

Our point is, however, that this truly holds only if there is just one market maker or dealer (perhaps threatened by an entry so that he or she quotes competitively) like the specialist at NYSE. In the case of a competitive dealer system (for instance NASDAQ, SEAQ in London, or our SPAD), market reaction due to inventory reasons will be different, a point which was not recognized in earlier papers. For instance, Stoll (1989) argues that NASDAQ can be treated as a one-dealer market. Stoll argues, “Competition among dealers, the desire of investors to trade with the dealer at the ‘inside’, knowledge by all dealers of quotes of other dealers, and the knowledge of transaction prices causes the inside quote [best bid and best offer] and transaction prices to behave as if there were one dealer”. [Stoll (1989) p. 123]. We disagree. Quote revisions behave differently when there are multiple dealers as compared to the case of only one specialist. Not only is the same inventory risk spread among a larger number of market makers and thus also across a broader capital base, which itself would make the inventory-induced reaction weaker, but also the behavior of dealers lowers the reaction to inventory.<sup>14</sup> First, dealers with extreme inventory positions tend to form the inside quote since they want their inventory positions to be reversed. Second, even if they had the best quotes and took an undesired inventory, they would revise the quote and thus cease to have the best quotes, but the second best quote (which is supposedly rather close in a competitive market and in the case when the inventory position of the first best quote is not so extreme that he would mind changes in his inventory) would not be affected by the change of inventory, but the

quote would change only due to the information content of the trade, assuming all market makers observe all trades.

Thus, even if the market makers who have the best quotes did not like the inventory change, only the affected market makers would revise their quotes due to inventory reasons. All other dealers would revise quotes only due to adverse selection, in other words, only due to a change in the expected fair value of the stock. They do not obtain any undesired inventory position. An argument might be made that the others would revise quotes just because they expect the dealer who obtained his undesired inventory position to try to get rid of it. This, however, is not a valid argument, since it would not be rational for them to revise their quotes differently than in line with the new expected fundamental price; why not let him unwind the position. Of course, if one knows there is a seller, one might be tempted to lower the price further and pick up the stock more cheaply. However, competition among market makers should prevent this.

The view that dealers set quotes according to their relative inventory positions has been present in the literature, though it has not been used in the estimation of the bid-ask spread components. The first paper that dealt with the issue of quote setting by competitive market makers was Ho and Stoll (1983). This model implies that inventory position will affect the placement of competing quotes, with the market makers that are at the inside quote having extreme inventory positions.<sup>15</sup> Lack of detailed data made empirical testing of this prediction impossible for almost 15 years. Only recently, two

---

<sup>14</sup> The risk-sharing feature of multiple dealer systems is also recognized in Affleck-Graves et al. (1994). We assume that dealers do not view the undesired inventory of other dealers as their own.

<sup>15</sup> In fact, Ho and Stoll (1983) solve the model explicitly only for two dealers. Also, the best quotes are not the reservation prices of the dealers with the extreme positions, but equal to the second best positions (or only marginally better). This does not affect our analysis here, though.

papers, Reiss and Werner (1998) and Hansch et al. (1998) tested the crucial predictions of this and other inventory models explicitly. These two papers use detailed data from the London Stock Exchange, which has been working as a multiple dealer market, to test the hypothesis that the positioning of quotes by one market maker relative to that of other market makers is a function of their inventories. Hansch et al. (1998) found that changes in quotes and inventories are strongly correlated and that standardized and relative inventories are mean reverting. They also report that market makers posting competitive quotes execute a significantly larger proportion of public trades (the data they use offer the distinction between interdealer and public trades) so competitive quoting indeed attracts more trade flow. This supports the view that the dealers with the best quotes are those with extreme inventory positions. Further evidence in Hansch et al. (1998) and the paper by Reiss and Werner (1998) suggest that dealers engage in interdealer trading primarily when their inventory position becomes extreme. This lends support to the basic view that a dealer with an undesired inventory position has two options: either to make his quotes the best on the market and wait for a client order (which is more profitable, but uncertain) or to accept a worse price, but unwind the position quickly in an interdealer trade. When the undesired position becomes large, it is too risky to wait for a public order.

In fact, this weak inventory-induced response of the competitive dealer system might explain the results of Stoll (1989), who used NASDAQ data and found only a small inventory component (10% of the spread). This finding was also one of the arguments George et al. (1991) use to defend their assumption that the inventory component of the spread does not exist—however, they use this assumption also for NYSE stocks. This

would be inappropriate if indeed (as we argue) the competitive dealer system response to inventory pressure is different from that of a specialist system.

In view of the above discussion on inventories, our approach to the inventory component of the spread is the following. We assume that under normal circumstances (what is normal is yet to be defined) the revision of quotes due to inventory reasons in a response to a trade is negligible since those who have the best quotes actually want the inventory and there is no reason to expect pressure from interdealer trading. If, however, the circumstances are not normal and there is thus substantial selling or buying pressure, several dealers accumulate undesired inventory and use interdealer trades to unwind the positions. Then, inventory reasons for quote revision might become important also for those who are not directly affected by the current trade (even though in a world with no frictions this should not happen). We thus create a binary variable (denoted as PRESS), which indicates whether a particular trade happened in a period of selling or buying pressure. We take a moving window of 10 trades prior to the particular trade and, given the number of sells and buys, determine whether there was trading pressure. We have chosen three different levels for this variable—we require the cumulative variable to exceed 4, 6 and 8 sells or buys, which selects almost 50%, 20–30%, and approximately 10% of all trades, respectively, as being conducted under selling or buying pressure.<sup>16</sup> It should be noted that these frequencies do not fully correspond to the frequencies we would expect if the trades came from an independent binomial distribution with equal

---

<sup>16</sup> This method is similar to the second approach of Huang and Stoll (1997) to the three way decomposition of the spread, which is based on portfolio trading pressure. They use an aggregate (cumulative) sell-buy indicator. Their approach is used for NYSE stocks, however, and is thus questionable, since it is difficult to argue that the specialist should react to inventory changes only when there is overall selling and buying pressure. Also, the selection of stocks that should be indicative of trading pressure is questionable and Huang and Stoll do not support this approach very strongly.



probability of buy and sell, which would be 51%, 9%, and only 0.2%, respectively. This is caused by the fact that our “draws” are not independent and, if there is indeed something like trading pressure, we would expect trades in a similar direction to be clustered (which is indeed the case since the frequencies in our data are higher than the theoretical frequencies).

Further, we assume that the traded spread is constant, which we believe is a reasonable assumption given that the time period under study is only two months. Therefore,

$$P_t - M_t = S/2 * Q_t + \varepsilon_t, \quad (4)$$

where  $S$  is the constant traded spread,  $P_t$  is the transaction price at time  $t$ ,  $M_t$  is the quote midpoint (average between best bid and ask quotes at the time of trade) at time  $t$ , and  $Q_t$  is the trade indicator variable, which equals 1 if the trade is a buy, -1 if it is a sell, and 0 if it occurs exactly at the midpoint. We assume that the error term  $\varepsilon$  has zero mean conditional on  $\mathbf{Q}$ .

The key part of the model we estimate is based on the model of Huang and Stoll (1997). In order to save notation (below we estimate the actual Huang and Stoll (1997) model for comparison) we use the same notation as in their paper. First, the unobservable fundamental value of the security at time  $t$ ,  $V_t$ , is driven by the new information in the most recent trade (from time  $t-1$ ) as indicated by the trade indicator variable  $Q_t$  and by several additional characteristics of the trade:

$$V_t = V_{t-1} + \alpha * S/2 * Q_{t-1} + \delta * \text{CROSS}_t + \lambda * \text{ACTIVE}_t + \eta_t .$$

The variable  $\text{CROSS}$  indicates whether the trade was a so-called cross trade, that is, a trade between the own accounting book of a dealer and its client. Such trades naturally

are not originated in the SPAD system, but they must be reported such that the market is aware of the order flow. The rationale for including this variable into the model is that such trades might be viewed by the market as having a different information value relative to normal trades. It is not clear, however, what sign the coefficient  $\delta$  should have, since the market might view cross trades as transactions that will not affect the market (then cross trades would have lower information value) or as an indication of stronger order flow (then the information value would be higher). We have also identified the most active market makers and created the variable ACTIVE. This variable equals 1 if the trade was initiated by one of the most active market makers—that is, if the trade is a buy, one of the most active market makers was on the buy side of the trade and vice versa. We consider the top five market makers in terms of trading volume to be the most active. They account for almost 50% of trading volume. We include this variable because a trade initiated by the most active market makers might have a higher information value for other market participants. We assume that the error term  $\eta$  has zero mean conditional on the right hand side variables.

The actual quotes that we observe are affected by inventory effects, however, as discussed above, we assume that inventory matters only when there is significant buying or selling pressure.<sup>17</sup> Thus, the quote midpoint  $M_t$  is given by

$$M_t = V_t + \sum_{i=0}^{t-1} \beta \frac{S}{2} Q_i PRESS_i .$$

Taking first differences and substituting for  $V_t$  yields

---

<sup>17</sup> The treatment of the inventory component is the main difference of our model from that of Huang and Stoll (1997). It is important, since it allows us to use a rather simple estimation technique. Our estimation

$$\Delta M_t = \alpha * S/2 * Q_{t-1} + \beta * S/2 * Q_{t-1} * PRESS_{t-1} + \delta * CROSS_t + \lambda * ACTIVE_t + \eta_t. \quad (5)$$

We estimate the equations (4) and (5) to obtain the estimates of  $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\lambda$ , and  $S/2$ . We estimate the two equations simultaneously by the SUR routine in the TSP 4.4 package. We also test for both autocorrelation and heteroskedasticity.

#### 4.2. *Huang and Stoll (1997) Model*

For the sake of comparison, we also estimate one of the two models of a three-way decomposition of the spread presented in Huang and Stoll (1997). Our model presented above is based on the general Huang and Stoll model, and we retain most of the notation to save space and lower confusion. Huang and Stoll (1997) base their model on three equations. First, the unobservable fundamental value of the stock  $V_t$  develops according to

$$V_t = V_{t-1} + \alpha * S/2 * Q_{t-1} + \varepsilon_t. \quad (6)$$

Second, they assume that every trade affects the inventory position of the market maker (they use NYSE data, so they apparently have the specialist in mind) and thus the position of the quote midpoint relative to the current fundamental value of the stock is

$$M_t = V_t + \sum_{i=0}^{t-1} \beta \frac{S}{2} Q_i. \quad (7)$$

Third, they assume the spread to be constant, with the random factor in the following equation reflecting the deviation of the observed spread from its constant value:

---

results and their comparison with the results of the Huang and Stoll (1997) approach, which are both presented below, appear to support our method.

$$P_t - M_t = S/2 * Q_t + \eta_t \quad (8)$$

Combining equations (6), (7) and (8) and taking first differences yields the basic regression model

$$\Delta P_t = S/2 * (Q_t - Q_{t-1}) + \lambda * S/2 * Q_{t-1} + e_t \quad (9)$$

where  $\lambda = \alpha + \beta$  and thus the combined inventory and adverse selection component of the model is estimated. This is the first model of Huang and Stoll (1997) that we estimate on our data.

Then, Huang and Stoll offer two ways of fully decomposing the spread, i.e., two ways to separate the inventory and adverse selection components. First, they use the serial correlation in trade flows. Second, they use the concept of portfolio trading pressure that we discussed above and that is similar to our approach. Since the concept of portfolio trading pressure as described by Huang and Stoll is disputable, especially as far as the specialist trading model is concerned, we have chosen to estimate the first model, which is based on the serial correlation of trade flows.

This way of differentiating between the adverse selection and inventory components is based on the observation that, given past order flow, the market forms an expectation about the next order that is to come to the market. The conditional expectation of the trade indicator at time t-1 given  $Q_{t-2}$  is

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

where  $\pi$  is the probability of trade reversal, i.e., the probability that the sign of the next trade will be opposite. If we assume that the market knows equation (10), it will take this expectation into account and the development of the fundamental value will be driven by

$$\Delta V_t = \alpha * S/2 * Q_{t-1} - \alpha * S/2 * (1-2\pi) Q_{t-2} + \varepsilon_t . \quad (11)$$

Combining this with equation (7), which describes the impact of inventory on the position of the mid-quote, yields

$$\Delta M_t = (\alpha + \beta) * S/2 * Q_{t-1} - \alpha * S/2 * (1-2\pi) Q_{t-2} + \varepsilon_t . \quad (12)$$

Furthermore, adding the constant spread assumption yields the final expression

$$\Delta P_t = S/2 * Q_t + (\alpha + \beta - 1) * S/2 * Q_{t-1} - \alpha * S/2 * (1-2\pi) Q_{t-2} + e_t \quad (13)$$

For the purposes of estimation, equation (10) can be rewritten as

$$Q_{t-1} = (1-2\pi) Q_{t-2} + \xi_t \quad (14)$$

and we assume that the error term  $\xi$  has zero mean conditional on  $Q_{t-2}$ .

By estimating equations (13) and (14) we are able to distinguish between the adverse selection ( $\alpha$ ) and the inventory component of the spread ( $\beta$ ). We estimate these equations simultaneously, again, using the SUR routine of the TSP 4.4 package.

## **5. Size of Adverse Selection Component—Results of Estimation**

### *5.1. The Adverse Selection Component of the Bid-Ask Spread—Our Model*

We report the results for the middle choice of the variable PRESS, that is, the case in which we require the cumulative trade indicator to exceed a value of 6, in Table 4. The estimates for the other two choices of the PRESS variable were not substantially different (except for the coefficient at the inventory variable itself). A full set of results can be found in the Appendix. We present estimates for the two-month period—for September and October 1999, since here we have the advantage of being able to use the ACTIVE

variable for those trades initiated by the most active dealers. We have also estimated the model for a four month period–August through November 1999–without this variable and the results were not significantly different.

The adverse selection component, which is the primary focus of this paper, ranges from 8% in the case of Unipetrol to 30% for IPS. It is not surprising that the highest share of adverse selection is found for the stock (IPS) that has the highest, almost 50%, probability of informed trading. The most frequently traded stock, SPT Telecom, exhibited a 12% share of adverse selection component in the spread. It might be surprising, but, on average, the size of the adverse selection component is relatively low, at 14%. In all eight cases (with the exception of RIF), the adverse selection component was significant at the 5% significance level.

**Table 4: Results of Estimation (for PRESS 6)**

Company (standard errors in parentheses)	S/2	$\alpha$	$\beta$ (inv. 6)	$\delta$ (Cross)	$\Lambda$ (Active)	NOB	R <sup>2</sup> (eq. 4)	R <sup>2</sup> (eq. 5)
ČESKÁ SPORITELNA	1.56 <sup>a</sup> (0.04)	0.17 <sup>a</sup> (0.03)	0.01 (0.05)	0.14 <sup>c</sup> (0.07)	-0.03 (0.05)	630	0.67	0.09
ČESKÉ RADIOKOMUNIKACE	6.11 <sup>a</sup> (0.16)	0.12 <sup>a</sup> (0.03)	0.07 (0.05)	0.15 (0.36)	-0.16 (0.23)	1,084	0.59	0.03
ČEZ	0.30 <sup>a</sup> (0.01)	0.13 <sup>a</sup> (0.02)	0.04 (0.03)	0.02 <sup>c</sup> (0.01)	0.03 <sup>a</sup> (0.01)	1,429	0.54	0.09
IPS	1.36 <sup>a</sup> (0.08)	0.30 <sup>a</sup> (0.07)	-0.08 (0.09)	0.36 <sup>b</sup> (0.14)	0.002 (0.10)	121	0.70	0.24
KOMERČNÍ BANKA	4.78 <sup>a</sup> (0.12)	0.10 <sup>a</sup> (0.01)	-0.001 (0.03)	0.12 (0.17)	0.30 <sup>a</sup> (0.09)	1,043	0.59	0.08
SPT TELECOM	2.04 <sup>a</sup> (0.03)	0.12 <sup>a</sup> (0.01)	0.08 <sup>a</sup> (0.02)	-0.08 (0.05)	0.10 <sup>a</sup> (0.03)	1,822	0.66	0.12
UNIPETROL	0.48 <sup>a</sup> (0.01)	0.08 <sup>a</sup> (0.02)	0.10 <sup>a</sup> (0.03)	-0.01 (0.02)	0.02 <sup>c</sup> (0.01)	791	0.67	0.08

Note: PRESS 6 denotes the case when we require the cumulative trade indicator to exceed 6 for the selling pressure to exist. The results for other definitions of the PRESS variable can be found in the Appendix. <sup>a</sup> Stands for rejection of the null hypothesis (that the coefficient equals zero) at the 1% level of significance, while <sup>b</sup> and <sup>c</sup> denote rejection at the 5% and 10% levels, respectively. In the case of RIF, we have encountered problems with the convergence of estimates, the reason might be rather low number of observations; we obtained implausible estimates and do not report them here.

As for the inventory component,  $\beta$ , the results suggest that our assumptions on inventory impact were justified. It increases from the least strict choice (which counts almost one half of all trades as occurring under trading pressure) to the most strict choice (which classifies only approximately 10% of trades in this way). Moreover, the statistical significance increases in this direction as well. While some of the coefficients at the least strict variables are not statistically significant, those at the most strict are mostly highly significant. The adverse selection component is small, on average it amounted to only 5%. All these facts suggest that the inventory component is not very important in the Czech market. The vast majority of the spread—on average some 80%—is thus formed by the order processing component.

As for the other explanatory variables, it appears that in half of the cases (SPT Telecom, ČEZ, Komerční banka and, to a limited extent, Unipetrol) it matters to the market that the trade is initiated by an active market maker. Interestingly, these are the four stocks with the highest turnover. The significant coefficients are positive, which implies that a trade initiated by a major market maker is viewed to have higher information value than other trades. The coefficient at the CROSS variable is significant only occasionally and it therefore appears that the market does not attach much significance to the information that a trade is of the cross type.

The traded spread is smaller than the posted spread, which is in line with the findings of previous studies. The traded spread as a fraction of the quoted spread is rather uniform across the stocks in our sample—in all cases (except for RIF) the traded spread fits into the range of 70–80% of the quoted spread.

## 5.2. Huang and Stoll (1997) Model

The results of estimating the Huang and Stoll (1997) models are depicted in the following table. In the table, we have combined the estimate of  $\lambda$  from equation 9 with the results of from joint estimation of equations 10 and 13.

**Table 5: Results of the Huang and Stoll (1997) model**

Company	$\lambda$ (from eq.9)	$\alpha$	$\beta$	$R^2$ (eq.10)	$R^2$ (eq.13)	N
ČESKÁ SPORITELNA	0.18 <sup>a</sup> (0.05)	-0.02 (0.10)	0.20 <sup>b</sup> (0.08)	0.29	0.39	629
ČESKÉ RADIOKOMUNIKACE	0.14 <sup>a</sup> (0.04)	0.13 (0.09)	0.05 (0.08)	0.24	0.35	1,083
ČEZ	0.23 <sup>a</sup> (0.03)	0.06 (0.08)	0.18 <sup>a</sup> (0.07)	0.23	0.36	1,428
IPS	0.38 <sup>a</sup> (0.13)	-0.03 (0.26)	0.41 <sup>b</sup> (0.19)	0.47	0.21	120
KOMERČNÍ BANKA	0.17 <sup>a</sup> (0.03)	0.09 (0.08)	0.10 (0.07)	0.19	0.47	1,042
RIF	0.25 <sup>a</sup> (0.09)	-0.25 (0.20)	0.40 <sup>a</sup> (0.15)	0.35	0.30	205
SPT TELECOM	0.20 <sup>a</sup> (0.02)	0.14 <sup>a</sup> (0.05)	0.10 <sup>b</sup> (0.05)	0.24	0.48	1,821
UNIPETROL	0.17 <sup>a</sup> (0.04)	0.28 (0.09)	-0.04 (0.08)	0.23	0.48	790

Note: <sup>a</sup> Stands for rejection of the null hypothesis (that the coefficient equals zero) at the 1% level of significance, while <sup>b</sup> and <sup>c</sup> denote rejection at the 5% and 10% levels, respectively.

While the estimates of the combined share of adverse selection and inventory components appear to be reasonable and correspond to the estimates implied by our model above (they average 0.22), the distinction between these two components corresponds to our results in only a few cases (one of them is SPT Telecom). Moreover, some of the estimates of  $\alpha$  are negative, which is impossible in our model.<sup>18</sup> Adding more

---

<sup>18</sup> In fact, Huang and Stoll (1997) also obtained negative estimates of  $\alpha$  initially. They argued, then, that it is due to large orders being executed as a series of small trades executed in a short time period. They thus combined all sequential trades at the same price and quotes into one trade. We do not consider this to be a problem here, especially since the trading frequency is considerably lower for our stocks.



data points by extending the sample by another two months does not change the results to any significant extent. Overall, we believe the estimates of  $\lambda$ , the combined share, as it is robust since it comes as a product of a simple and straightforward model, and view the results as a confirmation of the fact that adverse selection is not a major component of the spread—its combined share with the inventory component is not estimated to be very large.

## **6. Conclusion**

Abundant literature links informed trading and the size of the bid-ask spread and some papers express the concern that a large amount of informed trading might substantially limit the liquidity of the market or even lead to the shutdown of the market. The current literature is not settled on the issue of the relative importance of the three components—adverse selection, inventory and order processing ones—of the bid ask spread. We provide evidence on this issue, with a focus on the adverse selection component, from an emerging market that has been known to be driven by insider, or at least informed, trading. Our main objective is to examine the share of the adverse selection component with the eventual aim to establish the upper limit of its real-world significance. Should this share also be small on a market, where informed trading is substantially more widespread than on developed markets, it would confirm that the adverse selection component is relatively unimportant.

First we test whether the Czech equity market is indeed plagued by informed trading. It turns out that the probability of informed trading is indeed considerably higher on the Czech market as compared to the developed markets. Moreover, if an information event

occurs, the arrival rate of informed traders is significantly larger than the underlying arrival rate of uninformed traders, which further increases the risk market makers face. They should thus be expected to require higher compensation for this risk and increase the adverse selection component of the spread.

We then estimate the components of the bid-ask spread. We use recent findings of studies that tested the inventory behavior of dealers in multiple dealer trading systems to modify the general Huang and Stoll (1997) model. Earlier studies treated the systems of market makers the same as single-dealer systems. However, there is a substantial difference between the inventory impact in a single dealer (specialist) system and a multiple dealer system. First, multiple dealers have a higher capital base and thus also higher risk-taking capacity than a single dealer. Second, the positioning of quotes relative to other dealers depends on the current inventory level, which means that many trades (even though not initiated by the market maker) do not bring undesired inventory, but rather bring the inventory to the desired level. Therefore, in our model, inventory impacts quotes only in times of trading pressure. In fact, our results confirm that inventory is of rather low importance even in times of trading pressure.

Our estimates of the bid-ask spread components suggest that the adverse selection component amounts only to 14%, on average, which is rather low and comparable with the recent estimates from developed markets. It appears that the relative importance of the adverse selection component is low—even on a market, which is significantly driven by informed trading. Our findings thus confirm that adverse selection is relatively unimportant as a determinant of the bid-ask spread across markets.

## References

- Affleck-Graves, John, Shantaram P. Hedge and Robert E. Miller. 1994. "Trading Mechanisms and the Components of the Bid-Ask Spread". *Journal of Finance*. September 1994, 49 (4), pages 1471–1488.
- Easley, David, Nicholas M. Kiefer and Maureen O'Hara. 1996a. "Cream-Skimming of Profit-Sharing? The Curious Role of Purchased Order Flow". *Journal of Finance*. July 1996, 51 (3), pages 811–833.
- Easley, David, Nicholas M. Kiefer, Maureen O'Hara and Joseph B. Paperman. 1996b. "Liquidity, Information, and Infrequently Traded Stocks". *Journal of Finance*. September 1996, 51 (4), pages 1405–1436.
- Easley, David, Nicholas M. Kiefer and Maureen O'Hara. 1997. "One Day in the Life of a Very Common Stock". *Review of Financial Studies*. Fall 1997, 10 (3), pages 805–835.
- George, Thomas J., Gautam Kaul and M. Nimalendran. 1991. "Estimation of the Bid-Ask Spread and Its Components: A New Approach". *Review of Financial Studies*. 4 (4), pages 623–656.
- Glosten, Lawrence and Paul Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders". *Journal of Financial Economics*. 14, pages 71–100.
- Hansch, Oliver, Narayan Y. Naik, and S. Viswanathan. 1998. "Do Inventories Matter in Dealership Markets? Evidence from the London Stock Exchange". *Journal of Finance*. October 1998, 53 (5), pages 1623–1656.
- Ho, Thomas and Hans R. Stoll. 1980. "On Dealer Markets Under Competition". *Journal of Finance*. May 1980, 35 (2), pages 259–267.
- Ho, Thomas and Hans R. Stoll. 1983. "The Dynamics of Dealer Markets under Competition". *Journal of Finance*. September 1983, 38 (4) , pages 1053–74.
- Huang, Roger D. and Hans R. Stoll. 1997. "The Components of the Bid-Ask Spread: A General Approach". *Review of Financial Studies*. Winter 1997, 10 (4), pages 995–1034.
- Němeček, Libor. 1998. "Asymmetric Information and Insider Trading". In: Libor Němeček, PhD. Dissertation, CERGE-EI. Prague, August 1998.
- O'Hara, Maureen. 1995. *Market Microstructure Theory*. Basil Blackwell: Cambridge MA.
- Reiss, Peter C. and Ingrid M. Werner. 1998. "Does Risk Sharing Motivate Interdealer Trading?". *Journal of Finance*. October 1998, 53 (5), pages 1657–1703.
- Stoll, Hans R. 1989. "Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests". *Journal of Finance*. March 1989, 44 (1), pages 115–134.
- World Bank Country Study: Czech Republic – Toward EU Accession*. 1999. World Bank: Washington, D.C.

## Appendix

**Table A1: Results of Estimation (for PRESS 4)**

Company (standard errors in parentheses)	S/2	$\alpha$	$\beta$ (inv. 6)	$\delta$ (Cross)	$\lambda$ (Active)	N	R <sup>2</sup> (eq. 4)	R <sup>2</sup> (eq. 5)
ČESKÁ SPORITELNA	1.56 <sup>a</sup> (0.04)	0.18 <sup>a</sup> (0.03)	-0.02 (0.04)	0.14 <sup>c</sup> (0.07)	-0.03 (0.05)	630	0.67	0.09
ČESKÉ RADIOKOMUNIKACE	6.11 <sup>a</sup> (0.16)	0.09 <sup>a</sup> (0.03)	0.09 <sup>c</sup> (0.05)	0.18 (0.36)	-0.18 (0.24)	1,084	0.59	0.03
ČEZ	0.30 <sup>a</sup> (0.01)	0.13 <sup>a</sup> (0.02)	0.03 (0.03)	0.02 <sup>c</sup> (0.01)	0.03 <sup>a</sup> (0.01)	1,429	0.54	0.09
IPS	1.36 <sup>a</sup> (0.08)	0.30 <sup>a</sup> (0.08)	-0.07 (0.09)	0.37 <sup>b</sup> (0.14)	0.01 (0.10)	121	0.70	0.24
KOMERČNÍ BANKA	4.78 <sup>a</sup> (0.12)	0.11 <sup>a</sup> (0.02)	-0.03 (0.03)	0.13 (0.16)	0.31 <sup>a</sup> (0.09)	1,043	0.59	0.08
SPT TELECOM	2.04 <sup>a</sup> (0.03)	0.13 <sup>a</sup> (0.01)	0.04 <sup>b</sup> (0.02)	-0.08 (0.05)	0.10 <sup>a</sup> (0.03)	1,822	0.66	0.11
UNIPETROL	0.48 <sup>a</sup> (0.01)	0.07 <sup>a</sup> (0.02)	0.07 <sup>b</sup> (0.03)	-0.01 (0.02)	0.02 <sup>c</sup> (0.01)	791	0.67	0.08

**Table A2: Results of Estimation (for PRESS 8)**

Company (standard errors in parentheses)	S/2	$\alpha$	$\beta$ (inv. 6)	$\delta$ (Cross)	$\lambda$ (Active)	N	R <sup>2</sup> (eq. 4)	R <sup>2</sup> (eq. 5)
ČESKÁ SPORITELNA	1.56 <sup>a</sup> (0.04)	0.15 <sup>a</sup> (0.02)	0.12 <sup>b</sup> (0.06)	0.13 <sup>c</sup> (0.07)	-0.03 (0.05)	630	0.67	0.10
ČESKÉ RADIOKOMUNIKACE	6.11 <sup>a</sup> (0.16)	0.12 <sup>a</sup> (0.03)	0.13 <sup>c</sup> (0.07)	0.12 (0.36)	-0.17 (0.23)	1,084	0.59	0.03
ČEZ	0.30 <sup>a</sup> (0.01)	0.13 <sup>a</sup> (0.02)	0.08 <sup>b</sup> (0.04)	0.02 <sup>c</sup> (0.01)	0.03 <sup>a</sup> (0.01)	1,429	0.54	0.10
IPS	1.36 <sup>a</sup> (0.08)	0.25 <sup>a</sup> (0.06)	0.02 (0.09)	0.37 <sup>b</sup> (0.14)	0.01 (0.10)	121	0.70	0.23
KOMERČNÍ BANKA	4.78 <sup>a</sup> (0.12)	0.10 <sup>a</sup> (0.01)	0.05 (0.05)	0.12 (0.16)	0.30 <sup>a</sup> (0.09)	1,043	0.59	0.08
SPT TELECOM	2.04 <sup>a</sup> (0.03)	0.13 <sup>a</sup> (0.01)	0.16 <sup>a</sup> (0.03)	-0.08 (0.05)	0.09 <sup>a</sup> (0.03)	1,822	0.66	0.12
UNIPETROL	0.48 <sup>a</sup> (0.01)	0.10 <sup>a</sup> (0.02)	0.10 <sup>b</sup> (0.05)	-0.01 (0.02)	0.02 <sup>c</sup> (0.01)	791	0.67	0.08

Note: PRESS 4 and 8 denotes the cases, when we require the cumulative trade indicator to exceed 4 and 8, respectively, for the selling pressure to exist. <sup>a</sup> Stands for rejection of the null hypothesis (that the coefficient equals zero) at the 1% level of significance, while <sup>b</sup> and <sup>c</sup> denote rejection at the 5% and 10% levels, respectively. In the case of RIF, we have encountered problems with the convergence of estimates, the reason might be rather low number of observations; we obtained implausible estimates and do not report them here.