Price Impact and Spread: Application of Bias-Free Estimation Methodology to Portfolio Transitions

Anna Obizhaeva*

First Draft: February 24, 2008
This Draft: September 25, 2008

Abstract

The analysis of price impact and spread is a challenging task due to endogeneity issues and data limitations. To get around these problems, I develop a bias-free methodology for estimating the parameters of trading costs and apply it to a unique data set of portfolio transition trades. I provide the estimates of price impact and effective spread in traditional markets as well as in crossing networks for a period 2001 through 2005. These estimates vary across securities. I find that price impact increases with the stocks' overall trading volume and volatility; in contrast, effective spread decreases with these characteristics. For thinly traded securities, trading costs are almost invariant with respect to a trade size and spread-related payments account for their largest fraction. For actively traded securities, trading costs are very sensitive to a trade size and spread-related payments are less significant. Positive association between price impact and trading volume is counterintuitive and not easily explained within existing market microstructure models. I outline potential explanations that might underly these patterns.

Keywords: price impact, effective spread, trading costs.

JEL classification: G14

---

*Robert H. Smith School of Business, University of Maryland, e-mail:obizhaeva@rhsmith.umd.edu. This paper benefited a lot from my discussions with Pete Kyle. I am also very grateful to Ross McLellan and Simon Myrgren for their valuable help as well as to Mark Kritzman for his everlastingly kind support. All errors are my own.

1 Introduction

Price impact and spread are important variables that are of interest to both researchers and practitioners. Researchers are interested in learning about these variables because they reflect a complicated and not fully understood interplay between main blocks of market microstructure: trades, prices, and information. Practitioners are keen to know about these variables because they account for the largest part of their trading costs. Yet, the estimation of price impact and effective spread is a challenging task. The key question is how to account for endogeneity in trading strategies, i.e. price-dependent strategies or canceled orders, which may induce biases in conventional estimates. In this paper, I develop a bias-free estimation methodology and apply it to a unique data set of portfolio transition trades to get the unbiased estimates of trading costs. Then, I study how these estimates vary across securities with different characteristics.

In a conventional approach, the price impact and the effective spread are estimated using ex post trading costs of *executed* trades (e.g. Almgren et al. (2005), Breen, Glosten and Harris (1988), Glosten and Harris (1988), and Keim and Madhavan (1997)). This approach disregards the opportunity cost of *unexecuted* trades. Since trading strategies of market participants usually depend on the recent price dynamics and, consequently, the ex ante expected opportunity cost of unexecuted trades is not equal to zero, ignoring unexecuted trades may lead to biased estimates. This endogeneity problem is relevant in the following cases: (1) if only parts of entire ex ante orders rather than total ex ante orders are considered, e.g., when canceled trades are ignored or when the trading costs of a given day in multi-day trading packages are evaluated, (2) if the trading strategy is price contingent, e.g., when market participants use limit orders. For instance, if traders slow down their trading when security prices run away from them, then conventional estimates of the price impact are biased downwards. In contrast, if traders speed up their trading during unfavorable market conditions, then these estimates are biased upwards.

I develop a bias-free estimation procedure. It is based on the mark-to-market implementation shortfall that includes not only ex post trading costs of *executed* trades but also the opportunity costs of *unexecuted* trades (Perold (1988)). The main idea behind this method is that expected mark-to-market implementation shortfall is represented as the difference between expected execution costs of a total ex ante orders and the expected execution costs of unexecuted part. This representation allows me to separate price impact from opportunity cost and to obtain unbiased estimates of trading costs parameters. Comparing these estimates to the conventional ones, I show that biases of conventional estimates might be substantial.

The key input variable into my estimation procedure is a total ex ante orders. Initial intentions of market participants are usually unavailable for researchers, who focus on a sample of realized trades, e.g. TAQ data. However, I employ a data set that contains information on the size and direction of ex ante orders. My analysis is based on the execution data for orders that have been carried out by a leading provider of portfolio transition services during a period 2001 through 2005. Portfolio transitions are sizable and expensive transfers of funds from legacy portfolio to target portfolios, which are often undertaken by institutional investors, e.g. pension plans, insurance funds, endowments, and foundations. The composition of securities in legacy and target portfolios is fixed prior to the beginning of portfolio transition, and the ex post transition orders coincide with ex ante ones. Effectively, I know ex ante transition orders and can apply a bias-free estimation methodology.

I distinguish between trading costs at various trading platforms. Liquidity pools are known to be fragmented, and market participants, including portfolio transition managers, usually split their orders over different trading platforms such as conventional markets, external and internal crossing networks. I provide separate estimates of the effective spread for those trading platforms, however my methodology does not allow me to provide separate estimates of the price impact for these trading venues.

I define the price impact as the percentage change in the security prices (in percents of the daily return standard deviation) in response to a trade equivalent to 1% of the average daily volume. My unconditional estimates of this variable can be summarized as follows. On average, 1% of the average daily trading volume induces the price impact equal to 0.33% of the daily returns standard deviation. The price impact of purchases tends to be slightly larger than that of sales. The price impact does not significantly differs for the NYSE/Amex-listed and the Nasdaq-listed securities.

I define the effective spread as the doubled fixed trading costs (in percents of the daily return standard deviation) of establishing either a long or a short $1-position. Its unconditional estimates can be summarized as follows. For both open markets and external crossing networks, the effective spread correspond to a fixed cost of roughly 13% of the daily returns standard deviation. The existence of the effective spread on the crossing networks contradicts a claim that these networks help avoid the bid-ask spread. I also find that trading against internal pools of liquidity does not involve statistically significant spread-related costs. Moreover, the effective spread tends to be larger for the Nasdaq-listed securities rather than for the NYSE/Amex-listed ones. This finding is consistent with the overall conclusion of the literature that trading costs for Nasdaq issues are historically higher than those for NYSE issues (e.g. Christie and Schultz (1994), Huang and Stoll (1996), Keim and Madhavan (1997)). Although these differences in bid-ask spread have declined in a post-reform period, according to the NYSE Market Quality report (2006), they still exist during the period under consideration.

I also study how the parameters of trading costs vary across securities. First, I document a relation between the trading costs and the average daily volume: transactions in more actively traded stocks are associated with lower effective spread and with higher price impact. The effective spread on both open markets and external crossing networks (in percents of the return standard deviation) drops from roughly 30% for the thinly traded stocks to about 6% for the actively traded ones. In contrast, the price impact increases with average daily volume as much as tenfold. For example, 1% of the average daily volume triggers a percentage price change of 0.27% of a daily return standard deviation for thinly traded stocks and 2.99% for actively traded ones. For the former, trading costs are almost invariant with respect to a trade size and spread-related payments account for their largest fraction. For the latter, trading costs are very sensitive to a trade size and spread-related payments are less significant. Positive association between price impact and trading volume is in a sharp contrast to a conventional intuition that trading in larger stocks involves less risk of information asymmetry and, consequently, requires slighter price concessions. I discuss further potential explanations for these patterns.

Second, I analyze the relation between the trading costs and the volatility. I find that transactions in more volatile stocks are associated with larger effective spread, when controlling for the differences in trading volume. For instance, if I consider open markets for the securities with the largest trading volume, the most volatile stocks (the least volatile stocks) have the effective spread of about $2 \times 15\%$ ($2 \times 2\%$) of the daily return standard deviation. The relation between the price impact and volatility is less unambiguous. However, the cross-sectional regression analysis suggests that it is weakly positive.

Third, I document how the price impact and the effective spread depend on the conventional proxies of information asymmetry. For instance, there is an inverse relation between the price impact and the effective spread. As the average daily volume goes up, the former increases whereas the latter decreases. These findings call for a deeper understanding of the connection between various aspects of liquidity and furthermore their association with the asset prices.

Fourth, the market microstructure literature has recognized the fact that the trading costs might asymmetrically depend on the past price dynamics (e.g., Saar(2001), Chiyachantana et al. (2006)). I do not observe a statistically significant relation between the pre-transition stock returns and the price impact. At the same time, I find that the effective spread on purchases decreases and the effective spread on sales increases during transactions in stocks with high pre-transition returns. I discuss potential explanations for these patterns later in the paper; for instance, pre-transition returns might indicate the novelty of information contained in transition orders.

I outline several mechanisms that could explain a counterintuitive positive relation between the dollar trading volume and the price impact. The first explanation is a different degree of competition among market participants. For actively traded securities, the total trading volume is large but its substantial part is two-sided, buys alternates with sells; thus, the resulting order imbalances are small. For these stocks, a one-sided trade equivalent to 1% of the average daily volume represents a distinctive and informative event. In contrast, for thinly traded securities, the order flow is small but it tends to be one-sided, and so the resulting order imbalance are large. For these stocks, a 1%-trade does not correspond to a particularly informative event. The second explanation is a potential concavity of price impact functions. Trading in more actively traded securities might seem to be more expensive, since a typical magnitude of portfolio transition orders (as a fraction of order flow) in these stocks is significantly lower (e.g., Almgren et al. (2005), Hasbrouck (1991a), Hausman, Lo, and MacKinlay (1992), and Barra (1997)). Obizhaeva (2008b) shows that even after adjusting for non-linearity, price impact still increases with trading volume. Alternatively, it might be misleading to compare the price impact of securities with different trading volume and volatility as measured over the same time horizon; instead, the proper scaling adjustment has to be used. This line of research is further developed in Kyle and Obizhaeva (2008).

This paper is related to the strand of literature that examines the estimates of the trading costs and their cross-sectional properties. The examples include Almgren et al. (2005), Breen, Hodrick and Korajczyk (2002), Chan and Lakonishok (1993, 1995), Chen, Stanzl and Watanabe (2001), Dufour and Engle (2000), Glosten and Harris(1988), Hasbrouck (1991 a,b), Holthausen, Leftwich and Mayers (1987, 1990), Lillo, Farmer and Mantegna (2003), Madhavan, Richardson, and Roomans (1997). Aforementioned papers rely on ex post trading data and, consequently, report potentially biased estimates. The comparison of estimates across various studies is difficult due to diverse estimation designs and different time periods. However, my estimates of trading costs seem to slightly lower than those of earlier studies. For instance, Breen, Glosten and Harris (2002), Hasbrouck (1991 a,b), Glosten and Harris (1988) find that the average price impact of a 1000-share trade ranges from 30 bps to 18 bps, whereas in my study, the average price impact of this trade is 28 bps for thinly traded stocks and 8 bps for actively traded stocks.[1]

The paper is structured as follows. Section 2 introduces the parameters of interest,

[1]These numbers are derived from my estimates of price impact and a typical half spread for small stocks (0.27% and 14%) and for large stocks (2.99% and 1%) given that, in the sample of portfolio transitions, 1000 shares roughly corresponds to 2% of the average daily trading volume and the average daily standard deviation is about 2%.

the price impact and the effective spread. Section 3 presents a bias-free methodology for estimation of these parameters. Section 4 describes the data of portfolio transitions and econometric specifics of empirical tests. Section 5 reports the unconditional estimates of trading costs. Section 6 focuses on the cross-sectional properties of these estimates. Section 7 concludes.

2 Variables of Interest

I start with defining the price impact and the effective spread and showing how they relate to the price dynamics around portfolio transition trades. The portfolio transition manager has to sell securities from a legacy portfolio and to acquire securities into a target portfolio. His goal is to design a cost-efficient execution scheme. Disregarding the issues involved into portfolio trading, I consider his trading strategy in a particular security.[2] To fix the ideas, I assume that transition manager has to execute $\bar{X}$ units of a stock over a fixed time period $[0, T]$. I define a variable $\mathbb{I}_{BS}$ that accounts for the trading direction.

$$\mathbb{I}_{BS} = \begin{cases} 1, & \text{for buy orders,} \\ -1, & \text{for sell orders.} \end{cases}$$

The transition manager usually splits his order over $[0, T]$. The execution strategy can be then defined in terms of a variable X_t, which stands for the number of shares executed by time t. Clearly, $X_0 = 0$ and $X_T = \bar{X}$. For simplicity, I consider only continuous trading policies.[3] The trading rate at time t is defined as X'_t, and $X'_t dt$ is the total trading volume during $[t,\ t + dt)$. Formally, the set of admissible execution strategies for an order is

$$\left\{ X_{t, t\in[0,\,T]} :\ X'_t \geq 0,\ \int_0^T X'_t dt = \bar{X} \right\}.$$

Besides traditional open markets (e.g., the NYSE, the AMEX and the NASDAQ), the transition manager has an access to alternative sources of liquidity such as various crossing networks. Internal pools of liquidity encompass either orders generated by other contemporaneously executed portfolio transitions, if a transition company has succeeded in capturing a large fraction of the market, or orders from an affiliated passive asset management desk.

[2] I also do not model the issues related to the agency-based approach of managing portfolio transitions, during which transition managers do not provide additional capital or guarantee pre-specified execution prices but commit to act in the best interests of their clients.

[3] In general, both continuous and discrete trading strategies might be considered. Admissibility of discrete trading will not change my main conclusions.

Presumably, transactions in internal crossing networks incur neither trading costs nor information leakage. Another valuable source of liquidity are external crossing networks such as ITG's POSIT, LiquidNet or PipeLine. These networks facilitate crossing orders at prices that are derived from primary exchanges and are adjusted to alleviate gaming and manipulation concerns. Although not insuring a full anonymity, they claim to substantially reduce information leakage and help avoid the bid-ask spread.[4] Since the transition manager might split this order over different trading venues, the total trading at time t, $X_t'dt$, is further split into three groups. The fraction of orders executed in each of trading venue at time t is $\mathbb{I}_{t,omt}, \mathbb{I}_{t,ec}$, and $\mathbb{I}_{t,ic}$, respectively. For each t, they sum up to one, $\mathbb{I}_{t,omt} + \mathbb{I}_{t,ec} + \mathbb{I}_{t,ic} = 1$. During $[t,\ t+dt)$, $\mathbb{I}_{t,ic} \times X_t'dt$ shares are executed in internal crossing networks, $\mathbb{I}_{t,ec} \times X_t'dt$ in external crossing networks, and $\mathbb{I}_{t,omt} \times X_t'dt$ shares are transacted in traditional markets. Let X_t be a total number of shares executed during time interval $[0,\ t)$. I also denote the number of shares executed in open markets, external and internal crossing networks during time interval $[0,\ t)$ as

$$X_t^{omt} = \int_0^t \mathbb{I}_u^{omt} X_u' du, \quad X_t^{ec} = \int_0^t \mathbb{I}_u^{ec} X_u' du, \quad X_t^{ic} = \int_0^t \mathbb{I}_u^{ic} X_u' du.$$

I next specify the price dynamics around portfolio transition orders. Since the magnitude of portfolio transition orders is large, the price dynamics certainly depend on the executed trades of transition managers. I fix the pre-transition security price at level P_0 and specify the following price dynamics as:

$$dP_t = \lambda \times (\mathbb{I}_{BS} \times dX_t) + \sigma_P \times d\widetilde{Z}_t, \qquad \widetilde{Z}_0 = 0, \tag{1}$$

where σ_P is the standard deviation of security prices, $\mathbb{I}_{BS}$ is a variable equal to 1 for buy orders and -1 for sell orders, and $d\widetilde{Z}_t$ are news about security's fundamentals modeled as $N(0, dt)$. Parameter λ is the price impact. It reflects how strongly security prices change in response to trades.

The specification of price dynamics is inspired by a measure of information asymmetry, first introduced by Kyle (1985) and defined as the price change over the trade size. However, in order to make the price impact variables more meaningful for a cross-section of stocks, it is natural to use its scaled version. Therefore, I normalize trade size dX_t by a fraction of trading volume (e.g. the average daily volume) and express the price impact as a fraction of the return standard deviation (e.g. the daily return standard deviation). Then, the price

[4]For the recent overview of crossing networks see Ramistella (2006).

dynamics (1) is:

$$\frac{dP_t}{P_t \times \sigma_r} = \tilde{\lambda} \times \left(\mathbb{I}_{BS} \times \frac{1}{ADV} \times dX_t \right) + d\widetilde{Z}_t, \qquad \widetilde{Z}_0 = 0, \tag{2}$$

where σ_r is the daily return standard deviation, ADV is the average daily volume (in shares), $\mathbb{I}_{BS}$ is a variable equal to 1 for buy orders and -1 for sell orders, and $d\widetilde{Z}_t$ are news about security's fundamentals modeled as $N(0, dt)$. The scaled price impact $\tilde{\lambda}$ is the variable of interest in this study. It means by how much (in percents of the daily return standard deviation σ_r) 1% of the average daily volume will change the security prices. Thus, the underlying assumption is that a certain fraction of "normal" daily trading volume is expected to trigger a certain level of "normal" price motion (e.g., Almgren et al.(2005)). This scaling procedure assures the proper specification of the price impact for the stocks with the same market capitalization and return volatility but with different numbers of shares outstanding. This measure is also stable with respect to splits.

Another variable of interest is the effective spread. While trading, manager incurs a fixed transaction cost due to bid-ask spread. More precisely, I assume that he pays on a 1-share buy order and receives on a 1-share sell order α_{ic} dollars, if this order is executed as an internal cross, α_{ec} dollars if it is executed as an external cross, and α_{omt} dollars if it is traded in open markets. Presumably, $\alpha_{omt} \geq \alpha_{ec} \geq \alpha_{ic}$. Moreover, α_{ec} and α_{ic} might be equal to zero, since crossing networks are believed to help avoid bid-ask spread. In other words, variables $\alpha_{omt}, \alpha_{ec}, \alpha_{ic}$ are the effective half spread in open markets, in internal and external crossing networks, respectively. Again, it is useful to specified the effective spread in relative terms. I define $\tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$ that quantify the spread payment on \$1 dollar invested into stock in percents of the daily return standard deviation. In other words, these variables characterize the percentage effective spread, which is equal to $2 \times \tilde{\alpha}_{omt} \times \sigma_r$, $2 \times \tilde{\alpha}_{ec} \times \sigma_r$ and $2 \times \tilde{\alpha}_{ic} \times \sigma_r$ for various trading platforms.

To summarize, I estimate the scaled variables of the price impact and the effective spread:

$$\tilde{\lambda}, \quad \tilde{\alpha}_{omt}, \quad \tilde{\alpha}_{ec} \qquad \text{and} \quad \tilde{\alpha}_{ic}. \tag{3}$$

However, my results can be interpreted in terms of the unscaled variables, $\lambda, \alpha_{omt}, \alpha_{ec}$, and α_{ic}, because they can be expressed as the simple transformations of the scaled variables (3):

$$\lambda = \tilde{\lambda} \times \frac{\sigma_r P}{ADV}, \quad \alpha_{omt} = \tilde{\alpha}_{omt} \times \frac{\sigma_r P}{100}, \quad \alpha_{ec} = \tilde{\alpha}_{ec} \times \frac{\sigma_r P}{100}, \quad \alpha_{ic} = \tilde{\alpha}_{ic} \times \frac{\sigma_r P}{100}. \tag{4}$$

To further clarify my estimation procedure, I provide an illustrative example of plausible values for variables in this study. Variable dX_t is in the number of shares, e.g. 1000 shares.

Variable σ_r is the daily return standard deviation, e.g. 0.04 for volatile stocks and 0.01 for not volatile stocks. Variable ADV is the average daily volume in shares, e.g. 50 millions for large stocks and 1 million for small stocks. Variable P is the stock price, e.g. \$20 dollars.

If the price impact variable $\tilde{\lambda} = 2$, then it means that 1% of the average daily volume results in the positive return of $2\% \times \sigma_r$ or, in other words, in the absolute price deviation of $2\% \times \sigma_r \times P$. In a given above example, for large and volatile stocks, the purchase of $500,000$ shares of stocks (1% of ADV) at price \$20 is expected to induce a positive return of $2\% \times 0.04 = 8$ bps or move the price by $2\% \times 0.04 \times \$20 = 1.6$ cents.

If the half effective spread in open markets $\tilde{\alpha}_{omt} = 15$, then it means that the percentage effective spread in open markets is equal to $2 \times 15\% \times \sigma_r$ or, in other words, the absolute effective spread in open markets is equal to $2 \times 15\% \times \sigma_r \times P$. In a given above example, for volatile stocks with the current price of \$20, the percentage effective spread is equal to $2 \times 15\% \times 0.04 = 1.2\%$, and the absolute effective spread is equal to $2 \times 15\% \times 0.04 \times \$20 =$ 24 cents.

3 Methodology

In this section, I discuss a conventional method to estimate the parameters of trading costs and show that it delivers biased estimates in cases when trading strategies are price dependent and when only fractions of ex ante orders are considered. Then, I describe a bias-free method. The intuition behind both methods is illustrated in Figure 1.

3.1 Implementation Shortfall

The calibration of the parameters of trading costs is usually based on the reported costs. A standard metrics of execution quality is the implementation shortfall defined in Perold (1988). This measure is equal to the difference between two returns: (1) the dollar return of a paper portfolio where all shares are assumed to be transacted at the prevailing market prices at the time of trading decision and (2) the actual dollar return of the portfolio. Despite a straightforward intuition behind this measure, its definition involves a few subtle issues, therefore I discuss next its adaptation for a case of portfolio transition trading in more detail.

The benchmark implementation assumes that an entire package is transacted momentarily at pre-trade prices. The "hypothetical" cost of acquiring position is $\bar{X}P_0$. Its final value is a total desired position $\bar{X}$ evaluated at the stock price at time t. Thus, if $\mathbb{I}_{BS}$ is a buy/sell indicator, then the return on a paper portfolio by time $t \in [0,\ T)$ is

$$\mathbb{I}_{BS} \times \bar{X} \times (P_t - P_0).$$

The actual implementation includes trading costs due to various market imperfections. The return on a real portfolio is typically affected by two factors. The first one reflects the shifts in the market's assessment of stock fundamentals and the compensation for liquidity provision, whereas the second one is due to bid-ask spread. Thus, the actual portfolio return by time $t \in (0,\ T)$ is computed as

$$\mathbb{I}_{BS} \times \left(\int_0^t (P_t - P_u) X'_u du \right) - \int_0^t (\alpha_{omt} \mathbb{I}_u^{omt} + \alpha_{ec} \mathbb{I}_u^{ec} + \alpha_{ic} \mathbb{I}_u^{ic}) X'_u du.$$

Consequently, the total implementation shortfall, marked-to-market at time $t \in [0,\ T)$ and denoted as IS_t, can be determined as

$$\mathbb{I}_{BS} \times \left(\bar{X}(P_t - P_0) - \int_0^t (P_t - P_u) X'_u du \right) + \int_0^t (\alpha_{omt} \mathbb{I}_u^{omt} + \alpha_{ec} \mathbb{I}_u^{ec} + \alpha_{ic} \mathbb{I}_u^{ic}) X'_u du. \quad (5)$$

The implementation shortfall IS_t depicted as a shaded area in Figure 1. It can be also represented as a sum of two components $IS_t = \Pi_t + OC_t$ where

$$\Pi_t = \mathbb{I}_{BS} \times \int_0^t (P_u - P_0) X'_u du + \int_0^t (\alpha_{omt} \mathbb{I}_u^{omt} + \alpha_{ec} \mathbb{I}_u^{ec} + \alpha_{ic} \mathbb{I}_u^{ic}) X'_u du, \quad (6)$$

$$OC_t = \mathbb{I}_{BS} \times (P_t - P_0) \times \left(\bar{X} - X_t \right). \quad (7)$$

The first component, Π_t, corresponds to *the realized trading costs* of trades transacted during $[0,\ t)$. It comprises the explicit costs, such as bid-ask spread of various trading platforms, and the implicit costs, such as the price impact of transacting at what may be disadvantageous prices. Due to typically adverse price changes and the bid-ask spread, Π_t tend to be positive for both purchases and sales.

The second component, OC_t, stands for *the opportunity costs*, or losses of potential gains due to the transactions that are unexecuted by time t, $\bar{X} - X_t$. The unconditional expected value of this component, $E_0(OC_t)$, is non-zero whenever only a part of trading package rather than an entire ex ante order is executed. Moreover, its expected value tends to be positive, since predominantly adverse changes in security prices during the trading make $\mathbb{I}_{BS} \times (P_t - P_0)$, on average, greater than zero. Thus, non-zero opportunity costs is likely to be observed in many situations. For instance, it might be caused by canceled orders. Indeed, if a portion of an order is canceled and only the completed by time t trades are taken in account, then the opportunity costs, OC_t, is non-zero and most likely positive. Similarly, the opportunity cost, OC_t, differs from zero if the implementation shortfall IS_t is defined for a subinterval $[0,\, t) \in [0,\, T)$ rather than a whole trading period $[0,\, T)$, for instance, for a given day in a multi-day trading packages.

Parameters of the price impact and the effective spread (3) are often calibrated based exclusively on the realized trading costs of executed trades Π_t, since unexecuted trades are usually unobservable. I review this methodology in the next section and show mathematically when it leads to biased estimates. According to my further analysis, the opportunity cost OC_t of unexecuted trades is an essential part of trading costs, and disregarding it, while estimating (3), is not inconsequential.

3.2 Conventional Methodology

A conventional way to proceed with the calibration of transaction costs parameters (3) is to estimate them from a regression equation that relates these parameters to the realized trading costs, Π_t, without taking in account the opportunity costs, OC_t. I next illustrate this approach in more detail. The variable Π_t defined in (6) is a total realized costs of trades executed during time interval $[0,\, t)$ given a pre-trade benchmark price P_0. Using (1), I express the security price at time u as,

$$P_u = P_0 + \mathbb{I}_{BS} \times \lambda \times (X_u - X_0) + \sigma_P \times (\widetilde{Z}_u - \widetilde{Z}_0). \tag{8}$$

Plugging (8) into (6), I get

$$\Pi_t = \frac{1}{2}\lambda X_t^2 + \int_0^t \left(\mathbb{I}_{BS} \times \sigma_P (\widetilde{Z}_u - \widetilde{Z}_0) + \alpha_{omt}\mathbb{I}_u^{omt} + \alpha_{ec}\mathbb{I}_u^{ec} + \alpha_{ic}\mathbb{I}_u^{ic} \right) X'_u du. \tag{9}$$

Then, if I use equations (4) and $\sigma_P = \sigma_r P_0$, then the average trading costs per \$1 traded (in percents of the return standard deviation) can be written as

$$\frac{\Pi_t}{\sigma_r P_0 X_t} = \frac{1}{2}\tilde{\lambda}\frac{1}{ADV}X_t + \frac{\tilde{\alpha}_{omt}}{100}\frac{X_t^{omt}}{X_t} + \frac{\tilde{\alpha}_{ec}}{100}\frac{X_t^{ec}}{X_t} + \frac{\tilde{\alpha}_{ic}}{100}\frac{X_t^{ic}}{X_t} + \tilde{\epsilon}_t \tag{10}$$

where

$$\tilde{\epsilon}_t = \mathbb{I}_{BS} \times \int_0^t (\widetilde{Z}_u - \widetilde{Z}_0)\frac{X'_u}{X_t}du.$$

The equation (10) has an intuitive sense. The price "slippage", or the difference between the average execution price and the decision price, is determined by several factors. The first one is due to the price impact of a given order, which is usually linear in order size. The second one is due to the non-zero spread on various trading platform. The term $\tilde{\epsilon}_t$ is often assumed to capture the noise. Therefore, the trading costs variables (3) can be estimated from the regression equation (10) with $\tilde{\epsilon}_t$ assumed to be independent of regressors and $E(\tilde{\epsilon}_t) = 0$. Different versions of regression equation (10), sometimes with slight modifications, are used in the work of Almgren et al. (2005), Breen, Hodrick, and Korajczyk (2002), among others. I refer to regression equation (10) as a conventional methodology for estimating the parameters of trading costs (3).

A conventional approach relies on the assumption that ϵ_t can be approximated by a random variable with a zero mean. This assumption is not an innocuous one. Indeed, integrating by parts, I find that $E(\epsilon_t)$ can be presented as follows,

$$\mathbb{I}_{BS} \times E\left(\frac{1}{X_t}(\widetilde{Z}_u - \widetilde{Z}_0)X_u|_{u=0}^t\right) - \mathbb{I}_{BS} \times E\left(\int_0^t \frac{X_u}{X_t} dZ_u\right) = -E\left(\int_0^t \frac{X_u}{\mathbb{I}_{BS} \times X_t} dZ_u\right). \quad (11)$$

Therefore, the first moment of ϵ_t can be safely assumed to be zero, only if (11) is equal to zero for a given time t. This condition holds if $t = T$ and X_T is an entire ex ante order $\bar{X}$ with no trades remained unexecuted, since $E\left(\int_0^T X_u dZ_u\right) = 0$. In this case, conventional estimates are unbiased regardless of interim trading strategies of market participants, which may be price-dependent, for instance, it may include the limit orders. Indeed, even if conditioning the strategy on the security prices, the cost might be deferred in time, it will be inevitably realized by the end of a trading period $[0, T)$. To summarize, a conventional approach is valid for assessing the trading costs of an *entire ex ante orders* traded over a period $[0, T)$.

At the same time, if the regression equation (10) is used to estimate the trading costs for only a part of ex ante order, then the estimates are likely to be biased. Indeed, traders usually try to influence the magnitude of the variable trading costs by adapting their trading strategies to the conditions encountered when a trade is brought to the marketplace. Often, market participants slow down their trading when the market runs away. Frequently, some trades are canceled, since the available market prices are not considered favorable anymore, thus leaving out of an ex post sample the most expensive trades that have never been executed. This endogenous dependence of trading strategies on the security prices leads to the negative correlation between $\mathbb{I}_{BS} \times X_t$ and dZ_u for $u \in [0, t)$ and induces a downward bias in the estimates of price impact. Thus, if traders slow down their trading when security prices run away from them, then conventional estimates of price impact are biased downwards. In contrast, if traders speed up their trading during unfavorable market conditions, then these estimates are biased upwards. In other words, Π_t is not equal to $\frac{\lambda}{2}X_t^2$ (see Figure 1). For instance, this selection-bias problem is relevant to any calibrations based on the TAQ dataset, in which only realized rather than ex ante orders are observed.

To summarize, the conventional method can be applied only for estimation of the trading costs of total ex ante orders. In other situations, improper evaluation metrics will likely induce a bias in conventional estimates and lead to erroneous conclusions. Thus, a more robust estimation procedure is of particular interest, especially considering how prevailing are cancelations and limit orders in the markets. Next, I describe a bias-free estimation methodology that can be applied in a general case.

3.3 Bias-Free Methodology

A bias-free methodology is based on the total implementation shortfall, IS_t, marked-to-market at time t, which includes both components Π_t and OC_t, the realized trading costs of executed trades and the opportunity costs of unexecuted trades. The implementation shortfall IS_t, defined in (5), is equal to

$$\mathbb{I}_{BS}\times\left(\bar{X}(P_t-P_0)+X_t(P_T-P_t)+\int_0^t(P_u-P_T)X'_u du\right)+\int_0^t(\alpha_{omt}\mathbb{I}_u^{omt}+\alpha_{ec}\mathbb{I}_u^{ec}+\alpha_{ic}\mathbb{I}_u^{ic})X'_u du.$$

Using the expression (8) for security prices and equation (4), it is easy to show that the implementation shortfall IS_t normalized by $\sigma_r P_0 \bar{X}$ can be further decomposed into the sum of a deterministic and random components,

$$\frac{IS_t}{\sigma_r P_0 \bar{X}} = \frac{1}{2}\tilde{\lambda}\frac{1}{ADV}\frac{\bar{X}^2-(\bar{X}-X_t)^2}{\bar{X}} + \frac{\tilde{\alpha}_{omt}}{100}\frac{X_t^{omt}}{\bar{X}} + \frac{\tilde{\alpha}_{ec}}{100}\frac{X_t^{ec}}{\bar{X}} + \frac{\tilde{\alpha}_{ic}}{100}\frac{X_t^{ic}}{\bar{X}} + \tilde{\epsilon}_t, \tag{12}$$

where

$$\tilde{\epsilon}_t = \mathbb{I}_{BS}\times\left(\widetilde{Z}_t-\widetilde{Z}_0-\int_0^t(\widetilde{Z}_t-\widetilde{Z}_u)\frac{X'_u}{\bar{X}}du\right).$$

This equation shows explicitly what determines the average change in the implementation shortfall for one-dollar transaction (in percents of the return standard deviation σ_r). The intuition behind this method is that expected mark-to-market implementation shortfall IS_t is represented as the difference between expected execution costs of a total ex ante order $\bar{X}$ and the expected execution costs of unexecuted part $\bar{X}-X_t$ (see Figure 1). In other words, the total implementation shortfalls IS_t is written in a way that the random part $\tilde{\epsilon}_t$ does not involve the interaction terms between X_t and any past changes in prices. Since the trading strategy is independent from the future changes in prices, the expectation of $\tilde{\epsilon}_t$ is equal to zero regardless of t. These considerations suggests that unbiased estimates of trading cost parameters (3) can be obtained from the regression equation (12). Regardless of time interval $[0,t)$ under investigation, the key input is a total size of ex ante order $\bar{X}$.

To further clarify the distinction between two methodologies for estimating the parameters of trading costs, I compare two equations (10) and (12). The left-hand side variables in these equations are different: $\Pi_t/\sigma_r P_0 X_t$ and $IS_t/\sigma_r P_0 \bar{X}$, respectively. Both variables are stated in percents of the return standard deviation. However, in the conventional methodology, the left-hand side variable is the realized trading cost Π_t for trades that are executed during $[0, t)$ *per one dollar traded.* At the same time, in the bias-free methodology, the left-hand side variable is also the realized trading cost Π_t for trades executed during $[0, t)$ but augmented with the opportunity cost for the remaining orders, OC_t, and stated *per one dollar of total ex ante order.* Furthermore, the right-hand side of the equation (12) is properly adjusted for the expected costs of yet unexecuted trades.

The suggested method provides a flexible framework for estimating the parameters of the trading costs. It allows to assess trading costs of any part of executed orders without incurring biases due to endogenous split of executed orders *over time* in response to the observed price dynamics or order cancelations. For example, I can calibrate the estimates of (3) by analyzing the variable IS_t for any time t regardless of whether some trades are still incomplete at time t. Thus, the suggested methodology (12) is a good tool to disentangle price impact and opportunity costs and to obtain unbiased estimates of trading costs (3).

3.4 Comparison of Methodologies

I highlight the difference between two approaches on a simulated data for trades and price process defined in (1). For simplicity, I assume that there is no bid-ask spread. I fix the value of price impact $\widetilde{\lambda} = 2$ to be estimated and generate 5000 scenarios, each of which evolves according to the following schema. The average daily volume $ADV = 1000$, the daily standard deviation of returns $\sigma_r = 0.02$. The initial price P_0 is uniformly distributed between 1 and 100. A noise term $\widetilde{Z}$ is drawn from a normal distribution with zero mean and standard deviation equal to $\sigma_r \times P_0$. The size of ex ante buy order, $\bar{X}$, is uniformly distributed between -1000 and 1000. I also simulate the dependence of trading strategies on price dynamics. If security prices move favorably ($\widetilde{Z} \times \bar{X} < 0$), then total ex ante order $\bar{X}$ is executed. If security prices move adversely ($\widetilde{Z} \times \bar{X} > 0$), then a part of ex ante order is canceled and only a random fraction of $\bar{X}$ is executed. This schema reproduces the general practice to cut off trading during unfavorable market environment. The realized price trajectory is defined by (1).

Next, I estimate $\widetilde{\lambda}$ using two methodologies. I discuss a particular scenario to illustrate how to construct regression equations in both cases. The initial price P_0 is \$20. A trader plans to buy 1000 shares. While trading, he faces adverse market conditions as $\sigma_P \times \widetilde{Z}$ turns out to be \$0.5. The available prices are so high that a trader decides to cancel the execution of remaining 100 shares and thus only 900 shares are being purchased. From (4) and (1), $\lambda = 0.0008$, the post-trade price is $P_0 + \lambda \times X + \widetilde{Z} = 20 + 0.0008 \times 900 + 0.5 = \$21.22$ and the average execution price is $P_0 + \frac{1}{2}(\lambda \times X + \widetilde{Z}) = 20 + \frac{1}{2}(0.0008 \times 900 + 0.5) = \$20.61$. In conventional approach, this observation leads to the equation:

$$\frac{900 \times 20.61 - 900 \times 20}{\sigma_r \times 20 \times 900} = \frac{1}{2}\tilde{\lambda}\frac{1}{ADV}900 + \tilde{\epsilon}.$$

In bias-free approach, this observation gives the equation:

$$\frac{900 \times 20.61 + 100 \times 21.22 - 20 \times 1000}{\sigma_r \times 20 \times 1000} = \frac{1}{2}\tilde{\lambda}\frac{1}{ADV}\frac{1000^2 - (1000 - 900)^2}{1000} + \tilde{\epsilon}.$$

When a parameter $\widetilde{\lambda}$ (equal to 2 in my experiment) is estimated based on 5000 simulations, the results are the following. Using a conventional approach, the estimates of $\widetilde{\lambda}$ are 1.56 with standard errors equal to 0.03 and R-squared equal to 37%. Using a bias-free approach, the estimates of $\widetilde{\lambda}$ are 1.95 with standard errors equal to 0.04 and R-squared equal to 37%. As predicted, the conventional estimates are biased downwards, since by design a trader slows down his trading when the market runs away. At the same time, a new methodology delivers unbiased estimates.

3.5 Different Trading Venues

Methodologies are also robust to endogenous split of orders *over different trading venues* in response to current market environment. In fact, the selection of trading system might affect ex post trading costs. Usually, transition managers follow a pecking-order while implementing their strategies. The priority, in which they are considering trading platforms, is the following: internal crossing networks, external crossing networks and then traditional open markets, presumably from the cheapest alternatives to the most expensive ones.

This practice might lead to biased estimates of trading costs obtained from the analysis of ex post execution data. To illustrate, I consider two cases when a portfolio transition is executed during favorable and unfavorable market conditions. In the first scenario, many market participants are trading in the opposite direction to the transition manager, and since liquidity is easy to get, most orders are executed internally. Hence, ex post price impact in internal crossing networks are found to be close to zero, and the price impact in other trading venues are not be observed. In the second scenario, many market participants are trading in the same direction as the transition manager, and since liquidity is scare in crossing networks, most orders are directed to open markets. Hence, the ex post records of the price impact in the traditional exchanges are particularly significant, whereas the transaction data on crossing networks is not observed. Thus, if the estimates of the price impact are obtained separately for trading venues, then the pecking-order in selecting trading venues might trigger underestimation of trading costs for the external and internal crosses and their overestimation for the open markets trades.

To get around this problem, I bundle orders executed in various systems together. Thus, I provide only unconditional estimates of the price impact. At the same time, I assume that fixed costs of trading can not be influenced by strategic behavior of transition managers and, therefore, I estimate the implied effective spread for each of trading platform separately. Next, I discuss a data set of portfolio transition trades and other details of estimation procedure.

4 Estimation

4.1 Portfolio Transition Data

I use a proprietary database of portfolio transitions from a leading provider of portfolio transition services, who supervises more than 30% of transitions in the U.S. The samples includes about 2,680 portfolio transitions with a trading volume of roughly $630 billion (out of which $450 billion are traded in the U. S. markets), which are executed on behalf of U.S. institutions from January 2001 to December 2005.

This unique data set is based on the actual post-transition reports prepared by transition managers for their clients. For each transition, each security, each trading day and each trading venue (open markets, external or internal crossing networks), the number of shares traded, the average execution price, the pre-transition benchmark price as well as the information on transaction costs are specified.

The data set of portfolio transitions provides a unique opportunity to observe ex ante orders, i.e. their size and direction, which is a crucial input in the regression equation (12). In fact, the trading intentions of portfolio transition managers are set in advance when they get mandates with pre-specified lists of securities in legacy and target portfolios. This property of portfolio transitions is a valuable one, since usually ex ante orders cannot be identified, and observed orders might be influenced by a variety of investment styles and order placement strategies.

Trading data is available for each trading day in trading packages. Instead of aggregating information at order level, I utilize data on individual trades in a trading sequence and treat them as separate observations to enlarge the size of a sample under investigation. Namely, I estimate the parameters of trading costs $\lambda, \tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$ from regression equations (10) and (12) for days when trades were executed. For instance, if a purchase is executed over five days but transactions are done only on the first and the last days, then my sample includes two observations for these two days. For each day t and a security in a given transition, I assume that $\bar{X}$ is the total number of shares left to be executed at the beginning of day t, X_t is the number of shares executed during that day, Π_t is the realized trading costs for that day, and IS_t is mark-to-market implementation shortfall for that day (with a previous day close price taken as a benchmark price).

It is worth noting that since I consider execution of daily parts of orders rather than that of total ex ante orders and since trading strategies are most likely price dependent, conventional method might deliver biased estimates, whereas bias-free method delivers unbiased estimates. I also run tests on a sample of total orders. For this sample, both methods are mathematically identical and deliver identical estimates. Moreover, they are close to bias-free estimates for a sample of daily trades (not reported).

Additionally, I use the CRSP database to get data on stock prices, returns, volume, and shares outstanding. My sample includes common stocks (with CRSP share codes of 10 and 11) listed on the New York Stock Exchange (NYSE), the American Stock Exchange (Amex), and NASDAQ in the period of January 2001 through December 2005. I exclude any ADRs, REITS, or closed-end funds. I remove stocks with missing CRSP information, necessary to construct variables for the tests. I also exclude low-priced stocks and further check data to avoid any potential typographical errors and inaccuracies. Moreover, since the treatment of dividends and splits, occurred during portfolio transitions, is unclear from the data, I also exclude all observations with non-zero payouts during the first week following the starting date of portfolio transitions. I use a fixed period of one week rather than an actual duration of implementing the strategies in order to avoid any potential sample selection bias, as the duration of execution might be endogenously related to various properties of traded securities. In total, I consider 702,406 observations (334,035 buys and 368,371 sells).

4.2 Econometric Details

Beside the execution-related variables (IS_t, Π_t, X_t and $\bar{X}$), the regression equations (10) and (12) involve variables such as the trading volume, ADV, the return standard deviations, σ_r, and the security price, P_0. I use the variables known before portfolio transition trades in order to avoid any spurious effects of using contemporaneous variables. Next, I describe how I define these variables.

The trading volume, ADV, is the average daily trading volume, denoted in the number of shares, for a month prior to that during which a portfolio transition is implemented. It is a good proxy for a "normal" trading volume during portfolio transitions. The security price, P_0, is the close price in the previous day.

I also estimate the volatility of daily returns, σ_r. For each security, I first find the estimates of the monthly return variance as a the sum of the squared daily returns:

$$\sigma_{i,t}^m = \sum_{k=1}^{N_t} r_{i,t}^2$$

where there are N_t daily returns $r_{i,t}$ in month t. I do not subtract the average daily return in month t because it has little effect on the results. I also do not adjust the estimates for autocorrelation of returns adding a cross-product of adjacent returns, as it might result in the negative estimates of volatility for some stocks. Then, the historical daily volatility, $\sigma_{i,t}^h$, of stock i in month t is defined as

$$\hat{\sigma}_{i,t}^h = \frac{1}{\sqrt{N_t}} \sigma_{i,t}^m. \tag{13}$$

The volatility of stock returns is changing over time. Therefore, I also estimate the ARIMA model and get the conditional forecasts of the daily return standard deviations for each stock i and month t. I follow the procedure suggested in French et al. (1987). To account for the positive skewness of the standard deviation estimates, I use the logarithmic transformation for the volatility instead of its raw estimates. I estimate a third-order moving average process for the changes in $\ln \sigma_{i,t}^m$ over the whole sample from 2001 to 2005:

$$(1 - L) \ln \sigma_{i,t}^m = \Theta_0 + (1 - \Theta_1 L - \Theta_2 L^2 - \Theta_3 L^3) u_t$$

Then, the conditional forecast for the volatility of daily returns is

$$\hat{\sigma}_{i,t}^e = \exp \left[\ln \sigma_{i,t}^m + 0.5 \hat{V}(u) \right] \tag{14}$$

where $\hat{V}(u)$ is the variance of the prediction errors of ARIMA model. In estimation of the regression equations (12), I use proxies $\hat{\sigma}_{i,t-1}^e$ and $\hat{\sigma}_{i,t-1}^h$ for σ_r. I only report the results for $\hat{\sigma}_{i,t-1}^e$, as these results are quantitatively similar to the results for the other proxy.

It is important to adjust standard errors for the potential cross-sectional correlation. Without correcting for these effects, the precision of estimates can be largely overestimated. This interdependence across close in time observations may arise due to several reasons. First, splitting of trades over several days and considering them as separate observations might lead to high correlation of their standard errors. Second, this correlation might be also either due to the fact that the portfolio transitions involve trading portfolios of securities with often similar characteristics or due to effects of the overall market dynamics.

To correct for potential cross-sectional correlation of standard errors, I cluster standard errors of observations for each week and each industry. I consider the conventional definitions of 17 industries. For robustness check, I also implement a number of slightly modified tests and run my regression analysis for various numbers of industries and monthly-grouped data as well as using the Fama-McBeth procedure instead of the pooled regression. The resulting estimates are very stable across various econometric tests. This fact assures the robustness of my results.

5 Main Results

5.1 Portfolio Transition Orders

I describe next the size of portfolio transition orders and the characteristics of securities in legacy and target portfolios. To capture cross-sectional differences, I consider the groups of securities with different average daily volume and volatility. I divide observations into 10×5 groups based on the average daily trading volume in dollars and the historical standard deviation of daily returns, $\hat{\sigma}^h_{i,t}$ defined in (13), in the previous month. The volume thresholds are 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th quintiles, and the volatility thresholds are 20th, 40th, 60th, and 80th quintiles of the corresponding variables for the universe of the NYSE-listed stocks with CRSP share codes of 10 and 11. Each month, I recalculate thresholds and reshuffle stocks across bins. I choose not to consider equally-spaced thresholds for the trading volume and instead focus on larger stocks, which are economically more important.

Panel A of Table 1 shows that the number of observations is relatively stable across groups. Typically, there are more than 10,000 observation in each bin. Significantly more observations is placed only in the lowest volume and the highest volatility bins. Panel B and Panel C of Table 1 present the median values for the average daily trading volume and the daily return standard deviations, respectively. The daily trading volume ranges from \$1 million for small stocks to \$200 million for large stocks, whereas the historical daily volatility varies from 1% for non-volatile stocks to almost 4% for volatile stocks.

The magnitude of portfolio transition orders varies significantly across bins. Figure 2 shows the mean and median values of order sizes as a percentage of the average daily volume for 50 groups of securities, defined on the basis of the trading volume and volatility. Order sizes are inversely related to trading volume. For small stocks, the average order size is about 20%; for large stocks, it is only 0.50%. This difference remains, if the median values are considered. The relation between order sizes and volatility tend to be inverse as well.

5.2 Estimates of Trading Costs

Table 2 presents the estimates of the price impact $\tilde{\lambda}$ and the effective spread $\tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}$ and $\tilde{\alpha}_{ic}$, obtained using conventional method (10) and bias-free methods (12). I consider a total sample of daily portfolio transition trades. Since trading costs might be different for purchases and sales as well as for securities listed in different exchanges, I also supplement my analysis with the analysis of trading costs for corresponding sub-samples.

When a bias-free method is applied to a total sample, the unconditional estimate of the price impact $\tilde{\lambda}$ is equal to 0.33. Therefore, on average, 1% of the average daily volume has the price impact equal to 0.33% of the daily return standard deviation. These estimates tend to be higher for purchases than for sales, however, this difference is statistically significant only for the Nasdaq-listed stocks. Also, I do not find any significant difference between the magnitude of the price impact for Nasdaq and NYSE/Amex stocks. These patterns remain if estimates are obtained through the conventional method.

For both methods, the unconditional estimate of the effective spread for open market trades, $2 \times \tilde{\alpha}_{omt}$, is about 13, which corresponds to a fixed cost of 13% of the daily standard deviation per \$1-trade. These estimates are higher for transition purchases than for transition sales. Moreover, I document larger estimates of the effective spread for the Nasdaq-listed stocks. This finding is consistent with a report of the U.S. Securities and Exchange Commission that analyzes the effective spreads on various trading platforms and concludes that NYSE provides a better order execution than Nasdaq.

The estimated effective spread for the crosses on external crossing networks, $2 \times \tilde{\alpha}_{ec}$, is also about 13% of the daily standard deviation. Despite a belief that crossing networks help avoid bid-ask spread, there is no evidence that spread in crossing networks is significantly different from the one on the traditional markets. Indeed, the point estimates of the effective spread are sometimes even higher for the external crosses than for the open market transactions. The estimates of the effective spread on internal crossing networks, $2 \times \tilde{\alpha}_{ic}$, is usually insignificant.

I examine the difference between the estimates of conventional and bias-free methods. Table 2 reports the percentage difference between these estimates, $\%\Delta$. For purchases, the conventional estimates tend to be weakly biased downwards ($\%\Delta < 0$). For example, the magnitude of this bias is -13% for purchases of the NYSE-listed securities. For sales, the conventional estimates of $\tilde{\lambda}$ are biased upwards ($\%\Delta > 0$). The magnitude of this bias is substantial. It is equal to about 20%. The bias in the estimates $\tilde{\alpha}_{omt}$ is 31% for the NYSE/Amex-listed stocks and 16% for Nasdaq-listed stocks. The sign of bias reveals indirect evidence of price dependence of transition strategies. Negative bias for sales and positive bias for purchases indicate that when shifting funds in unfavorable market conditions, transition managers tend to speed up sales of securities in legacy portfolios and somewhat slow down purchases of securities in target portfolios. These patterns might be explained by the agency-based approach of managing portfolio transitions, during which transition managers do not employ their own capital but have to use proceeds of sales to purchase new securities.

6 Cross-Sectional Properties of Trading Costs

In the previous sections, I reported unconditional estimates of price impact and effective spread that can be summarized as follows. For conventional method, the estimate of regression equation (10) is

$$\frac{\Pi_t}{\sigma_r P_0 X_t} = \frac{0.39}{2} \times \frac{1}{ADV} X_t + \frac{6.25}{100} \times \frac{X_t^{omt}}{X_t} + \frac{6.53}{100} \times \frac{X_t^{ec}}{X_t} + \frac{1.33}{100} \times \frac{X_t^{ic}}{X_t} + \tilde{\epsilon}_t$$

For bias-free method, the estimate of regression equation (12) is

$$\frac{IS_t}{\sigma_r P_0 \bar{X}} = \frac{0.33}{2} \times \frac{1}{ADV} \frac{\bar{X}^2 - (\bar{X} - X_t)^2}{\bar{X}} + \frac{6.55}{100} \times \frac{X_t^{omt}}{\bar{X}} + \frac{6.94}{100} \times \frac{X_t^{ec}}{\bar{X}} + \frac{0.22}{100} \times \frac{X_t^{ic}}{\bar{X}} + \tilde{\epsilon}_t.$$

My next goal is to analyze cross-sectional properties of the parameters of interest, $\tilde{\lambda}, \tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$. In this section, I consider how these parameters depend on stocks' average daily volume, their volatility, the proxies of information asymmetry, and the pre-transition price dynamics.

6.1 Trading Costs and Volume

I begin with the analysis of how the parameters of trading costs change across stocks with different average daily volume.[5] I split all observations into 10 volume groups according to the trading volume of securities in the previous month. The thresholds are the same as described in Section 5.1. Then, I estimate price impact and effective spread for each of these groups.

Effectively, I augment the regression equations (10) and (12) with the interaction terms of

[5]Since the market capitalization and the dollar trading volume are highly correlated across stocks, the relations between trading costs and stock size is very similar to the ones discussed in this section.

existing explanatory variables and dummy variables for each volume groups. This procedure allows me to obtain the estimates of the price impact and the effective spread for securities with different trading volume. I focus on a pooled regression rather than on a set of separate regressions for each group of observations, since this procedure enables me to take into account potential correlation of residuals across volume groups. Additionally, I redo my tests for the sub-samples of observations, namely, purchases and sales of the NYSE/Amex-listed and the Nasdaq-listed securities. Table 3 reports the estimates of the bias-free trading costs for three out ten volume groups, Groups 1, 5 and 10, obtained through bias-free (complete results as well as conventional estimates can be provided upon request). Figure 3 graphically shows the estimates for all groups and their 95%-confidence intervals for the bias-free approach. The observed patterns are worth discussing.

First, the price impact $\tilde{\lambda}$ increases with the average daily volume. For instance, if a total sample is considered, then estimates $\tilde{\lambda}$ from bias-free method range from 0.27% for thinly traded stocks to 2.99% for actively traded stocks. These point estimates have the following interpretation. On average, 1% of trading volume triggers the price changes of $0.27\% \times \sigma_r$ and $2.99\% \times \sigma_r$ for securities with the low and high dollar trading volume, respectively. Thus, the price impact is ten times higher for the latter than for the former. Likewise, all sub-samples exhibit similar relation between the price impact and the trading volume. The only exception is a relatively small price impact induced by purchases of the largest Nasdaq-listed stocks.

Second, the effective half spread $\tilde{\alpha}_{omt}$ decreases with the dollar trading volume, and it is particularly significant for the least actively traded stocks (Group 1). For both methods, while being roughly $2 \times 15\% \times \sigma_r$ for the securities with the lowest dollar trading volume, the effective spread drops to only $2 \times 3\% \times \sigma_r$ for the securities with the highest dollar trading volume. I also document similar patterns of inverse relation between the trading volume and the effective spread for external crossing networks. At the same time, internal crosses do not involve statistically significant spread-related trading costs.

Figure 4 shows the implied price impact functions for the most thinly and actively traded stocks (Groups 1 and 10). Plots are based on the estimates of trading costs in Table 3. I also assume that orders are split across trading venues according to the average proportions in my sample. For stocks with the lowest (highest) volume, these proportions are 31% (24%) for open markets, 45% (38%) for external crossing networks, and 17% (38%) for internal crossing networks. The signatures of price impact functions are different for these securities. For thinly traded securities, trading costs are almost invariant with respect to a trade size and spread-related payments account for their largest fraction. For actively traded securities, trading costs are very sensitive to a trade size and spread-related payments are less significant. It is worth emphasizing that high price impact for stocks with large trading volume reveals high sensitivity of price changes to order sizes. However, trading in these stocks is cheaper in dollar terms.

To summarize, I find that transactions in stocks with *higher* volume tend to be associated with *lower* effective spread but with *more substantial* price impact. While the first finding is consistent with existing market microstructure models, the second one seems to contradict conventional intuition. Indeed, it is commonly believed that trading in large stocks involves less risk of information asymmetry and requires smaller compensation for liquidity providers which, consequently, leads to slighter price concessions. Moreover, the price impact for trading in large stocks is mitigated by the easiness of market-making in their markets: it is less strenuous to find a hedge for these securities or reallocate them in the future to other market participants.

My findings might be counterintuitive at first sight but they also appear in work of others. For instance, Breen, Hodrick, Korajyzyk (2002) mention that the price impact, estimated with transaction data from the TAQ database, increases with the market capitalization, if trade size in its definition is scaled by the shares outstanding and, consequently, these estimates quantify the price impact for one percent of turnover rather than one dollar. To avoid these "unreasonable" results, they choose to focus on the unscaled measure of price impact (as λ in (1)) that "reasonably" decreases with the market capitalization (see also Hasbrouck (1991b) and Chen, Stanzl and Watanabe (2005), among others)). Thus, the choice of specification for the price impact parameters is crucial. Furthermore, for a more meaningful scaled specification of the price impact (2), its estimates increase with the market capitalization and the dollar trading volume.

I outline several reasons that may potentially explain the documented patterns. My first argument is that trading in large stocks is a subject to a significant competition among market participants, which might further result in high estimates of the price impact. On one hand, when stocks are traded intensively and their trading volume is large, its substantial part is two-sided: buys alternates with sells, as market participants trade on their information.[6] Consequently, for these stocks, a one-sided trade equivalent to 1% of the average daily volume represents a distinctive and informative signal. On the other hand, when stocks are not traded a lot and their trading volume is small, a large fraction of this trading volume constitutes one-sided trading. Therefore, for these stocks, the information content of a trade, equivalent to 1% of the average trading volume, might be less significant. This mechanism

[6]For example, Chordia, Huh, and Subrahmanyam (2007) find that the stock size is negatively related to the absolute order imbalances, or the net selling or buying pressure, and positively related to the trading volume.

might lead to higher price impact for securities with larger trading volume.

Potential concavity of the price impact functions might also manifest itself in the counterintuitive relation between the trading volume and the price impact. As I show in Section 5.1, the average magnitude of transition orders (in percents of the average trading volume) is much lower for stocks with high trading volume and market size. If large order are "cheaper" than small orders, then trades in thinly traded stocks might appear less expensive that trades in large stocks. Obizhaeva (2008b) explores the non-linearity of price impact functions and shows that even after adjusting for non-linearity, price impact still increases with trading volume.

Finally, it might be misleading to compare the estimates of the price impact and the effective spread of different securities by scaling them using the trading volume, ADV, and the standard deviation, σ_r, over the same time interval (a day in). An incorrect scaling procedure might lead to the observed patterns, whereas if a properly adjusted scaling procedure is implemented, then the estimates of trading costs might be constant across securities in different groups. This line of research is further developed in Kyle and Obizhaeva (2008).

6.2 Trading Costs and Volume/Volatility

Volatility is another stock characteristic that might help explain the cross-sectional variation in the estimates. To analyze the relation between this variable and parameters of trading costs, I split all observations into 10×5 groups based on the average daily trading volume in dollars and the standard deviation of daily returns, $\sigma^h_{i,t}$, defined in (13). The rule of assigning securities to different bins is the same as in Section 5.1 and Section 6.1. I augment the regression equation (12) with the interaction terms of explanatory variables and dummy variables for each bin. I focus on the estimates obtained using the bias-free methodology. Figure 5 shows estimated parameters of price impact ($\tilde{\lambda}$) and effective half spread in open markets, external crossing networks and internal crossing networks ($\tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$), in Panel A, B, C and D, respectively.

First, the relation between the price impact and the volatility is not clearly defined. Perhaps, the split of sample into 50 groups decreases the number of observations available to get individual estimates for each subset and, consequently, amplifies their standard errors. Second, the relation between the effective half spread and the volatility is less ambiguous. For both open markets and external crossing networks, the effective spread tend to be smaller for more volatile stocks. These patterns are particularly pronounced for the securities with the lowest average daily volume. When these securities are traded in open markets, the effective half spread drops from 19% for the least volatile stocks to 11% for the most volatile stocks. To summarize, the association between the price impact and the volatility is weak. Yet, the effective spread tends to decrease with volatility.

6.3 Cross-Sectional Regressions

In the previous section, I analyze how price impact and effective spread depend on the average daily volume and the standard deviation of daily returns. It is known that parameters of trading costs may be also explained by other stocks' characteristics. For instance, trading costs may be related to the measures of information asymmetry (e.g. Breen, Hodrick, Korajyzyk (2002), Hasbrouck (1991 a,b), Glosten and Harris (1988)). Also, Saar (2001) presents theoretical arguments why trading costs may be affected by past returns dynamics, and Chiyachantana et al. (2006) tests this hypothesis. Since these characteristics are systematically related to each other, I study these relations using the cross-sectional bias-free regression analysis.

I modify the regression equation (12) and include interaction terms of $\tilde{\lambda}$, $\tilde{\alpha}_{omt}$, $\tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$ with the scores of the following stocks' characteristics. I consider $ADV_{\$}$, the average daily dollar trading volume in the previous month in million of dollars, and σ_r^h, the standard deviation of daily returns in the previous month. As proxies for information asymmetry, I include *Spread*, the quoted percentage spread in basis points, and $\#ANAL$, the number of analysts following stocks in the previous month. To capture the past price dynamics, I analyze the effect of $Ret_{-5,0}$, the stock return over five days prior to portfolio transition. I also include the interaction terms with a dummy $Nasdaq$, which is equal to 1 for the Nasdaq-listed stocks and 0 otherwise. I cluster the standard errors at weekly levels for 17 industries. Table 4 reports the estimates for a sample of portfolio transition purchases; Table 5 presents the results for a sample of portfolio transition sales.

The cross-sectional patterns of the price impact, $\tilde{\lambda}$, can be summarized as follows. The unconditional estimates of the price impact are positive and statistically significant for both purchases and sales. They increase with the dollar trading volume, $ADV_{\$}$, and slightly increase with volatility, σ_r^h, but this association is weak. For portfolio transition purchases, the coefficient on $\sigma_r^h \times \tilde{\lambda}$ is positive and significant, if explanatory variables $ADV_{\$}$, σ_r^h, *Spread*, and $Nasdaq$ are considered, and insignificant for other specifications. For portfolio transition sales, this coefficients is positive but statistically insignificant. Moreover, my analysis does not reveal any statistically significant difference between $\tilde{\lambda}$ for the NYSE/Amex-listed securities and that of the Nasdaq-listed ones. I also find that $\tilde{\lambda}$ decreases with the quoted spread, *Spread*. These findings might represent a manifestation of aforementioned patterns, i.e. price impact increases with the trading volume, and spread decreases with the trading volume. These findings contradict a conventional intuition and imply that the relation between different facets of liquidity, i.e. spread and price impact, might be more complicated than a one-to-one correspondence. For instance, trading in securities with the largest spread may be associated with the lowest price impact. Therefore, it will be interesting to do more research on how different aspects of liquidity are related to each other as well as to asset prices. Furthermore, I do not find any statistically significant relation between $\tilde{\lambda}$ and past returns, $Ret_{-5,0}$.

I next describe the cross-sectional properties of the effective half spread, $\tilde{\alpha}_{omt}$, $\tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$. The unconditional estimates of $\tilde{\alpha}_{omt}$ and $\tilde{\alpha}_{ec}$ are positive and statistically significant, whereas that of $\tilde{\alpha}_{ec}$ is usually indistinguishable from zero. Other results are, by and large, consistent with the conclusions in the previous sections. For instance, the effective half spread, $\tilde{\alpha}_{omt}$ and $\tilde{\alpha}_{ec}$, tend to decrease with the trading volume and the return standard deviation. Also, trading in Nasdaq-listed stocks usually involves larger losses due to bid-ask spread than trading in the similar NYSE/Amex-listed securities. Furthermore, the effective spread is positively related to the quoted spread, *Spread*, and negatively related to the number of analysts reporting their forecasts, $\#ANAL$. These patterns are consistent with a conventional intuition that the effective spread is larger for securities with higher risk of adverse selection. Also, it is worth discussing the relation between the effective spread and the stock returns over one week before portfolio transitions. Following high stock returns in the pre-transition period, the effective spread on transition purchases decreases and the effective spread on transition sales increases. This asymmetric relation between the realized impact of trades and the price history is similar to the one documented in Chiyachantana et al. (2006). The authors report that the institutional purchases (sales) of stocks with several days of price run-up induce smaller (larger) permanent price changes.

Several explanations may be suggested for the documented relation between past returns and trading costs. First, it might be a mechanical manifestation of bid-ask bounce. Second, these patterns might be due to the negative autocorrelation of stock returns, potentially caused by the market resiliency, as the market accommodates pre-transition order imbalances, consequently, sometimes the estimates of trading costs seem to be lessening by the reversals in stock prices. Third, the relation between trading costs and the recent price dynamics can be explained by the degree of new information content underlying transactions (e.g., Obizhaeva (2008a)). For example, if the positive information has been already partially incorporated into security prices in the pre-transition period and thus reflected in the upward price dynamics, then new coming buy orders might contain "stale" rather than "new" information. On the other hand, after an increase of security prices, new coming sale orders might potentially reflect the new negative information and induce significant price reduction.

To summarize, I find that the price impact increases with the dollar trading volume and the return standard deviation. At the same time, the effective spread on the traditional trading platforms and external crossing networks decreases with the dollar trading volume and the return standard deviations. It is also larger for Nasdaq-listed securities. Finally, the parameters of trading costs depend on the conventional proxies of information asymmetry and the price dynamics in pre-transition periods.

7 Conclusion

This paper is a study of price impact and effective spread. Its contribution to the existing literature is threefold. First, I develop a bias-free methodology for the trading costs estimation, which allows to obtain unbiased estimates even when trading strategies are price dependent or trades are canceled. Second, I apply this methodology to a unique dataset of portfolio transition trades. Third, I document a number of new findings about the parameters of trading costs for a period from 2001 to 2005 as well as their cross-sectional properties.

I find that, on average, 1% of the average daily trading volume has price impact equal to 0.33% of the daily returns standard deviation. For both open markets and external crossing networks, the estimates of the effective spread are roughly 13% of the daily returns standard deviations, thus questioning a claim that crossing networks help avoiding spread costs. At the same time, the effective spread in internal crossing networks is insignificant.

I also investigate the cross-sectional properties of trading costs. In a sharp contrast with conventional intuition, the price impact increases with the dollar trading volume: it is about tenfold higher for actively traded stocks than for thinly traded ones. The effective spread is lower for securities with higher trading volume. Furthermore, the price impact weakly decreases with the volatility, whereas the effective spread sharply increases with this characteristic. I also document that the price impact does not statistically differ for various trading platforms, whereas the effective spread is more significant for the Nasdaq-listed securities rather than for the NYSE/Amex-listed ones. Most aforementioned results call for further theoretical and empirical investigation.

References

Almgren, Robert, Chee Thum, Emmanuel Hauptmann, and Hong Li, 2005, Equity market impact, *Risk*, 57–62.

Almgren, Robert, Execution costs, 2008, *Encyclopedia of Quantitative Finance.*

Barra, 1997, Market impact model handbook.

Breen, William, Laurie Hodrick, and Robert Korajczyk, 2002, Predicting Equity Liquidity, *Management Science*, 48, 470-483.

Chan, Louis, and Josef Lakonishok, 1993, Institutional trades and intra-day stock price behavior, *Journal of Financial Economics* 33, 173–199.

Chan, Louis, and Josef Lakonishok, 1995, The behavior of stock prices around institutional trades, *Journal of Finance* 50(4), 1147–1174.

Chen, Zhiwu, Werner Stanzl, and Masahiro Watanabe, 2001, Price impact costs and the limit of arbitrage, *ICF Working Paper* No. 00–66.

Chiyachantana, Chiraphol, Pankaj Jain, Christine Jiang, and Robert Wood, 2006, Price history and asymmetry in the impact of institutional trades, *Working paper.*

Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2007, The cross-section of expected trading activity, *Review Financial Studies.*

Christie, William, and Paul Schultz, 1994, Why do NASDAQ market makers avoid odd-eighth quotes?, *Journal of Finance* 49(5), 1813–1840.

Dufour, Alfonso and Robert Engle, 2000, Time and the price impact of a trade, *Journal of Finance*, 55(6), 2467-2498.

French, Kenneth, William Schwert, and Robert Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics*, 19(1), 3–29.

Glosten, Lawrence, and Lawrence Harris, 1988, Estimating the components of the bid/ask spread, *Journal of Financial Economics*, 21(1), 123–142.

Hasbrouck, Joel, 1991a, Measuring the information content of stock trades, *Journal of Finance* 46(1), 179–207.

Hasbrouck, Joel, 1991b, The summary informativeness of stock trades: An econometric analysis, *Review of Financial Studies* 4(3), 571–595.

Hausman, Jerry, Andrew Lo, and Craig MacKinley, 1992, An ordered probit analysis of transaction stock prices, *Journal of Financial Economics*, 31, 319–379.

Holthausen, Robert, Richard Leftwich, and David Mayers, 1987, The effect of large block transactions on security prices: A cross-sectional analysis, *Journal of Financial Economics* 19(2), 237–67.

Holthausen, Robert, Richard Leftwich, and David Mayers, 1990, Large-block transactions, the speed of response, and temporary and permanent stock-price effects, *Journal of Financial Economics* 26(1), 71–95.

Huang, Roger, and Hans Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE, *Journal of Financial Economics*, 41(3), 313–357.

Keim, Donald, and Ananth Madhavan, 1995, Anatomy of the trading process empirical evidence on the behavior of institutional traders, *Journal of Financial Economics* 37(3), 371–398.

Keim, Donald, and Ananth Madhavan, 1997, Transactions costs and investment style: an inter-exchange analysis of institutional equity trades, *Journal of Financial Economics* 46(3), 265–292.

Kyle, Albert, and Anna Obizhaeva, 2008, *University of Maryland, Working paper.*

Kyle, Albert, 1985, Continuous auctions and insider trading, *Econometrica*, 53(6), 1315–1335.

Lillo, Fabrizio, Doyne Farmer, and Rosario Mantegna, 2003, Master curve for price-impact function, *Nature*, 421, 129–130.

Madhavan, Ananth, Matthew Richardson, Mark Roomans, 1997, Why do security prices change? A transaction-level analysis of NYSE stocks, *Review of Financial Studies*, 10(4), 1035–1064.

NYSE Research, 2006, NYSE market quality: Evidence from the 67 NASDAQ transferring stocks during 2002-2005.

Obizhaeva, Anna, 2008a, Information vs. liquidity: Evidence from portfolio transitions, *University of Maryland, Working paper.*

Obizhaeva, Anna, 2008b, Are ex ante price impact functions linear, *University of Maryland, Working paper.*

Perold, Andre, 1988, The implementation shortfall: Paper vs. reality, *Journal of Portfolio Management* 14(3), 4–9.

Ramistella, Alex, 2006, Crossing networks: Bringing back large block trades to institutional trading, *TowerGroup report.*

Saar, Gideon, 2001, Price impact asymmetry of block trades: An institutional trading explanation, *Review of Financial Studies* 14, 1153–1181.

U.S. Securities and Exchanges Commission, 2001, Report on the comparison of order executions across equity market structures.

Table 1: The Information on ADV and SD Bins

Panel A: Number of Observations in Each Bin (in thousands)

	adv1	adv2	adv3	adv4	adv5	adv6	adv7	adv8	adv9	adv10
sd1	12.84	18.32	11.51	12.82	7.87	9.16	9.16	10.25	11.83	14.53
sd2	16.05	17.86	10.89	12.39	6.91	7.56	7.66	9.74	11.49	13.39
sd3	21.87	21.38	12.03	13.50	7.44	7.09	8.02	9.14	11.93	12.76
sd4	29.15	26.77	14.49	16.18	8.08	8.60	8.26	9.51	11.22	14.07
sd5	37.55	36.38	19.66	22.29	12.07	11.32	10.58	11.12	11.11	12.52

Panel B: Median ADV in Each Bin (in millions of dollars)

	adv1	adv2	adv3	adv4	adv5	adv6	adv7	adv8	adv9	adv10
sd1	1.05	4.84	9.07	15.49	24.43	31.23	43.43	62.94	104.00	232.90
sd2	1.20	4.76	9.65	15.85	23.33	30.53	42.32	59.77	100.40	216.20
sd3	1.24	4.82	9.39	14.90	23.09	29.60	40.10	56.98	99.47	196.30
sd4	1.18	4.75	9.27	14.17	21.38	29.16	38.59	55.49	96.81	198.20
sd5	1.21	4.72	9.27	14.37	20.74	28.23	37.92	54.36	92.31	199.10

Panel C: Median SD in Each Bin

	adv1	adv2	adv3	adv4	adv5	adv6	adv7	adv8	adv9	adv10
sd1	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
sd2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
sd3	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
sd4	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
sd5	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04

Table shows the main characteristics of stocks in 50 volume-standard deviation bins. These bins are based on the average daily volume of securities, $ADV_\$$, and the volatility of daily returns, σ_r, in the previous to transition month. Each month, thresholds are recalculated and stocks are reshuffled across the bins. For the average daily volume, the thresholds are 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th quintiles for common NYSE-listed stocks; for standard deviation, the thresholds are 20th, 40th, 60th, and 80th quintiles. Panel A presents the number of observations (in thousands). Panel B presents the median values of the dollar average daily trading volume of each stock (in million of dollars). Panel C presents the median values of the daily standard deviations. The sample ranges from January 2001 to December 2005.

Table 2: The Estimates of Price Impact and Effective Spread

			NYSE		NASDAQ	
		All	Buy	Sell	Buy	Sell
$\tilde{\lambda}$	M1	0.39***	0.49*	0.40**	0.96***	0.32***
		(6.06)	(2.35)	(3.05)	(7.18)	(6.11)
	M2	0.33***	0.56*	0.34***	0.91***	0.26***
		(5.80)	(2.58)	(3.53)	(8.59)	(7.44)
	%Δ	18%	-13%	18%	5%	23%
$\tilde{\alpha}_{omt}$	M1	6.25***	7.36***	3.22***	8.97***	6.24***
		(11.04)	(7.40)	(3.50)	(7.52)	(4.83)
	M2	6.55***	8.47***	2.46*	11.62***	5.39***
		(9.58)	(7.25)	(2.49)	(8.89)	(3.48)
	%Δ	-5%	-13%	31%	-23%	16%
$\tilde{\alpha}_{ec}$	M1	6.53***	6.94***	4.23***	7.91***	7.40***
		(13.53)	(8.69)	(5.30)	(6.32)	(5.45)
	M2	6.94***	6.94***	4.50***	7.80***	9.01***
		(12.19)	(6.99)	(5.10)	(5.21)	(5.93)
	%Δ	-6%	0%	-6%	1%	-18%
$\tilde{\alpha}_{ic}$	M1	1.33	2.64	-1.03	2.69	2.32
		(1.77)	(1.82)	(-0.97)	(1.05)	(1.73)
	M2	0.22	0.92	-2.18	1.42	2.63
		(0.25)	(0.56)	(-1.83)	(0.44)	(1.76)
	%Δ	505%	187%	-53%	89%	-12%
	#Obs	702,406	210,194	228,934	123,841	139,437

Table shows the trading costs estimates $\tilde{\lambda}, \tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}, \tilde{\alpha}_{ic}$ based on the conventional methodology (10) as M1 and on the bias-free methodology (12) as M2. For each estimate, %Δ is the percentage difference equal to $(M1 - M2)/M2 \times 100\%$. Results are also presented for stocks listed at the NYSE/Amex and the Nasdaq as well as for buy and sell orders, separately. The estimates of $\tilde{\lambda}$ quantifies by how much (in percent of the daily standard deviation) the 1-% of the average daily volume changes the security price. The estimates $\tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}, \tilde{\alpha}_{ic}$ are the effective half spread in traditional markets, external and internal crossing networks (in percent of the daily standard deviation). The standard errors are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005.**is significance at 1% level, *is significance at 5% level, †is significance at 10% level.

Table 3: Price Impact and Spread for ADV Groups

			NYSE		NASDAQ	
		All	Buy	Sell	Buy	Sell
$\tilde{\lambda}$	x $adv1$	0.27***	0.43**	0.21**	0.80***	0.23***
		(7.84)	(2.61)	(3.05)	(7.60)	(10.30)
	x $adv5$	2.22***	1.54*	2.77***	0.65	3.56**
		(5.83)	(2.29)	(4.69)	(0.48)	(2.60)
	x $adv10$	2.99***	3.36*	2.76**	-2.06	7.52*
		(3.60)	(2.33)	(3.01)	(-0.81)	(2.02)
$\tilde{\alpha}_{omt}$	x $adv1$	14.79***	15.52***	10.96***	16.19***	13.75***
		(10.23)	(8.85)	(4.21)	(7.87)	(5.79)
	x $adv5$	3.36**	9.41***	-1.76	9.19***	-2.56
		(2.90)	(4.93)	(-0.94)	(3.62)	(-0.88)
	x $adv10$	2.79**	2.63	2.94	5.15	0.46
		(2.72)	(1.37)	(1.74)	(1.79)	(0.17)
$\tilde{\alpha}_{ec}$	x $adv1$	18.10***	17.15***	17.08***	18.54***	17.19***
		(15.96)	(9.88)	(8.37)	(9.60)	(9.10)
	x $adv5$	1.53	2.50	-1.51	3.39	3.73
		1.53	2.50	-1.51	3.39	3.73
	x $adv10$	3.45***	3.63	4.98**	3.92	-2.00
		(3.50)	(1.87)	(3.17)	(1.53)	(-0.66)
$\tilde{\alpha}_{ic}$	x $adv1$	-0.18	-4.52	-1.25	-5.56	6.75**
		(-0.06)	(-0.92)	(-0.47)	(-0.81)	(2.64)
	x $adv5$	0.57	4.84	-4.86*	7.14	-0.87
		(0.32)	(1.62)	(-1.96)	(1.37)	(-0.24)
	x $adv10$	-2.60*	-3.59	-2.63	-3.25	1.23
		(-2.50)	(-1.89)	(-1.46)	(-1.00)	(0.39)
Adj. $R2$		0.007	0.006	0.004	0.013	0.011

Table shows the results of regression (12) augmented with the interaction terms of $\tilde{\lambda}$, $\tilde{\alpha}_{omt}$, $\tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$ with ten dummy variables for each of ten trading volume groups. These groups are based on the average daily volume of securities, $ADV_{\$}$, in the previous to transition month. The thresholds are 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th quintiles of the average daily volume for common NYSE-listed stocks. Each month, thresholds are recalculated and stocks are reshuffled across the bins. Results are presented for stocks listed at the NYSE/Amex and the Nasdaq as well as for buy and sell orders separately. Only estimates for groups 1, 6, and 10 are reported. The standard errors are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005.**is significance at 1% level, *is significance at 5% level, †is significance at 10% level.

Table 4: The Cross-Sectional Regressions, Buy Orders

			Model 1		Model 2		Model 3		Model 4	
$\tilde{\lambda}$	x	1	1.55***	(5.53)	1.56***	(5.61)	1.72***	(6.12)	1.70***	(5.62)
	x	$ADV_{\$}$	1.43**	(3.01)	1.31**	(2.79)	0.99*	(2.05)	1.20*	(2.38)
	x	σ_r^h	0.23*	(2.41)	0.18*	(2.32)	-0.08	-0.94)	-0.08	(-1.12)
	x	$Nasdaq$	0.15	(1.14)	0.17	(1.22)	-0.17	(-1.07)	-0.05	(-0.35)
	x	$Spread$			-0.07**	(-2.86)				
	x	$\#ANAL$					0.09	(0.56)		
	x	$Ret_{-5,0}$						-0.06	(-0.98)	
$\tilde{\alpha}_{omt}$	x	1	7.66***	(6.82)	8.02***	(7.22)	7.21***	(6.37)	6.67***	(5.92)
	x	$ADV_{\$}$	-1.62**	(-2.74)	-0.97	(-1.61)	-0.44	(-0.61)	-2.12***	(-3.54)
	x	σ_r^h	-2.07*	(-2.29)	-2.43**	(-2.67)	-1.80*	(-1.99)	-1.35	(-1.50)
	x	$Nasdaq$	4.26***	(3.39)	3.38**	(2.82)	4.19***	(3.33)	4.75***	(3.70)
	x	$Spread$			2.92***	(4.88)				
	x	$\#ANAL$					-1.31	(-1.94)		
	x	$Ret_{-5,0}$							-2.49**	(-3.25)
$\tilde{\alpha}_{ec}$	x	1	5.52***	(5.45)	5.83***	(5.76)	5.13***	(5.10)	5.30***	(5.12)
	x	$ADV_{\$}$	-1.20*	(-2.50)	-0.76	(-1.61)	0.51	(0.98)	-1.40**	(-2.86)
	x	σ_r^h	-4.08***	(-5.25)	-4.33***	(-5.66)	-3.88***	(-4.96)	-3.56***	(-4.46)
	x	$Nasdaq$	3.72**	(3.12)	2.96*	(2.53)	3.68**	(3.06)	3.40**	(2.74)
	x	$Spread$			2.02***	(3.85)				
	x	$\#ANAL$					-2.20***	(-3.65)		
	x	$Ret_{-5,0}$							-2.01***	(-3.33)
$\tilde{\alpha}_{ic}$	x	1	1.21	(0.66)	1.31	(0.72)	1.13	(0.62)	0.47	(0.25)
	x	$ADV_{\$}$	-1.33	(-1.68)	-1.34	(-1.74)	-1.77*	(-2.52)	-1.09	(-1.38)
	x	σ_r^h	1.80	(1.14)	1.82	(1.13)	2.42	(1.54)	2.70	(1.66)
	x	$Nasdaq$	-0.94	(-0.40)	-0.82	(-0.36)	-0.97	(-0.42)	-1.26	(-0.52)
	x	$Spread$			0.27	(0.32)				
	x	$\#ANAL$					1.07	(1.14)		
	x	$Ret_{-5,0}$							-2.56*	(-2.33)
		Adj. $R2$		0.008		0.009		0.008		0.009
		#Obs		333,467		331,593		311,251		314,888

Table shows the results of regression (12) augmented with the interaction terms of $\tilde{\lambda}$, $\tilde{\alpha}_{omt}$, $\tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$ with various explanatory variables. The sample includes only buy order. Four regression models are considered. Additional explanatory variables are $ADV_{\$}$, the average daily volume in the previous month in million of dollars, σ_r^h, the standard deviation of daily returns in the previous month, $Sprd$, the percentage spread in basis points, $\#ANAL$, the number of analysts following stocks in the previous month, $Ret_{-5,0}$, the stock return over five days before portfolio transition starts. Scores of these variables are interacted with trading costs variables (3). Interaction terms with a dummy $Nasdaq$, which is equal to 1 for Nasdaq stocks and to 0 otherwise, is also included. The standard errors are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005.**is significance at 1% level, *is significance at 5% level, †is significance at 10% level.

Table 5: The Cross-Sectional Regressions, Sell Orders

			Model 1		Model 2		Model 3		Model 4	
$\tilde{\lambda}$	x	1	1.93***	(6.50)	1.94***	(6.52)	2.19***	(7.71)	1.79***	(6.11)
	x	$ADV_{\$}$	2.76***	(5.45)	2.78***	(5.45)	1.81***	(3.41)	2.61***	(5.22)
	x	σ_r^h	0.01	(0.31)	0.00	(0.11)	0.03	(0.42)	0.01	(0.15)
	x	$Nasdaq$	-0.05	(-0.49)	0.04	(0.42)	-0.13	(-1.00)	-0.01	(-0.02)
	x	$Spread$			-0.02***	(-4.41)				
	x	$\#ANAL$					0.34*	(2.37)		
	x	$Ret_{-5,0}$							-0.02	(-0.38)
$\tilde{\alpha}_{omt}$	x	1	2.02*	(2.07)	2.69**	(2.77)	1.30	(1.31)	2.13*	(2.14)
	x	$ADV_{\$}$	-1.17	(-1.94)	-0.30	(-0.50)	1.10	(1.77)	-1.16	(-1.88)
	x	σ_r^h	-0.19	(-0.28)	-0.92	(-1.38)	-0.13	(-0.19)	-0.47	(-0.69)
	x	$Nasdaq$	2.88*	(2.22)	1.24	(0.96)	1.87	(1.42)	3.40*	(2.54)
	x	$Spread$			4.67***	(8.88)				
	x	$\#ANAL$					-2.52***	(-3.75)		
	x	$Ret_{-5,0}$							1.07	(1.65)
$\tilde{\alpha}_{ec}$	x	1	4.14***	(4.51)	4.69***	(5.03)	3.31***	(3.64)	3.88***	(4.17)
	x	$ADV_{\$}$	-2.40***	(-4.89)	-1.27**	(-2.59)	0.57	(1.06)	-2.20***	(-4.43)
	x	σ_r^h	-0.08	(-0.12)	-0.84	(-1.30)	-0.20	(-0.30)	0.08	(0.12)
	x	$Nasdaq$	4.11**	(3.09)	1.54	(1.18)	2.81*	(2.10)	4.44**	(3.23)
	x	$Spread$			5.00***	(10.23)				
	x	$\#ANAL$					-3.66***	(-6.34)		
	x	$Ret_{-5,0}$							3.17***	(4.58)
$\tilde{\alpha}_{ic}$	x	1	-2.04	(-1.70)	-1.88	(-1.57)	-1.99	(-1.68)	-1.68	(-1.35)
	x	$ADV_{\$}$	-0.53	(-1.01)	-0.53	(-0.99)	0.92	(1.53)	-0.37	(-0.68)
	x	σ_r^h	2.73***	(3.46)	2.73***	(3.42)	2.41**	(2.99)	3.00***	(3.67)
	x	$Nasdaq$	3.21*	(2.27)	3.04*	(2.11)	2.39	(1.68)	3.50*	(2.46)
	x	$Spread$			0.67	(0.79)				
	x	$\#ANAL$					-2.64***	(-3.44)		
	x	$Ret_{-5,0}$							4.88***	(5.70)
		Adj. $R2$		0.005		0.006		0.005		0.005
		#Obs		368,358		365,336		342,702		348,324

Table shows the results of regression (12) augmented with the interaction terms of $\tilde{\lambda}$, $\tilde{\alpha}_{omt}$, $\tilde{\alpha}_{ec}$, and $\tilde{\alpha}_{ic}$ with various explanatory variables. The sample includes only sell order. Four regression models are considered. Additional explanatory variables are $ADV_{\$}$, the average daily volume in the previous month in million of dollars, σ_r^h, the standard deviation of daily returns in the previous month, $Sprd$, the percentage spread in basis points, $\#ANAL$, the number of analysts following stocks in the previous month, $Ret_{-5,0}$, the stock return over five days before portfolio transition starts. Scores of these variables are interacted with trading costs variables (3). Interaction terms with a dummy $Nasdaq$, which is equal to 1 for Nasdaq stocks and to 0 otherwise, is also included. The standard errors are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005.**is significance at 1% level, *is significance at 5% level, †is significance at 10% level.

Figure 1: Conventional and Bias-Free Methodologies

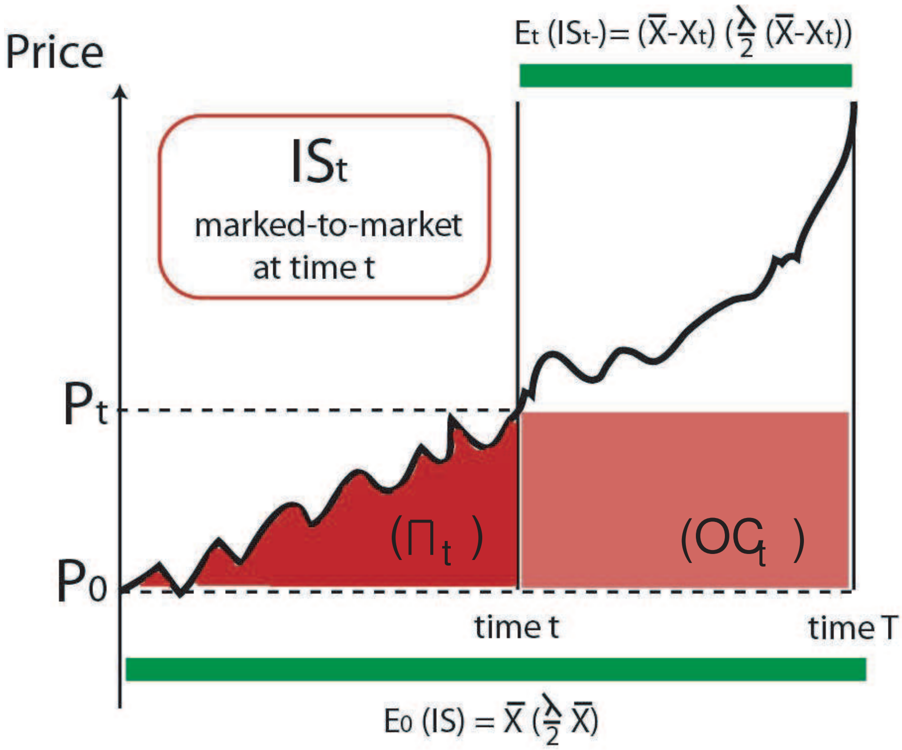
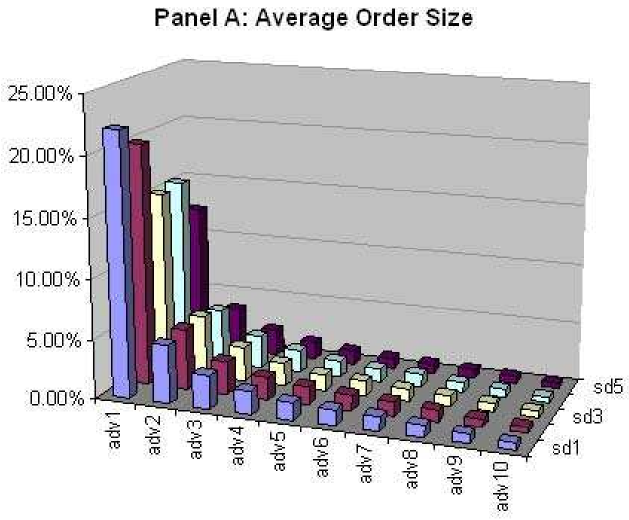


Figure highlights intuition behind bias-free and conventional methodologies of estimation the parameters of trading costs. Implementation shortfall IS_t mark-to-market at time t consists of two parts: the trading costs, Π_t, denoted as (1), and the opportunity costs, OC_t, denoted as (2). The mark-to-market implementation shortfall IS_t is the difference between expected implementation cost of total ex ante order $\bar{X}$, i.e. $\frac{\lambda}{2}\bar{X}^2$, and the part $\bar{X} - X_t$ still to be executed at time t, i.e. $\frac{\lambda}{2}(\bar{X} - X_t)^2$. This intuition underlies a bias-free methodology. At the same time, for any $t \in [0, T)$, due to dependence of trading strategy on the price dynamics, the expected value of realized trading costs, Π_t, is not equal to $\frac{\lambda}{2}X_t^2$. Consequently, the invalidity of conventional method can not be justified.

Figure 2: Order Sizes across ADV and SD Bins



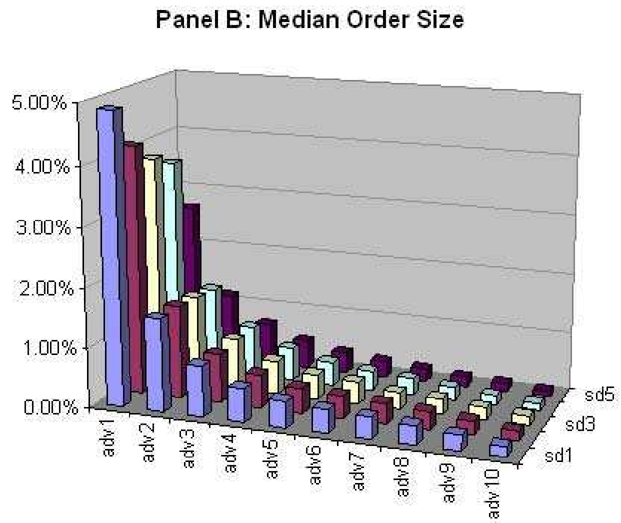


Figure shows the order sizes as the percentage of the average daily volume in the previous. Panel A presents the average values of order sizes. Panel B presents the median values of order sizes. The observations are split into 10×5 bins. These bins are based on the average daily volume of securities, $ADV_{\$}$, and the volatility of daily returns, σ_r, in the previous to transition month. Each month, thresholds are recalculated and stocks are reshuffled across the bins. For the average daily volume, the thresholds are 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th quintiles for common NYSE-listed stocks; for standard deviation, the thresholds are 20th, 40th, 60th, and 80th quintiles. The sample ranges from January 2001 to December 2005.

Figure 3: Price Impact and Effective Spread Across ADV Bins

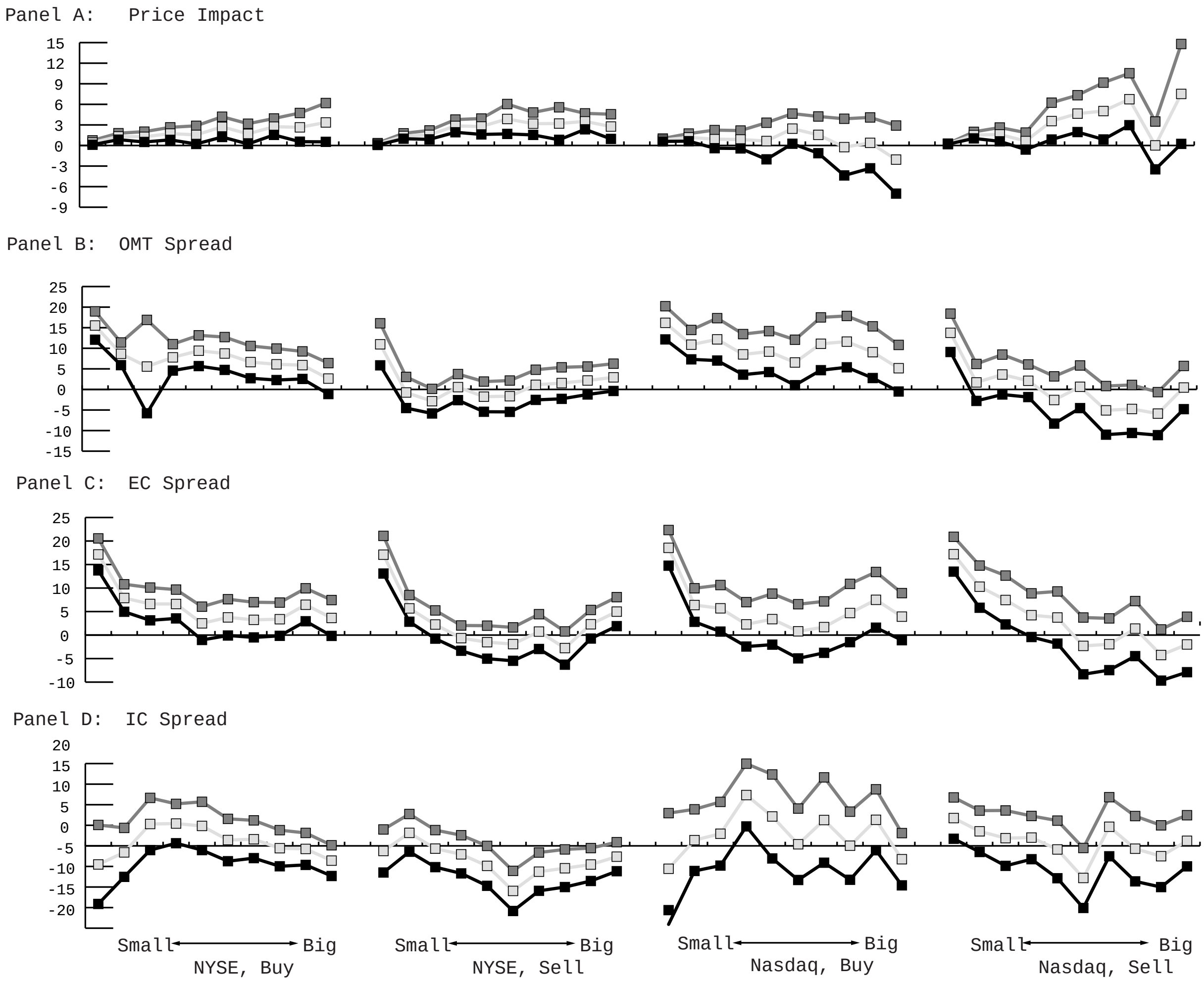
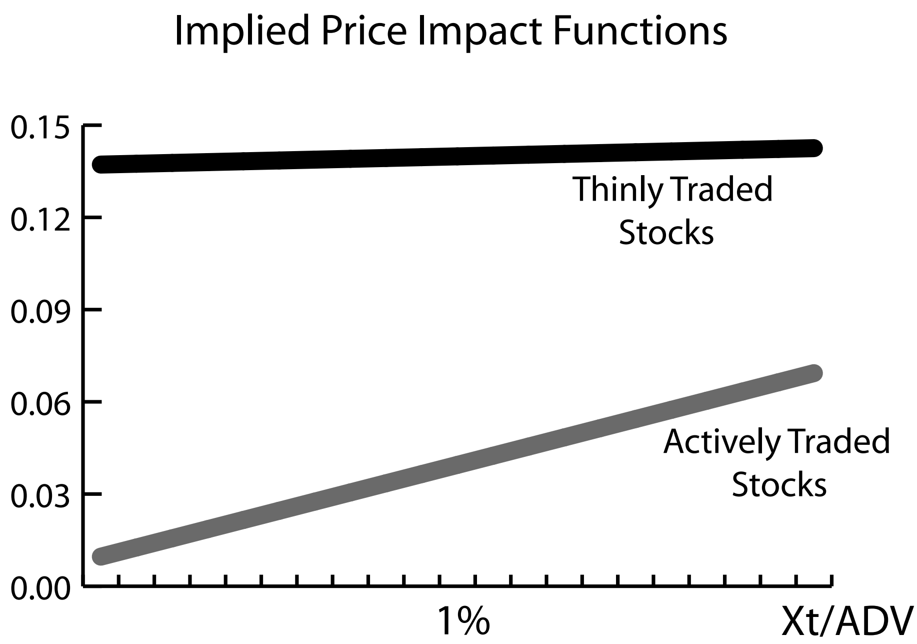


Figure shows the estimates of the price impact, $\tilde{\lambda}$, and the effective half spread in open markets, external crossing networks and internal crossing networks, $\tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}, \tilde{\alpha}_{ic}$ in Panel A, Panel B, Panel C and Panel D, respectively, for stocks with different dollar trading volume. Each month, observations are split into 10 bins according to stocks' average daily volume, $ADV_{\$}$, in the previous month (from Small stocks to Big stocks). The thresholds are 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th quintiles of the average daily volume for common NYSE-listed stocks. Each month, thresholds are recalculated and stocks are reshuffled across the bins. The estimates of the coefficients from regression (12), which is augmented with the interaction terms between the trading costs parameters (3) and the dummy variables for each group, are depicted together with their 95%-confidence intervals. The observations for the NYSE/Amex-listed and the Nasdaq-listed stocks as well as buy and sell orders are considered separately. The standard errors are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005.

Figure 4: Cross-sectional Properties of Trading Costs



On the y-axe, I plot the difference between average execution price and pre-trade benchmark price, $\frac{dP_t}{P_t}$, in percents of the daily standard deviation σ_r. On the x-axe, I plot the trade size, X_t in percents of average daily volume, ADV. There are two components in this price change. The first component is due to the price impact. It is calculated for price dynamics (2). The second component is due to bid-ask spread. It is calculated for a typical split between trading venues. For thinly (actively) traded stocks, these proportions are 31% (24%) for open markets, 45% (38%) for external crossing networks, and 17% (38%) for internal crossing networks. The estimates of parameters $\tilde{\lambda}, \tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}, \tilde{\alpha}_{ic}$ from Table 3 are used. Results are presented for 30% of the least actively traded stocks (thinly traded stocks, Group 1) and 5% of the most actively traded stocks (actively traded stocks, Group 10). The sample ranges from January 2001 to December 2005.

Figure 5: Price Impact and Spread Across ADV and SD Bins

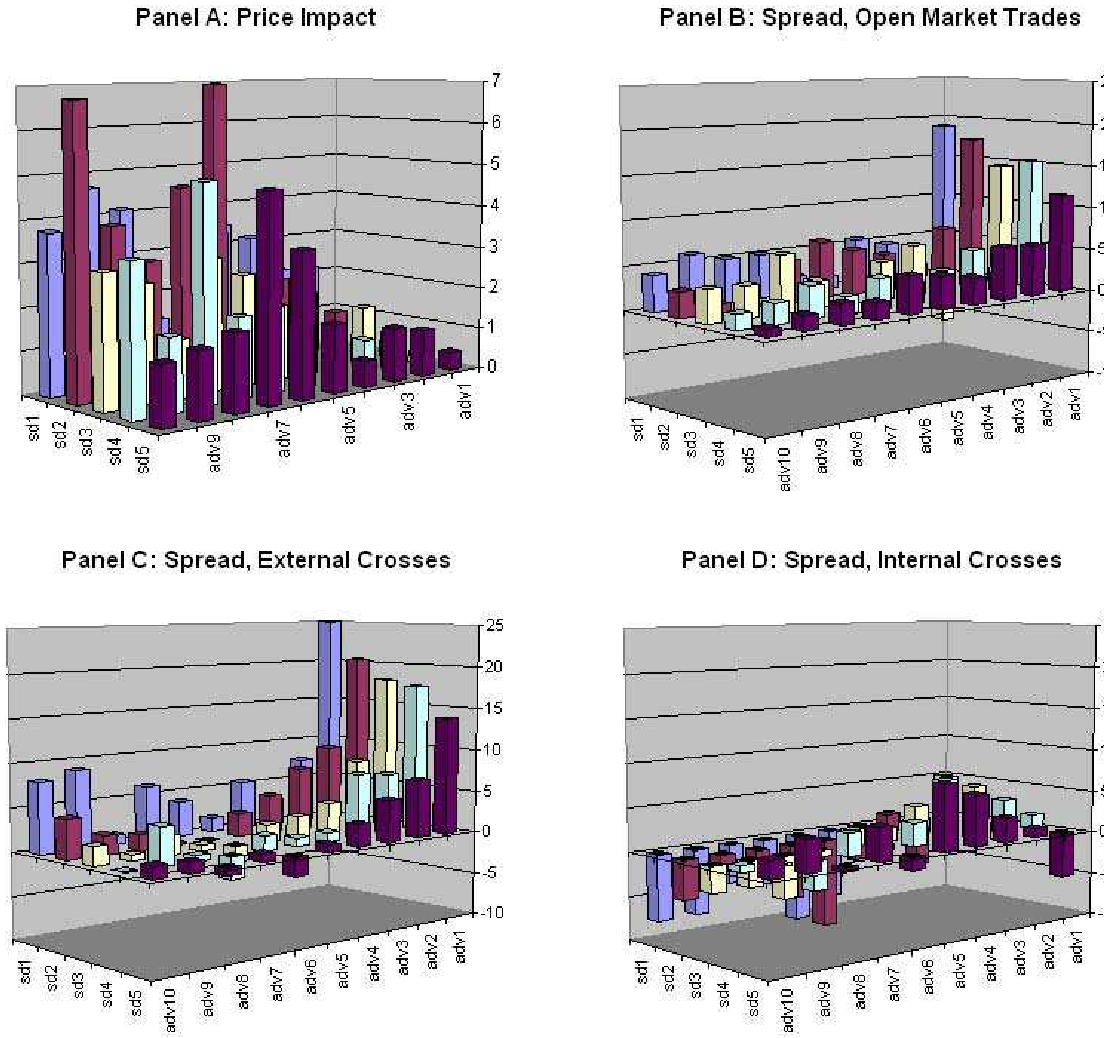


Figure shows the estimates of the price impact, $\tilde{\lambda}$, and the effective half spread in open markets, external crossing networks and internal crossing networks, $\tilde{\alpha}_{omt}, \tilde{\alpha}_{ec}, \tilde{\alpha}_{ic}$ in Panel A, Panel B, Panel C and Panel D, respectively, for stocks with different dollar trading volume, $ADV_{\$}$, and return standard deviation, $\hat{\sigma}^h_{i,t}$. Bias-free methodology is used to obtain these estimates. The observations are split into 10×5 bins according to the average daily volume of securities, $ADV_{\$}$, and the volatility of daily returns, σ_r, in the previous to transition month. Each month, thresholds are recalculated and stocks are reshuffled across the bins. For the average daily volume, the thresholds are 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th quintiles for common NYSE-listed stocks; for standard deviation, the thresholds are 20th, 40th, 60th, and 80th quintiles. The estimates of the coefficients from regression (12), which is augmented with the interaction terms between the trading costs parameters (3) and the dummy variables for each group, are reported. The standard errors are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005.