

Liquidity Estimates and Selection Bias

Anna A. Obizhaeva*

July 5, 2012

Abstract

Since traders often employ price-dependent strategies and cancel expensive orders, conventional estimates tend to overestimate available liquidity. This paper studies trading costs using the sample of portfolio transition trades. The known exogeneity of these trades eliminates the selection bias problem. We estimate a piece-wise linear price impact functions with the intercept corresponding to fixed spread costs and the slope corresponding to variable price impact costs. Buy orders are more expensive than sell orders due to specific institutional features of portfolio transitions. For high-volume stocks, small trades are executed at a discount relative to a piece-wise linear specification. Since the size of this discount is comparable to bid-ask spread, we attribute documented non-linearity to ability of portfolio transition managers sometimes to earn bid-ask spread instead of paying it by providing liquidity to other market participants.

Keywords: market impact, spread, trading costs, selection bias.

JEL classification: G14

*Robert H. Smith School of Business, University of Maryland, obizhaeva@rhsmith.umd.edu.

Introduction

It is difficult to obtain good estimates of transaction costs. As discussed by Hasbrouck (2007), traders usually employ price-dependent strategies and often choose not to execute their orders entirely. When quantities traded and price changes are used to obtain the estimates of transaction costs, these estimates tend to be biased due to the endogeneity of trading strategies.

The following example illustrates that liquidity estimates can be significantly distorted by a selection bias. Consider a trader who intends to buy 100,000 shares of stock. At the time he places his order, the price is \$40 per share. After the trader purchases 80,000 shares at an average price of \$40.20, the price moves adversely to \$45 per share, at which point the trader decides to cancel the remaining 20,000 shares. Usually, a database contains the 80,000 shares executed at an average price of \$40.20 but no indication of 20,000 shares not executed at a price which would have been about \$45 per share. In this situation, the trading costs of executed trades are calculated as 50 basis points for 80,000 shares. A 50 basis point number is a biased estimate of actual trading costs, because it fails to take account of the cost of at least 1250 basis points that would have been incurred on the 20,000 share portion of the order that was canceled. Underestimation of trading costs have implications for the actual limits to arbitrage in financial markets, performance of asset managers, and viability of different trading strategies.

The endogeneity problem is significant, because a large fraction of contemplated orders is canceled. Hasbrouck and Saar (2005) report that only 13% of non-marketable limit orders, which account for 90% of all incoming orders for Nasdaq-listed stocks, are filled. It is difficult, however, to solve this problem: Researchers usually do not observe unexecuted trades, and it is difficult to reconstruct trading intentions from sequences of executed trades.

Prior literature adopts a number of valuable strategies to handle the selection bias problem. One line of research relies on multivariate linear models for joint dynamics of trades and prices. The examples include Hasbrouck (1988, 1991a, 1991b), Dufour and Engle (2000), and Madhavan, Richardson and Roomans (1997). The resolution of contemporaneous effects in these papers is achieved by imposing a restriction that innovations in trades can affect innovations in price changes, but not vice versa. The restriction though leaves out economically plausible situations when quantities executed depend on contemporaneous price dynamics (e.g., limit orders). Alternative approach is to model specific order-submission strategies (e.g., Lesmond, Ogden, and Trzcinka (1999)) or to predict quantities traded. Cheng and Madhavan (1997) and Conrad, Johnson and Wahal (2003), for example, use endogenous switching regressions to adjust for selection bias. These procedures involve predicting the choice of trading mechanism and quantities traded in a first-stage regression, and then using the fitted values in a second-stage regression. Yet, as the authors point out, a diverse price-path dependency of trading strategies confounds the estimates of trading costs and makes it difficult to assign a structural interpretation to the estimates as ex ante trading costs. Other papers suggest examining proprietary transaction data to better identify intended orders (e.g., Almgren et al. (2005), Chan and Lakonishok (1993, 1995), Keim and Madhavan (1997)). A different approach is to altogether disregard the selection bias problem. The representative examples include Breen, Glosten and Harris (2002), Glosten and Harris (1988), Holthausen, Leftwich and Mayers (1987, 1990), and Lillo, Farmer and

Mantegna (2003). All these methods help to obtain more precise estimates, but they offer only loose bounds on the true trading costs. The non-zero costs of foregone trades cannot be accurately measured without detailed understanding of the individual trader's trading strategies and, in particular, how these strategies depend on prices.

The prominent study of Amihud (2002) advocates for another convenient way to obtain liquidity estimates, referred to as illiquidity ratios, from prices and trading volume in the CRSP data set. Under the assumption that markets are sufficiently resilient and any random price changes dissipate by the end of the trading day, these estimates might avoid the selection bias problem. The illiquidity ratio is calculated as the ratio of price change to daily volume. Unfortunately, these estimates implicitly assume that the same number of orders is executed during the day for any stock and any time period. Since the number of orders executed per day is much bigger for high-volume stocks, the illiquidity ratios tend to overestimate liquidity available in markets for active stocks.

We address the selection bias problem in a different way. We exploit a proprietary database of portfolio transitions. Portfolio transitions are economically significant transactions initiated by institutional sponsors transferring funds from one portfolio to another. These transfers occur when sponsors replace their fund managers, rebalance their asset classes, or accommodate large cash inflows and outflows. Institutional sponsors usually delegate portfolio transitions to a transition manager. The transition manager replaces a legacy portfolio with a new portfolio by selling a portfolio held by the incumbent manager and buying a portfolio chosen by the new manager. The portfolio transition database includes 2,680 portfolio transitions that have been carried out by a leading vendor of portfolio transition services. It involves more than 400,000 individual orders executed over the period from 2001 to 2005.

The data provides a unique laboratory for the estimation of transaction costs. Because of several unique institutional properties, the data does not suffer from the selection bias problem. The transition manager gets the lists of orders to be sold and to be bought the night before the transition begins. He executes entire transition orders and cannot cancel requested transactions. This implies that there are no unexecuted trades. The feedback effect between quantities traded and price changes is thus broken because the former do not depend on the price dynamics during execution. Also, the timing of transitions is determined by a schedule of investment committee meetings of institutional sponsors, who make decisions to undertake transitions; portfolio transition follow shortly after the decision is made. This implies that portfolio transition trades are unlikely to be correlated with short-term price dynamics of individual securities during the transition period. Since there are many large portfolio transition orders in the dataset, our empirical tests also avoid the problem of low statistical power, associated with using execution data for estimation of transaction costs, because the price impact of transactions is usually very small relative to overall uncertainty in stock returns.

We estimate a piece-wise linear price impact functions with intercepts corresponding to bid-ask spread and slopes corresponding to price impact spread. On average, for a stock with daily volatility of two percent, the half price impact and the half spread are estimated to be 0.30 and 19.20 basis points, respectively. Buy orders appear to be more expensive than sell orders. A possible explanation of this asymmetry is a double-selection mechanism of how securities are selected to be traded during portfolio transitions. The securities to be

purchased are chosen by asset managers, who are in turn selected by institutional sponsors, as the best out of many candidates to get a mandate. The securities to be sold happened to be in legacy portfolios held by presumably less skilled asset managers. Portfolio transition sales are exogenous liquidations, which most likely do not have any negative information content.

The estimates vary significantly across stocks with different levels of trading activity. Per dollar traded, trading is usually more expensive in low-volume stocks than in high-volume stocks. Per percent of daily volume traded, by contrast, trading is usually less expensive in low-volume stocks than in high-volume stocks. Keeping daily volatility at two percent, the price impact costs of trading one percent of daily volume increase ten-folds from 0.25 basis points for low-volume stocks to 2.37 basis points for high-volume stocks. The spread costs drop from 39.29 basis points for low-volume stocks to 7.85 basis points for high-volume stocks. The intuition is that one percent of trading volume is much bigger relative to typical order imbalances for active stocks than for inactive stocks. Total trading costs have different properties. For inactively traded stocks, the costs are almost invariant with respect to trade sizes as a percent of daily volume, and spread-related payments accounts for a large part of the total costs. For actively traded stocks, the total trading costs are highly sensitive to trade sizes as a percent of daily volume, with spread-related payment being less significant.

The comparison of our estimates with those in previous studies is problematic due to a diversity of estimation procedures, definitions, and time periods. Directly comparing our estimates with those in the previous literature, we find that our estimates are similar. Breen, Glosten and Harris (2002), Hasbrouck (1991 a,b), Glosten and Harris (1988) find, for instance, that the average price impact of a 1000-share trade ranges from 18 to 30 basis points. The average price change after a comparable portfolio transition trade is 13 basis points for actively traded stocks and 40 basis points for inactively traded stocks (1000 shares roughly corresponds to two percents of average daily volume). However, most estimates in the previous work are obtained using pre-decimalization data, while our estimates are based on post-decimalization data when trading costs have decreased significantly. The similarity of our estimates therefore indicates that previous estimates probably overestimated liquidity in a pre-decimalization period, because they could not properly account for a selection bias.

We also examine non-linearity of price impact functions. In most theoretical studies, these functions are linear in equilibrium, whereas most empirical studies suggest that these functions are concave. The concavity is often attributed to a selection bias, because inexpensive orders are executed as a whole and expensive orders are broken into a sequence of smaller trades. Even though this argument is not valid for price impact functions based on portfolio transition orders, which are executed completely, we still find that price impact functions are concave. Even though the price impact functions are usually well approximated by a piece-wise linear functions that have two parts, a linear price impact part and a constant bid-ask spread, the approximation is not very accurate for high-volume stocks, for which small orders are executed at a “discount” relative to an approximation. Since the size of discount is comparable to the magnitude of bid-ask spread, we attribute this non-linearity to ability of traders sometimes to postpone execution of their orders and wait for the opportunities to earn the spread by filling out their orders on the opposite side of the book.

The remainder of this paper is structured as follows. Section 1 explains the selection

bias problem. Section 2 describes the data. Section 3 reports the estimates. Section 4 concludes.

1 Estimation of Trading Costs

This section describes how to estimate trading costs using quantities traded and explains why this procedure delivers biased estimates when trading decisions are endogenous.

1.1 Model with Linear Price Impact and Bid-Ask Spread

The parameters of trading costs are usually estimated using quantities traded and realized trading costs. Let \bar{Q}_i denote unsigned number of shares in order i intended to be executed over a fixed time period $[0, T_i]$. Let \mathbb{I}_i be an indicator of trading direction for order i , which is equal to 1 for buy orders and -1 for sell orders. Let $Q_{t,i}$ denote number of shares actually traded during time interval $[0, t)$. If the order is only partially executed by time T_i , then $Q_{T_i,i} < \bar{Q}_i$; if it is executed completely, then $Q_{T_i,i} = \bar{Q}_i$.

The execution of order i moves the midquote from pre-trade price $P_{0,i}$ to post-trade price $P_{t,i}$,

$$P_{t,i} = P_{0,i} + \lambda \cdot \mathbb{I}_i \cdot Q_{t,i} + \sigma_{P,i} \cdot (\tilde{Z}_{t,i} - \tilde{Z}_{0,i}), \quad (1)$$

where the price impact λ is the sensitivity of price changes to shares traded, $\sigma_{P,i}$ is the volatility of daily price changes, $\tilde{Z}_{t,i} - \tilde{Z}_{0,i}$ is the price change due to arrival of new information or contemporaneous transactions of other traders. In most theoretical models, the equilibrium relation between trades and permanent price changes is linear. Linear functions are consistent with the stylized model of Kyle(1985), in which an informed trader trades on his information with a risk-neutral market maker in the presence of a noise trader. Many other theoretical models rely on linear price impact functions (e.g., Admati and Pfleiderer (1988), Back (1992), Foster and Vishwanathan (1996)). Huberman and Stanzl (2000) show that only linear functions rule out a certain form of arbitrage. Almgren et al. (2005) find that linear specification is also consistent with empirical data.

Let $P_{t,i}^*$ be the average execution price of shares during time period $[0, t)$. Since the trader is a perfectly discriminating monopolist, he gradually executes order by walking up or down the demand curve. The average execution price is therefore a mid-point between pre-trade price $P_{0,i}$ and post-trade $P_{t,i}$. Since market makers may not be perfectly competitive, the actual transaction price may also deviate from the midquote by a half bid-ask spread $\kappa/2$,

$$P_{t,i}^* = (P_{t,i} + P_{0,i})/2 + \mathbb{I}_i \cdot 1/2 \cdot \kappa. \quad (2)$$

Our price impact model is a piece-wise linear function with two parts, a fixed bid-ask spread and a permanent linear price impact. This function is concave, as suggested by many studies (e.g., Almgren et al. (2005), Coppejans, Domowitz and Madhavan (2001), Chen, Stanzl and Watanabe (2001), Hasbrouck (1991), Hausman et al. (1992), Keim and Madhavan (1996), Kempf and Korn (1999), and Lilo et al. (2003)).

Let $\Pi_{t,i}$ denote the realized trading costs by time t of executing order i , in dollars,

$$\Pi_{t,i} = \mathbb{I}_i \cdot (P_{t,i}^* - P_{0,i}) \cdot Q_{t,i}. \quad (3)$$

For example, a trader decides to buy 100,000 shares, when a stock price is equal to \$40 ($\bar{Q}_i = 100,000$, $\mathbb{I}_i = 1$, $P_{0,i} = \$40$). If he executes the first 80,000 shares at the average price of \$42 and decides to cancel the remaining 20,000 shares because the price went up to \$45 ($Q_{t,i} = 80,000$, $P_{t,i}^* = \$42$), then the realized trading costs $\Pi_{t,i}$ are equal to $80,000 \cdot (42 - 40)$ or \$160,000.

Plugging (1) and (2) into (3), we obtain the following relation between the realized trading costs $\Pi_{t,i}$ per one dollar traded by time t and the parameters of price impact λ and bid-ask spread κ ,

$$\frac{\Pi_{t,i}}{P_{0,i} \cdot Q_{t,i}} = \frac{\lambda/2}{P_{0,t}} \cdot Q_{t,i} + \frac{\kappa/2}{P_{0,t}} + \sigma_i \cdot \tilde{\epsilon}_{t,i}. \quad (4)$$

The daily volatility of returns $\sigma_i = \sigma_{P,i}/P_{0,i}$ and a random variable $\tilde{\epsilon}_{t,i} = \mathbb{I}_i \cdot (\tilde{Z}_{t,i} - \tilde{Z}_{0,i})/2$. The regression equation (4) and its variations were extensively used in the previous work for estimation of transaction costs from data on quantities traded and realized trading costs. The examples include Breen, Hodrick, and Korajczyk (2002) and Glosten and Harris (1988).

1.2 Model with Non-Linear Price Impact and Bid-Ask Spread

The piece-wise linear specification is a special case in a class of more general non-linear price impact functions. One way to model non-linear effects is to use power functions, as discussed in Lillo et al. (2003) and Almgren et al. (2003):

$$P_{t,i} = P_{0,i} + \lambda^p \cdot \mathbb{I}_{BS,i} \cdot (Q_{t,i})^{z+1} + \sigma_{P,i} \cdot (\tilde{Z}_{t,i} - \tilde{Z}_{0,i}). \quad (5)$$

Plugging (5) into (3) and re-scaling parameters, we derive the regression equation:

$$\frac{\Pi_i}{P_{0,i} Q_{t,i}} = \frac{1}{z+2} \cdot \frac{\lambda^p}{P_{0,t}} \cdot Q_{t,i}^{z+1} + \frac{\kappa^p/2}{P_{0,t}} + \tilde{\epsilon}_i. \quad (6)$$

There are three parameters to estimate: the market impact λ^p , its curvature z , and the bid-ask spread κ^p . The superscript ‘‘p’’ emphasizes that these are estimates for the power specification. The linear specification (4) is a special case of equation (6). The market impact is convex if $z > 0$, concave if $z < 0$, and linear if $z = 0$.

Power specification restricts the marginal cost of infinitesimal trade to be infinite. A less restrictive specification is an inverse-quadratic model with market impact λ_0 and λ_1 specified at two levels of order sizes x_0 and x_1 . This model can be derived from equation $x = g(y) = ay^2 + by + c$, where x stands for the order size and y stand for the price change, with the following restrictions:

$$\begin{cases} g^{-1}(0) = 0, \\ \frac{\partial}{\partial x} g^{-1}(x)|_{x=x_0} = \lambda_0, \\ \frac{\partial}{\partial x} g^{-1}(x)|_{x=x_1} = \lambda_1. \end{cases} \quad (7)$$

The first equation ensures that market impact of a zero trade is equal to zero. The last two equations are definitions of λ_0 and λ_1 quantifying the slopes of functions at x_0 and x_1 , respectively. Parameters x_0 and x_1 are arbitrary scaling constants, we will later fix them at 10 and 100 basis points of average daily volume. If $\lambda_0 = \lambda_1$, then the model is linear. Solving

this system and finding constants a , b , and c , we can write down the inverse-quadratic price impact function,

$$P_{t,i} = P_{0,i} + \mathbb{I}_{BS,i} \cdot \lambda_0 \cdot \alpha \cdot Q_{t,i} \left(\frac{1}{2} \sqrt{1 + \beta \cdot Q_{t,i}} + \frac{1}{2} \right)^{-1} + \sigma_{P,i} \cdot (\tilde{Z}_{t,i} - \tilde{Z}_{0,i}), \quad (8)$$

with parameters

$$\alpha = \sqrt{\frac{\lambda_1^2(x_1 - x_0)}{x_1\lambda_1^2 - x_0\lambda_0^2}} \quad \text{and} \quad \beta = \frac{\lambda_0^2 - \lambda_1^2}{x_1\lambda_1^2 - x_0\lambda_0^2} \quad (9)$$

Plugging (8) into (3) and re-scaling parameters, we derive the regression equation for the inverse-quadratic specification:

$$\frac{\Pi_i}{P_{0,i}Q_{t,i}} = \frac{\lambda_0^q}{P_{0,t}} \cdot Q_{t,i} \frac{\alpha(1 + 4/3\beta Q_{t,i})}{(1 + \beta Q_{t,i})^{3/2} + 3/2\beta Q_{t,i} + 1} + \frac{\kappa^q/2}{P_{0,t}} + \tilde{\epsilon}_i. \quad (10)$$

The variables λ_0^q and λ_1^q denote the market impact at points x_0 and x_1 . The superscript “q” emphasizes that these are estimates from the inverse quadratic specification. The linear specification (4) is a special case of equation (10). The permanent price impact function is linear if $\lambda_0^q = \lambda_1^q$.

1.3 Selection Bias Problem

The regression equations (4), (6) and (10) result in unbiased estimates of price impact and bid-ask spread only if the corresponding orthogonality conditions are satisfied. These conditions, however, rarely holds in the data, because investment decisions and trading decisions are usually price dependent.

Indeed, investment strategies are often related to price dynamics. For example, some traders construct their strategies based on short-term information. They buy stocks if information is positive and sell stocks if information is negative. As traders implement their strategies, prices tend to change in the same direction exacerbating transaction costs. The estimates of trading costs will appear to be too high.

Trading strategies may also be price dependent. For example, many traders tend to slow down trading or even cancel orders, when the market runs away from them. Relatively expensive orders are canceled, and relatively cheap orders are executed. The estimates of trading costs will appear to be too low.

The direction of a selection bias ultimately depends on the correlation between price changes $\tilde{\epsilon}_{t,i}$ and quantities traded $Q_{t,i}$. For example, if traders tend to cancel expensive orders, then we will underestimate the price impact λ and overestimate the bid-ask spread κ (to match the averages of the left-hand side and right-hand side in equation (4)).

The selection bias is the big problem in most available datasets. For example, the Trades and Quotes (TAQ) dataset contains data on executed trades, generated by a complicated mixture of presumably price-dependent strategies of all market participants. The proprietary datasets such as Plexus and Abel Noser datasets also have only quantities traded, and intended orders are usually reconstructed based on quantities traded over some time window.

Sometimes we observe not only executed quantities but also intended quantities, as for example, with data on execution of limit orders. In this case, it is possible to mitigate a selection bias by taking into account unrealized trading costs of unexecuted shares. For example, Handa and Schwartz (1995) as well as Harris and Hasbrouck (1996) suggest to take into account unrealized bid-ask spread costs of unexecuted trades by imputing to the canceled portion of a limit order a fill at the opposite-side quote prevailing at the time of cancelation (e.g. the ask price for buy order). The methodology can be adjusted to account for potential price impact costs of unexecuted trades. It is impossible, however, to eliminate a selection bias problem completely.

As we explain next, the data on portfolio transitions provides a unique laboratory for transaction costs estimation. Portfolio transition trades are ideal for estimation of transactions costs, because the known exogeneity of the size of the trades eliminates the problem of selection bias. There are no canceled orders. The opportunity cost is always equal to zero. Trading costs coincide with the implementation shortfall defined in Perold (1988).

2 Data

2.1 Portfolio Transition Data

We exploit a proprietary database of portfolio transitions from a leading vendor of portfolio transition services. During the evaluation period, this portfolio transition vendor supervised more than 30 percent of outsourced U.S. portfolio transitions. The sample includes about 2,680 portfolio transitions executed over the period from 2001 to 2005. This database is derived from the post-transition reports prepared by transition managers for their U.S. clients. These reports contain detailed information on execution of portfolio transitions, thoroughly verified and discussed by both transition managers and institutional sponsors after their implementation.

Portfolio transition data have several important properties which make it particularly advantageous for estimating trading costs.

First, quantities traded are not influenced by price changes between the time orders are placed and the time they are executed. The composition of legacy and target portfolios is fixed in the mandates that transition managers receive the night before portfolio transitions begin. Transition managers then execute entire orders regardless of price dynamics. There are no order cancelations or modifications.

Second, quantities intended are not affected by short-term price dynamics. The timing of portfolio transitions is likely determined by a schedule of investment committee meetings of institutional sponsors, who make decisions to undertake transitions. The investment committee meets regularly on schedules set well in advance of the meetings. Among the issues boards discuss are the replacement of fund managers and the changes of asset mix. If a decision is made to replace a portfolio manager, then a portfolio transition is arranged shortly after the meeting. These decisions are unlikely to be correlated with short-term price dynamics of individual securities during the transition period.

These properties are not often shared by other data. Consider a database built up from trades by a mutual fund, a hedge fund, or a proprietary trading desk at an investment bank. In such samples, the trading intentions are usually not recorded, they change over time

and often depend on price dynamics, as traders follow contrarian or momentum strategies. Moreover, they often condition trading strategies on prices by using limit orders or canceling parts of their orders, thus hard-wiring the selection bias problem into the data.

The portfolio transitions data also make it possible to deal with another major problem associated with using implementation shortfall to estimate trading costs, the problem of low statistical power. Suppose, for example, that a trade of one percent of average daily volume has a transactions cost of 20 basis points, but the stock has a price volatility of 200 basis points per day. If we think of the 20 basis points as a random variable which could be positive or negative depending on whether the underlying transition order is a buy or a sell, then the transition order adds about 1% to the variance of the stock's return. This implies that a properly specified regression to estimate trading costs using implementation shortfall is going to have an R^2 of about 0.01, and statistical power is going to be low. Clearly, larger trades with higher trading costs reduce this problem and make the transactions cost easier to estimate.

The portfolio transition database addresses the problem of low statistical power in two ways. First, the data has the large number of degrees of freedom increasing the statistical power of the estimates. Second, some orders are large enough to induce relatively significant trading costs and increase statistical power as well.

The portfolio transitions database contains the data on individual transactions. Each observation has the following fields: a trade date, an identifier of a portfolio transition, its starting and ending dates, the name of the stock traded, the number of shares traded, buy or sell indicator, the average execution price, the pre-transition benchmark price, commissions, and fees. The data is given on separate lines for three trading venues: internal crossing networks, external crossing networks, and open market transactions. It is also given separately for each trading day in a trading package. Old and new portfolios usually overlap. For example, both portfolios may have positions in some large and therefore widely held securities. Instead of first selling overlapping holdings from legacy portfolios and then acquiring them into target portfolios, these positions are transferred from one account to another one as "in-kind" transactions, which do not incur any trading costs. For example, if old portfolio had 10,000 shares of IBM and new portfolio had 4,000 shares of IBM in a portfolio transition, then 4,000 shares are transferred in-kind and recorded as in-kind transactions. The rest 6,000 shares will be sold. If transition manager sells these shares in two days with open market trades on the first day (1,000 shares) and both external crosses and open market trades on the second day (2,000 and 3,000 shares), then there will be four lines in the database corresponding to IBM stock in that portfolio transition.

The portfolio transition data is then matched with the CRSP to obtain data on stock prices, returns, and volume. Only common stocks (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 are included in the sample. ADRs, REITS, and closed-end funds were excluded. Also excluded were stocks with missing information necessary for our empirical tests, low-priced stocks defined as stocks with prices less than 5 dollars, and transition observations which appeared to contain typographical errors and obvious inaccuracies. Since it was unclear from the data whether adjustments for dividends and stock splits were made in a consistent manner across all transitions, all observations with non-zero payouts during the first week following the starting date of

portfolio transitions were excluded from statistical tests as well.

After exclusions, there are 702,406 daily observations (334,035 buy orders and 368,371 sell orders).

2.2 CRSP Data: Prices, Volume, and Volatility

Our estimation procedures require variables such as the expected daily volume, the volatility of daily returns, and pre-trade share price. We calculate these variables using the Center for Research in Security Prices (CRSP) data.

As a pre-trade price, denoted $P_{0,i}$ for order i , we use the closing price of the corresponding security on the evening before portfolio transition began. As expected trading volume during portfolio transitions, denoted V_i for order i , we use the average daily trading volume (in the number of shares) of the corresponding security in the pre-transition month.

As a measure of expected volatility, it would be reasonable to use the implied volatility for our estimation, but this measure is not available for most securities in our sample. We therefore estimate the expected volatility of daily returns, denoted σ_i for order i , using historical daily returns. We use two estimates of volatility, historical volatility during the pre-transition month and a more complicated ARIMA estimate.

For each security, we calculate the monthly standard deviation of daily returns. We denote this volatility as $\sigma_{i,t}^m$ for month t and stock in order i . We do not de-mean the returns since the mean return in a month is very small relative to its standard deviation. We also do not make any adjustments for autocorrelation (e.g., by adding a cross-product of adjacent returns), since this might result in negative estimates for volatility of some stocks.

One proxy for daily volatility is the monthly standard deviation converted to daily units given the number of observations $N_{i,t}$, $\sigma_{i,t}^h = \sigma_{i,t}^m / N_{i,t}^{1/2}$.

Another proxy for daily volatility is forecasted volatility from ARIMA model. To reduce effects from the positive skewness of the estimates, we use a logarithmic transformation and estimate a third-order moving average process for the changes in $\ln \sigma_{i,t}^m$ over the whole sample from 2001 to 2005: $(1 - L) \ln \sigma_{i,t}^m = \Theta_0 + (1 - \Theta_1 L - \Theta_2 L^2 - \Theta_3 L^3) \cdot u_t$. The conditional forecast is $\sigma_{i,t}^e = \exp \left[\ln \sigma_{i,t}^m + \frac{1}{2} \hat{V}(u) \right] / N_{i,t}^{1/2}$, where $\hat{V}(u)$ is the variance of the prediction errors of the ARIMA model.

To the extent that expected volatility is estimated with errors, there might be an error-in-variables problem. We use the pre-transition variables known before portfolio transition trades in order to avoid any spurious effects from using contemporaneous variables, except to the extent that the ARIMA model itself uses in-sample data to estimate model parameters. Since our results are quantitatively similar for both proxies, we report our estimates only for expected volatility based on the estimated ARIMA model.

2.3 Descriptive Statistics

Table 1 reports summary statistics of individual transition orders and securities traded during portfolio transitions. Statistics are calculated for all securities in aggregate as well as separately for ten groups based on average daily dollar volume. Instead of dividing the securities into ten deciles with the same number of securities, volume break points are set at the 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of trading

volume for the universe of stocks listed on the NYSE with CRSP share codes of 10 and 11. Group 1 contains stocks in the bottom 30th percentile by dollar trading volume. Group 10 contains stocks in the top 5th percentile. Finer percentiles for the more active stocks make it possible to focus on the stock which are most important economically. For each month, the thresholds are recalculated and the stocks are reshuffled across bins.

Panel A of Table 1 reports statistical properties of stocks in our sample. For the entire sample, the median trading volume of stocks is \$19.99 million per day, ranging from \$1.22 million for the lowest volume decile to \$212.55 million for the highest volume decile. Since the average dollar volume ranges over more than two orders of magnitude, this variation in the data should create statistical power helpful in determining how trading costs vary with dollar volume. The median volatility for all stocks is 1.85 percent per day. Volatility tends to be slightly higher in the lower volume groups than the higher ones. The volatility for the lowest volume group is 2.04 percent and 1.76 percent for the highest volume group. The average quoted bid-ask spread is 23.67 basis points. From low-volume group to high-volume group, the bid-ask spread declines monotonically across groups from 64.05 basis points to 7.46 basis points. We will see later that quoted spreads are similar to our statistical estimates of spreads from portfolio transition data.

Panel B of Table 1 reports properties of portfolio transition orders. On average, the order is 3.90 percent of average daily volume. These values decline monotonically across the ten volume groups from 15.64 percent in the smallest group to 0.49 percent in the largest. The median order is only 0.56 percent of average daily volume, also declining monotonically across the ten volume groups, from 3.48 percent in the low-volume group to 0.14 percent in the high-volume group. Since the medians are much smaller than the means, the distribution of order sizes is skewed to the right. This is to be expected, since the order size is a non-negative number, and there may be some very small orders from highly diversified portfolios and small transitions as well as very large orders from less diversified portfolios and large transitions. The larger order sizes in lower deciles generate more statistical power for using implementation shortfall in estimating market impact. Our estimates for the entire sample will be therefore more representative of those for securities in lower volume groups.

2.4 Details of Estimation Procedure

To implement our estimation procedures, we make several adjustments to regression equation (4). Two of them are based on economics, and one of them is based on statistics.

First, in order to assign an intuitive interpretation to the estimates, we define an arbitrary “benchmark” stock with volatility of 2% per day and re-scale parameters λ and κ so that they represent trading costs (in basis points) for trading one percent of average daily volume for that stock. To be consistent with how most traders think about trading costs, we report our estimates for orders as a percentage of volume rather than for a dollar traded. We denote the re-scaled parameters as $\bar{\lambda}$ and $\bar{\kappa}$.

Second, transition managers have access to different pools of liquidity. Transition orders can be executed through internal crossing networks, through external crossing networks, or in open market transactions. Trading costs may be different across trading venues. Some of the portfolio transitions are the result of internal crosses. In an internal cross, one of

the transition managers customers buys from the other at some price. Both the buyer and the seller may represent different portfolio transitions being implemented simultaneously. Internal crosses with other types of customers also occur, for example, crosses against flows from a passive investment management unit affiliated with the same firm as the transition management unit. Since the buyer and the seller pay the same price, we assume that neither customer pays the spread for internal crosses, but there is spread for external crosses and open market transactions.

Concerning market impact for crosses and open market transactions, we assume that transition managers optimally choose the percentages of the orders to execute via these trading venues. To the extent that crosses are cheaper than open market transactions, this will show up as a larger percentage of the orders being crossed than executed in open markets, not as lower market impact and spread costs on crosses. The fact that both crosses and open market transactions are used in a significant proportion of orders suggests that there are significant pools of liquidity in both crossing networks and open markets. Neither dominates the other. We therefore assume that there is market impact associated with internal crosses that is equal in magnitude to the impact of external crosses and open market trades.

Third, since errors in the regression equation (4) are likely to be proportional in size to volatility, we divide both sides of the equation by price volatility to correct for heteroscedasticity.

To be more precise, we estimate price impact $\bar{\lambda}$ and bid-ask spread \bar{k} in a piece-wise linear specification using the regression equation (4) with the three adjustments,

$$\frac{\Pi_i}{P_{0,i}Q_i} \cdot 10^4 \cdot \frac{(0.02)}{\hat{\sigma}_i^e} = \frac{1}{2}\bar{\lambda} \cdot \frac{Q_i}{(0.01)\hat{V}_i^e} + \frac{1}{2}\bar{k} \cdot \frac{Q_{omt,i} + Q_{ec,i}}{Q_i} + \tilde{\epsilon}_i. \quad (11)$$

Each observation corresponds to a triple of a transition, a stock, and a trading day. For multi-day portfolio transitions, we assume that each day corresponds to a new observation i , because the transition manager effectively faces a new task of executing Q_i shares left to be executed at the beginning of a given day. For example, if an order was executed over two days, there are two observations. For $i = 1$, the order Q_1 corresponds to the total number of shares executed in day one and in day two. For $i = 2$, the order Q_2 corresponds to the number of shares left to be executed in day two. Variables $Q_{omt,i}$, $Q_{ec,i}$ and $Q_{ic,i}$ specify the number of these shares executed in open market transactions, external crosses and internal crosses, respectively, $Q_i = Q_{omt,i} + Q_{ec,i} + Q_{ic,i}$. Benchmark price $P_{0,i}$ is established the night before the trading period. The term $(0.02)/\hat{\sigma}_i^e$ adjusts for heteroscedasticity. In this regression, the observed data items have subscript i .

We also estimate price impact $\bar{\lambda}^p$ and bid-ask spread \bar{k}^p in non-linear power specification using the regression equation (6) with similar adjustments,

$$\frac{\Pi_i}{P_{0,i}Q_i} \cdot 10^4 \cdot \frac{(0.02)}{\hat{\sigma}_i^e} = \frac{1}{z+2} \cdot \bar{\lambda}^p \cdot \left(\frac{Q_i}{(0.01)\hat{V}_i^e} \right)^{z+1} + \frac{1}{2}\bar{k}^p \cdot \frac{Q_{omt,i} + Q_{ec,i}}{Q_i} + \tilde{\epsilon}_i. \quad (12)$$

Finally, we estimate price impact $\bar{\lambda}^q$ and bid-ask spread \bar{k}^q in non-linear inverse-quadratic

specification using the regression equation (10) with similar adjustments,

$$\frac{\Pi_i}{P_{0,i}Q_i} \cdot 10^4 \cdot \frac{(0.02)}{\hat{\sigma}_i^e} = \bar{\lambda}_0^q \cdot \frac{Q_i}{\hat{V}_i^e} \frac{\alpha \left(1 + 4/3\beta \frac{Q_i}{\hat{V}_i^e}\right)}{\left(1 + \beta \frac{Q_i}{\hat{V}_i^e}\right)^{3/2} + 3/2\beta \frac{Q_i}{\hat{V}_i^e} + 1} + \frac{1}{2}\bar{k}^q \cdot \frac{Q_{omt,i} + Q_{ec,i}}{Q_i} + \tilde{\epsilon}_i, \quad (13)$$

where parameters α and β are defined in (9). Parameters x_0 and x_1 are fixed being equal to 10 and 100 basis points of average daily volume, $x_0 = 0.0010 \cdot \hat{V}_i^e$, and $x_1 = 0.0100 \cdot \hat{V}_i^e$.

To adjust standard errors for positive correlation in contemporaneous returns, we pool 702,446 observations by weeks over the period from 2001 to 2005 and across 17 industry categories into 4,389 clusters.

3 Results

3.1 Estimates Based on Linear Specification

Table 2 reports the estimates of market impact and spread. The estimate for half market impact is $\bar{\lambda}/2 = 0.30$ with standard error of 0.05, and the estimate for half spread is $\bar{k}/2 = 19.20$ with standard error of 1.42. It means that, for the benchmark stock, an order of one percent of daily volume is expected to have the market-impact cost of 0.30 basis points and the spread cost of 19.20 basis points. The total trading costs is 19.50 basis points. The implied bid-ask spread of 38.40 basis points is slightly higher than average quoted spread of 23.67 basis points in table 1.

These estimates imply a simple formula for the expected trading costs for the post-decimalization period. Let X denote the number of shares in the order. Let $C(X)$ denote the expected cost of trading X shares of a stock with the expected daily volatility σ and the expected daily volume in shares V . Given the reported estimates, we can write $C(X)$, measured in basis points, as follows:

$$C(X) = 0.30 \cdot \frac{\sigma}{0.02} \cdot \frac{X}{(0.01)V} + 19.20 \cdot \frac{\sigma}{0.02}, \quad (14)$$

The first term of trading costs quantifies the size of the variable costs due to market impact, and the second term quantifies the size of fixed costs due to bid-ask spread.

The estimated costs are broadly consistent with the estimates in earlier studies. For instance, Jones and Lipson (2001) report in table 3 that the effective spread is equal 51 basis points for transactions with less than 1000 shares traded (or less than 0.35% of average daily volume) and 70 basis points for transactions with more than 100,000 shares traded (or more than 35% of average daily volume). This implies that the half market impact of a trade equivalent to one percent of average daily volume is about $1/2(70 - 51)/(35 - 0.35) = 0.30$ basis points. The fixed part of trading costs is equal to about $51/2 = 25.5$ basis points. Both numbers are similar to our estimates of 0.30 and 38.40 basis points, respectively. Since the estimates of Jones and Lipson (2001) are based on the pre-decimalization period, when trading costs are known to be higher than after decimalization in 2001, the similarity of our results suggests their estimates may be too small because of a selection bias.

The disaggregated results for four sub-samples NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells suggest that buying is more expensive than selling. For both

NYSE and NASDAQ, market-impact costs and spread costs are larger for buy orders than for sell orders by margins that are economically meaningful, if not statistically significant. For NYSE buy orders, the estimate of half price impact is $\bar{\lambda}/2 = 0.87$ with standard error of 0.21, and the estimate of half bid-ask spread is $\bar{k}/2 = 21.80$ with standard error of 2.20. For NYSE sell orders, the estimate of half price impact is $\bar{\lambda}/2 = 0.32$ with standard error of 0.07, and the estimate of half bid-ask spread is $\bar{k}/2 = 9.11$ with standard error of 2.41. For NASDAQ buy orders, the estimate of half price impact is $\bar{\lambda}/2 = 0.87$ with standard error of 0.87, and the estimate of half bid-ask spread is $\bar{k}/2 = 22.30$ with standard error of 3.38. For NASDAQ sell orders, the estimate of half price impact is $\bar{\lambda}/2 = 0.24$ with standard error of 0.03, and the estimate of half bid-ask spread is $\bar{k}/2 = 22.35$ with standard error of 3.68.

These differences are consistent with the potential information asymmetry of portfolio transition buy and sell orders. In the sample of portfolio transition orders, transition managers have to purchase securities chosen by newly hired fund managers, who are in turn selected by institutional sponsors as being presumably skillful candidates. This double-selection mechanism implies that buy orders can be informative about future price changes, and their execution may be somewhat expensive. In contrast, transition managers have to sell securities from portfolios of managers, whose mandates are simply terminated by institutional sponsors. Trading costs for sell orders therefore may be lower.

The spread tends to be larger for the NASDAQ-listed securities than for the NYSE/Amex-listed ones. The previous literature reported similar patterns (e.g., Christie and Schultz (1994), Huang and Stoll (1996), Keim and Madhavan (1997), and NYSE Market Quality report (2006) for more recent data).

Estimates For Ten Trading Volume Groups Table 3 reports the estimates of trading costs for the ten volume groups. Volume groups are based on the pre-transition trading volume with thresholds 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of the dollar volume for common NYSE-listed stocks. To obtain these estimates, we modify regression equation (4) by replacing $\bar{\lambda}$ and \bar{k} with the dummy variables $\bar{\lambda}_j$ and \bar{k}_j for each of the ten volume group j . The result is a regression with twenty coefficients, two coefficients for each volume bin, with one coefficient for half market impact and one coefficient for half spread,

$$\frac{\Pi_i}{P_{0,i}Q_i}10^4\frac{(0.02)}{\hat{\sigma}_i^e} = \left(\sum_{j=1}^{10} \mathbb{I}_{j,i} \cdot \frac{1}{2}\bar{\lambda}_j \right) \cdot \frac{Q_i}{(0.01)\hat{V}_i^e} + \left(\sum_{j=1}^{10} \mathbb{I}_{j,i} \cdot \frac{1}{2}\bar{k}_j \right) \cdot \frac{Q_{omt,i} + Q_{ec,i}}{Q_i} + \tilde{\epsilon}_i \quad (15)$$

where $\mathbb{I}_{j,i}$ is an indicator equal to one if order i is executed in a stock from volume group j , $j = 1, \dots, 10$. We chose to consider one regression rather than a set of separate regressions for each volume group, since this procedure enables us to take into account potential correlation of residuals across different volume groups.

We find that stocks with higher volume tend to have lower spread and larger market impact per percent of volume traded. The market impact increases with trading volume from $\bar{\lambda}_1 = 0.25$ for inactively traded stocks to $\bar{\lambda}_{10} = 2.37$ for actively traded stocks. Controlling for volatility, the market impact is approximately ten times larger for actively traded stocks than for inactively traded stocks. Similar patterns are also observed for disaggregated subsamples of NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells. The spread

decreases with trading volume from $\bar{k}_1 = 39.29$ basis points for inactively traded stocks to $\bar{k}_{10} = 7.85$ basis points for actively traded stocks. The spread thus drops almost by a factor of five.

These results may seem to be counter-intuitive. It is important, however, to emphasize that market impact $\bar{\lambda}_j$ is scaled to quantify market impact costs per one percent of daily volume, not per one dollar traded. Because of a large difference in volume between groups, one percent of volume $V_{10} \cdot P_{10}$ in high-volume stocks is hundreds of times bigger than one percent of volume in low-volume stocks. Trading one percent of volume in high-volume stocks is more expensive than trading one percent of volume $V_1 \cdot P_1$ in low-volume stocks, $\bar{\lambda}_{10} > \bar{\lambda}_1$. Trading one dollar in high-volume stocks is usually less expensive than trading one dollar in low-volume stocks, $\lambda_{10}/(V_{10} \cdot P_{10}) < \lambda_1/(V_1 \cdot P_1)$.

The intuition can be also explained in terms of differences in a degree of competition. Trading volume of large stocks is high, but it is usually “two-sided,” because numerous traders with different views are active in that market. Many buy and sell orders cancel with each other, and the absolute value of daily order imbalances is very small relative to daily volume. Execution of one percent of daily volume for large stocks is therefore a big distinct trade in the two-sided order flow, and its market impact is expected to be substantial. In contrast, trading volume of small stocks is low, but it is usually “one-sided,” because only a few traders are active in that market. Buy orders tend to follow buy orders, and sell orders tend to follow sell orders. The absolute value of daily order imbalances is often comparable to daily volume. Execution of one percent of daily volume would not be a very noticeable event in the one-sided order flow, and its market impact is expected to be relatively small.

A simple example illustrates our intuition about market impact and degree of competition. Let us assume that there are γ traders who have different (independent) views about the true value of a risky security and place buy or sell orders. For simplicity, we assume that $\tilde{\epsilon}_i$ is equal to 1 if trader i decides to buy and -1 if he decides to sell. Trades have the same size of Q shares, $Q = V/\gamma$, where V is the average number of shares traded per day. Given a linear price impact, the stock price changes by ΔP in response to these orders, $\Delta P = \sum_{i=1}^{\gamma} \tilde{\epsilon}_i \cdot Q \cdot \lambda$. The daily volatility σ_P is equal to $\sigma_P = \sqrt{\gamma} \cdot Q \cdot \lambda = V/\sqrt{\gamma} \cdot \lambda$. The price impact of one percent of daily volume expressed as a percent of volatility, $V \cdot \lambda/\sigma_P$, is equal to $\sqrt{\gamma}$. This example shows that the larger is the number of traders γ in the market, the bigger is price impact per percent of volume, as it is increasing with $\sqrt{\gamma}$.

Similar positive association between market impact and overall trading volume has been mentioned in Breen, Hodrick, Korajczyk (2002) who find that market impact increases with market capitalization, if market impact is defined as the price change after a trade equivalent to one percent of shares outstanding. To avoid these “unreasonable” results, authors instead estimate price changes after a one-dollar trade and find the estimates are reasonably decreasing with market capitalization (see also Hasbrouck (1991b), Chen, Stanzl and Watanabe (2005)). When thinking about price impact functions, it is therefore very important to understand how constants in their specifications are scaled.

Amihud (2002) suggests to estimate illiquidity as the ratio of absolute value of returns to dollar volume per day, $|r|/(V \cdot P)$, averaged over a month. High illiquidity ratios are expected to indicate limited amount of liquidity available in the market. Since $|r|$ is closely related to the volatility of returns σ , these ratios are proportional to $\sigma/(V \cdot P)$. Plugging

$\sigma = \sigma_P/P$ and $\sigma_P = V/\sqrt{\gamma} \cdot \lambda$ from above, we find that illiquidity ratios are proportional to $(\lambda/P^2) \cdot \gamma^{-1/2}$. They can be low for two reasons, because the price impact λ/P^2 per dollar traded is low or the number of traders γ is high. To compare liquidity using illiquidity ratios, one has to assume that the number of trades per day is the same across markets. In reality, the number of trades in high-volume markets differs from the number of trades in low-volume markets by a factor of more than a hundred, see Kyle, Obizhaeva and Tuzun (2012). What many papers attribute to lower transaction costs could be just a sign of higher degree of competition in the market. This is another example demonstrating importance of understanding how liquidity indicators are scaled.

3.2 Estimates Based on Non-Linear Specification

Table 4 shows our estimates of market impact and spread for non-linear specifications. Panel A reports the estimates for the power specification. Panel B reports the estimates for the inverse quadratic specification. There are estimates based on the total sample as well as disaggregated samples of stocks in ten different volume groups. To obtain these estimates, we modify regression equations (12) and (13) by analogy with regression equation (15). For power specification, we replace $\bar{\lambda}^p$ and \bar{k}^p with the sum of $\bar{\lambda}_j^p$ and \bar{k}_j^p , multiplied by indicator $\mathbb{I}_{j,i}$, which is equal to one if order i is executed in a stock from volume group j . For quadratic inverse specification, we replace $\bar{\lambda}^q$ and \bar{k}^q with the sum of $\bar{\lambda}_j^q$ and \bar{k}_j^q , multiplied by indicator $\mathbb{I}_{j,i}$.

For the power specification, we find that the estimated curvature z is statistically significant. Its value is equal to -0.46 for pooled data. The estimated curvature parameter varies across volume groups from -0.46 to -0.21, with the highest values corresponding to group 3 and group 6. Since the hypothesis that $z = -0.5$ can not be statistically rejected, the power price impact function can be therefore well approximated by a square root specification, popular among practitioners (e.g., Barra model). The estimates of spread $\kappa/2$ range from -0.87 in volume group 4 to 9.08 in volume group 1. All of these estimates have very big standard errors, making them statistically insignificant. Since power specification restricts the slope of price impact functions to zero for trade sizes equal to zero, it can not cleanly separate the bid-ask spread costs from price-impact costs.

For the inverse-quadratic specification, we find that the price impact $\bar{\lambda}_0^q = 15.57$ basis points at $x_0 = 0.001 \cdot V$ (with standard error of 3.85) and the price impact $\bar{\lambda}_1^q = 9.49$ basis points at $x_1 = 0.01 \cdot V$ (with standard error of 1.01). Since the difference between two slopes is statistically significant, the price impact function is concave. The estimate $\kappa^q/2$ is equal to 3.67 (with standard error of 1.97).

The estimates of $\bar{\lambda}_{0,j}^q$ vary across volume groups. The price impact $\bar{\lambda}_0^q$ tends to increase from 11.45 for low-volume group (with standard error of 3.34) to 31.57 for high-volume group (with standard error of 9.51). The price impact $\bar{\lambda}_1^q$ tends to increase as well from 8.20 for low-volume group (with standard error of 1.59) to 12.04 for high-volume group (with standard error of 2.85). For all volume groups, the price impact functions are concave, $\bar{\lambda}_{0,j}^q > \bar{\lambda}_{1,j}^q$. The estimate of half-spread $\bar{k}^q/2$ decreases almost monotonically from 13.33 basis points for low-volume group (with standard error of 3.94) to 2.49 basis points for high-volume group (with standard error of 2.58).

Figure 1 shows the implied non-linear price impact functions based on calibrated equa-

tions (12) and (13) with the estimates from table 4. The estimated price impact function for different volume groups are presented on different subplots. The order size in percents of daily volume is on x-axis. The price impact in basis points is on y-axis. Since estimated functions depend on volatility, we plot them under the assumption that volatility is equal to the average volatility for stocks in the corresponding volume group in table 1. Solid lines track the inverse quadratic specification. Dashed lines track the power specification. On each chart, we also superimpose a shaded area, with the width equal to the average quoted spread for that group from table 1.

Several observations are important. First, price impact functions tend to be concave. The curve for inverse quadratic price impact function approaches y-axis at a significantly smaller angle than the curve for power price impact function, suggesting that the later specification may be too restrictive. At the same time, both specifications make economically similar predictions, because implied functions deviate from each other only slightly.

Second, even though statistical estimates imply that price impact functions are concave, their non-linearity is economically insignificant for many volume groups. For example, the price impact functions are almost linear for volume groups 1 through 7. Deviations from piece-wise linear approximations are small, especially comparing to the magnitude of bid-ask spread. For low-volume group, for example, the deviation of an inverse-quadratic specification from a piece-wise linear approximation is about 8 basis points, while the quoted bid-ask spread is equal to 64 basis points.

Third, the concavity is more pronounced for volume group 9 and volume group 10, where small orders seem to be executed at a significant “discount” relative to a piece-wise linear model. For small trades, deviations from this specification are comparable to the magnitude of bid-ask spread.

We attribute these patterns to order execution algorithms that portfolio transition managers may follow. When time is not pressing, traders often postpone their trades hoping to achieve a more favorable execution by providing liquidity to other market participants, for example, by placing limit orders on the opposite side of the book rather than executing market orders. If limit orders are picked off, then not only do they execute original orders, but also gain the spread. This strategy is more reasonable when traders have to execute orders of relatively small size, because large orders usually have to be executed more aggressively. This strategy also works better in more active markets of high-volume stocks, where buy orders frequently alternate sell orders. As a result, execution of small orders in active markets may appear to be relatively cheap. Consistent with our hypothesis, Griffiths et al. (2000) report that the implementation shortfall of small limit orders is negative for securities with high market capitalization. Since this strategy is less reasonable for trading low-volume stocks, when traders usually have to pay both price impact costs and bid-ask spread costs, price impact functions for low-volume stocks do not differ much from piece-wise linear specifications with a fixed bid-ask spread and linear price impact.

4 Conclusion

Portfolio transition trades are ideal for estimation of transactions costs, because the exogeneity of these trades eliminates selection bias, which is endemic for most other samples with quantities traded being endogenously determined.

For benchmark stocks with daily volatility equal to two percent, the half market impact is equal to 0.30 basis points for trading of one percent of daily volume and the half spread is equal to 19.20 basis points. Buy orders are more expensive than sell orders, which may reflect institutional specifics of how securities are selected to be purchased or sold during portfolio transitions.

Trading costs vary significantly across stocks with different levels of volatility and trading volume. Per dollar traded, trading in low volume stocks is more expensive than trading in high-volume stocks. Per one percent of daily volume traded, price impact costs is smaller for low-volume stocks than for high-volume stocks. The implied expected trading costs $C_j(X)$ in basis points for volume group j , $j = 1, \dots, 10$, can be calculated using the corresponding estimates $\bar{\lambda}_j$ and $\bar{\kappa}_j$ from table 3,

$$C_j(X) = \frac{1}{2} \bar{\lambda}_j \cdot \frac{\sigma}{0.02} \cdot \frac{X}{(0.01)V} + \frac{1}{2} \bar{\kappa}_j \cdot \frac{\sigma}{0.02}. \quad (16)$$

The half market impact of trading one percent of daily volume increases tenfold from $\bar{\lambda}_1 = 0.25$ basis points for low-volume stocks to $\bar{\lambda}_{10} = 2.37$ basis points for high-volume stocks. At the same time, spread costs decrease fivefold from $\bar{\kappa}_1 = 39.29$ basis points to $\bar{\kappa}_{10} = 7.85$ basis points.

Price impact functions are concave in a sense that estimated exponents are statistically different from levels that would correspond to piece-wise linear functions with a linear price impact part and a constant spread part. This non-linearity turns out to be insignificant economically for most stocks, except for high-volume stocks. For these stocks, small trades executed at significant “discounts” relative to a piece-wise linear approximation. Since these discounts are comparable to the magnitude of bid-ask spread, the concavity of price impact functions is likely to be a result of strategic execution of orders by portfolio transition managers, who may often decide to postpone execution and wait for the opportunities to earn bid-ask spread when more impatient traders arrive on the other side of the market.

Reliable liquidity estimates can be used to study whether seemingly profitable trading strategies remain profitable after adjusting their returns for trading costs. These studies have important implications for the assessment of the limits of arbitrage in financial markets and viability of trading strategies designed by asset managers. The existing academic studies on this subject usually have to rely on the transaction costs estimates provided by Keim and Madhavan (1997), based on the execution data for 21 institutions from January 1991 through March 1993. The examples are the study on performance of mutual funds by Wermers (2000), the study on the profitability of long-short strategies based on earning surprises by Chordia et al. (2009), and the study on profitability of strategies based on book-to-market ratios, firm sizes, and lagged returns by Cooper, Gutierrez, and Marcum (2005). Our study provides alternative liquidity estimates. These estimates are not distorted by a selection bias. They better represent transaction costs in the recent post-decimalization period and provide more accurate predictions for large transactions, for which non-linearity effects are important.

References

- Almgren, Robert, Chee Thum, Emmanuel Hauptmann, and Hong Li, 2005, Equity market impact, *Risk*, 57–62.
- Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.
- Back, Kerry, 1992, Insider trading in continuous time, *Review of Financial Studies* 5(3), 387–409.
- 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.
- Cheng, Minder, and Ananth Madhavan, 1997, In search of liquidity: Block trades in the upstairs and downstairs markets, *Review of Financial Studies* 10, 175–203.
- Chordia, Tarun, Amit Goyal, Ronnie Sadka, Gil Sadka, and Lakshmanan Shivakumar, 2009, Liquidity and the post-earnings-announcement drift, *Financial Analysts Journal* 65, 18–32.
- Christie, William, and Paul Schultz, 1994, Why do NASDAQ market makers avoid odd-eighth quotes?, *Journal of Finance* 49(5), 1813–1840.
- Conrad Jennifer, Kevin Johnson, and Sunil Wahal, 2003, Institutional trading and alternative trading systems, *Journal of Financial Economics*, 70, 99–134.
- Cooper, Michael, Roberto Gutierrez, and William Marcum, 2005, On the predictability of stock returns in real time, *Journal of Business*, 78, 469–499.
- Coppejans, Mark T., Ian H. Domowitz, and Anath Madhavan, 2004, Resiliency in an automated auction, *ITG Working Paper*.
- Dufour, Alfonso and Robert Engle, 2000, Time and the price impact of a trade, *Journal of Finance*, 55(6), 2467–2498.
- Foster, Douglas, and Viswanathan, S, 1996, Strategic trading when agents forecast the forecasts of others, *Journal of Finance*, 51(4), 1437–78.

- 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.
- Handa, Puneet, and Robert Schwartz, 1996, Limit order trading, *Journal of Finance*, 51(5), 1835–1861.
- Harris, Lawrence, and Joel Hasbrouck, 1996, Market vs. limit orders: The SuperDOT evidence on order submission strategy, *Journal of Financial and Quantitative Analysis*, 31(2), 213–231.
- Hasbrouck, Joel, 1988, Trades, quotes, inventories and information, *Journal of Financial Economics* 22, 229–252.
- 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.
- Huberman, Gur, and Werner Stanzl, 2004, Price manipulation and quasi-arbitrage, *Econometrica*, 72(4), 1247–1275.
- Jones, Charles, and Marc Lipson, 2001, Sixteenths: direct evidence on institutional execution costs, *Journal of Financial Economics* 59(2), 253–278.
- 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, 1996, The upstairs market for large-block transactions: Analysis and measurement of price effects, *Review of Financial Studies*, 9(1), 1–36.
- 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.
- Kempf, Alexander, and Olaf Korn, 1999, Market depth and order size, *Journal of Financial Markets*, 2(1), 29–48.
- Kyle, Albert, 1985, Continuous auctions and insider trading, *Econometrica*, 53(6), 1315–1335.
- Lesmond, David, Joseph Ogden, and Charles Trzcinka, 1999, A new estimate of transaction costs, *Review of Financial Studies*, 12(5), 113–1141.
- 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.
- Perold, Andre, 1988, The implementation shortfall: Paper vs. reality, *Journal of Portfolio Management* 14(3), 4–9.
- U.S. Securities and Exchanges Commission, 2001, Report on the comparison of order executions across equity market structures.
- Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses, *Journal of Finance* 55(4), 1655–1695.

Table 1: The Descriptive Statistics.

Panel A: Properties of Securities

	All	1	2	3	4	5	6	7	8	9	10
Med(V) (m \$)	19.99	1.22	5.14	9.97	15.92	23.92	31.45	42.11	60.16	101.51	212.55
Med(σ)	1.89	2.04	2.00	1.92	1.95	1.88	1.85	1.79	1.78	1.76	1.76
Med($Sprd$) (bps)	11.54	38.16	18.34	13.53	11.81	10.12	9.34	8.09	7.16	5.92	4.83
Mean($Sprd$) (bps)	23.67	64.05	31.27	21.83	18.40	15.65	13.86	12.14	11.00	9.02	7.46

Panel B: Properties of Orders

	All	1	2	3	4	5	6	7	8	9	10
Avg(X/V) (%)	3.90	15.64	4.58	2.63	1.82	1.36	1.18	1.07	0.88	0.69	0.49
Med(X/V) (%)	0.56	3.48	1.39	0.80	0.54	0.40	0.35	0.30	0.25	0.20	0.14
Avg OMT Share	0.31	0.38	0.33	0.32	0.32	0.31	0.31	0.30	0.29	0.27	0.24
Avg EC Share	0.40	0.42	0.42	0.41	0.41	0.41	0.40	0.40	0.39	0.38	0.36
Avg IC Share	0.29	0.20	0.25	0.26	0.27	0.28	0.29	0.30	0.32	0.35	0.40
# Obs	441,865	65,081	68,545	41,559	49,532	28,621	30,087	30,710	35,733	42,331	49,666

Table reports characteristics of securities and transition orders in the sample. Panel A shows the median of average daily dollar volume V (in millions of \$), the median of the daily returns volatility σ (in percents), the median and the mean of the percentage spread $Sprd$ (in basis points). Panel B shows the average order size (in percents of V), the median order size (in percents of V), the average fraction of transition order executed in open market (Avg OMT Share), external and internal crossing networks (Avg EC and IC Shares), as well as the total number of observations. Results are presented for ten volume groups. Each month, the observations are split into 10 bins according to stocks' dollar trading volume in pre-transition month. The thresholds correspond to 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of dollar trading volume for common stocks listed on the NYSE. Group 1 (Group 10) contains orders in stocks with lowest (highest) dollar trading volume. The sample ranges from January 2001 to December 2005.

Table 2: The Estimates of The Piece-wise Linear Specification.

	All	NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
$\frac{1}{2}\bar{\lambda}$	0.30*** (0.05)	0.87*** (0.21)	0.32*** (0.07)	0.87*** (0.87)	0.24*** (0.03)
$\frac{1}{2}\bar{k}$	19.20*** (1.42)	21.80*** (2.20)	9.11*** (2.41)	22.30*** (3.38)	22.35*** (3.68)
R^2	0.010	0.015	0.004	0.019	0.017
# Obs	702,406	210,194	228,934	123,841	139,437

Table presents the estimates of price impact $\bar{\lambda}$ and bid-ask spread \bar{k} from the regression equation (11):

$$\frac{\Pi_i}{P_{0,i}Q_i} \cdot 10^4 \cdot \frac{(0.02)}{\hat{\sigma}_i^e} = \frac{1}{2}\bar{\lambda} \cdot \frac{Q_i}{(0.01)\hat{V}_i^e} + \frac{1}{2}\bar{k} \cdot \frac{Q_{omt,i} + Q_{ec,i}}{Q_i} + \tilde{\epsilon}_i.$$

Each observation corresponds to portfolio transition order i . Π_i is the implementation shortfall in basis points, $P_{0,i}$ is the pre-trade benchmark price. The term $(0.02)/\hat{\sigma}_i^e$ adjusts for heteroscedasticity. Q_i is the number of shares in the order i with $Q_{omt,i}$ and $Q_{ec,i}$ shares executed in open market and external crossing networks, respectively; $\mathbb{I}_{BS,i}$ is the buy/sell indicator; $\frac{1}{2}\bar{\lambda}$ estimates in basis points the market impact costs of a trade of one percent of expected daily volume V_i^e in a benchmark stock with daily volatility 2%, and $\frac{1}{2}\bar{k}$ estimates in basis points the effective spread cost. The results are presented for stocks listed at the NYSE/Amex and the NASDAQ as well as for buy and sell orders. The standard errors (in parentheses) are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005. ***, **, *denotes significance at 1%, 5% and 10% levels, respectively.

Table 3: The Estimates of The Piece-wise Linear Specification For Ten Volume Groups.

	All	j: 1	2	3	4	5	6	7	8	9	10
$1/2\bar{\lambda}_j$	0.30*** (0.05)	0.25*** (0.03)	1.24*** (0.17)	1.72*** (0.27)	1.80*** (0.50)	2.04*** (0.43)	2.79*** (0.56)	2.38*** (0.46)	2.54*** (0.63)	2.08*** (0.47)	2.37*** (0.57)
$1/2\bar{k}_j$	19.20*** (1.42)	39.29*** (2.89)	17.66*** (2.34)	10.93*** (2.32)	12.01*** (2.20)	7.79*** (2.15)	6.41** (2.11)	8.26*** (2.08)	8.14*** (2.20)	9.31*** (1.79)	7.85*** (1.84)

Table presents the estimated price impact $\bar{\lambda}_j$ and spread \bar{k}_j from the regression (15) for ten volume groups j . $\bar{\lambda}_j/2$ quantifies the market impact costs of a trade of one percent of expected daily volume V_i^e in a benchmark stock with daily volatility 2% based on data, $\bar{k}_j/2$ quantifies the effective spread cost (in basis points). Volume groups are based on the pre-transition dollar trading volume with thresholds 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of the dollar volume for common NYSE-listed stocks. Group 1 (Group 10) contains stocks with the lowest (highest) trading volume. The standard errors (in parentheses) are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005. ***, **, * denotes significance at 1%, 5% and 10% levels, respectively.

Table 4: The Estimates of Non-linear Price Impact Functions.

	All	1	2	3	4	5	6	7	8	9	10
$\bar{\lambda}^p$	16.59*** (1.88)	15.59*** (2.51)	11.83*** (2.61)	8.82** (2.80)	15.84*** (4.64)	13.28** (4.74)	13.14** (4.37)	15.09** (4.62)	12.14** (4.15)	17.00*** (4.01)	18.49*** (3.80)
z	-0.46*** (0.02)	-0.46*** (0.03)	-0.32*** (0.06)	-0.21* (0.10)	-0.39*** (0.11)	-0.32* (0.14)	-0.22 (0.14)	-0.32** (0.10)	-0.34 (0.19)	-0.36*** (0.10)	-0.40*** (0.09)
$1/2\bar{\kappa}^p$	0.51 (2.03)	9.08* (4.02)	0.65 (3.46)	1.15 (3.61)	-0.87 (3.87)	0.39 (3.57)	0.33 (3.23)	2.22 (3.48)	3.08 (3.19)	2.95 (2.94)	1.98 (2.63)
Adj. R^2	0.011	0.030	0.014	0.005	0.005	0.003	0.004	0.004	0.002	0.003	0.002

	All	1	2	3	4	5	6	7	8	9	10
$\bar{\lambda}_0^q$	15.57*** (3.85)	11.45*** (3.34)	9.43*** (2.07)	7.61*** (1.79)	14.56* (6.22)	10.33* (4.14)	12.88* (5.12)	14.92** (5.76)	14.76* (7.32)	22.12* (9.09)	31.57*** (9.51)
$\bar{\lambda}_1^q$	9.49*** (1.01)	8.20*** (1.59)	8.22*** (1.41)	7.15*** (1.47)	10.38*** (2.12)	9.07*** (2.58)	11.36*** (2.69)	11.61*** (2.81)	9.81*** (2.32)	12.73*** (1.86)	12.04*** (2.85)
$1/2\bar{\kappa}^q$	3.67 (1.97)	13.33*** (3.94)	2.49 (3.12)	1.85 (3.34)	1.70 (3.09)	2.47 (2.88)	0.99 (2.84)	3.64 (3.01)	3.76 (2.74)	4.06 (2.63)	2.49 (2.58)
Adj. R^2	0.011	0.030	0.013	0.005	0.004	0.002	0.001	0.002	0.000	0.001	0.002
p-val ($\bar{\lambda}_0^q = \bar{\lambda}_1^q$)	0.036	0.07	0.08	0.17	0.33	0.44	0.56	0.29	0.40	0.25	0.06
F-stat ($\bar{\lambda}_0^q = \bar{\lambda}_1^q$)	4.40	3.37	3.12	1.87	0.96	0.59	0.33	1.12	0.72	1.34	3.66
# Obs	421,505	62,827	65,537	39,524	47,001	27,096	28,642	29,162	33,991	40,327	47,398

Table presents the estimates of the power parameter z , the price impact $\bar{\lambda}^p$, and the spread $\bar{\kappa}^p$ in the power specification from the regression (6) and the estimates of the price impacts $\bar{\lambda}_0^q$ and $\bar{\lambda}_1^q$ as well as the spread $\bar{\kappa}^q$ in the inverse-quadratic specification from the regression (10). Results are presented for stocks with different dollar trading volume. Each month, the observations are split into 10 bins according to stocks' dollar trading volume in pre-transition month. The thresholds correspond to 30th, 50th, 60th, 70th, 80th, 90th, and 95th percentiles of dollar trading volume for common stocks listed on the NYSE. Group 1 (group 10) contains orders in stocks with lowest (highest) dollar trading volume. The sample ranges from January 2001 to December 2005. ***, **, * denotes significance at 1%, 5% and 10% levels, respectively.

Figure 1: The Implied Non-Linear Price Impact Functions.

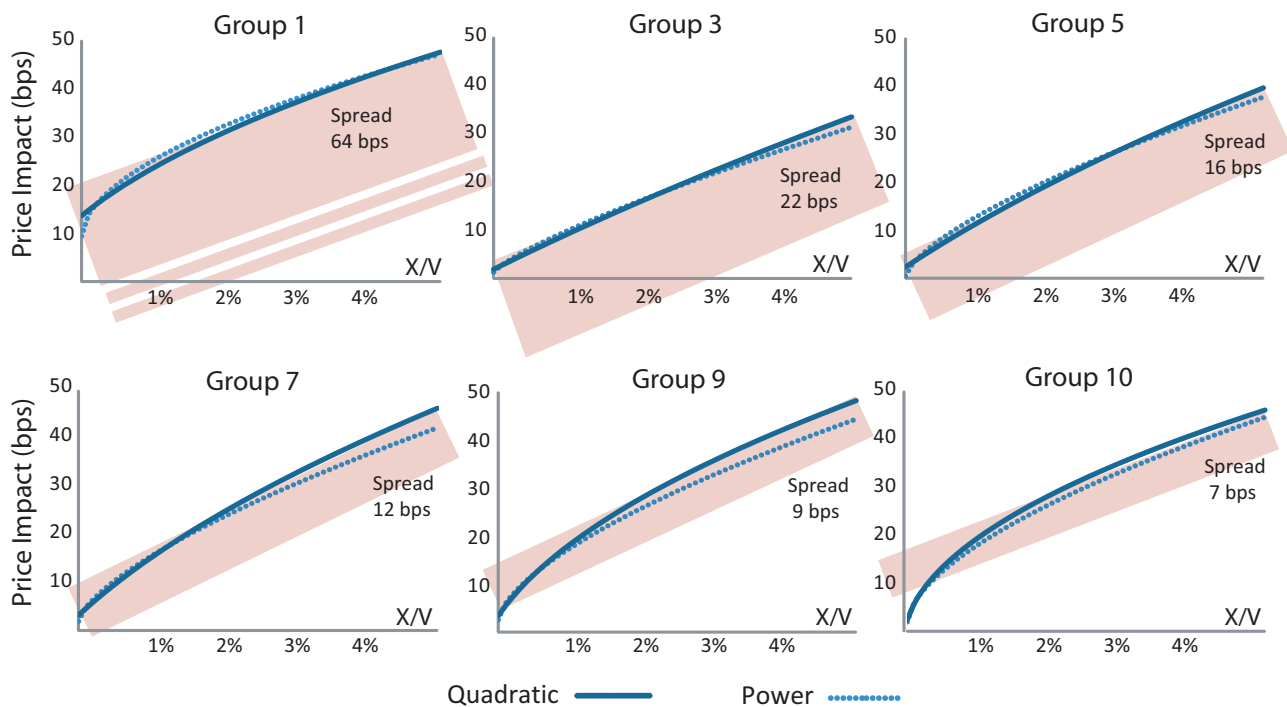


Figure shows the expected trading costs (in basis points) for volume groups 1, 3, 5, 7, 9, and 10. The expected cost $C_j(X)$ for volume group j is plotted as a non-linear function of a trade size X as percent of daily volume V . Solid lines correspond to the implied inverse-quadratic specification (13) of the price impact functions with the estimated market impact $\bar{\lambda}_j$ and the spread $\bar{\kappa}_j$ from table 4. Dashed lines correspond to the implied power specification of the price impact functions (12) with the estimated market impact $\bar{\lambda}_j$ and the spread $\bar{\kappa}_j$ from table 4. The thresholds for the ten volume groups are 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of the dollar volume for common NYSE-listed stocks. Group 1 (group 10) contains stocks with the lowest (highest) trading volume. The width of the shaded area is equal to the average quoted spread from table 1 for each volume group.