The Magic of Hindsight: Creating a Post-Trade Transaction Cost Estimate Based on Realized Market Conditions

MILAN BORKOVEC AND HANS G. HEIDLE

June 29, 2007

Abstract

The crucial assumption in most pre-trade transaction cost models is market neutrality. Consequently, expected costs for these models are entirely based on one's own trading strategy and direct market impact. Market effects due to other market participants are generally completely ignored. This paper describes how to improve transaction cost estimates (benchmarks) for post-trade analysis. After deriving the expected transaction costs from a specific pre-trade model, we incorporate the general market effects at the time when the trades actually took place. The proposed model incorporates market returns, spread variables, and trade imbalances. The model also allows for the decomposition of the cost of a transaction into two components: the cost due to one's own trading and the cost due to general market effects. The framework is applied to ITG's Agency Cost Estimator (ACE) model.

MILAN BORKOVEC is Head of the Financial Engineering Group at Investment Technology Group, Inc., 44 Farnsworth Street, Boston MA 02210, Tel: (617) 692-6733; e-mail: milan.borkovec@itg.com

HANS G. HEIDLE is Manager of Financial Engineering's Pre-Trade Analytics Group at Investment Technology Group, Inc., 44 Farnsworth Street, Boston MA 02210, Tel: (617) 692-6828; e-mail: hans.heidle@itg.com

The authors would like to thank Gabriel Butler, Ian Domowitz, Joe Emanuelli, Jon Fatica, Brett Jiu, Mark Kijesky, Charles Qin, Vitaly Serbin, and Robert Sinclair for their support and comments.

1. OVERVIEW

Understanding the components behind transaction costs has become an integral part of the investment process. Over recent years, the market has seen an unusual bombardment of pre-trade transaction cost models which identify trading strategies that weigh expected trading costs against execution risk due to stock price volatility. Most of these models are based on the papers by Bertsimas and Lo (1998), Almgren and Chriss (2000), and Huberman and Stanzl (2005), in which best execution is defined as the trading strategy that provides the minimum execution costs for trading over a fixed period of time while taking into account the volatility of the stock price associated with different strategies.

Domowitz, Glen, and Madhavan (2002) identify transaction costs as a key element in evaluating portfolio performance. Large enough execution costs substantially reduce or even eliminate the notional return. Monitoring and minimizing these costs has become the industry norm. In a recent paper, Edelen, Evans, and Kadlec (2007) find that transaction costs not only reduce fund return performance but rather they also predict fund return performance. They argue that transaction costs are a better predictor of fund return performance than expense ratios. In addition, they suggest that trading costs, rather than fund size, are the primary source of diseconomies of scale for funds.

Our paper revisits the structural pre-trade transaction cost models and reconciles the estimated pre-trade cost estimates with what is actually observed by traders after executions are completed ("post-trade cost estimates"). In reconciling, we capture the actual market dynamics.

Pre-trade transaction cost models typically measure the average price impact costs of institutional investors. The embedded crucial assumption is market neutrality. Consequently, the estimated costs are entirely based on one's own trading strategy and the associated price impact. Pre-trade models do not necessarily account for market effects due to other market participants, such as short-term serial correlation in price movements,[1] news events/announcements, and the underlying investor sentiment in the market. In this paper, we introduce a significant improvement to estimating transaction costs by explicitly incorporating market dynamics and hence moving from a pre-trade to a post-trade cost model, such as ITG's ACE model. The paper outlines the factors that drive execution costs and how the costs vary systematically with market liquidity and return volatility - both factors that are relevant for any investment strategy.

The paper is organized as follows. Section 2 outlines the general framework of most pre-trade models and our procedures to enhance post-trade transaction cost estimates. Section 3 describes the data that have been used to estimate ACE for post-trade analysis. In Section 4, we highlight the empirical evidence associated with how the market dynamic variables affect the prediction of transaction costs. The paper concludes in Section 5.

[1] Price movements exhibit serial correlation across various time horizons (Lo and MacKinlay (1988)) and reflect changes in market conditions and the presence of privately informed traders (Bertsimas and Lo (1998)).

2. MODEL FRAMEWORK

This section outlines the framework and necessary assumptions underlying most existing theoretical pre-trade transaction cost models, and maps this class of pre-trade models to a post-trade model setting that allows enhancements to trading cost estimation.

Most theoretical pre-trade transaction cost models divide each trading day into N periods of equal duration (bins). For example, for the U.S. market, we can break the trading day into thirteen 30-minute bins. The trading horizon may consist of several days with arbitrary starting and ending bins on the first and last day, respectively. An order in a security is defined by

- the trading horizon T (in days) with starting bin s on the first day and ending bin e on the last day,
- the side δ, where $\delta = 1$ (-1) for buy (sell),
- the size S, and
- the trading strategy $\left(n_{ij}\right)_{i=1,\dots,T,\, j=1,\dots,N}$,

where n_{ij} is the number of shares of the security traded in bin *j* on day *i* and *N* is the number of bins on a given day. It is assumed that trading of all share quantities is completed within their respective bin.

The average transaction costs (per share) of an order with the above characteristics is defined as the signed difference between the price $p_{1,s-1}$ of the security at order placement time (i.e. the end of bin $s-1$ of day 1) and the volume-weighted average execution price. Specifically,

$$Pre\_Cost\left(S\ ,\left(n_{ij}\right)_{i=1,..,T;\, j=1,...,N}\right) = \delta \cdot \left(\left[\sum_{i=1}^{T}\sum_{j=1}^{N} \tilde{p}_{ij} n_{ij} \big/ S\right] - p_{1,s-1}\right), \tag{1}$$

where

$$S = \sum_{i=1}^{T}\sum_{j=1}^{N} n_{ij}\ \text{, with } n_{1j} = 0 \text{ for } j < s \text{ and } n_{T,j} = 0 \text{ for } j > e, \tag{2}$$

and $\tilde{p}_{ij}$ is the execution price in bin *j* on day *i*. Note that the actual trade start date $1 \le T_{Start} \le T$ and trade start bin s_{Start} do not have to match with the order placement date, 1, and time, *s*, i.e., we have

$$n_{kj} = 0 \quad \text{for} \quad \left(k < T_{Start},\ \ j = 1,\dots,N\right) \quad \text{and} \quad \left(k = T_{Start},\ \ j = 1,\dots,s_{Start} - 1\right). \tag{3}$$

What distinguishes among the existing pre-trade transaction cost models is how $\tilde{p}_{ij}$ is forecasted to include price impact and spread costs. In general, these two cost components are modeled separately and thus the cost formula can be subdivided into

$$
\begin{aligned}
&Pre\_Cost\left(S,\left(n_{ij}\right)_{i=1,..,T;\,j=1,...,N}\right)= \\
&\qquad Pre\_CostSpread\left(\left(n_{ij}\right)_{i=1,..,T;\,j=1,...,N}\right)+Pre\_CostPI\left(S,\left(n_{ij}\right)_{i=1,..,T;\,j=1,...,N}\right),
\end{aligned}
\tag{4}
$$

where *Pre_CostSpread* is the pre-trade transaction cost estimate due to the spread and *Pre_CostPI* is the pre-trade transaction cost estimate due to price impact.

Typically, the price impact costs are decomposed into a temporary and a permanent component. The temporary price impact is of transitory nature and is purely an inventory effect where market imbalances are adjusted with price incentives. The permanent or persistent price impact reflects changes in the market participants' views about the value of the security due to one's trading. Intuitively, demanding liquidity with a buy reveals to the market that the security may be undervalued, whereas demanding liquidity with a sell signals that the security may be overvalued (see e.g., Fabozzi, Facardi, and Kolm (2006)).

In order to forecast the expected price impact associated with a trade for S units of a particular security with a trading horizon of T days, mid quote prices at the end of each bin are often modeled iteratively. Specifically, the mid quote price p_{ij} at the end of bin j of day i associated with the executed trade volume n_{ij} is generally modeled as a function of the previous bin's last mid quote price p_{ij-1}, the trade volume n_{ij}, the volume V_{ij}, the volatility σ_{ij} and the market sentiment ms_{ij}, i.e.

$$
p_{ij} = f\left(V_{ij}, \sigma_{ij}, ms_{ij}, n_{ij}, p_{ij-1}\right). \tag{5}
$$

Given that the actual trade volume, volatility, and market sentiment are not known prior to trading, most pre-trade cost models optimize their trading strategies using historical intra-day means or medians as volume and volatility estimates. As for market sentiment, it is typically either ignored or modeled as a function of past trade imbalances or returns. Consequently, the "true" equation (5) is approximated by the estimate

$$
\hat{p}_{ij} = f\left(E\left(V_{ij}\right), E\left(\sigma_{ij}\right), 0, n_{ij}, \hat{p}_{ij-1}\right), \tag{6}
$$

where *E( )* denotes the expected value and is estimated by the historical mean or median. Therefore, instead of estimating transaction costs based on the "true" price dynamics in equation (5), pre-trade models use equation (6).

In principle, post-trade models can utilize all information from the trading process. However, attention should be paid to the fact that just replacing the estimated variables in (6) with the true variables in (5) is problematic. First, the above mentioned pre-trade models are structural models and require variable input which is relatively smooth and free of outliers. Unusual volume or volatility can provide unintuitive results. Second, using equation (5) does not solve the problem of possible model misspecification. If the model is wrong, better input variables do not necessarily provide better cost estimates. Third, all the input variables such as intra-day volume, volatility, and trade imbalances are affected by one's own trading. However, a post-trade model is supposed to be a benchmark and not gameable. In addition, the econometric problem of endogeneity arises, which we discuss in more detail below.

In what follows, we pursue a different approach. After deriving the expected transaction costs from the pre-trade model, we incorporate the general market effects at the time when the trades actually took place into the post-trade model. The new model incorporates factors such as market returns and trade imbalances. The post-trade cost model is given by the equation

$$\begin{aligned} &Post\_Cost\left(S,\left(n_{ij}\right)_{i=1,..,T;\, j=1,...,N}\right)= \\ &\quad Pre\_Cost\left(S,\left(n_{ij}\right)_{i=1,..,T;\, j=1,...,N}\right)+\gamma_1\cdot X_1(S,T)+...+\gamma_N\cdot X_N(S,T) \end{aligned}, \tag{7}$$

where S is the order size, T is the trading horizon (in days) and $X_j(S,T)$ are factors such as

- the standardized actual volume over the trading period $\left(V(T)-E\left(V(T)\right)\right)/\sigma\left(V(T)\right)$,
- the standardized actual volatility over the trading period $\left(\sigma(T)-E\left(\sigma(T)\right)\right)/\sigma\left(\sigma(T)\right)$,
- the standardized actual spread over the trading period $\left(s(T)-E\left(s(T)\right)\right)/\sigma\left(s(T)\right)$,
- a proxy of the signed intra-day stock-specific momentum over the trading period $m\left(\left(n_{ij}\right),T\right)$.

The coefficients $\gamma_1,\gamma_2,...,\gamma_N$ are estimated for different exchanges and liquidity groups using the Panel data regression

$$\begin{aligned} &Realized\_Cost\left(S,\left(n_{ij}\right)_{i=1,..,T;\, j=1,...,N}\right)-Pre\_Cost\left(S,\left(n_{ij}\right)_{i=1,..,T;\, j=1,...,N}\right)= \\ &\qquad \gamma_1\cdot X_1(S,T)+...+\gamma_N\cdot X_N(S,T)+\varepsilon \end{aligned}. \tag{8}$$

Note that we do not use stock-specific intra-day momentum directly since stock-specific momentum and transaction costs are highly correlated and co-dependent. Ignoring endogeneity between costs and stock-specific momentum leads to biased estimates. One might actually obtain a large R^2 when regressing costs against stock-specific momentum since both variables are co-dependent. However, the associated regression parameters would be misleading since minimizing the distance function would not yield stable parameter estimates. To avoid endogeneity, we approximate stock-specific momentum within the trading period T with an instrumental variable that is determined by factors completely independent of the pre-trade model. Specifically, for the most liquid stocks, we estimate stock-specific momentum by the intra-day market return and the stock-specific trade imbalances during the trading period. For liquid stocks, we use the sector return and trade imbalances. For the least liquid stocks, we use the sector return, the industry return, and trade imbalances. The use of the various returns is necessary since the most liquid stocks would drive the industry return and we would introduce the endogeneity problem we try to address. The trade imbalances are defined as the intra-day signed share volume imbalances. The trades are classified as buys and sells using a generalized version of the Lee and Ready (1991) algorithm. Trades above (below) the mid quote are classified as buys (sells). Trades at the mid quote are classified using the tick test, i.e., up ticks are buys and down ticks are sells.

The stock-specific intra-day momentum is defined as the strategy-weighted return starting at the order decision time and ending when the order is fully executed. Specifically, for the stock-specific intra-day momentum we have

$$m\left(\left(n_{ij}\right)_{i=1,..,T;\, j=1,...,N},T\right)=10,000\cdot\sum_{i=1}^{T}\sum_{j=1}^{N}\frac{n_{ij}}{S}\cdot\left(p_{ij}-p_{1s-1}\right)\Big/ p_{1s-1} \tag{9}$$

Notice that when $T_{Start} > 1$, the definition above can be re-written, after some simple arithmetic manipulation[2] as

$$m\left(\left(n_{ij}\right)_{i=1,..,T;\, j=1,...,N},T\right)=10,000\cdot\left(\frac{p_{T_{start},s_{start}-1}}{p_{1s-1}}\cdot\underbrace{\left(\sum_{i=T_{Start}}^{T}\sum_{j=1}^{N}\frac{n_{ij}}{S}\left(p_{ij}-p_{T_{start},s_{start}-1}\right)\Big/ p_{T_{start},s_{start}-1}\right)}_{(*)}+\frac{\left(p_{T_{start},s_{start}-1}-p_{1,s-1}\right)}{p_{1,s-1}}\right), \tag{10}$$

Note that equation (10) incorporates the stock-specific momentum between order placement time $(1,s)$ with associated prevailing mid quote price $p_{1,s-1}$, and the time when the order starts to be executed at (T_{Start},s_{Start}), i.e. $p_{T_{start},s_{start}-1}$ is the last mid quote price before or equal to the starting time of bin s_{Start} in day T_{Start}. The endogeneity problem occurs only when one starts trading. Consequently, the stock-specific intra-day momentum component (*) is the only part that needs to be approximated by our instrumental variable.

The intra-day market, sector, and industry momentums are defined and calculated the same way as in (*) of equation (10), i.e., as the strategy-weighted returns from (T_{Start},s_{Start}) to (T,e). The stock-specific trade imbalance is defined similarly using intra-day trade imbalances instead of returns in (*) of equation (10).

The strategy (n_{ij}) in (9), (10) and in the post-trade estimate in (7) can be either a pre-trade strategy (e.g., an optimal strategy based on a certain risk aversion parameter) or the actual trading strategy. Note that these two strategies will measure two different things: choosing the pre-trade strategy evaluates your actual realized costs versus the costs that you would have had if you stayed with the pre-trade strategy. Choosing the realized strategy evaluates your execution against your peers who used the same strategy. With both options, inclusion of the strategy in the momentum calculations adds more strategy-dependence in the post-trade cost estimate.

To illustrate various strategies, Table 1 reports an example for stock XXX[3]. For a hypothetical order in stock XXX of 25,000 shares which is approximately 18% of median daily share volume (MDV), we report four different pre-trade strategies. The pre-trade strategies are based on the information

---

[2] One has to write $\left(p_{ij}-p_{1s-1}\right)$ out as $\left(p_{ij}-p_{T_{start},s_{start}-1}+p_{T_{start},s_{start}-1}-p_{1,s-1}\right)$.

[3] Stock XXX is a randomly chosen mid-cap stock. The stock's market capitalization is about $980 million, the median daily share volume (MDV) is about 140,000 shares and the daily volatility is 2% as of August 1, 2006.

set at the time of order placement, here 9:10am on August 1, 2006 and computed using ITG's ACE model. The first strategy assumes a risk aversion of zero, that is, we ignore risk associated with a trading strategy and only minimize expected transaction costs. ACE gives us a two-day strategy as optimal strategy. The shares in each trading bin are reported in Table 1. The second strategy assumes a risk aversion of 0.3, which is considered as being neutral. Now, ACE's optimal strategy is a one-day strategy that is somewhat front-loaded. The third strategy assumes a risk aversion of 0.9, which is considered as being aggressive. ACE's optimal strategy is a one-day strategy with heavy trading early in the day. Finally, the fourth pre-trade strategy is a one-day volume-weighted average price (VWAP) strategy. The strategy mirrors the average intra-day volume distribution of the stock. However, traders do not usually follow their pre-trade strategies completely, they rather adjust their trading behavior during the course of the day to current market conditions. For our example, trading for the order actually does not start until 11:30am. Table 1 reports two such strategies that utilize all available information. Strategy 5 is based on the actual empirical VWAP on that day. A trader just trades with the order flow of the stock. Finally, Strategy 6 is based on a VWAP strategy put in place at 11:30am when trading starts. That means, we trade according to the volume distribution estimated at 11:30am. It is obvious from Table 1, that different strategies yield quite different trading patterns. These trading patterns enter equations (9), (10) and thus also (7) through the trading quantities $\left(n_{ij}\right)_{i=1,..,T;\, j=1,...,N}$.

3. DATA

To model post-trade costs, we utilize ITG's proprietary Peer Group Database which consists of execution data (market and limit orders) from more than 80 large investment management firms. The study in this paper is based on U.S. execution data from April 2004 to March 2006. The data are collected from 84 institutions. To minimize transaction costs, investment managers break large orders into multiple smaller orders. The cost associated with the fragmented elements of the initial intended trade is then reported in the database. In order to capture the price impact and execution costs of institutional trading associated with the initial order, we use a clustering technique which is well known in the transaction cost literature (see e.g., Chan and Lakonishok (1995)). A buy (sell) "cluster" is the succession of purchases (sales) of a particular stock by the same manager. The order cluster ends when

(a) the manager stays out of the market for at least one day,
(b) the manager does not execute more than 2% of MDV within a day.

After reconstructing the initial order clusters, we identify the market conditions associated with each cluster. Generally, the execution time stamps are not reported. In order to establish a trading timeline for order clusters that do not contain trade time information, we assume that the investment manager used a VWAP trading strategy.[4] Using ITG's ACE model, we derive the pre-trade cost estimates of each cluster which is based on historical market conditions and neutral market sentiment. Consequently, the ACE pre-trade costs are entirely based on one's own trading strategy and direct market impact. ACE does not assume market effects due to other market participants.[5]

[4] Large institutions often use VWAP as their benchmark.
[5] We assume a VWAP trading strategy with trading horizon in days. This strategy reflects the benchmark costs for an average (typical) trader during the trading horizon.

We distinguish between Listed and Over-the-Counter (OTC) stocks to take into account cost differences for different market structures.[6] Stocks then are grouped based on their 21-day median dollar volume. We rank all available stocks (approximately 7,000) according to their 21-day median dollar volume at the beginning of each month during the sample period. For Listed and OTC stocks separately, we divide the stocks into eleven liquidity groups. Liquidity group 0 represents the least liquid stocks and liquidity group 10 represents the most liquid stocks. Table 2 presents the liquidity group thresholds for Listed and for OTC stocks, respectively.

Table 3 reports descriptive statistics for Listed and Table 4 for OTC stocks. For Listed stocks, Table 3 reports almost 1.6 million orders in our sample with more of the orders being concentrated in the more liquid stocks. Share volume ranges from a low of 10 million shares for liquidity groups 0-2 to 8.15 billion shares for liquidity group 9 with the total share volume of executed orders being 22.27 billion. Dollar volume totals almost $750 billion and ranges from $150 million for liquidity groups 0-2 to almost $292 billion for liquidity group 10. The average execution price across all orders is $33.68, but the average execution price raises from $10.64 for the least liquid stocks to $39.80 for the most liquid stocks. The average order size is about 14,000 shares with a range from 3,910 to more than 18,000 shares. The most liquid stocks have the largest average order size and the standard deviation also is largest for the most liquid stocks. The average market capitalization is $29.5 billion. For liquidity groups 0 through 8, the firm size is relatively small between $400 million and $4.8 billion. Only for the liquidity groups 9 and 10 is the market capitalization substantial at $15.2 billion and $89.5 billion, respectively. The average days to completion is about 1.3 days for all liquidity groups. The time horizon of orders does not seem to depend on the liquidity groups. Overall, the average order executes 5.5% of MDV. The participation rate ranges from 39.5% for the least liquid stocks to only 1% for the most liquid stocks. Obviously, for less liquid stocks, any order constitutes a substantial amount of daily trading volume.

For OTC stocks, Table 4 reports almost 690 thousand orders in our sample with more of the orders being concentrated in the more liquid stocks. For liquidity groups 0-2, there are no observations at all. Share volume ranges from a low of 4 million shares for liquidity group 3 to 5.19 billion shares for liquidity group 10 with the total share volume of executed orders being 12.80 billion. Dollar volume totals over $300 billion and ranges from $34 million for liquidity group 3 to over $146 billion for liquidity group 10. Compared to the Listed stocks in Table 2, we have fewer observations and less trading activity for OTC stocks. The average execution price across all orders is $23.46, but the average execution price raises from $8.67 for the least liquid stocks to $28.22 for the most liquid stocks. The OTC stocks in our sample tend to be lower-priced stocks compared to the Listed stocks. The average order size is about 18,600 shares with a range from 3,620 to almost 32,000 shares. The most liquid stocks have the largest average order size and the standard deviation also is largest for the most liquid stocks. Compared to the Listed stocks, orders in the OTC stocks tend to be larger. The average market capitalization is $16.8 billion. For liquidity groups 3 through 9 the firm size is relatively small between $300 million and $3.6 billion. Only for the liquidity group 10 is the market capitalization substantial at $64 billion. As expected, the OTC stocks are smaller compared to the Listed stocks in our sample. The average days to completion is about 1.3 days for all liquidity groups. The time horizon of orders does not seem to depend on the liquidity groups. This finding is identical to the Listed stocks. Overall, the average order executes 9.7% of MDV. The participation rate ranges from 33.5% for the least liquid stocks to only 1% for the most liquid stocks. The average

[6] Listed stocks are listed on the New York Stock Exchange (NYSE) or the American Stock Exchange (Amex). All other stocks are considered OTC stocks.

participation rate is greater for the OTC stocks since the OTC stocks are less liquid compared to the Listed stocks.

Figure 1 plots the average realized transaction costs for all liquidity groups. Average costs decrease as the liquidity of a stock increases for both Listed and OTC stocks. They range from almost 25 bps to about 2 bps for Listed stocks and from almost 35 bps to about 4 bps for OTC stocks. The pattern in transaction costs is partially attributed to the fact that less liquid stocks have larger bid-ask spreads (the majority of clusters have small order sizes). Note that average costs for Listed and OTC stocks are not directly comparable because of different liquidity group thresholds.

Figures 2 and 3 display average realized costs by relative order size (relative to MDV) for different liquidity groups. The charts show that average costs increase in relative order size due to price impact. Most liquid stocks as in liquidity group 10 have higher realized costs due to higher price impact. However, for lower liquidity groups, there seems to be little difference in average realized costs between groups. One also needs to keep in mind that the same relative order sizes for liquidity group 10 and liquidity group 5 mean very different actual order sizes. This may explain the higher price impact costs for the most liquid stocks. OTC stocks appear to be more expensive than Listed stocks when controlling for order size only.

For each liquidity group and for Listed and OTC, the parameters in equation (8) are estimated separately. In the following, we consider the model

$$\begin{aligned} &Post\_Cost\left(S,\left(n_{ij}\right)_{i=1,..,T;\,j=1,...,N}\right)-Pre\_Cost\left(S,\left(n_{ij}\right)_{i=1,..,T;\,j=1,...,N}\right) \\ &\qquad = \gamma_1 \cdot m_{proxy}\left(\left(n_{ij}\right)_{i=1,..,T;\,j=1,...,N},T\right)+\gamma_2 \cdot stand.spread+\varepsilon \end{aligned}, \qquad (11)$$

where m_{proxy} is the signed proxy for stock-specific intra-day momentum and $stand.spread$ is the standardized actual spread[7] over the trading period defined in Section 2. The dependent variable $Post\_Cost - Pre\_Cost$ and the momentum proxy m_{proxy} are normalized with the stock-specific volatility to control for heteroskedasticity. This regression model is motivated by its mere simplicity. Modeling the impact associated with deviation from expected volume, volatility, and/or spread is only significant during unusual and unexpected stock-specific events. The proxy for stock-specific intra-day momentum is estimated based on a 60-day rolling window. Note that it is independent of one’s own trading since we factor in only market, sector or industry movements and trade imbalances net of one’s own trading.

Equation (11) and our discussion above show that our approach has decomposed transaction costs into two components: the costs due to one’s own trading and the costs due to general market effects.

4. RESULTS

[7] We standardize the actual spread by the corresponding 60-days mean and standard deviation.

When empirically modeling equation (11), there are two potential problems. As discussed above, the first problem relates to the concern of endogeneity where stock-specific momentum is correlated with the error term. To alleviate this problem we use the strategy-weighted market, the strategy-weighted sector, and the strategy-weighted industry return along with the stock-specific trade imbalances excluding one's own trading as instrumental variables which we described in Section 2. Note, for the time between order decision and actual start of trading, we could use the stock-specific intra-day momentum without introducing endogeneity as in equation (10). However, in the current analysis, we do not use this approach. Consequently, the results may improve by using the stock-specific momentum in equation (10).

The second problem relates to the fact that the model coefficients for the difference in pre- and post-trade costs may depend on liquidity group, Listed vs. OTC,[8] and order size. For this purpose, we estimate the model for each exchange and liquidity group separately. In addition, we investigate the effect of the momentum proxy for different order size buckets (relative to MDV)[9] separately.

We assess the performance of our instrumental variables by analyzing the prediction errors between stock-specific momentum and the instrumental variable prediction.[10] For liquidity group 10, we find that the prediction error is within 50 bps for the majority of cases with extremes of as much as 120 bps. This compares to stock-specific momentum of as much as about 210 bps. For liquidity group 3, a large portion of the distribution of the prediction error is again within 50 bps. However, in the extreme, the prediction error is as large as about 200 bps which compares to stock-specific momentum of more than 350 bps. Overall, our results indicate that the instrumental variable approach in ITG's ACE model for post-trade analysis explains a considerable amount of stock-specific momentum.

The instrumental variable approach has another nice feature. Since we estimate the coefficients of the instrumental momentum proxy variable daily based on the most recent 90 days (similarly to ITG ACE) we ensure that the model performs out-of-sample reasonably well as well. In fact, extensive research has shown that in- and out-of-sample results are almost indistinguishable. In what follows, we therefore present in-sample results only. Out-of-sample results are available from the authors upon request.

Figure 4 reports the average adjusted R^2 for regression (11) over all order sizes for different liquidity groups of Listed and OTC stocks. The R^2s are slightly lower for OTC stocks than for Listed stocks. They are greatest for liquidity groups 8 at about 38% and 37%, and lowest for liquidity groups 3 at about 27% and 24% for Listed and OTC stocks, respectively.[11] Overall, the R^2s are of considerable magnitude. Figure 4 confirms the intuition that costs for illiquid stocks are harder to assess. The lower R^2s are mostly attributed to the lower explanatory power of the momentum proxy. Sector and/or industry momentum have not much impact on very illiquid stocks. The trade imbalance measure can be very noisy and not very informative for stocks with few trades per day.

---

[8] See Table 1 for the definition of each liquidity group for Listed and OTC stocks.

[9] Order size buckets are 0-1%, 1-5%, 5-10%, 10-25%, 25-50%, and >50% of MDV.

[10] For brevity of exposition, we do not report the results, but they are available from the authors upon request.

[11] R^2s for liquidity groups 0, 1, and 2 are not reported since we do not have any or enough observations (see Table 2 and Table 3)

Figures 5 and 6 show the estimates of coefficient γ_1 in regression (11) for different order size buckets for the selected liquidity groups 5, 7, and 10 for Listed and OTC stocks, respectively. The results are qualitatively similar for other liquidity groups. The two graphs indicate that γ_1 is relatively constant for smaller relative order sizes and then sharply decreasing. The decreasing part is intuitive. For larger order sizes, the permanent price impact due to one's own trading should dominate. The declining trend is liquidity group and exchange-specific. For OTC stock, it seems that the trend begins around 5-10%. For Listed stocks, it starts later especially for more liquid securities.

Figure 7 reports the estimates for coefficient γ_2 in regression (11) for different liquidity groups for Listed and OTC stocks, respectively. As one would expect, all γ_2 coefficients are positive, i.e., market conditions that had higher than normal spread values result in higher transaction costs. For both Listed and OTC, we observe that the spread coefficients are declining as the stocks become more and more liquid. This is intuitive since transaction costs for illiquid stocks depend hugely on the underlying spreads while the spread values for very liquid stocks are almost always one cent during the trading day.

Figures 8 and 9 plot average realized transaction costs, both ACE (pre-trade) and ACE for post-trade analysis transaction cost estimates for Listed and OTC stocks, respectively. In both charts realized and ACE transaction costs match very well. This is no surprise since ITG's ACE transaction cost estimates are calibrated to realized transaction costs. However, the ACE estimate is much smoother than the average realized costs. This is to be expected since we construct a smooth estimator that does not take into account market conditions. ITG's ACE model for post-trade analysis transaction cost estimates are also very similar to the realized costs. Compared to the ACE estimates, they are more volatile and closer to the realized costs. Again, this is to be expected, since for ITG's ACE model for post-trade analysis, we take market conditions into account and are able to explain average realized costs better.

Figures 10 and 11 plot the distributions of the prediction errors of ACE and ITG's ACE model for post-trade analysis transaction cost estimates for Listed and OTC stocks, respectively. Both charts show that the prediction error of ACE is much more fat-tailed. ITG's ACE model for post-trade analysis estimates fit the realized costs better.

5. CONCLUSIONS

In this paper, we develop a post-trade transaction cost model that incorporates general market conditions into the estimation of transaction costs. We then apply the general model to ITG's ACE model. Incorporating these market factors gives a more accurate estimate of transaction costs. We are able to decompose transaction costs into transaction costs due to one's own trading and transaction costs due to market effects. The proposed ACE model for post-trade analysis explains between 25% and 37% of the variation in transaction costs for all exchanges and liquidity groups. This indicates that our model is performing well.

To address potential endogeneity problems, we do not use stock-specific momentum, but rather employ an instrumental variable approach. While the instrumental variable approach yields reasonable predictions for stock-specific momentum for most cases, it does not always do so. For the cases in which the instrumental variable approach does not yield a good predictor, it seems to make sense to apply some heuristic rules in the implementation of the model. Furthermore, for very small order sizes, endogeneity is not really an issue and there is probably no need to use the instrumental variable approach but just use the stock-specific momentum.

Finally, note that we focus only on some determinants of trading costs. There are many other transaction costs factors that are important. A word of caution is therefore in order. As with all transaction cost models, the proposed transaction cost model for post-trade analysis should not be viewed as a tool to analyze trades one by one, rather it is necessary to aggregate across trades, e.g., consider averages. Only aggregation allows for meaningful analyses and comparisons.

6. REFERENCES

ALMGREN, ROBERT, AND NEIL CHRISS, 2000, “Optimal Execution of Portfolio Transactions,” *Journal of Risk*, 3 (Winter 2000/2001), 5-39.

BERTSIMAS, DIMITRIS, AND ANDREW LO, 1998, “Optimal Control of Execution Costs,” *Journal of Financial Markets*, 1, 1-50.

BERTSIMAS, DIMITRIS, PAUL HUMMEL, AND ANDREW LO, 2000, “Optimal Control of Execution Costs for Portfolios,” *Computing in Science & Engineering*, 40-53.

CHAN, LOUIS, AND JOSEF LAKONISHOK, 1995, “The behavior of stock prices around institutional trades,” *Journal of Finance*, 50, 1147-1174.

DOMOWITZ, IAN, JACK GLEN, AND ANANTH MADHAVAN, 2002, “Liquidity, Volatility and Equity Trading Costs Across Countries and Over Time,” *International Finance*, 4, 221-256.

EDELEN, ROGER M., EVANS, RICHARD B. AND KADLEC, GREGORY B., “Scale Effects in Mutual Fund Performance: The Role of Trading Costs" (March 17, 2007). Working Paper. Available at SSRN: http://ssrn.com/abstract=951367

FABOZZI, FRANK, SERGIO FOCARDI, AND PETTER KOLM, 2006, Financial Modeling of the Equity Market. From CAPM to Cointegration. John Wiley and Son, Inc., Hoboken, New Jersey.

HUBERMAN, GUR, AND WERNER STANZL, 2005, “Optimal Liquidity Trading,” *Review of Finance*, 9, 165-200.

ITG INC., 2007, ITG ACE – AGENCY COST ESTIMATOR.

LEE, CHARLES M.C., AND MARK J. READY, 1991, “Inferring trade direction from intraday data,” *Journal of Finance*, 46, 733-746.

LO, ANDREW, AND A. CRAIG MACKINLAY, 1988, “Stock market prices do not follow random walks: Evidence from a simple specification test,” *Review of Financial Studies*, 1, 203-238.

PEROLD, ANDRE, 1988, “The Implementation Shortfall: Paper Versus Reality,” *Journal of Portfolio Management*, 14, 4-9.

TABLE 1

EXAMPLE FOR DIFFERENT STRATEGIES

Example: Stock XXX[12]
Order Date and Time: August 1, 2006 at 9:30am
Order Size: 25,000 shares (approximately 18% of MDV)

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9	Bin 10	Bin 11	Bin 12	Bin 13
Pre-Trade Strategies (9:30am)													
1. ITG ACE optimal strategy, risk aversion r = 0 (minimize expected costs only)													
Day 1	700	600	500	800	900	1000	900	900	800	800	1200	1500	1200
Day 2	700	600	500	700	900	1000	900	900	800	800	1200	1500	1900
2. ITG ACE optimal strategy, risk aversion r = 0.3 (neutral)													
Day 1	2300	2000	2100	2200	2300	2000	1800	1500	1400	1300	1600	2000	2500
3. ITG ACE optimal strategy, risk aversion r = 0.9 (aggressive)													
Day 1	14400	4900	1800	400	400	400	400	400	300	400	500	700	0
4. Volume-Weight Average Price Strategy (VWAP) - one day													
Day 1	2800	200	1700	1500	1600	1400	1400	1200	1200	1400	1700	2700	4400
Post-Trade Strategies (trade start time = 11:30am)													
5. Actual (using empirical VWAP)													
Day 1	0	0	0	0	955	250	325	15535	665	800	910	1600	3960
6. VWAP from start bin (11:30am)													
Day 1	0	0	0	0	2300	2100	2000	1900	1800	1900	2500	4000	6500

Table 1 reports possible strategies that can be chosen in equations (9) and (10) and thus as input variable for post-trade performance evaluation for a hypothetical order in stock XXX of 25,000 shares which is approximately 18% of MDV. The strategy $\left(n_{ij}\right)_{i=1,..,T\,;\,j=1,...,N}$ in equations (9) and (10) can be either a pre-trade strategy or the actual trading strategy. Note that these two strategies will measure two different things: choosing the pre-trade strategy evaluates your actual realized costs versus the costs that you would have had if you had stuck with the pre-trade strategy. Choosing the realized strategy evaluates your execution against your peers who used the same strategy. With both options, inclusion of the strategy in the momentum calculations adds more strategy-dependence in our post-trade cost estimates.

[12] Stock XXX is a randomly chosen mid-cap stock. The stock's market capitalization is about $980 million, the median daily share volume is about 140,000 shares and the daily volatility is 2% as of August 1, 2006.

TABLE 2

LIQUIDITY GROUP THRESHOLDS

	Listed		**OTC**	
	Lower	***Upper***	***Lower***	***Upper***
Group 0 (least liquid)	0	34254	0	9098
Group 1	34254	119920	9098	37681
Group 2	119920	311484	37681	110391
Group 3	311484	725338	110391	253117
Group 4	725338	1613774	253117	508304
Group 5	1613774	3145090	508304	998401
Group 6	3145090	6010025	998401	1879740
Group 7	6010025	11592707	1879740	3581698
Group 8	11592707	24676326	3581698	7995956
Group 9	24676326	101104928	7995956	81474182
Group 10 (most liquid)	101104928	10000000000	81474182	10000000000

This table shows the liquidity group thresholds for Listed and OTC stocks for August 2006. The thresholds are in US$. The liquidity group of a stock is determined by comparing its 21-day median daily share volume with the above thresholds.

TABLE 3

DESCRIPTIVE STATISTICS FOR LISTED STOCKS

	ALL	GROUP 0-2	GROUP 3	GROUP 4	GROUP 5	GROUP 6	GROUP 7	GROUP 8	GROUP 9	GROUP 10
NUMBER OF ORDERS (in 000's)	1,580.1	3.5	17.5	34.8	65.3	102.0	152.1	239.6	559.7	405.5
SHARE VOLUME (in bill)	22.27	0.01	0.09	0.23	0.52	1.01	1.75	3.17	8.15	7.33
DOLLAR VOLUME (in bill)	749.92	0.15	1.09	3.51	10.11	22.12	43.63	89.65	187.88	291.77
AVG. EXECUTION PRICE (in dollars)	33.68	10.64	12.65	15.34	19.32	21.98	24.90	28.28	35.30	39.80
AVG. ORDER SIZE (in 000's)	14.09 (77.25)	3.91 (8.31)	4.94 (11.81)	6.57 (17.44)	8.01 (23.80)	9.86 (3.14)	11.52 (42.81)	13.23 (54.66)	14.57 (77.61)	18.08 (109.70)
AVG. MARKET CAP (in bill)	29.5 (55.5)	0.4 (0.5)	0.7 (2.4)	0.8 (2.0)	1.1 (3.5)	1.7 (2.6)	2.4 (2.3)	4.8 (5.0)	15.2 (11.4)	89.5 (82.7)
AVG. DAYS TO COMPLETION (in days)	1.3 (1.0)	1.3 (0.7)	1.3 (0.8)	1.3 (0.8)	1.3 (0.8)	1.3 (0.9)	1.3 (0.9)	1.3 (0.9)	1.3 (1.0)	1.2 (0.9)
AVG. PARTICIPATION RATE (in percent)	5.5 (29.3)	39.5 (117.0)	31.1 (94.2)	22.3 (73.3)	16.6 (56.9)	12.4 (44.4)	9.1 (35.2)	6.0 (26.5)	2.9 (15.0)	1.0 (6.2)

Table 3 reports descriptive statistics for all Listed stocks in our sample. The sample consists of all executions in U.S. stocks in ITG's Peer Group Database between April 2004 and March 2006. Average days to completion is defined in days and average participation rate is the ratio of order size and sum of total composite share volume during the days to completion. The numbers in parentheses are the standard deviations of the respective variables.

TABLE 4

DESCRIPTIVE STATISTICS FOR OTC STOCKS

	ALL	GROUP 0-2	GROUP 3	GROUP 4	GROUP 5	GROUP 6	GROUP 7	GROUP 8	GROUP 9	GROUP 10
NUMBER OF ORDERS (in 000's)	688.6	0	1.1	16.4	30.5	46.2	66.7	102.0	262.2	163.5
SHARE VOLUME (in bill)	12.80	0	0.004	0.08	0.19	0.39	0.66	1.23	5.06	5.19
DOLLAR VOLUME (in bill)	300.18	0	0.034	0.73	2.10	4.69	9.80	21.76	114.65	146.43
AVG. EXECUTION PRICE (in dollars)	23.46	N/A	8.67	9.40	10.81	12.11	14.95	17.74	22.65	28.22
AVG. ORDER SIZE (in 000's)	18.58 (106.50)	N/A	3.62 (9.15)	4.77 (11.58)	6.36 (15.01)	8.38 (21.94)	9.82 (26.35)	12.02 (35.75)	19.30 (87.28)	31.74 (184.31)
AVG. MARKET CAP (in bill)	16.8 (44.9)	N/A	0.3 (0.2)	0.3 (0.2)	0.4 (0.3)	0.5 (0.3)	0.6 (0.4)	0.9 (0.7)	3.6 (3.6)	64.0 (74.5)
AVG. DAYS TO COMPLETION (in days)	1.3 (0.8)	N/A	1.2 (0.6)	1.3 (0.7)	1.3 (0.8)	1.3 (0.8)	1.3 (0.8)	1.3 (0.8)	1.3 (0.8)	1.2 (0.7)
AVG. PARTICIPATION RATE (in percent)	9.7 (46.1)	N/A	33.5 (93.9)	33.2 (99.7)	28.1 (97.2)	22.0 (73.8)	17.4 (64.0)	12.4 (45.7)	6.2 (28.4)	1.0 (6.7)

Table 4 reports descriptive statistics for all OTC stocks in our sample. The sample data consists of all executions in U.S. stocks in ITG's Peer Group Database between April 2004 and March 2006. Average days to completion is defined in days and average participation rate is the ratio of order size and sum of total composite share volume during the days to completion. The numbers in parentheses are the standard deviations of the respective variable.

FIGURE 1

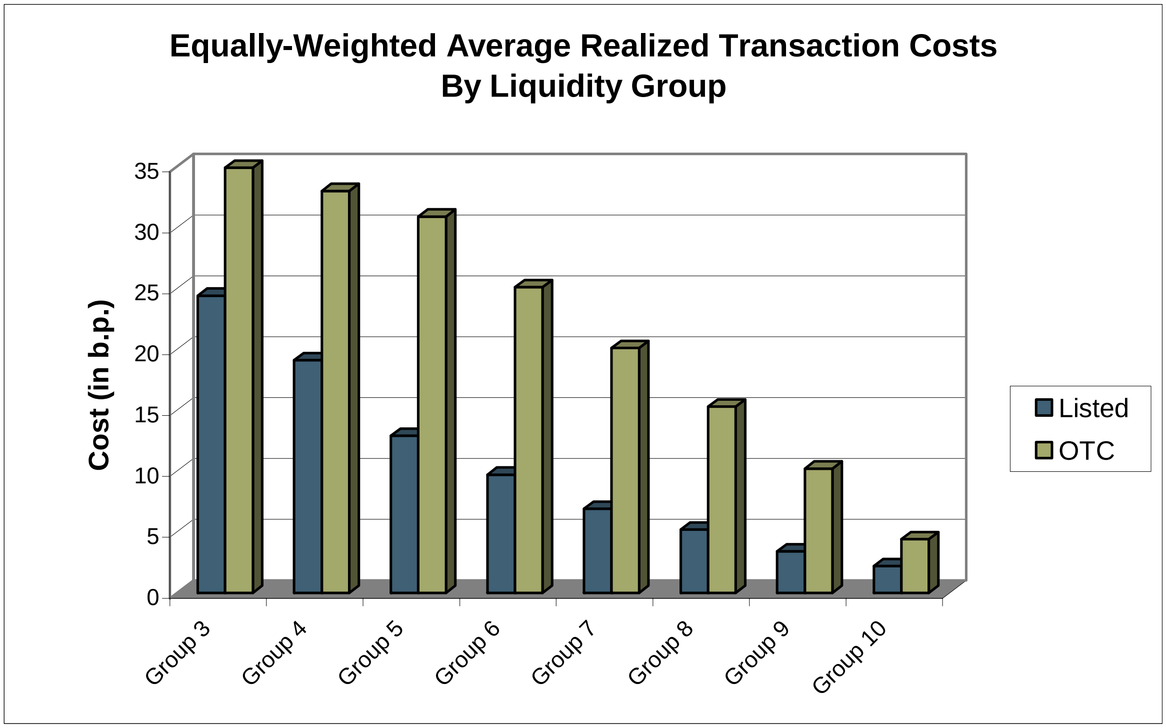


Figure 1 presents the equally-weighted average realized transaction costs for Listed and OTC stocks and different liquidity groups. Average costs for Listed and OTC stocks are not comparable because of different liquidity group thresholds (see Table 2).

FIGURE 2

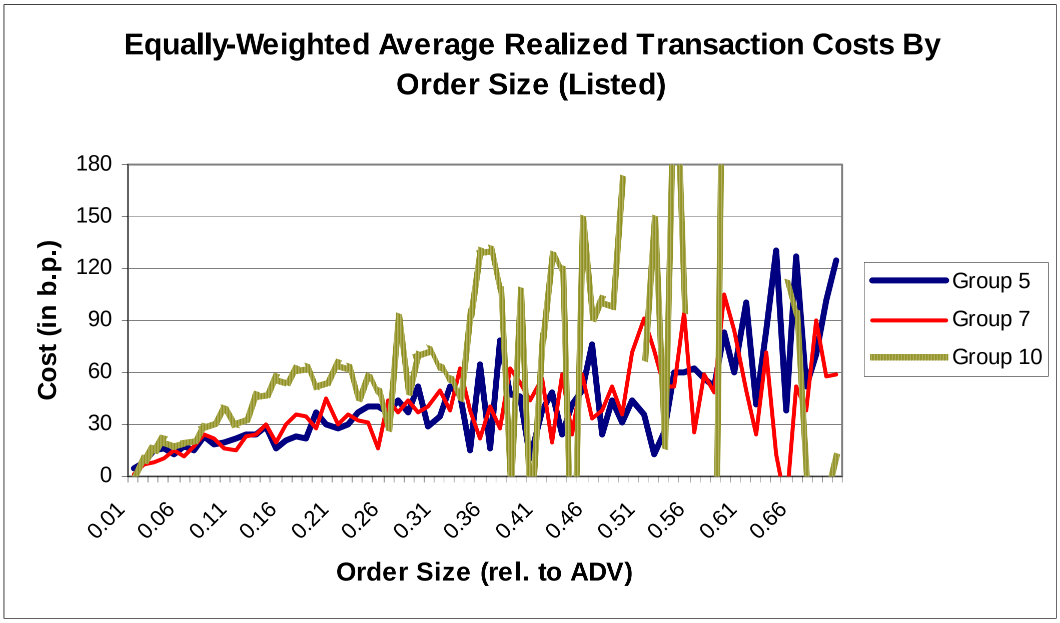
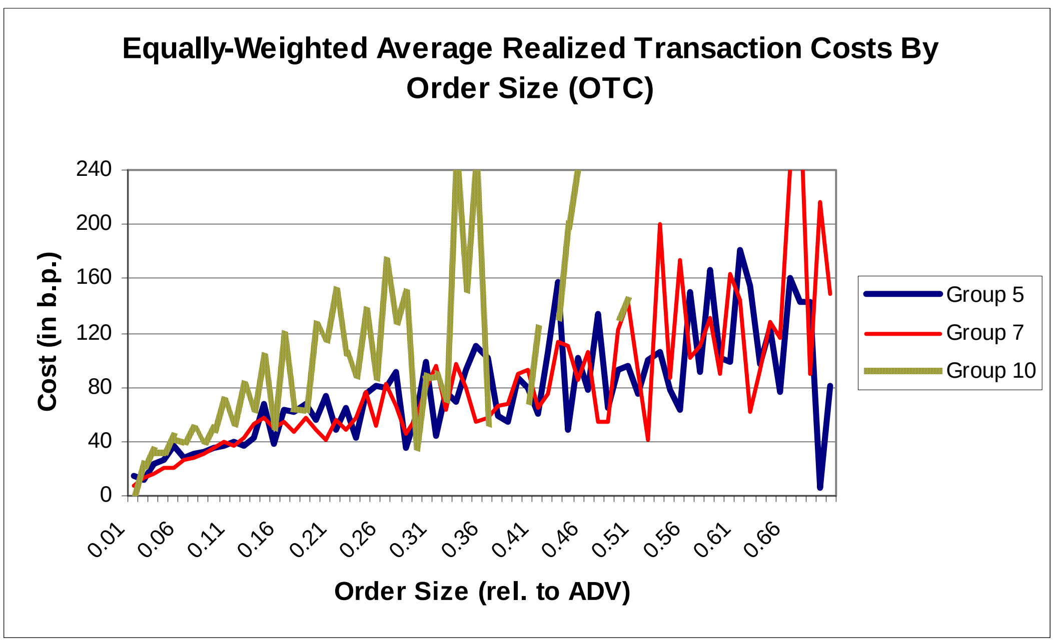


FIGURE 3



Figures 2 and 3 display average realized costs by relative order size (relative to MDV) for different liquidity groups.

FIGURE 4

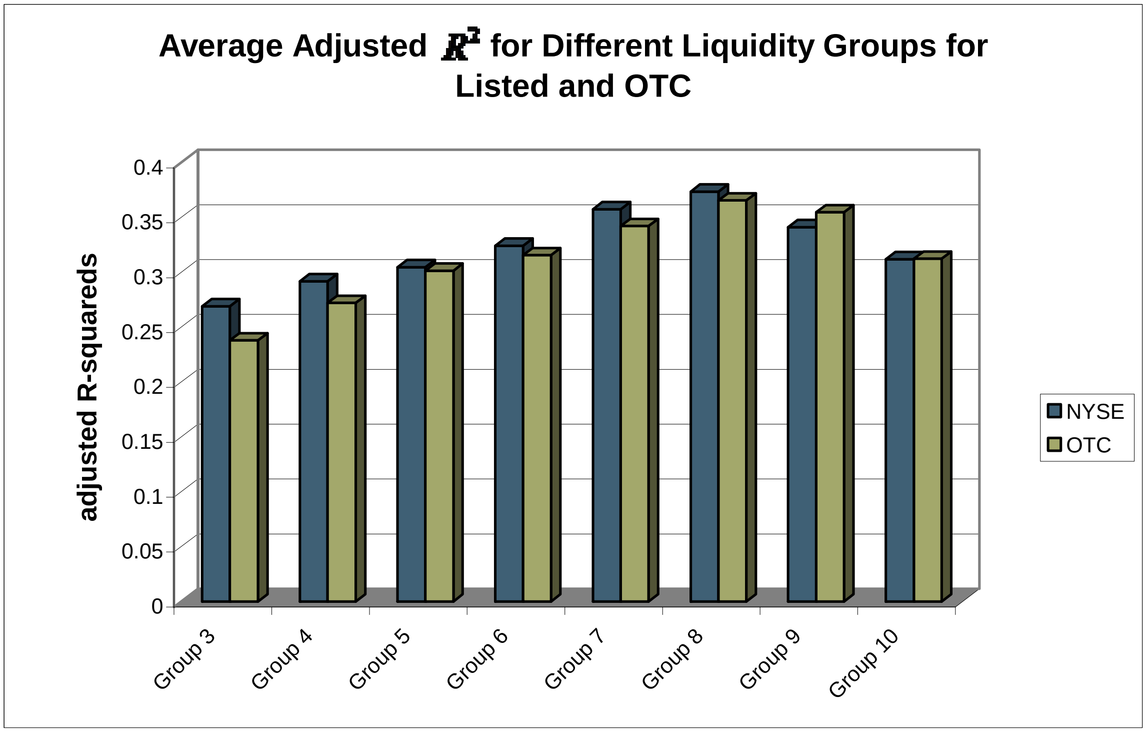


Figure 4 reports average adjusted R^2 s for regression (11) over all order sizes for different liquidity groups for Listed and OTC stocks, respectively.

FIGURE 5

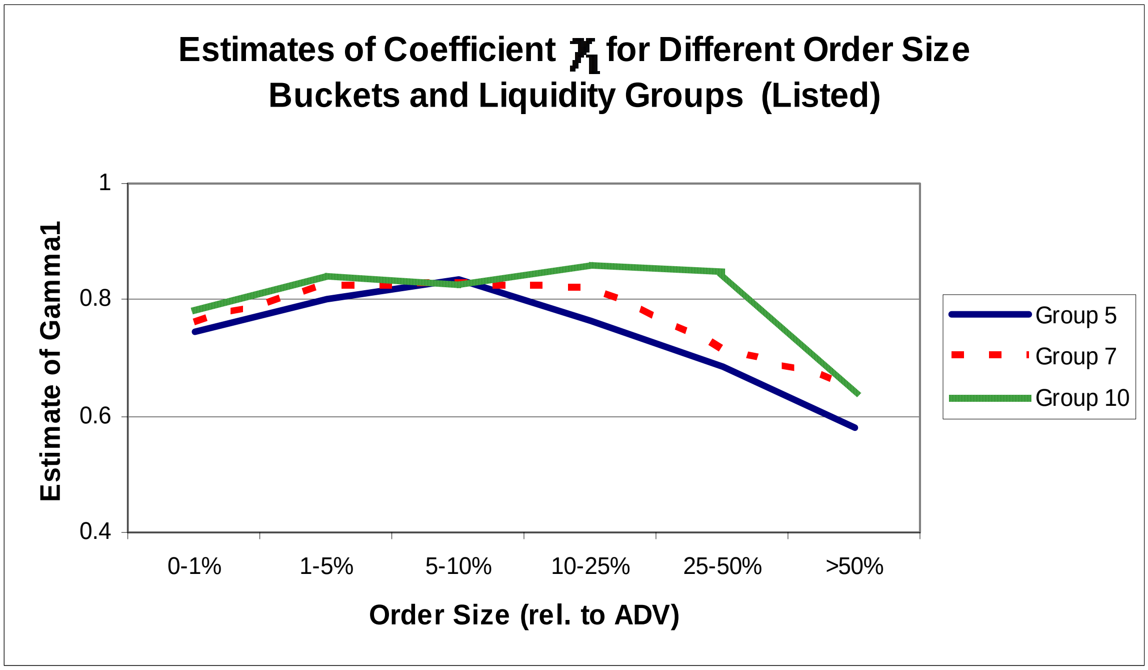
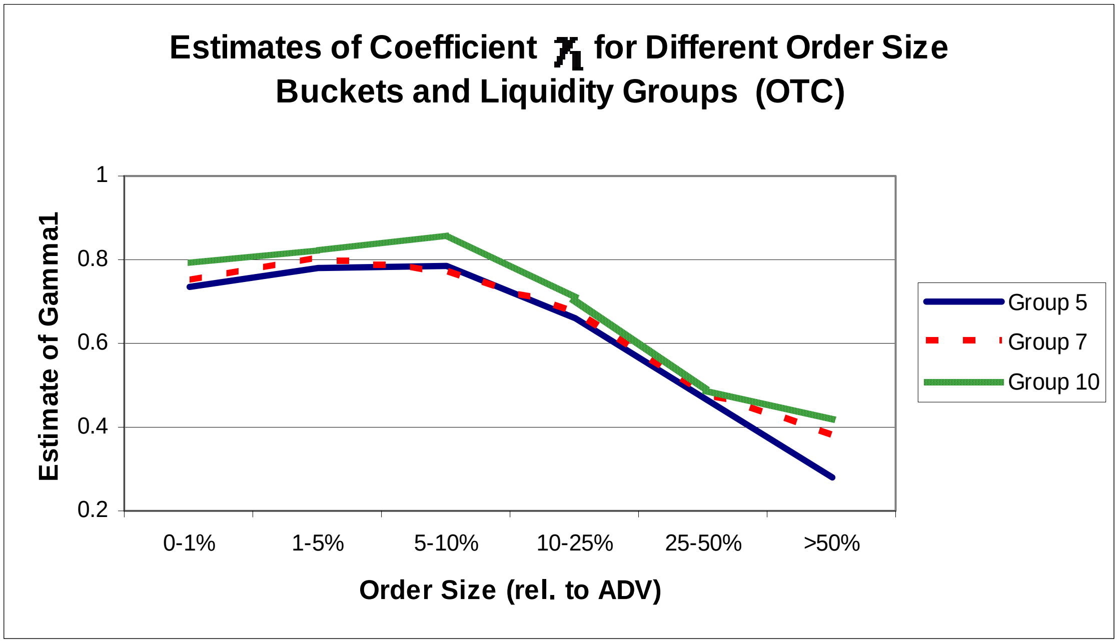
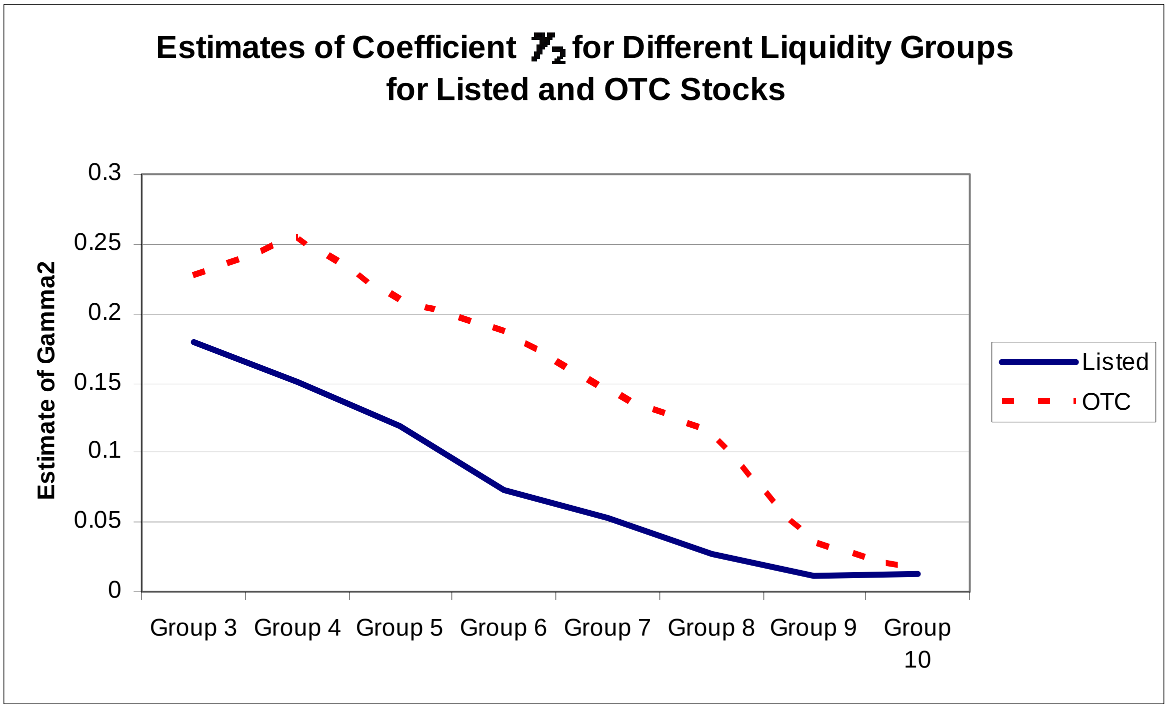


FIGURE 6



Figures 5 and 6 show the estimates of coefficient γ_1 in regression (11) for different order size buckets for selected liquidity groups of the Listed and OTC stocks, respectively.

FIGURE 7



Figures 7 reports the estimates of coefficient γ_2 in regression (11) for different liquidity groups of Listed and OTC stocks, respectively.

FIGURE 8

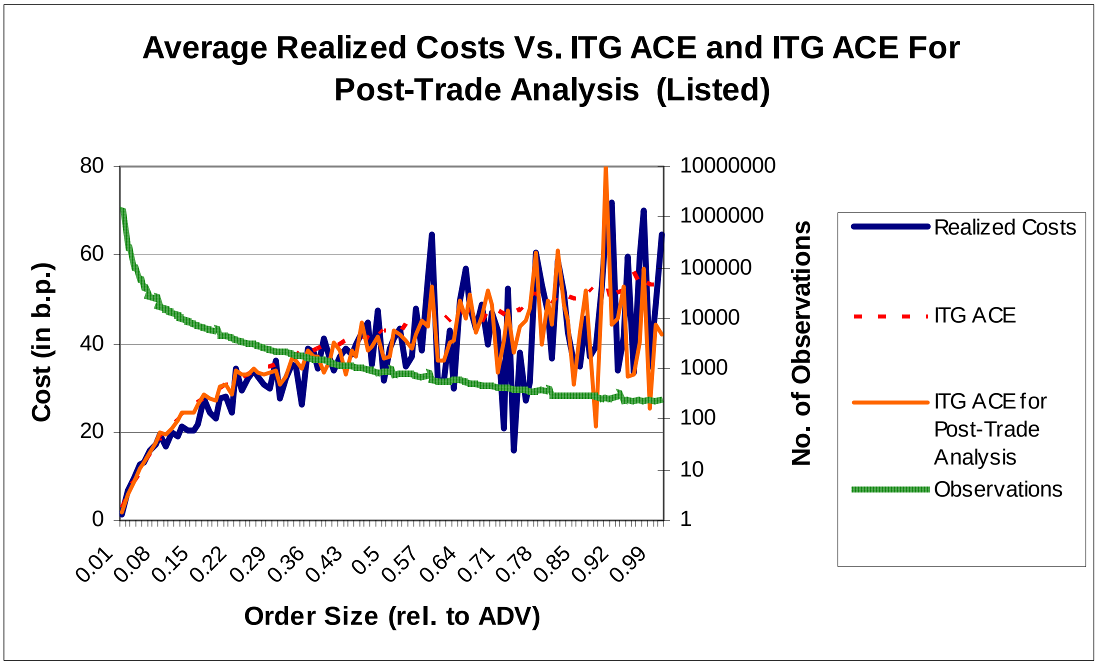
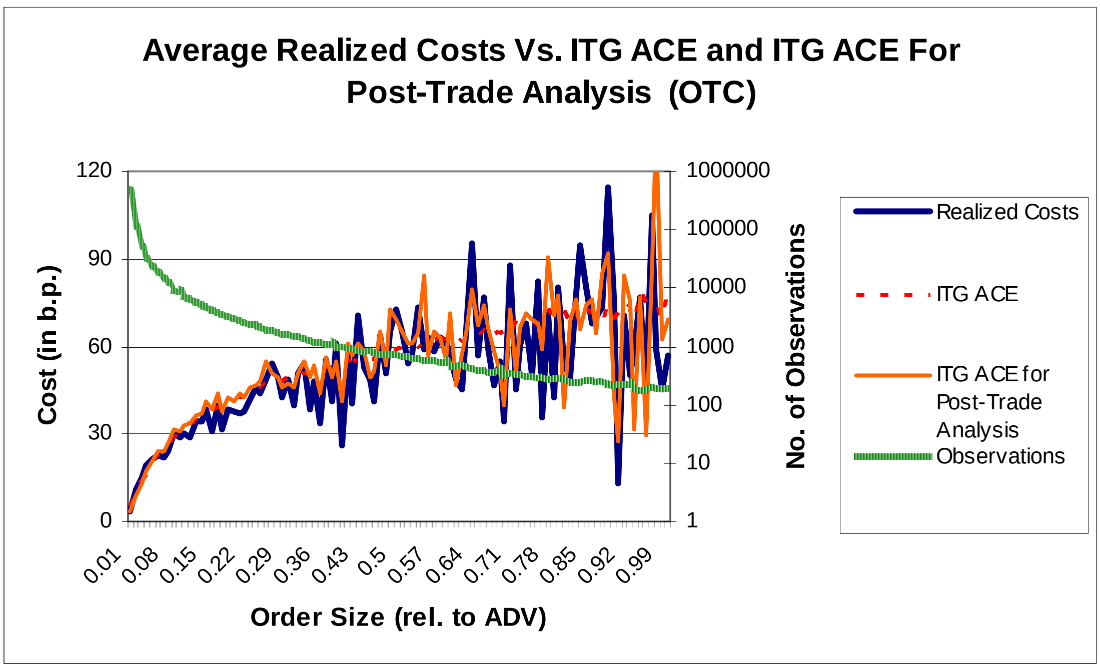


FIGURE 9



Figures 8 and 9 plot average realized transaction costs, ITG ACE, and ITG ACE for Post-Trade Analysis transaction cost estimates for Listed and OTC stocks, respectively.

FIGURE 10

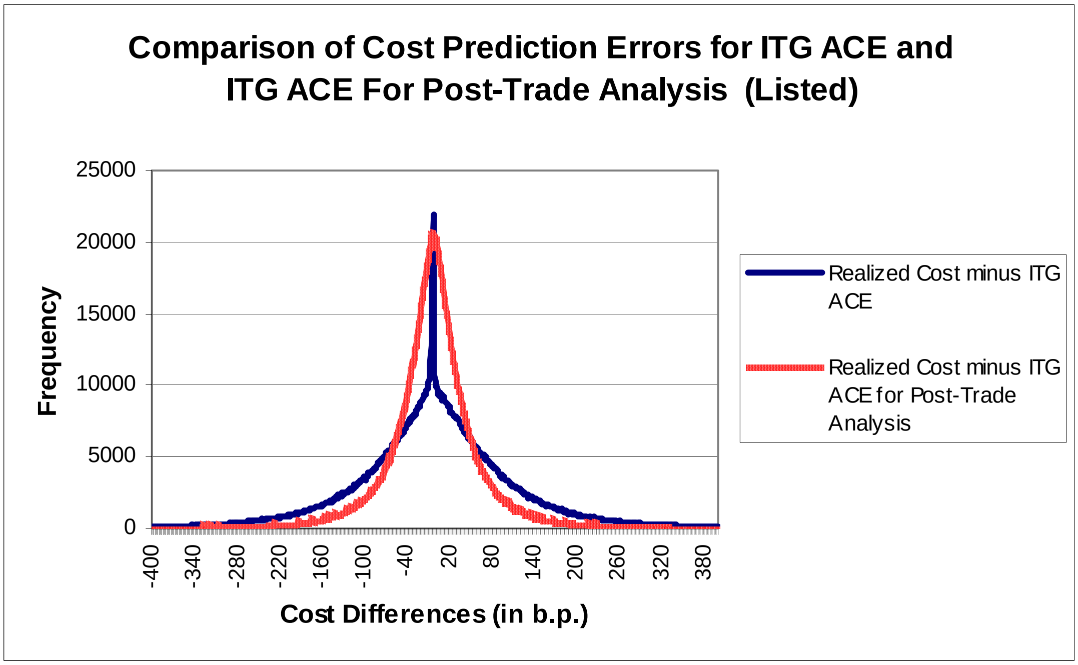
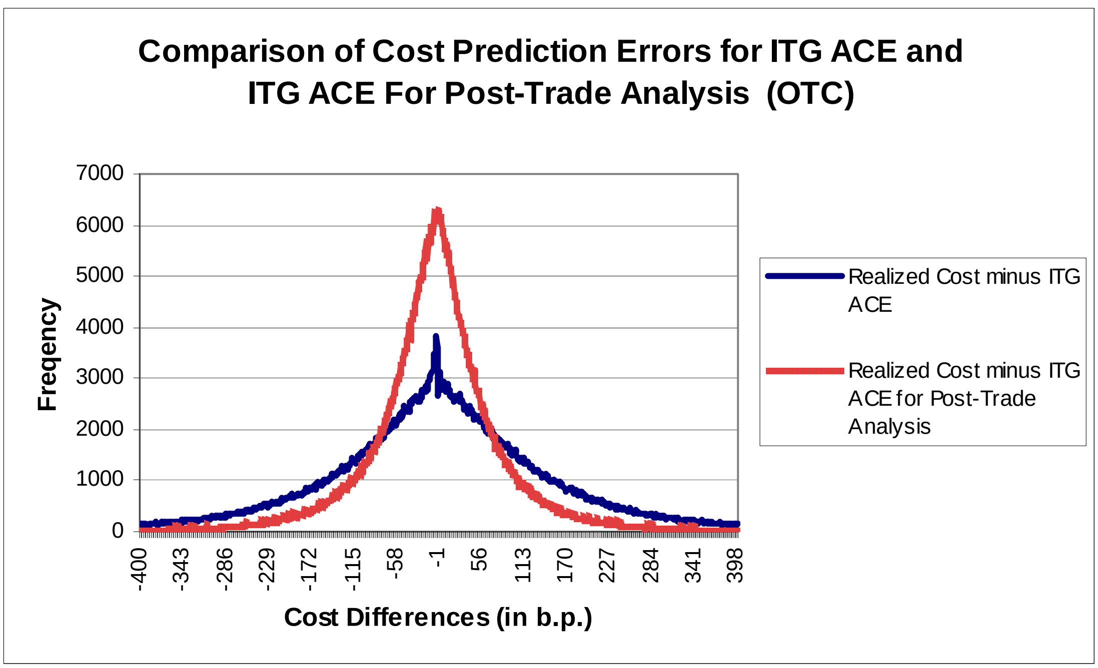


FIGURE 11



Figures 10 and 11 plot the distributions of the prediction errors of ITG ACE and ITG ACE for Post-Trade Analysis transaction cost estimates for Listed and OTC stocks, respectively.