Trading Costs

ANDREA FRAZZINI, RONEN ISRAEL, AND TOBIAS J. MOSKOWITZ*

First draft: April 2018
This draft: August 2018

Abstract

Using 1.7 trillion dollars of live trade execution data from a large institutional money manager across 21 developed equity markets over a 19-year period, we measure the real-world trading costs and price impact function of a large trader. We provide a novel description of how costs vary across trade type, stock characteristics, trade size, time, and exchanges globally to test various theories of price impact. We find actual trading costs to be an order of magnitude smaller than previous studies suggest, and describe the trading process leading to these costs. A model calibrated to match the distribution of actual costs across trade size, stocks, and time outperforms other models from the literature in out of sample tests that attempt to describe independent costs from brokers and realized costs of traded index funds. Our model based on realized costs from live trades portrays very different implementation costs than previous studies suggest.

---

*Andrea Frazzini is at AQR Capital, e-mail: andrea.frazzini@aqr.com. Ronen Israel is at AQR Capital, e-mail: ronen.israel@aqr.com. Tobias Moskowitz is at the Yale School of Management, Yale University, NBER, and AQR, email: tobias.moskowitz@yale.edu. This paper is an extension of some of the material in an earlier paper titled, "Trading Costs of Asset Pricing Anomalies," that has been separated into another paper. We thank Cliff Asness, John Campbell, Josh Coval, Darrell Duffie, Eugene Fama, John Heaton, Bryan Kelly, John Liew, Michael Mendelson, Stefan Nagel, Robert Novy-Marx, Lasse Pedersen, Amit Seru, Ken Singleton, Brian Weller, and seminar participants at the University of Chicago, Harvard Business School, Stanford University, University of California Berkeley, and the NBER Asset Pricing Meetings in Chicago for helpful comments. Moskowitz thanks the IFC at Yale University for financial support. AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR.

Understanding the role of liquidity and trading costs in asset pricing is an empirical challenge in academia due to lack of access to order and execution data from a live trading platform. We aim to fill this void using a unique database of orders and live executions from a large institutional money manager. Doing so, we find that real-world costs based on live trades are very different than those estimated in the literature from daily or intra-day data and provide new insights into the distribution of trading costs along several dimensions not previously studied.

Using a unique dataset of $1.7 trillion of live executed trades from a large money manager from August 1998 to June 2016, across 21 developed equity markets, we measure the real-world trading costs and price impact function of a large money manager (who trades portfolios based on many of the anomalies discovered in the academic literature). The data offer a singular look into the real-time costs of an investor who resembles the theoretical "arbitrageur" in asset pricing models. The data also offer a unique look into how trading costs vary globally across trade type, size, and exchange.

The trade execution data provide a unique laboratory to test many theories of transactions costs. Our live trading data contains several novel features not previously studied. We can identify the real-time price impact of a trade at various trade sizes. We observe whether the trade was a buy or sell, the market price at trade initiation, the amount traded, and the execution price for each share traded. These data allow us to calculate a precise measure of price impact. We can also differentiate between types of trades (e.g., to initiate a position – buy long or sell short – versus cover a position – sell long or buy to cover) which is unique to the literature and provides a more accurate picture of real trading costs to long-short portfolios common in the literature. In addition to estimating these costs for U.S. stocks, our data also covers 20 other developed equity markets, providing the first look at trading costs in different markets and a wealth of out of sample evidence.

We begin by describing the trading process of how trades in our data are generated and executed to better understand the trading cost estimates we get from the live trade data. We illustrate how our manager can execute trades to limit their market impact and describe in detail the anatomy of a live trade that produces the execution costs we estimate. This process includes patiently setting a series of limit orders that seek to provide rather than demand liquidity, and is in part why our cost estimates are much lower than those found in the literature from trade and quote (TAQ) data, which is an aggregation of many trade types across many traders.

We examine the realized trading costs from the execution database using an implementation shortfall and price impact method (Perold (1988)). A key feature of our data is the rich heterogeneity in trade size executed by our manager, which enables us to estimate price impact as a function of trade size (fraction of daily trading volume traded in a stock). Our methodology accounts for all costs associated with a trade, which include explicit costs, such as commissions and bid/ask spreads, and, most importantly, price impact which tends to dominate the full costs of trading (Kyle (1985), Easley and O'Hara (1987), Glosten and Harris (1988), Hasbrouck (1991a, 1991b), Huberman and Stanzl (2001), Breen, Hodrick, and Korajczyk (2002), Loeb (1983), Keim and Madhavan (1995, 1997), and Knez and Ready (1996)). Since in our data (as described below) most of the orders are executed passively, the effective bid-ask spread across all trades averages less than 0.015% per year in the U.S. These costs are very small relative to the price impact costs facing a large institutional trader. Hence, cost estimates that ignore price impact, such as those based solely off of effective spreads (e.g., Novy-Marx and Velikov (2017)) are not a very accurate measure of the true costs facing a large trader in markets, which is a fact we confirm in our tests.

Armed with these cost estimates, we provide a novel description of how trading costs vary across trade type (e.g., long vs. short positions), stock characteristics, trade size, over time, and across international markets and exchanges. We further decompose trading costs into temporary and permanent components across these dimensions.

One caveat of our data is that we only have costs for one particular investor, albeit a large one, that may not generalize to other investors or other trading strategies not deployed by our manager. We argue that the trading costs we estimate are generalizable to other trades and other managers for several reasons. First, the costs estimated are exogenous to the portfolios being traded. Our data only examines the live trades of longer-term strategies, where the portfolio formation process generating the desired trade is *separate* from the trading process executing it. Trades are executed to their intended positions in a manner seeking to lower execution costs using a proprietary trading algorithm that, importantly, does not make any buy or sell decisions. The algorithm decides how patient to trade (minutes versus days), but not *what* to trade (or not to trade). In our sample, 99.9% of all intended trades are completed, with the average realized trade horizon for completion being slightly less than one day.[1] Second, to further rule out concerns over the endogeneity of portfolio weights and trades to (expected) transactions costs, we also examine trading costs using only the first trade from new inflows of long-only mandates, which typically specify adherence to a benchmark portfolio like the S&P 500 or Russell 1000. The initial trades from these new inflows are essentially exogenous to trading costs since there is little scope for deviation due to the tight tracking error mandate. We find

[1] The independence of portfolio formation from trading execution is a feature only of the longer-term models and would not apply, for instance, to intra-day or higher frequency trading strategies that are not a part of this database. The only endogenous choice in security selection our manager makes with respect to trading costs in our database is the universe of stocks covered, which does not include extremely small and illiquid microcap or penny stocks, or stocks with very limited daily trading volume.

that trading costs from these "forced" initial trades are no different than those from other trades on average, consistent with the exogeneity of trading costs to the portfolio/strategies generating the trades. Finally, we collect trading cost data from three different brokers (ITG, Deutsche Bank, and JP Morgan), and a consulting firm (ANcerno, formerly Abel-Noser), who collectively cover trades from more than 2,000 institutions across 2,000 brokers globally, and find our estimates match average costs across trade size, time, and country from these other sources.

We next examine what drives trading costs by looking at the characteristics of the market, the stock being traded, and the trade itself as determinants of costs. We first find that market structure has intuitive, but limited, explanatory power for price impact costs. Examining variation in trading rules across the 21 markets in our sample, we find that median market impact costs rise with intraday triggered auctions and decline with the number of competitive trading venues available to investors in that market, but we find no significant effect from transactions taxes or short-sale uptick rules on costs. For the largest trades, exchange rules have a miniscule effect on costs.

Looking within a market, we find that price impact costs rise with market volatility, consistent with the theoretical microstructure literature. We also show that price impact costs have exhibited a steady decline over time across markets. Hence, the documented decline in trading costs associated with bid-ask spreads (Jones (2001), Hasbrouck (2009)) is also evident for price impact costs. We also find that costs are larger for smaller stocks and stocks with greater idiosyncratic risk, consistent with theories of market maker inventory risk raising price impact.

Finally, we find that the most important variable determining price impact is the size of the trade – measured as the fraction of daily volume traded in a stock – where larger trades generate greater price impact, consistent with theory (Kyle (1985)). We find that the relationship between price impact and trade size is non-linear, with impact rising with trade size at a decreasing rate. Theories propose either concave or linear relationships between price impact and trade size (Hasbrouck (1991), Hausman, Lo, and MacKinlay (1992), Keim and Madhavan (1995), Almgren (2003), and Almgren, Thum, Hauptmann, and Li (2005)), with limited empirical verification. A formal test of this relationship using our data reveals a power-law function with a coefficient close to ½. We therefore approximate price impact with a square root function in trade size, which is consistent with models by Kyle (1985), Barra (1997), Kahn (1993), Grinold and Kahn (1999), and Loeb (1983), and matches Almgren (2003) and Almgren, Thum, Hauptmann, and Li (2005).

Using these insights, we estimate a price impact model based on 1) market conditions, 2) stock characteristics, and 3) trade size calibrated to both U.S. and international live executed trades. We compare the estimated costs from our model to other measures and models in the literature, including out of sample tests on data obtained from brokers, consultants, and live portfolios of index funds.

We first examine a set of measures of liquidity and trading costs proposed in the literature from daily or intra-daily trade and quote (TAQ) data.[2] While each of these measures is positively related to our price impact costs from executed trades in univariate tests, controlling for the variables in our model subsumes the information in these measures. Notably, trade size is a critical determinant of live trading costs that none of these other measures capture, including the TAQ price impact measure which is based off of changes in prices and traded quantities over fixed time intervals. These other measures are not able to capture the variation in trading costs in our database.

Second, we compare our price impact model to other models in the literature estimated from TAQ data. Following Korajczyk and Sadka (2004) and Sadka (2006), we use TAQ data to estimate a linear price impact function. We also estimate a square root function for trade size using the TAQ data. In both cases, the TAQ-calibrated models deliver trading cost estimates that are several times larger than our model implies or than the actual trading costs experienced by our manager. In particular, the linear cost models employed by Korajczyk and Sadka (2004), Lesmond, Schill, and Zhou (2003), and Novy-Marx and Velikov (2017) imply trading costs that are an order of magnitude larger than our model or live costs suggest. The reason our price impact function produces trading cost estimates that are substantially smaller than those in the literature is that aggregated daily TAQ data captures the *average* trader, which includes informed insiders, retail traders, liquidity demanders, and impatient traders, who face much higher costs than those of a patient trader.

Third, we use our model to predict the trading costs of live funds based on passive stock indices where we have good estimates of actual fund (and therefore trade) size and transactions costs. Examining Vanguard's S&P 500 index fund and Blackrock's iShares Russell 2000 ETF, our model predicts their costs accurately, suggesting that our cost estimates are in line with other large traders. However, the TAQ-estimated models produce cost estimates for these index funds that are several

[2] Studies attempting to estimate trading costs typically rely on theoretical models and one of two types of data: daily spread and price information or intra-day transaction level data aggregated across all traders over fixed time intervals. Examples of trading cost measures using daily price and spread data include Roll (1984), Huang and Stoll (1996), Chordia, Roll, and Subrahmanyam (2001), Amihud (2002), Acharya and Pedersen (2005), Pastor and Stambaugh (2003), Watanabe and Watanabe (2007), Fujimoto (2003), Korajczyk and Sadka (2004), Hasbrouck (2009), and Bekaert, Harvey, and Lundblad (2007). Examples using transaction level data from the trade and quote (TAQ) data from the NYSE, Rule 605 data (e.g., Goyenko, Holden, and Trzcinka (2009)), or proprietary broker data (e.g., Engle, Ferstenberg, and Russell (2008)), include Goyenko (2006), Sadka (2006), and Holden (2009). For a comparison of daily versus intra-daily measures, see Lesmond, Ogden, and Trzcinka (1999), Lesmond (2005), Engle, Ferstenberg, and Russell (2008), Lehmann (2003), Werner (2003), Hasbrouck (2009), and Goyenko, Holden, and Trzcinka (2009). Hasrbrouk (2009) and Goyenko, Holden, and Trzcinka (2009) summarize these measures.

times larger than their actual costs. Fourth, as mentioned above, our trading cost model also matches the live costs from brokers and the ANcerno database, while models based on TAQ data do not. Finally, we examine the cross-sectional distribution and time-series properties of estimates from our trading cost model and find they are more reliable than other measures proposed in the literature.

While a few studies examine proprietary trade data from institutions (e.g., Keim (1999), Keim and Madhavan (1997), Engle, Ferstenberg, and Russell (2008)), these studies use much more limited data in terms of breadth and depth, typically covering only a couple of years of U.S.-only transactions and a small cross-section of stocks. Our sample spanning 19 years of trading activity across almost 10,000 stocks in 21 different equity markets allows for a richer and more reliable trading cost model. Moreover, we estimate a price impact function with respect to trade size. Our data provides the first look at trade orders and execution on a trade-by-trade basis, allowing us to estimate a precise price impact model as a function of trade size. Our calibrated model provides more reliable trading cost estimates out of sample that subsume other measures in the literature, match the costs of live portfolios, other institutional sources of costs, and deliver less extreme and more stable estimates over time and across firms in a way other models fail to accomplish.

The out of sample performance of our trading cost model and its ability to capture real trading costs of large institutions makes it useful for assessing costs in other settings. The model is based on observable characteristics about the firm and market, with the trade size being a choice variable. Using this model, Frazzini, Israel, and Moskowitz (2017) examine how various trading strategies based on asset pricing anomalies survive trading costs at different fund sizes. The model may also be useful for other settings where a measure of trading costs for a large arbitrageur or institution is desired, such as event studies, mutual fund or hedge fund trading, and possibly for evaluating the impact of correlated trading effects (e.g., Lo (2007), Lou and Polk (2014), Alquist, Jiang, and Moskowitz (2017)). Our model also provides cost estimates globally.

The paper proceeds as follows. Section I describes the trading process that produces the execution database containing our live trades, including a detailed trading example of how the trading algorithm executes. Section II outlines our methodology for measuring trading costs from the executed trades and presents results on realized trading costs over time, across international markets and exchanges, for different trade types, and decomposed into temporary and permanent components. Section III examines what drives trading costs, using the live executed trades to estimate a price impact model. Section IV compares our estimated costs to others in the literature and practice, and compares our calibrated price impact model to other models of trading costs, including how they perform in out of sample tests. Section V concludes.

I. Live Trading Data and Trading Process

We describe our live trading data, how trades are generated and executed by detailing the anatomy of a live trade, and present summary statistics. We start with a description of the portfolio formation and trading process used by our manager.

A. Portfolio generation and trading process

Given a portfolio, our data comes from a trading process that attempts to establish positions by executing trades in the cheapest possible manner using a proprietary trading algorithm that, importantly, cannot make any buy or sell decisions. The algorithm merely determines how best to execute the desired trades, including deciding how patiently to trade (minutes versus days). In our sample, 99.9% of all intended trades are completed, but not necessarily immediately. The average realized trade horizon for completion is slightly less than one day, indicating low tracking error to the intended position. The trading algorithm has no discretion on the position itself, but can optimally decide how long to take to trade to that position. The decision to trade and the size of the trade are determined independently.

We analyze the price impact costs that arise from the executed trades of the trading algorithm, which we describe below. Understanding how portfolios and trades are created and executed is important for interpreting the trading cost measures we obtain from these data. We only examine trades where the portfolio generation process is separate from the trading process. Specifically, we examine the live trades of longer-term signals/strategies, where the set of intended trades is created from longer-term models (such as value) and/or from specific client mandates that specify adhering to a benchmark with a tight tracking error constraint. The separation of portfolio decisions from trading decisions is a feature of the longer-term model trades, but would not apply, for instance, to high frequency intra-day trading strategies. For this reason, all trades from short-term (daily or intraday) signal/models are excluded from the analysis.

For the long-term models that we focus on, a set of desired trades is generated based on the optimal holdings of a particular model, often backtested using historical data. While the models employed are often more complex, a lot of the portfolio generation process is based on factors familiar to the academic literature such as value, momentum, and quality (e.g., Asness, Moskowitz, and Pedersen (2013), Fama and French (2015), and Asness, Frazzini, and Pedersen (2017)).

Once a theoretical portfolio is determined, an optimization process generates portfolio weights subject to various constraints that include risk and other client mandated constraints (e.g., tracking error to a benchmark, shorting constraints, industry and position limits, etc.). These portfolio weights imply a set of trades that move current portfolio positions to the desired portfolio positions. The trades are then loaded into the trading algorithm which is responsible for lowering execution costs and does not make any explicit portfolio decisions. The timing of individual trades can be completed in a matter of minutes or can take several days (the maximum trade horizon is 9.8 days in our sample, though most are executed within a day and 99% are executed within three days).

B. Trade execution data

Our trading cost data are drawn from the internal trade execution database maintained by AQR Capital, a firm managing global strategies ranging from benchmark-oriented long-only strategies to absolute return long-short strategies in a variety of asset classes. The database is compiled by the execution desk and covers all trades executed algorithmically (described below) in any of the firm's funds since inception, excluding trades associated with intra-day models as discussed above.[3]

The data contains orders, execution prices, and quantities. We exclude all non-equity and emerging markets trades, restricting the sample to developed market equity transactions, which includes cash equities and equity swap transactions. For each individual order we collect the trade-order-set identifier, stock identifier, order size, trade horizon, order type, country, portfolio type, reference benchmark, ex-ante predicted beta (described below), model-implied price, P^{theory}, arrival price, P^{start}, execution time, execution price, P^{ex}, and execution quantity, Q^{ex}.

The trade-order-set has a unique identifier that allows the mapping of individual trades to corresponding portfolios. A trade-order-set is simply a basket of trades submitted together. In a given day, the same stock can appear in multiple order sets. Therefore, we use a trade-order-set stock grouping as our cross-sectional unit of observation for each date. Hereafter, we use the term "order" to indicate a trade-order-set, stock, and submission date triplet.

The order size is the desired number of shares to trade when orders are submitted. For our data, the desired trade amount will almost always equal the sum of the actual execution quantities (desired quantities are filled 99.9% of the time). In the very rare cases where the desired amount of shares is not completely filled, it is typically due to trading suspensions in a security or other rare events. While the quantity being traded is not a choice variable for the trading algorithm, the duration (or

[3] AQR does not do any high-frequency trading, but does have a small number of "statistical arbitrage" portfolios based on short-term price pressure. The trades from these models are excluded in the analysis since they are necessarily and by design endogenous to trading costs – for example, trades for these strategies are only executed when costs are expected to be low. We focus exclusively on trades from long-term models, where the trading algorithm is not allowed to make portfolio decisions.

patience) of the trade is. The trade horizon is the target trade duration (in days) at the time of order submission, but since the actual total trade duration is endogenous and depends on realized fill rates, it may, and often does, differ from the target duration. Our estimate of trading costs, therefore, includes how our manager executes trades, which is designed to minimize costs of a given desired quantity, where the trade quantity is fixed. This feature of the execution data means that we can estimate trading costs for a given quantity traded, but where the patience of the trade execution is an endogenous characteristic that we consider part of our measure of trading costs.

As a result, our trading cost estimates will vary across different trade sizes (quantities), where "trading patience" is deemed part of the costs we measure. We do not provide measures of trading costs by trade size holding trade time fixed (e.g., estimating the cost of immediate execution across trade size). Because our trading cost estimates are independent from the portfolio weights associated with the trade, we can apply these costs to similar-sized trades generated from other portfolios, under the assumption that the patience of trading at each trade size is approximately the same and is embedded in our cost estimate. Assuming other large traders would/could execute trades similarly with the same patience, these costs can then be generalized to other large traders. We provide evidence in the paper supporting the generalizability of our cost estimates.

The order type indicator identifies the transaction as "buy-long," "buy-to-cover," "sell-long," and "sell-short." The portfolio type indicators classify each portfolio as "large" or "small" in terms of market capitalization (based on their relevant benchmark, such as the Russell 1000 or 2000 in U.S. equities, for example) and "long/short" or "long-only" based on the portfolio's mandate.[4] The benchmark identifies the relevant reference benchmark, while the ex-ante predicted beta is a forecasted beta for each stock to the benchmark at the time of order submission, which is based on a regression using past daily stock returns on the corresponding benchmark[5]. The model price is defined as the stock price at portfolio formation. For portfolios constructed during market hours this is equal to the current price. For portfolios constructed off market hours this is the latest closing price. The arrival price is defined as the current price at the time the first order is submitted to the market, which is recorded by the trading algorithm when orders are submitted. The execution prices are time-stamped and incorporate commissions.

Trades are executed using proprietary, automated trading algorithms designed and built by the manager. The trading algorithm breaks orders into smaller pieces and dynamically submits and

[4] We classify relaxed constraint portfolios, such as 130-30 or 140-40, as long-only. These represent a very small fraction of trades and excluding them or reclassifying them does not alter any of our results.
[5] The estimation of the ex-ante betas follows Frazzini and Pedersen (2015).

cancels a range of limit orders such that there are many executions per total order. To illustrate how the algorithm works, we first define the following terms:

1) **Trade basket**: *trade order set* (collection of orders submitted together).

2) **Parent order**: each trade basket is composed of a series of "parent" orders, which contain instructions for a given stock (e.g., buy to cover 10,000 shares of IBM). There can be multiple parent orders per trade basket for a given stock, especially if the underlying position is changing (e.g., buy to cover 10,000 shares of IBM and buy long 20,000 shares of IBM).

3) **Child order**: each parent order is split into smaller "child" orders that are submitted as limit orders, possibly on multiple exchanges. A fraction of these orders are filled as prices hit the limit prices, and the rest are cancelled.

4) **Executions:** distinct executions of the limit orders at the limit (execution) price.

The trading algorithm directly and anonymously accesses market liquidity through electronic exchanges and, in order to minimize market impact, tries to provide rather than demand liquidity by not demanding immediacy using a system of limit orders (with prices generally set to buy at the bid or below and sell at the ask or above) that dynamically break up total orders (parent orders) into smaller orders (child orders), where both the sizes of child orders and the time in which they are sent are randomized.[6]

C. *Example of a trade*

Figure 1 illustrates an example of an actual trade made by the trading algorithm. The trade is a live execution of a buy order for Microsoft common equity shares, executed on May 8, 2014 from a long-only account. The trade is for 1.4% of Microsoft's one-year rolling average daily trading volume (hereafter referred to as "DTV") and it was executed over a time interval of 6 minutes and 3 seconds, from 9:40:01 am to 9:46:04 am.

The first graph of Figure 1 plots the limit orders submitted by the trading algorithm at various prices, the best bid offered across all trading venues throughout the trading interval, and the executed buying price of the limit orders, all time-stamped to the millisecond. The graph illustrates how the algorithm works by submitting a series of limit buy orders below the best bid offer in the hopes that the bid price falls to the limit order price. This is the idea behind "passive execution," where the trading algorithm submits buy orders at or below the current bid. The algorithm will submit limit

[6] There are additional features the trading algorithm employs to handle corporate announcements, special events, etc., that are beyond the scope of this paper. Results are robust to excluding such events from the data.

orders at or below the bid, and then, as bid prices move or the algorithm's desire to be passive or aggressive changes, the algorithm will cancel stale orders whose limit prices are "too far" from the ideal limit price and reset new orders closer to the new limit price, which will be a function of how passive the algorithm would like to be and the bid-ask spread of the stock. The reason for the continual cancelling and resetting of orders is twofold. On the one hand, the algorithm seeks to execute the trade quickly to minimize implementation shortfall by putting new limit orders closer to the current bid. On the other hand, the reason stale orders are cancelled is to limit exposure in the case of a major news event or shock suddenly moving prices significantly. Since the U.S. is also a fragmented market containing multiple trade venues, orders are placed across venues to seek the best price and execution.

The first graph in Figure 1 illustrates the trading algorithm's execution for the Microsoft buy order over the six minutes and three seconds trading interval. The graph shows the movement in the best bid price over the interval, as well as the new limit order prices set by the trading algorithm and the points at which the limit orders were executed as the market bid price crosses the limit price. The dashed line indicates how long the limit orders are left on the trading venue until a new limit order is placed and the old one is cancelled, as the best bid prices change over time. For this particular trade, following the initial limit order price of 39.25, 13 different additional limit prices were set as the best bid prices changed, and 10 of those limit orders were executed at the bid prices that crossed those limit prices. For example, after the initial 39.25 limit order, placed at 9:40:01 am when the national best bid offer (NBBO) was 39.26, bid prices on Microsoft rose. By 9:40:32 the best bid offer rose to 39.29, causing the algorithm to cancel the old limit order (at a price limit of 39.25) and place a new limit order at a price of 39.27. The NBBO continued to rise and by 9:41:12 the old limit order at 39.27 was cancelled and a new one was set at a limit order price of 39.30. At that point, the best bid offer crossed the 39.30 price and the buy order was executed. From there, the best bid dropped to 39.29 and a new limit order was placed and executed at that price at 9:41:25. This process continued until the entire size of the order completed.

Aggregating all of these trades, the trade-weighted sum of these executions at the executed prices determines the weighted average executed price of the trade, which when compared to the initial price right before the beginning of trading (39.25), measures the total price impact of the trade, which in this case is 4.3 bps. In the next section, we detail our price impact cost calculation.

The second graph of Figure 1 plots the best bid and ask prices at the time of order submission, the prices at which the limit orders are executed, and the best bid and ask prices at the time of order execution. The blue diamonds represent the best ask prices and the red squares the best bid offers at each point in time, with the green triangles representing the execution prices. The first set of bid and ask prices are those at the time of order *submission* and the second set of bid and ask prices highlighted in a shadow adjacently are those at the time of *execution*. As the graph shows, the algorithm executed at prices consistently at or below the best bid price offered at the time of order submission. Execution below the best bid offer at the time of submission occurs because the bid and ask prices at the time of execution might move down between the submission time and execution time. For this particular stock on this day, the algorithm consistently sets price limit buy orders slightly at or below the bid at the time of submission and then patiently waits until the next bid offer crosses that limit price. If the price is not hit quickly enough or the bid price rises (or falls), then the algorithm sets a new limit order at the higher (or lower) price, but again at or below the new bid price. In each case the algorithm executed at the best bid offer which was consistently either at or below the initial bid offer at the time of submission. For example, for the first executed order at 9:41:24 am, the limit order is executed at the best bid offer at the time of submission because the NBBO did not move between the time of submission and the time of execution. However, for the order executed at 9:42:09 am, the limit order was set and executed below the best bid price at the time of submission because the national bid and ask moved slightly lower in the few milliseconds between the time of submission and the actual execution of the order.[7]

The same trading process is executed for sells using limit orders to sell at or above the current national best ask price, where patient limit orders are placed just above the current ask price and the algorithm waits until changes in the best ask price reaches the limit order price, or the algorithm posts new limit orders at new prices and cancels old ones as the posted best ask price changes.

While the trading algorithm is designed to patiently trade by posting passive limit orders at or below the current bid price in the case of a buy, or at or above the current ask price for a sell, some trades may demand more immediacy (when the opportunity cost of waiting to trade is very high.) In this case, limit orders to buy (sell) may be posted more aggressively at the current bid (ask) or even

[7] Two instances on the graph seem to show that the algorithm executed below the best bid at the time of execution, as well. In both cases, the bid offer does not appear to move between the time of submission and execution, yet the executed price appears to be below the bid. However, this apparent discrepancy is due to small variation in the millisecond time stamps between the TAQ data, where we draw the NBBO numbers to illustrate this example, and our execution data, which is drawn from direct data feeds from the exchanges. Different computer latency between the two datasets can create discrepancies (of a few milliseconds) between the NBBO computed ex-post on TAQ and prices recorded in our data. For very liquid stocks such as Microsoft, when plotting executions against TAQ data this could generate the impression of executing below the bid when the bid just moved, but the time stamps from the two data sources were not perfectly aligned. While it is theoretically possible to execute below the best bid offered by getting crossed internally by brokers or getting hit by a dark venue, in general the algorithm will not be able to execute at prices below the NBBO at the time of execution.

slightly above the bid (below the ask) in order to get more immediate execution, or in some cases will demand liquidity by buying at the ask or selling at the bid. These "aggressive" trades occur about 15% of the time on average in our sample (Table I).

The costs our manager faces are a function of these execution choices, which are designed to lower the costs of trading a given fixed quantity of shares. We believe these features are very similar to the way most large institutions trade (or could trade), and therefore are not unique to our data, but rather represent the costs facing any large institutional trader; a fact we confirm using outside data from brokers, institutional trading organizations, and other live trades from other managers. In addition, we argue that the costs can be generalized to trades outside of our execution database for several reasons. First, the trading costs we estimate are independent of the strategy generating the trades. The trading algorithm takes the desired portfolio weights as given, and the algorithm executes trades in the same manner given the inputs to the algorithm, regardless of the model or strategy generating portfolio weights that determine these trades. The reason the algorithm trades in this manner is because we are only looking at the trades associated with long-term strategies (such as value) where the returns/alpha of the strategy are not determined by the speed of execution, and hence the opportunity cost of patiently trading is low. Second, the trading algorithm controls trade execution, not asset selection or position size. Desired positions are fully completed 99.9% of the time and usually within a day. Third, many client mandates that form the basis for desired portfolios force the portfolio to trade to a benchmark (e.g., S&P 500) with low tracking error, leaving little scope for deviation. To rule out any remaining concerns over the endogeneity of portfolio weights to trading costs, we exploit these mandates by estimating trading costs using only their first trade from new inflows. The trading costs from these inflexible mandates are identical to those from all other trades, supporting the notion that the trade executions we use to estimate costs are independent of the portfolio generation process, and hence that these trading cost estimates are generalizable.

Our estimates are not, however, likely representative of the costs facing an individual investor or even the average investor in the market, a conjecture we provide some evidence for in Section IV. Rather, the cost estimates represent those of an average sophisticated institutional trader, who serves as an arbitrageur in markets, and therefore more closely resembles the costs of the marginal investor.

D. Summary statistics

Table 1 presents summary statistics of our trade execution data. The live execution database contains 11,044,700 parent orders, 4,368,100,000 child orders, and 691,600,000 executions across 9,543 stocks globally between August 1998 and June 2016, totaling US $1,701,390,000,000 in trades.

Panel A of Table I reports the amounts traded (in US $billions) each year from 1998 to 2016 in the U.S. equity market, across the 20 international markets, for large cap and small cap stocks separately, and for long-short and long-only portfolios separately. The amounts traded have grown substantially over time from $1.29 billion traded in the U.S. in 1998 to $167.39 billion in 2015 (the last full year in our data since the 2016 data ends in June with about $135 billion traded). Equities traded in international markets have grown from $1.67 to $95.87 billion. Large cap portfolios (Russell 1000 or MSCI universe) account for $1.65 trillion out of the total $1.7 trillion traded over the sample period. Nevertheless, $52.3 billion worth of trades were placed among small cap stocks. Finally, $1,246.11 billion, or roughly 73%, of trades occur for long-short strategies.

Panel B of Table I reports the market coverage of the execution database across seven regions: US, Europe, Japan, UK, Canada, Australia, and the Pacific. We obtain the universe of stocks from the union of the CRSP tapes and the Compustat North America and Global database from September 1998 to June 2016. We assign individual stocks to the corresponding market based on the location of the primary exchange. For international companies with securities traded in multiple markets, we use the primary trading vehicle identified by Compustat. The execution database covers 9,543 stocks, which represent 71.5% of the market capitalization of the 21 markets on average.

The average (median) size of firms traded is $15.4 ($4.9) billion. More than half of the stocks traded are on US exchanges (4,984 stocks), with another 1,252 in Europe (Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, Israel, Italy, Netherlands, Norway, Portugal and Sweden), 1,302 in Japan, 556 in the UK, 415 in Canada, 354 in Australia, and 680 in the Pacific (Hong Kong and Singapore).[8] Table A1 in the appendix lists coverage statistics by country. The database covers a maximum of 80.8% of the market capitalization in the UK and a minimum of 59.2% in the Pacific, with more than half of the countries having at least 75% of their market cap traded in our database. Table A2 in the appendix reports the fraction of market capitalization covered by our trading database by region and year. Coverage has obviously grown over time, reaching 81% in 2016 on average across all countries.

Panel C of Table I reports time-series means, medians, standard deviations, and ranges of the number of stocks traded per year, number of countries traded, and number of exchanges traded on. A maximum of 5,041 stocks in 21 markets across 32 exchanges are traded in a given year.

Also reported are summary statistics on the trades themselves. The number of trade baskets averages 1,352.7 per year for an average trade size of $607,200, and a fraction of daily trading

[8] We arbitrarily classify Israel as "Europe" for the regional summary statistics.

volume (DTV) of 0.9%. There is wide variation in trade size, ranging from less than 0.1% to 13.1% of DTV. We use this variation in trade size to estimate a price impact function, where trade sizes are pre-determined and exogenous to the trading process. The trading algorithm, however, has discretion over the duration of the trade. The average ex-ante expected trade horizon is 2.7 days, with the median trade taking place over 1.7 days and the maximum taking 9.8 days.

The remaining rows in Panel C report summary statistics on the trading algorithm and executions. On average there are 135 stocks per trade basket, with 154 parent orders per trade basket or about 581,000 parent orders per year. The parent orders are split into 223 million child orders, of which 36.4 million are executed on average per year. The ratio of orders filled to cancelled is 0.13 on average and about 15% of orders are executed "aggressively" (submitted buys at the ask or sells at the bid). The average bid-ask spread at the time of order arrival is 21.33 bps for our sample of stocks, which is consistent with spreads quoted in other studies from TAQ data (Korajczyk and Sadka (2004), Hasbrouck (2009), and Novy-Marx and Velikov (2017)). However, as we show, it is rare for our trades to incur the full or even half of the spread because of the passive limit orders. The main cost our trades face is price impact.

II. Realized Trading Costs

We describe our methodology for calculating trading costs, and present summary statistics of trading costs over time, for different types of stocks, across markets, and for various types of trades.

A. Measuring transactions costs: price impact and implementation shortfall

To measure transactions costs, we use an implementation shortfall (*IS*) methodology as defined in Perold (1988). Implementation shortfall measures the difference between a theoretical or benchmark price (e.g., a model price) and an actual traded price, scaled by the amount traded. We also define market impact (*MI*) by looking at the difference between an arrival price (i.e., the price that exists when a trade begins in the market) and an actual traded price, scaled by the amount traded. The difference between these two measures represents pre-trade moves that might occur from the time a model portfolio is generated and the time trading actually begins. The following equations illustrate the two costs:

$$IS = ret_{p,theory} - ret_{p,actual} \tag{1}$$

$$= ret_{p,theory} - \frac{cost_{execution}}{ptheory} - \frac{cost_{opportunity}}{ptheory}$$

where the opportunity cost is the return times the number of shares not traded, which in our case is essentially zero. The cost of execution is measured as

$$\text{cost}_{execution} = Q_{+}^{ex}(P^{ex} - P^{theory}) + Q_{-}^{ex}(-P^{ex} + P^{theory}) \tag{2}$$

with P^{ex} and P^{theory} representing the execution price and the theoretical price (the price used by the model prior to the trade being initiated) and Q_{+}^{ex} and Q_{-}^{ex} representing the quantity of shares bought and sold, respectively. The largest part of the cost of execution for a large institutional trader is market impact costs, as defined below:

$$MI = Q_{+}^{ex}(P^{ex} - P^{start}) + Q_{-}^{ex}(-P^{ex} + P^{start}) \tag{3}$$

where P^{start} is the price at the time trading starts. Equations (2) and (3) take into account bid-ask spread, market impact, and commissions. Commissions and effective bid-ask spreads are small by comparison to price impact at large trade sizes, since they do not scale with size.

Both *IS* and *MI* are effective measures of the costs of trading. Implementation shortfall measures the total amount of slippage a strategy might experience from its theoretical returns, including both the cost of trading and the cost of not trading (opportunity cost). For the trades we examine (coming from AQR's longer-term models), we primarily focus on the cost of trading as the trades coming from the long-term signals are completed 99.9% of the time, hence little information is provided about the cost of not trading. The theoretical price is defined as the closing price at the time the strategy's desired holdings and trades are generated, which is typically the prior day's closing price. The start price is similarly defined as the first recorded price at the time the strategy begins trading to the new desired holdings (e.g., current day's opening price). We express these costs in terms of returns (relative to the traded price) rather than dollars.

Since the time of order submission is a choice variable (i.e., at the open versus 10am) and since the trades are generated from the long-term models that do not form opinions about overnight price moves (and nearly always get fully executed), any difference between *IS* and *MI* should be random with an expected value of zero.[9] Empirically, we find little difference between *IS* and *MI*.

There are many ways to think about transactions costs depending upon the objective. Our notion of trading costs computes the difference between the results of a theoretical portfolio which has zero transactions costs and the results of a practical portfolio which attempts to track the theoretical portfolio but is subject to actual traded prices. Effectively, our cost computation captures how much of the theoretical returns to a strategy are actually achieved in practice. This measure seems to be the

[9] This of course may not, and likely would not, be true for high frequency or statistical arbitrage trades that are excluded from our analysis.

most relevant for addressing market efficiency questions like: what is the drag on performance of a particular strategy from trading costs? What are the net returns to a strategy? Does an anomaly survive trading costs and at what fund size?

We can decompose our measure of price impact further into transitory and permanent components. If general demand for a security has increased, due to either information or perhaps long-term sentiment, then the price may permanently move to reflect that information. Conversely, prices can also be impacted temporarily by, for instance, liquidity demanding trades or uninformed trades. In the former case, the price will remain at the new level while in the latter case any temporary price impact should reverse back to its normal level once trading stops. As a simple measure of the transitory component of trading costs, we calculate how much of the price impact reverses over the following trading day. The part that remains the following trading day is then the permanent component. More formally, we define the temporary and permanent components of *MI* from equation (3) as follows:

$$MI^{temp} = [Q_{+}^{ex,t}(P^{ex,t} - P^{t+1}) + Q_{-}^{ex.t}(-P^{ex,t} + P^{t+1})] \tag{4}$$

$$MI^{perm} = MI - MI^{temp} \tag{5}$$

where $P^{ex,t}$ is the price of the weighted average executed trade on day t and P^{t+1} is the closing price of the stock on the next trading day following execution.

There are other ways to measure transactions costs, such as to compare actual traded prices to other possible traded prices that existed during the trading period, like the volume weighted average price (VWAP) or the price that would have been achieved using market orders over the same period. We can define market impact costs relative to the VWAP as follows,

$$MI^{relative} = Q_{+}^{ex}(P^{ex} - VWAP\ ) + Q_{-}^{ex}(-P^{ex} + VWAP\ ). \tag{6}$$

This measure captures the effectiveness of a trade relative to other traders in the market at the same time. However, this measure does not help us understand what the costs of trading are for an investor nor the net of cost returns to a strategy. We report all of these measures for comparison.

B. Trading costs

Figure 2 shows the event-time evolution of the average price impact for U.S. trades in our database. The figure plots the price movement for the average trade from our live trading data over the 24 hours following the trade. We restrict the sample to U.S. trades (buys and sells are signed appropriately to aggregate the trades and total price impact) with a trading horizon of less than or equal to one day and compute average price impact during trading throughout the day and the average overnight and intraday return over the subsequent 24 hours. We estimate the average price impact during the day split equally across 30 minute trading windows. For overnight and next day returns, we split the event-time window into two-hour windows for simplicity. Since trades are broken up into smaller child orders, we estimate market impact as the integral of the function plotted in Figure 2 over the trade horizon – from the beginning of trading at the first traded price with the first child order, to the last child order filled for that trade (when the trade horizon exceeds one day we compute market impact over multiple days). From Figure 2, our estimate of market impact is just under 9 basis points on average for all trades completed within a day. Most (around 85%) of that price impact appears to be permanent, as only 1.26 basis points on average are reversed over the next 24 hours. While some may consider the 1.26 basis points as the price impact of our manager because it is the cost "caused" by the manager's trading, and hence reverses once the manager stops trading, this is an incomplete picture of the cost facing an investor. Even if the permanent component of price impact is beyond the manager's control, it is still a cost borne by the investor. For questions regarding the efficiency and after-cost returns to a strategy, the entire price impact (permanent + temporary) is the relevant metric, as it is the total cost the strategy incurs from implementation.

Table II reports the full sample mean, median, and value-weighted mean (weighted by dollar value of trades) of the market impact (*MI*) and implementation shortfall (*IS*) estimates we get from the live trading data, following equations (1) through (3). Specifically, the cross-sectional mean, median, and weighted mean are computed each month and the time-series average of these monthly measures are reported, where each month is weighted by the number of stocks traded in that month. Standard errors of the monthly estimates are reported in the style of Fama and MacBeth (1973) at the bottom half of each panel. Panel A reports results for the full sample period from August 1998 to June 2016. The first column reports the summary statistics for all trades, where the mean market impact measure is 9.97 basis points and the mean implementation shortfall measure is 11.02 basis points. The 1.05 basis points difference between the two represents the difference between the intended model prices and the actual prices at the start of trading, which, as expected, is small. The median *MI* and *IS* costs are quite a bit lower at 6.18 and 8.63 basis points, respectively, suggesting that trading costs are positively skewed by more expensive trades. Weighting trades by their dollar value, the value-weighted means are quite a bit higher at 15.14 basis points for market impact and 16.06 basis points for implementation shortfall, which indicates that the largest trades are the most expensive trades, consistent with trading cost models that argue that trading costs increase with trade size – a fact we confirm more formally in the data. The standard errors for these estimates are small and typically a bit less than one basis point, making the difference between the *MI* and *IS* measures within two standard errors of zero in most cases.

By exchange. The next three columns of Panel A of Table II report the same statistics for NYSE-AMEX and NASDAQ traded stocks as well as for all stocks traded on international exchanges, separately. Whether measured by market impact or implementation shortfall, the average costs of trading on the NYSE appear smaller than on NASDAQ or in international markets, which is intuitive since the NYSE on average contains larger and more liquid firms.

By size. The fifth and sixth columns of Panel A of Table II report results for large and small cap stocks separately. The average large cap stock trade generates 8.90 basis points of market impact costs compared to 18.95 basis points for small cap stocks. Of course, this does not control for the size of the trade, which we will address in the next section.

By portfolio type. The last two columns of Panel A report trading costs for trades made within long-short portfolios and long-only portfolios separately. The trading costs for long-only portfolios are larger than those made in long-short portfolios. The average long-only trade faces nearly 13.7 basis points of market impact costs compared to only 8.4 basis points for the average long-short trade. However, on a value-weighted basis, long-short portfolios face about the same costs.

Panel B of Table II reports the estimates over the most recent decade from 2006 to 2016. Since the time-series averages reported in Table II weight each month by the number of stocks traded, more recent periods are given more weight, so the 2006 to 2016 costs look very similar to the full period.

Finally, Panel C of Table II reports trading costs for different types of trades: buy long, buy to cover, sell long, and sell short. The first two columns report the fraction of trades, in both dollar terms and numbers, of each trade type. Buying long and selling long account for two-thirds of all trades, with the remaining one third split evenly between short selling and covered buys. Over the full sample, buying long generates about 12.5 basis points of price impact, but buying to cover has 15.5 basis points of price impact. Short selling is slightly more expensive by 0.62 basis points on average than selling long, but the difference is not statistically significant.[10] Although a large literature discusses the additional costs associated with short-selling, conditional on actually shorting, we see no marked difference in trading costs between selling a long position versus selling short. If short selling is indeed costlier, it is likely to be from opportunity cost (i.e., not being able to short) or from lending fees for stocks on special. The execution costs of short selling versus selling long

[10] Shorting costs and shorting revenues from lending fees are not included in these costs, which just capture price impact. The majority (around 99%) of short positions are in stocks that are not hard to borrow and hence the general collateral rate applies, which is small on average.

appear no different, even with uptick rules and other barriers to shorting. Results for international exchanges are similar. Figure 3 summarizes these results graphically. Table A3 in the Appendix also reports pooled means of the trading cost measures, rather than Fama and MacBeth averages. The estimates of trading costs are similar.

C. *Other trading cost measures*

Permanent and temporary costs. Table III reports the value-weighted (by dollars traded) average market impact decomposed into temporary and permanent components, following equations (4) and (5). Panel A reports averages over the full sample and Panel B from 2006 to 2016. Of the 15.14 basis points of total price impact, only 2.08 basis points are reversed the next day, which represent the temporary component of these costs. This decomposition is similar across exchanges, trade size, and portfolio type, where about 85-90% of the price impact is permanent across NYSE-AMEX, Nasdaq, international exchanges, large and small cap stock trades, and for long-short and long-only portfolios. Results are similar over the more recent decade of trading data in Panel B as well.

Relative cost measures. Table IV reports market impact costs relative to VWAP following equation (6), for our sample of U.S. trades. Once again, Panel A reports the full sample results and Panel B the results over the most recent decade of data. The trading costs when measured versus the VWAP are much lower than those in Table II that measure costs relative to the theoretical price. Instead of the 9.97 basis point average price impact from Table II, the average price impact measured relative to the VWAP is only 4.50 basis points, indicating that our manager paid 4.50 basis points more on average relative to the VWAP during the trading horizon. The median price impact relative to VWAP is 2.62 bps. These numbers are close to our estimate of the temporary price impact cost we measure with our live trade data. Intuitively, the additional amount our manager pays relative to the VWAP or best price achievable in markets at the time should be related to how much additional price pressure our manager is exerting. Since the price impact from permanent demand should be reflected in rising VWAPs, any price impact above the VWAP may be related to temporary price pressure.

For the same reasons we use the total cost (permanent + temporary price impact), we believe the theoretical price impact numbers are the most relevant for answering questions such as what are the implementation costs of a large investor? While the theoretical price impact measure delivers much larger cost estimates than the temporary price impact or the impact relative to VWAP, and therefore are conservative relative to these other measures, these are the actual costs an investor faces.

D. International trading costs

Figure 4 plots market impact costs by country. There is some heterogeneity in costs across markets, with Australia, Hong Kong, Japan, and Singapore having higher costs, and Sweden, Germany, and France having costs closer to those in the U.S.

Differences in trading costs across countries could arise from different market structure features across countries, such as execution mechanisms, regulation, and trading rules governing participation in these markets. In addition, these average costs do not account for trade size, which may differ across countries or other attributes of the stocks traded that may differ. We investigate in the next section what drives trading costs across time, stocks, and markets.

E. Exogenous trades from inflows

Finally, we argue that the trading costs we estimate are essentially exogenous to the portfolio weights and set of securities being traded. To show some direct evidence consistent with that claim, we estimate trading costs using only the first trade from new inflows coming from long-only mandates that adhere to a benchmark with a tight tracking error constraint (1% to 2% annualized). The initial trades from new inflows provide little scope for deviation or optimization and hence offer an estimate of trading costs exogenous to portfolio construction choices.

Table V reports the mean, median, and value-weighted mean market impact measures for long-only trades coming from new inflows only and reports the same summary statistics for all *other* trades for comparison. The estimates across the two samples are very similar and show no systematic differences in market impact costs. The last column performs a formal statistical test on the differences in trading costs that cannot reject that the two measures are the same. Cutting the trades by small and large cap stocks yields the same conclusions. All formal tests that trading costs differ across the two samples fail to find any statistical significance, with *t*-statistics that are substantially below one and oscillate in sign.

The evidence that new inflow trades, with little scope for discretion, yield similar cost estimates, indicates that our trading cost estimates are exogenous to the portfolios generating the trades, consistent with our earlier claim that use of the long-term models separates execution from portfolio formation. This evidence implies that our trading cost estimates can be applied more broadly to other portfolios or settings and are not uniquely tied to the models used to generate them in our data. In the next section we investigate what drives trading costs and build a trading cost model to use out of sample for other portfolios, other trade sizes, and even other stocks and time periods.

III. What Drives Trading Costs?

We investigate what drives trading costs across our sample of live executed trades. The $1.7 trillion of executed trades across nearly 10,000 stocks in 21 countries over an almost 19-year period provides a unique laboratory to test many proposed theories of transactions costs. Our data offer extremely rich heterogeneity across a multitude of characteristics, which we group into: market structure, market environment, stock-level, and trade-level characteristics, to examine what drives trading costs. Using these insights, we then build an empirical model for expected trading costs based on observable characteristics and compare them to other trading cost measures in the next section.

A. Market structure

We start by assessing how market structure affects trading costs. Our data provides trading cost estimates across 21 equity markets covering 32 different exchanges, each with its own regulations, rules, and execution mechanisms.

We characterize the different markets across a number of trading structure features, including short sale uptick rules; transaction taxes; fragmentation, which represents whether securities are traded on multiple venues within a country; number of "LIT" venues, where a LIT venue displays bid and ask prices at which market participants are willing to trade (the opposite of a "dark" pool venue); order protection rules, which are provisions that investors receive an execution price equivalent to what is being quoted on any other exchange at the same time within the country where the security is traded, possibly including OTC markets; an indicator for whether OTC trading is available; the use of daily auctions and intraday triggered auctions (e.g., after a trading halt or "circuit breaker"); and handling of separate odd lots, which specifies how odd lot shares – residual shares from the standardized minimum lot size for the exchange – are traded, which can be handled on different venues from the exchange, including OTC venues (e.g., for 100 share lots, a trade of 115 shares would have a 15 odd lot trade executed on a separate venue in a country that separates odd lot handling). Table A4 in the appendix details each of these market structure features for each of the 21 equity markets covered by our trading sample.

Table VI reports regression results for the average, median, and value (dollar trade)-weighted average of market impact costs in each country on dummies for various market features. The first column of Table VI reports univariate regression results on each variable for the median market impact in each country. Intuitively, median costs rise with separate odd lot handling, intraday triggered auctions, and short sale uptick rules, though the latter is statistically indistinguishable from zero, as these features place constraints on trading activity. Costs decline with OTC trading, number of LIT venues, and fragmentation, because these features help provide more competitive quotes. The second column reports multivariate regression results that indicate only intraday triggered auctions and number of LIT venues have a significant impact on trading costs. Exchanges that employ intraday triggered auctions typically experience 1.89 basis points more of price impact, whereas the number of LIT venues decreases price impact by 1.48 bps per dollar traded.

The third and fourth columns of Table VI repeat the regressions using the mean price impact in each country, which is skewed toward larger costs. The same variables remain significant with the same sign, although their coefficient estimates are larger due to mean price impact being larger than the median. Again, intraday triggered auctions raise average trading costs by about 3 bps, whereas OTC trading and number of other venues reduce average costs by more than 4 bps. Finally, the last two columns of Table VI report regression results using the value-weighted (by trade size) mean price impact in each country as the dependent variable. Here, nothing is significant, though both intraday triggered auctions and OTC trading are close to significant and maintain their respective signs. The weaker results and general lack of significance for the value-weighted mean indicates that for the largest trades, where price impact is of greatest concern, these features have little import.

Of course, it is important to keep in mind that the trading algorithm is designed to endogenously adjust to mitigate the influence of these market features. Hence, the small or weak effects from these market structure variables may not apply to simple market orders or liquidity-demanding trades. In fact, most of the market structure dynamics will not have a material impact on trading costs because the majority of the nuances across exchanges will be handled by order routing logic. For example, consider the lack of significance of the transaction tax for all specifications (including univariate regressions). It would seem that a transaction tax, like the one imposed in the U.K. on foreign investors, which is 50 bps for all buys plus £1 for buys exceeding £10,000, would affect the trading costs experienced in this market. However, these trades are typically executed through swaps with local counterparties, which are not subject to these taxes.

However, there are other features of the market that are not easily handled by order routing, but where the trading algorithm adjusts to mitigate the impact on costs. It is interesting, therefore, to consider how the trading algorithm responds when facing these frictions. One notable example is when an uptick rule is in effect, which prevents shorting in the stock if the price is declining. The trading algorithm when facing an uptick rule will take the rule into account and not issue a short if it does not satisfy the rule. This constraint could in theory impose higher shorting costs in these markets. While the point estimates in Table VI for the short sale uptick rule are consistently positive, consistent with this conjecture, the estimates are not reliably different from zero, suggesting the impact of the restriction is negligible. Likewise, the trading algorithm is less likely to participate in daily auctions, where the patient trading of limit orders may be less effective, and hence is not affected by the daily auction mechanism across markets. However, intraday auctions that are triggered by unforeseen events, such as large price drops, are harder to avoid and hence increase trading costs.

In general, these market features do not have a material effect on price impact costs because the trading mechanism is able to mitigate their influence. As further testament to that conclusion, Table VI shows that for the value-weighted averages that reflect the largest trades, where price impact is of greatest concern and therefore where the trading algorithm has the greatest incentive to reduce these costs, the effect of these market features is uniformly insignificant. Hence, to a first-order approximation, the effect of these market features on the trading costs of a large institutional trader using a patient trading algorithm is negligible and not the primary source of costs. We investigate now what else might be driving variation in these costs across trades.

B. Trading size and volume

Most theories of trading costs model price impact as a function of trade size (Kyle (1985)). While proportional trading cost models, where the cost of trading does not vary with the size of the portfolio traded, such as the effective spread, are often used to model trading costs for smaller trades or small investors, these costs do not match theory or data (or practice) when considering larger portfolios and trades. Non-proportional trading cost models, where the cost of trading varies with the size of the portfolio traded, are more realistic for larger traders, because their trades can move prices. We investigate below how well measures of proportional and non-proportional costs capture the live trading costs from our data.

The shape of the price impact function with respect to trade size is an important ingredient in modeling costs. Theories have proposed either concave or linear relationships between price impact and trade size (Hasbrouck (1991), Hausman, Lo, and MacKinlay (1992), Keim and Madhavan (1995), Almgren (2003), and Almgren, Thum, Hauptmann, and Li (2005)), but empirical estimates of this relationship are limited. Our unique 19-year sample of nearly 700 million executions and $1.7 trillion worth of trades provides a rich laboratory to investigate this relationship.

Figure 5 plots the average price impact measure from our live trading data as a function of trade size, measured by the fraction of daily trading volume (DTV) traded. The plot shows a concave relation between market impact and trade size, with costs growing significantly at very large sizes (as a percentage of DTV), but at a decreasing rate. A fitted curve (superimposed on the graph) seems to capture price impact better than a linear function, something we test formally below.

Models by Barra (1997), Kahn (1993), Grinold and Kahn (1999), and Loeb (1983) use a square root process to model the relationship between price impact and trade size, and Almgren (2003) and Almgren, Thum, Hauptmann, and Li (2005) find that a square root function does a good job describing their trading cost data from brokers. Consistent with these studies, we find that a square root function also describes our trading cost data well. Figure A1 in the appendix shows a log-log plot of market impact on trade size, where the regression coefficient of log market impact on log trade size is 0.35, which is close to the square root coefficient, and has an R-square of 95%. Given several theories and other studies advocate a square root function, and attempting to avoid overfitting, we model price impact as a square root function of trading size (%DTV), rather than a power law of 0.35. Results in the paper are not sensitive to this assumption, but what is important is that the price impact function is concave and not linear in %DTV.

C. Market and stock characteristics

In addition to trade size, we also examine how market and stock characteristics influence costs. For example, theory (Kyle (1985)) predicts that volatility increases costs, as market makers face more inventory risk. Costs will also rise with asymmetric information and the depth of the market.

To test these ideas, we estimate an econometric model for describing trading costs. We regress costs from each trade in our database on three sets of characteristics designed to capture the 1) size of a trade, 2) the underlying stock being traded, and 3) the market environment at the time of trading. For trade size we include the fraction of daily trading volume traded, measured as the dollar size of the trade divided by that stock's average daily trading volume over the past six months, and the square root of the fraction of daily volume traded.

For firm characteristics that may be related to the cost of trading, we include the size of the firm, which is the log of one plus the market value of equity in million USD [log(1+ME)], and the idiosyncratic volatility of the firm's equity return, which is the standard deviation of the residuals from a regression of one-year daily stock returns of the firm on the corresponding market index for the country in which that stock trades (annualized percentage). The size of a firm and its volatility may be related to the amount of information asymmetry of the firm's value and/or the depth of the firm's shares for trading, which many studies claim (see Hasbrouck (2009) for a review) are associated with higher trading costs. Idiosyncratic volatility may also capture market maker risk.

Finally, we also include a set of market-wide variables to capture variation in trading costs over time that result from different market conditions. We include the "VIX" to capture market volatility, which is the monthly variance of each country's value-weighted index computed using daily returns (expressed as an annualized percentage). We also include a linear time trend to capture changes in aggregate trading costs over time due to technology, changes in tick size, etc.[11] We also include the variable "Beta*IndexRet*buysell", which is the beta of the stock times the index return for that market times a variable "buysell" that equals 1 for buys and -1 for sells. "Beta" is the stock's predicted beta at the time of order submission (based on the coefficient estimate from a rolling regression of the stock's daily returns on the corresponding market's returns for the country in which the stock trades over the last year) and "IndexRet" is the market index return over the life of the trade. This variable is used to account for contemporaneous (beta-adjusted) market returns over the trade period and to sign the trades correctly, since our data contains both buys and sells. This variable is only used to describe historical transactions costs. It plays no role in estimating future expected transactions costs that we apply out of sample in the next section, since we impose the restriction that daily market returns are unforecastable (i.e., E[IndexRet] = 0). Thus, when estimating expected transaction costs, we simply drop this term while holding the other parameters fixed.

D. Trading cost model calibrated from live trades

We estimate a model containing all of the variables discussed in the prior section. Table VII presents results from estimating this regression separately for the full sample of trades, U.S. trades only, and international trades only, using pooled regressions with country fixed effects. The dependent variable is the realized market impact from the live trading data, in basis points.

The results are economically intuitive and match theory. The first column reports that trading costs have declined over time, as indicated by the negative coefficient on the time trend (Appendix Table A5 shows that this decline is largely driven by technological events like moving to decimalization in traded prices). In addition, larger stocks face lower price impact costs, which is also intuitive as the depth of trading in those stocks is typically much greater than for small stocks, consistent with Kyle (1985) and a large empirical literature (summarized by Hasbrouck (2009)). Column (2) adds trade size (fraction of daily volume traded) to the regression, which shows that larger trades have higher transactions costs, which is also consistent with price impact and trading

[11] Table A5 in the appendix looks at other time-varying market-wide variables, including dummies for events like decimalization, and finds that the estimates on the other variables are not affected by various choices for time-varying market conditions.

cost models such as Kyle (1985). Column (3) adds the square root term for trade size, which is highly significant in the regression, consistent with the results in Section III.B. An *F*-test of whether the square root specification of trade size is preferred to a linear specification easily rejects the linear model in favor of the square root model, consistent with our earlier results. The concavity of price impact costs with respect to trade size is supported by theory (Hasbrouck (1991), Hausman, Lo, and MacKinlay (1992), Keim and Madhavan (1996)) and other empirical evidence from institutional trading costs (Almgren (2003), and Almgren, Thum, Hauptmann, and Li (2005)). The fourth column adds the stock's idiosyncratic volatility and the level of market volatility (VIX) to the regression. More volatile firms have higher transactions costs, consistent with models of market maker inventory risk, and more volatile market environments are also associated with larger price impact costs, consistent with market makers needing to be compensated more in more volatile markets.

The *R*-squares of the regressions are just over 10%, which is quite large for these types of models given the noise in trade data. The dependent variable here, our market impact measure from equation (3), is essentially a daily (or more frequent) return, and given the substantial noise in high frequency returns, a 10% *R*-square is an accomplishment.

As an additional test, we also add a variable which is the return on the stock being traded minus the return on a similar stock with similar characteristics not traded in our data at that time (times the "buysell" indicator to match the trade sign). The matched stock is based on the procedure of Daniel, Grinblatt, Titman, and Wermers (1997) that matches stocks along three characteristics: size, BE/ME, and momentum (past 12-month return). The idea is to test whether a stock being traded in our data on a given day experiences different price movements than a similar stock that was not traded that day. As column (5) shows, the average trade experiences an additional 4 bps of market impact on the stocks traded relative to stocks of similar characteristics not traded by the algorithm that day. This number may represent our manager's demand for immediacy and its temporary price impact. Consistent with that intuition, the estimated coefficient is very similar to our previous estimate of price impact relative to the VWAP and its reversal the next day.

The coefficient estimates are all very similar across markets, with U.S. and international data giving similar parameterizations. The fact that the patterns and estimates are similar across 21 different equity markets is reassuring and indicates the model and findings are robust. The data confirm the implications of several trading cost theories that costs decline with firm size, rise with volatility (both idiosyncratic and market), and are a concave function of trade size.

E. Expected trading costs

While the trading cost model in Table VII provides a description of realized trading costs at the actual sizes traded, we can use this calibrated model for price impact to calculate *expected* trading costs for any stock and any portfolio size. The only unobservable variables from the model are the contemporaneous market or matched-characteristic benchmark (DGTW) return, which are naturally dropped from the expected trading cost model because at a daily frequency their expectation is zero. The other variables – time trend, market cap, daily *volume*, idiosyncratic volatility, and VIX – are all observable *ex ante* and the only remaining variable is the desired amount to be traded, or by extension, the size of the portfolio to which the trade belongs, which is a choice variable.

The model can be used to estimate the expected trading cost of a portfolio out of sample, assuming the relationship between the observable variables and trading costs is stable. Specifically, we calibrate our market impact model using the estimated coefficients from Table VII (using the parameters in column (5) for a global sample and in column (10) for the U.S.) as follows:

$$MI_t = \theta_0 + \theta_2\, r_t^{mkt} buysell + \theta_2 time\ trend + \theta_3 \log(me_{t-1}) + \theta_4 x_{t-1} + \theta_5 sign(x_{t-1})\sqrt{|x_{t-1}|} + \theta_6 \sigma_{t-1}^{IV} + \theta_7 VIX_{t-1} + \theta_8\, r_t^{dgtw} buysell + \varepsilon_t \quad (7)$$

where $x = 100 \times m/dtv$ is signed dollar volume (m) as a fraction (in %) of the stock's average past one-year dollar volume (dtv). Note that the right-hand side variables are known quantities at the beginning of period t with the exception of market r_t^{mkt} and DGTW r_t^{dgtw}returns which are measured over the life of the trade. To compute a measure of expected price impact as a function of the amount traded, we set expected market and DGTW matched returns to zero $E(r_t^{mkt}) = E(r^{dgtw}) = 0$ and evaluate the other right-hand side variables at their sample medians:

$$MI = a + b\, x + c\, sign(x)\sqrt{|x|} \quad (8)$$

where $a = \theta_0 + \theta_2 \overline{time\ trend} + \theta_3 \overline{\log(me)} + \theta_6 \overline{\sigma^{IV}} + \theta_7 \overline{VIX}$, where $\bar{z}$ denotes the sample median of variable z, $b = \theta_4$ and $c = \theta_5$. The constant a controls for the general level of trading costs as well for the components of trading costs that do not depend on trade size such as commissions (which are captured in our execution prices).

Figure 6 plots the expected price impact function across trade sizes (fraction of daily volume) using the model and parameter estimates in column (5) of Table VII, and where the other right-hand-side variables are evaluated at their full sample medians. Although this model is estimated over the entire sample period from 1998 to 2016, the time trend, firm size, idiosyncratic volatility, and VIX account for market conditions that may affect trading costs through time. As a test of the reasonableness of the model with respect to time variation, we examine the market impact function over four sub periods: 1998-2005, 2006-2016, 2007-2009, and 2010-2016, which allows for estimates of costs before, during, and after the financial crisis. The bottom plot in Figure 6 shows the price impact functions over the sub periods. Costs are higher earlier in the sample, then decline significantly, but jump up during the financial crisis from 2007-2009, and then recede again. This is consistent with anecdotal evidence that price impact, in addition to spreads, was greater during the global financial crisis. The patterns in international markets mirror those of the U.S.[12]

In the next section we compare the trading cost estimates from our model to a host of other measures and models proposed in the literature, including a variety of estimates from other data sources that include data from TAQ, brokers, and several institutional trading cost providers.

IV. Comparison To Other Trading Cost Models and Measures

We compare our trading cost estimates to others proposed in the literature from other data.

A. Comparison to other cost and liquidity measures

We first examine how other measures of trading costs and liquidity compare with our price impact measure. Table VIII repeats the regressions in Table VII that regress the price impact from each live trade on various cost and liquidity measures from the literature.

We examine the modified measure of Roll (1984), the illiquidity measure of Amihud (2002), the proportion of trading days with zero returns (examined by Hasbrouck (2009) and Goyenko, Holden, and Trzcinka (2009)), the effective bid-ask spread from the TAQ database, and an estimate of price impact, known as lambda from the Kyle (1985) model, which is estimated from signed trades from the TAQ database. We describe the TAQ estimates below, which are examined by Korajczyk and Sadka (2004), Sadka (2006), Hasbrouck (2009), and Goyenko, Holden, and Trzcinka (2009).[13] The calculations of these measures follow Hasbrouck (2009) and are detailed in Table A7 of the appendix. Since these other cost and liquidity measures have only been examined in the U.S., Panel A of Table VIII reports results from U.S. equity trades for comparison. We also compute these measures in the other 20 markets in our sample and report those results in Panel B. Since TAQ data only cover U.S. securities, we cannot compute the TAQ effective spread and lambda internationally.

[12] Table A6 in the Appendix re-estimates the trading cost model over the four sub periods for the full sample, U.S., and international trades only and reports all of the coefficient estimates.

[13] Korajczyk and Sadka (2004) use both measures to assess the trading costs of momentum strategies, Lesmond, Schill, and Zhou (2003) use the effective spread to measure trading costs of momentum strategies, and Novy-Marx and Velikov (2017) use the effective spread to measure the trading costs of a host of trading strategies.

The first column of Panel A of Table VIII regresses the price impact from live trades on the modified Roll measure, which is designed to capture the effective spread for a stock. We also include the "Beta*IndexRet*buysell," and time trend regressors to account for the direction of the trade, contemporaneous market movements, and time effects in trading costs. As the first column reports, price impact from our live trades is positively related to the modified Roll measure across stocks, with a coefficient of 0.03 and a *t*-statistic of 3.01. Higher effective spreads as measured by the modified Roll variable are associated with slightly higher price impact costs. However, the explanatory power of the Roll measure is small, as the adjusted *R*-square *after* accounting for the beta and time trend is only 0.01%. The second column of Table VIII shows that the coefficient on the modified Roll measure becomes insignificant (0.01 with a *t*-statistic of 1.39) once we control for the other stock and trade characteristics from Table VII. Specifically, both the idiosyncratic volatility of the stock and the fraction of daily volume traded seem to drive out the explanatory power of the modified Roll measure. These variables capture a lot more of the variation in price impact costs as the adjusted *R*-square (after accounting for beta and trend) jumps from 0.01% to 4.1%.

Columns (3) and (4) repeat the regressions using the Amihud (2002) illiquidity measure in place of the modified Roll measure. Amihud's measure does not explain much of the variation in price impact costs from our live trades. Columns (5) and (6) report results for the proportion of zero return trading days, which again exhibits no significant ability to explain our price impact costs. Columns (7) and (8) examine the TAQ effective spread, which is positively related to our price impact costs by itself, but is completely captured (changes sign and is insignificant) by the other model variables when we include them in the regression. The effective spread captures only 0.02% of the variation in price impact costs, while the other model variables capture 4.1%. Columns (9) and (10) show similar results for the TAQ lambda measure of price impact, which has the strongest relation to our actual price impact costs, with a *t*-statistic of 4.33, capturing 0.08% of the variation in our costs. But, adding our other model variables that include trade size subsumes the explanatory power of the TAQ lambda and captures substantially more variation in real costs, as the *R*-square rises to 4.1%.

Finally, columns (11) and (12) examine all of the other cost and liquidity variables simultaneously. Only the TAQ lambda appears to be related to our live trading cost estimates, which makes sense since this variable is supposed to capture price impact estimated from the TAQ data. However, including other variables from the model of Table VII, particularly idiosyncratic volatility and trade size, renders the TAQ lambda far less significant (its coefficient drops by 2/3) and captures a much larger fraction of actual price impact costs.

Panel B of Table VIII repeats the regressions for international markets. A similar picture emerges where the liquidity and cost measures from the literature are positively related to the price impact from live trades, but where our model of trading costs that includes trade size largely subsumes the explanatory power of these other measures. For the international markets, the measure of Amihud (2002) does the best job explaining price impact costs, but still only captures 0.12% of the variation in costs. Adding our variables from the model reduces the coefficient on Amihud's measure from 0.21 to 0.06 (though it is still significant) and captures 6.13% of the variation in live trading costs, even after market movements and time trends have been accounted for.

B. Comparison to other trading cost models

We next compare our price impact model, estimated from live trades, to other models estimated from TAQ data. To compare our market impact estimates to estimates obtained from TAQ data we calibrate the aggregate market impact model specified in equations (7) and (8), using the U.S. estimated coefficients of Table VII column (10), since TAQ data is only available in the U.S. We then compare this model to estimates from models used in the literature that are calibrated to TAQ data. Specifically, we examine the linear market impact model of Breen, Hodrick, and Korajczyk (2002) (used by Korajczyk and Sadka (2004)) and the square root price impact model of Kyle (1985).

We detail below how we calibrate each of these models with TAQ data, which essentially follows equation (8) but substitutes the functional form for each model's price impact function in lieu of coefficients b and c, and estimates those coefficients with TAQ data as described below. The constant a in equation (8), however, which captures the general level of trading costs as well as components of trading costs that do not depend on trade size such as commissions, is held constant across the various models. This is done for several reasons. First, neither the linear market impact model of Breen, Hodrick, and Korajczyk (2002) nor the square root impact model of Kyle (1985) includes a constant term. Second, prices on TAQ do not contain commissions, whereas our execution prices do. Third, we wish to focus on the price impact cost differences. Therefore, to isolate comparisons across the price impact costs coming from the models, we compare estimates over the same universe and time period and keep the parameter a constant across all models, focusing on the marginal cost of trade size only from each model.

Linear model: Breen, Hodrick and Korajczyk (2002). The price impact function of Breen, Hodrick and Korajczyk (2002) (and used by Korajczyk and Sadka (2004)) implies a linear relation between market impact and dollars traded over a given time interval. To calibrate this model, for each stock i and for each trading day t, we estimate market impact γ as:

$$10{,}000 \times \Delta \, logP_\tau = \gamma \, m_\tau + \varepsilon_\tau \qquad \tau \in t$$

where $10{,}000 \times \Delta \, logP_\tau$ is the log price change (in basis points) over time interval τ (5 minutes) and, m is signed dollar volume, defined as dollar volume signed by the order imbalance measure of Lee and Ready (1991) and aggregated over the same trading interval. We use the WRDS TAQ Consolidated Trades (WCT) and NBBO's Bid and Ask midpoints matched to each trade relative to their trade time. Because we do not know whether the trades are buy or sell initiated, the trades are signed based on the quote test (first) or tick test (second) following Lee and Ready (1991). We drop days with less than 10 intra-day return observations and allow for an intercept in the regression.[14] The coefficient gamma γ represents the market impact (in basis points) of m dollars traded. We express γ more conveniently as a fraction of daily volume (in %) to match the units in our model:

$$10{,}000 \times \Delta \, logP_\tau = b \, x + \varepsilon_\tau \qquad \tau \in t$$

where: $x = 100 \times m/dtv$ is signed dollar volume as a fraction (in %) of the stock's average past one-year dollar volume (dtv) and

$$b = \gamma \, dtv/100.$$

The coefficient b represents the market impact (in basis point) of x% of daily volume traded. To reduce noise, for each stock i we average the estimated b over the prior one year and require a minimum of 200 trading days. Our final sample includes a panel of stock i at day t estimated impact coefficients b_{it} covering 9,909 securities between October 1993 and June 2016.[15]

To compare TAQ market impact functions to our estimated market impact from our live trading database, we first restrict the sample to stocks in our execution data and then aggregate TAQ market impact estimates over the same sample period from August 1998 to June 2016. Similar to Table II, we compute monthly cross-sectional means (weighted by total dollars traded) and compute the time-series average of these monthly estimates. We then plug b into equation (8) and set $c = 0$ for this model (and where a remains the same as in our model), and compare this model's price impact estimates over the same universe and time period to our model's estimates.

Square root model: Kyle (1985). The price impact function of the Kyle (1985) model implies a linear relation between market impact and the square root of the dollars traded over a time interval.

For each stock i and for each trading day t we estimate market impact λ as:

$$10{,}000 \times \Delta \, logP_\tau = \lambda \, sign(m_\tau)\sqrt{|m_\tau|} + \varepsilon_\tau \qquad \tau \in t$$

[14] Not allowing for an intercept in the regression yields identical results.
[15] TAQ data starts in January 1993 and we require a minimum of 200 trading days.

where $\Delta\, logP_\tau$ is the log price change (in basis point) over time interval τ (5 minutes), and m is signed dollar volume, defined as dollar volume signed by the order imbalance measure of Lee and Ready (1991) and aggregated over the same trading interval, using the same TAQ data and allow for an intercept in the regression. The coefficient lambda λ represents the market impact (in basis points) of m dollars traded. We can express λ more conveniently as a fraction of daily volume (in %):

$$10{,}000\ \times\ \Delta\, logP_\tau = c\ sign(x_\tau)\sqrt{|x_\tau|} + \varepsilon_\tau \qquad \tau \in t$$

where: $x = 100 \times m/dtv$ is signed dollar volume as a fraction (in %) of the stock's average past one-year dollar volume (dtv) and

$$c = \lambda\sqrt{dtv/100}.$$

The coefficient c represents the market impact (in basis points) of x% of daily volume traded. To reduce noise, for each stock i we average the estimated c over the prior one year and require a minimum of 200 trading days.

Again, we restrict the sample of stocks to those in our trading cost database and aggregate TAQ market impact estimates over the same sample period. We then plug the time-series average of the monthly value-weighted cross-sectional mean of c into equation (8) and set $b = 0$ for this model (and where a remains the same as in our model), and compare this model's price impact estimates over the same universe and time period.

Linear plus square root model. To be consistent with our transaction cost model in Table VII and equation (8), we also estimate a specification allowing both a linear and square root component of price impact:

$$10{,}000\ \times\ \Delta\, logP_\tau\ = a + b\, x_\tau + c\ sign(x_\tau)\sqrt{|x_\tau|} + \varepsilon_\tau \qquad \tau \in t$$

where coefficients are estimated in the same manner as the TAQ linear and square root models, and the coefficients b and c are plugged into equation (8) (keeping a constant).

Comparing all four price impact models. Figure 7 plots our calibrated price impact function from Table VII (column (5)) for various trade sizes (%DTV), versus estimates from the linear, square root, and linear plus square root models using TAQ data at the same trade size. Since, the range of trading size our trade execution data provides is limited to 12% of DTV, we limit the range of trade sizes in Figure 7 to 12%, where we are most confident in the precision of our estimates.

A comparison of the price impact functions shows how the TAQ data leads to different costs versus our live trading data. First, the linear price impact model of Breen, Hodrick, and Korajczyk (2002) and used by Korajczyk and Sadka (2004) produces substantially larger trading costs than the models that use a square root term, especially at higher trade sizes. At 2% of DTV, the linear TAQ model implies costs of 44.89 bps, which is more than three times higher than the 13.73 bps from our calibrated model from live trades. At 10% of DTV, the linear TAQ model implies trading costs of 223.31 bps, which is almost an order of magnitude larger than the costs we estimate (or the actual costs we find) from trades of that size, which are 32.34 bps. The costs estimated from the linear TAQ model are massively higher, even though we hold fixed the same universe of stocks and time period and hold fixed the constant term in equation (8), related to costs associated with market conditions and firm characteristics. The only difference in the two models is the price impact as a function of trade size, where both the linear model and use of TAQ data grossly exaggerate these costs.

The TAQ models that employ a square root function deliver more reasonable cost estimates, though they are still quite a bit higher than our estimates based off of live transactions. The second graph in Figure 7 removes the linear model estimates for better scaling in order to compare the cost estimates from the TAQ models that use a square root term. The square root price impact function produces lower cost estimates, especially at large trade sizes. Still, the costs estimated from TAQ even under this specification are larger than those we get from our model or live trading.

The linear plus square root model uses the same functional form as our model, except that the parameters on % DTV coefficients are calibrated by TAQ instead of our execution database. As Figure 7 shows, the TAQ-estimated price impact functions still produce higher trading cost estimates than our model (and the actual trading costs experienced by our manager). At 2% of DTV, the square root TAQ model estimates 27.09 bps of price impact and the linear plus square root model estimates 29.07 bps, compared to our model of 13.73 bps. At 10% of DTV, our model implies 32.34 bps of trading costs, while the TAQ-calibrated models imply 60.23 and 64.21 bps, respectively.

The square root and linear plus square root models calibrated with TAQ data seem to produce price impact costs roughly twice those we get from our model. This result is intuitive given that aggregated TAQ data approximates the *average* trade, including informed insiders, retail traders, liquidity demanding trades, and impatient trades, all of which may face significantly higher costs than those of a large institution who patiently trades.

The question remains, however, which cost estimates are a better reflection of the real-world costs facing a generic large institutional trader? On the one hand, our data comes from a large institutional trader who executes trades patiently as opposed to TAQ data that is an average of all trades. This suggests our cost estimates may more closely resemble the real-world costs facing a large trader. On the other hand, our data comes from a single institutional manager. So, if our cost estimates are unique to our specific manager (e.g., if our manager simply experiences lower trading costs than the average trader), then they may not be a good estimate of costs more generally. To answer this question, we conduct several out of sample tests of our model and the TAQ models to: 1) explain the costs of live portfolios from other institutional managers, 2) match the costs reported by brokers and other institutional cost providers, and 3) examine the distribution of costs implied by each model across stocks and over time.

C. Out of sample test on live portfolios

We argue and show that the costs we estimate are much closer to real-world trading costs facing a large trader and match those from other sources, whereas the cost estimates from TAQ consistently and grossly overestimate these costs. As an out of sample test, we apply the trading cost models to live portfolios where we can obtain reasonable and independent information on trading costs. Specifically, we examine passive funds that closely track an index to see if the models can produce expected trading costs that approximate the actual costs these funds incur.

Table IX examines the expected trading costs of the S&P 500 and the Russell 2000 indices, which are benchmarks many fund managers are evaluated against and some funds track passively. A nice feature of both indices is that we can obtain fairly reliable estimates of actual costs from live funds benchmarked to each index, such as ETFs designed to track the index with low tracking error. To measure actual transactions costs from ETFs, we examine the difference in returns between the ETF and the underlying index, where (given the low tracking error) deviations in returns between the fund and the index should be due to transactions costs and stock loan revenue, which the index itself does not incur. (Also bearing in mind that these funds may face additional turnover from inflows and outflows, which the index does not.) Table A8 in the appendix details the calculations, where we assume the expected return to tracking error is zero, and hence treat the difference in returns between the fund and the index as the sum of trading costs and expenses minus any revenue such as lending fees from lending shares.[16] We collect data for two ETFs: The Vanguard 500 Admiral Index Fund tracking the S&P 500 (ticker: VFIAX) and the iShares Russell 2000 ETF tracking the Russell 2000 (ticker: IWM). From their most recent annual report we obtain the expense ratio of each fund, as well their reported annualized turnover, which allows us to calculate a transaction cost per year. To estimate the average return deviation from the benchmark we obtain monthly ETF and underlying index returns from Morningstar and estimate lending fee revenue per dollar of assets under

[16] The assumption that the expected return to tracking error is zero seems valid for passive index funds who merely try to match the index, but is almost certainly not true for active funds. Patton and Weller (2018) essentially make this assumption and try to back out implementation costs from active fund returns relative to their exposure to factor benchmarks such as the Fama and French (1993) factors. However, their estimates implicitly lump trading costs, lending fees, expenses, and, most importantly, any active performance differences together and call the sum "transactions costs."

management from Blocher and Whaley (2015), who calculate lending fees for all U.S. ETFs. Appendix Table A8 details the calculations, which produce an annual cost of 4.72 bps for the S&P 500 and 12.87 bps for the Russell 2000.

We then compute expected trading costs for each of these index funds using our and the TAQ-calibrated trading cost models. To compute the implied trading costs from these calibrated models for the S&P 500 and Russell 2000 indices, we use the following procedure. For a given fund size (NAV) we compute the average dollar volume per stock at each rebalance as a fraction of the stock's daily volume:

$$x = m/dtv$$

where $m = \left(\frac{turn \cdot NAV}{n}\right)/h$, $turn$ is the fund's annualized turnover, $turn \cdot NAV$ is the fund total trading, n is the number of stocks in the portfolio and h is the number of rebalances per year. We assume monthly rebalancing ($h = 12$) . We compute dtv as the portfolio weighted average daily volume of each stock:

$$dtv = \sum_i w_i v_i$$

where v_i is stock i's one-year median daily volume. For a given x, total costs are computed as

$$TC = turn \cdot MI(x)$$

where $MI(\cdot)$ is the calibrated market impact function from whichever model and data is chosen.

Panel A of Table IX reports the calibration parameters for the models. The S&P 500 index is a buy and hold cap-weighted index, but incurs 4% turnover per year due to new additions, deletions, net issuance activity, and corporate events such as mergers. The Russell 2000 has annual turnover of 17%. However, we also need an estimate of aggregate fund size, which determines the trade size for each stock when the fund trades. As a measure of aggregate fund size, we use the current dollar amount benchmarked to the S&P 500, which according to the Standard & Poor's and Dow Jones website was roughly $7.8 trillion in 2016 and the amount of capital estimated to be benchmarked to the Russell 2000 index (obtained from Russell), $856 billion. Based on these numbers, the fraction of daily volume traded for each index fund is calculated and input into the various trading cost models.

Panel B of Table IX compares the estimated trading costs under each calibrated model to the measures of actual trading costs from the ETFs. The first row reports the actual costs, which are 4.72 bps and 12.87 bps per annum, respectively for the S&P 500 and Russell 2000 as computed above. The second row reports the estimated trading costs of both funds under our model calibrated with the trade execution data ("trade data, FIM"). Our estimate of annual costs on the S&P 500 is 4.81 bps –

almost exactly those we obtain from Vanguard. For the Russell 2000, we estimate 12.36 bps, which also matches the iShares ETF. The third row reports costs from the linear TAQ model (used by Korajczyk and Sadka (2004)). The annual estimated trading costs from this model for the S&P 500 are 29.79 basis points or about 6 times larger than the actual costs, and for the Russell 2000 are 155.79 bps, which are more than 12 times greater than actual costs. The fourth row reports the estimated costs from the square root TAQ model, which produces 9.07 bps and 20.70 bps for the S&P 500 and Russell 2000, respectively. These numbers are about twice as large as actual costs. Finally, the last row reports cost estimates from the linear plus square root TAQ model, which also generates twice the actual costs. Overall, our trading cost model delivers cost estimates very close to actual fund costs from other managers, while the TAQ models used in the literature consistently overestimate these costs.

D. Comparison to other institutional trading cost providers

In addition to comparing our cost estimates to real portfolios from other managers (such as Vanguard and Russell), we also compare them to trading cost estimates provided by brokers and consulting firms who monitor institutional execution costs. We collect trading cost data from three different brokerage houses (ITG, Deutsche Bank, and JP Morgan), and a large consulting company, ANcerno, who collectively cover trades from more than 1,000 institutions. We compare our cost estimates to those from these other sources across trade size and time.

Panel A of Table X reports average realized price impact costs across different trade sizes (% DTV) from three different brokers: ITG, Deutsche Bank (DB), and JP Morgan (JPM). Six different levels of trade size are reported: 0.25-0.5% DTV, 0.5-1% DTV, 1-1.5% DTV, 1.5-2% DTV, 2-5% DTV, and 5-10% DTV. The common sample period for these trading costs is 2008 to 2011, where we could obtain data from all three brokers.

The last three columns of Panel A report estimated costs at the various trade sizes from our model calibrated from the execution database, as well as estimated trading costs at the same trade sizes from the linear, square root, and linear plus square root models calibrated using TAQ data. Our model produces trading cost estimates very close to the average costs reported by the three brokers, across the various trade sizes. However, the TAQ-based models produce cost estimates that are consistently higher, and are two to more than five times higher at larger trade sizes, especially for the linear TAQ model. Since the linear TAQ model has been a favorite of many trading cost studies (Lesmond, Schill, and Zhou (2003), Korajczyk and Sadka (2004), Novy-Marx and Velikov (2017)) that analyze the net of cost returns and break-even capacity of various trading strategies related to capital market anomalies, our results indicate that these studies seriously overestimate what live trading costs would be. Frazzini, Israel, and Moskowitz (2017) use our calibrated model to assess a host of trading strategies and find that trading costs are an order of magnitude smaller and break-even capacities an order of magnitude larger than suggested by prior studies.

Panel B of Table X reports a comparison of average realized trading costs over time, measured from our execution database, versus average realized costs from ANcerno (formerly Abel Noser), a consulting firm that monitors institutional investor execution costs and whose database contains more than 104 million executed trades from 750 institutions. We obtain cost estimates from ANcerno from two studies. The first is Anand, Irvine, Puckett, and Venkataraman (2012), who examine average realized costs for the period 1999 to 2008, and separately from 1999 to 2006 and 2007 to 2008 for U.S. stocks. The average trade size as a percentage of daily trading volume (%DTV) in their study is reported as 2.4%, hence we report average market impact costs from our execution database for comparison for trades of 2.4% of DTV. The first column reports the average price impact cost from our execution database over the common 1999 to 2008 time period as 15.4 bps per year. The average cost from ANcerno over the same time period and same trade size is 24.1 bps. From 2007 to 2008 we obtain an 18.2 bps cost estimate, and ANcerno provides an estimate of 24.5 bps. Over the 1999 to 2006 sample period, our data average 28.7 bps of price impact, while ANcerno averages 24.0 bps over the same period. Our market impact estimates seem to be in line with those of other large institutions. The second study using the ANcerno data is Di Maggio, Franzoni, Kermani, and Sommavilla (2017), who report price impact and trading fees for the top 30 brokers in the ANcerno database from 1999 to 2014. The average trade size in their sample is 0.5% of DTV, so we report average market impact costs from our execution database for comparison at this trade size. Over this longer period and at a different trade size, costs from the ANcerno database match our estimated costs (10.5 bps from ANcerno versus 7.1 bps from our database), suggesting our cost estimates are a reasonable proxy for the costs borne by large institutional traders generally.

While our average costs match the average costs from other sources of institutional trades, our data provide a unique look at trade-by-trade orders and executions that allow us to measure a precise price impact function. The broker data and data from ANcerno provide aggregated data across many trades and traders that make identifying a price impact function more difficult.

E. Distribution of trading costs implied by various models

As a final set of tests of our trading cost model, we examine the distribution of our implied costs out of sample to other stocks not traded in our database and in other time periods outside of our sample.

We then compare our model's estimates of trading costs to other models and measures of trading costs from the literature on these same stocks and time periods. This exercise serves a dual purpose as well. Not only does it provide an evaluation of competing trading cost models, but it also serves to highlight the generalizability and usefulness of our model to predict trading costs in other settings.

The advantages of using actual trading costs are clear. The disadvantage is that we only have these costs for the sample period covered by our data and for the stocks and trade sizes it contains. While our data covers a large cross section of securities (9,543 stocks internationally) over a long sample period (1998 to 2016), it is interesting to see how our model predicts trading costs out of sample to other stocks and other time periods.

Another interesting question is whether different models give rise to different distributions of trading costs across stocks and over time and whether those distributions are stable. A desirable property of a useful trading cost model should be to provide reliable trading cost estimates that vary in a reasonable way, with the goal that researchers can use the model in other settings without having to rely on micro-level trade execution data.

E.1 FIM model from trade data

We begin with our price impact model calibrated with live execution data. Using the estimated coefficients from the model for the U.S. only (column (10) of Table VII), we estimate trading costs for all common stocks on the CRSP database each month from September 1984 to September 2016. The model is therefore applied out of sample to *all* CRSP stocks, even those not in our execution database, and out of sample in the 14 years that pre-date our execution data sample period. The start date of 1984 is chosen to coincide with the international data so we can look at U.S. and international estimates of trading costs over a common period. All variables from our model are easily available at the stock or market level (e.g., market cap, idiosyncratic risk, VIX, or more precisely an estimate of market volatility when the VIX is not available), but since our model also requires a trade size, we choose two levels of trade size for all stocks for comparison: 1% of DTV, corresponding to the average trade size in our database, and 5% of DTV to examine estimated costs at larger sizes.

Figure 8 plots the time-series variation in estimated trading costs from our model monthly from 1984 to 2016. The period pre-dating our live execution sample (January 1984 to July 1998) is shaded to show how the model predicts costs in the out of sample period prior to the existence of our data. We plot the median cost from the model each month at both 1% and 5% of DTV, but also plot the 10^{th} and 90^{th} percentiles of costs across stocks each month (at each trade size) to provide a sense of the cross-sectional distribution of estimated costs and how those evolve over time. The blue lines represent the costs at 1% of DTV. The solid lines are the median costs over time and the dotted lines plot the 10th and 90th percentiles across stocks over time.

As Figure 8 shows, estimated trading costs vary intuitively over time – increasing following the market crash in 1987, rising steadily during the fall of technology stocks in the late 1990's and early 2000's. Our model predicts these changes through rising market and idiosyncratic volatilities during this time. These patterns are exhibited not just for the median stock, but also for the top and bottom 10th percentiles of the distribution of trading costs. For the 90% most expensive stocks, costs steadily rise leading up to the financial crisis and then soar during it.

The red lines plot the time-variation for the median and extreme 10th percentiles of stocks at a trade size of 5% of DTV. The time-series patterns are similar, but where higher trade sizes shift the level of costs upward. Interestingly, during volatile times like the global financial crisis, the 90th percentile of stocks with the highest price impact at 1% of DTV far exceed the median price impact of stocks at 5% of DTV. However, the lowest 10th percentile of stocks at 5% of DTV still face higher costs than the median stock at 1% of DTV. This suggests that trade size is a vital determinant of price impact for most of the distribution of stocks, but that there are some stocks that have large estimated price impact even at lower trade sizes.

E.2 TAQ models

Figure 9 plots the same time-series of the effective spread and Kyle's (1985) lambda estimated from TAQ data commonly used in the literature (Korajczyk and Sadka (2004), Novy-Marx and Velikov (2017)). The graphs show that the estimates from the TAQ data are highly variable over time and that the cross-sectional variation in trading costs based on the TAQ data is extremely wide. For example, the 90th percentile of effective spreads is highly volatile and many times larger than the median spread and an order of magnitude larger than the 10th percentile of spreads. In addition, time-series variation in the median, 90th, and 10th percentiles are not very correlated, though intuitively we would expect some degree of cointegration due to aggregate market conditions. This evidence is consistent with the effective spread from TAQ data being noisy and hence a poor measure of costs. Likewise, the price impact coefficient lambda from TAQ data is very noisy. The second graph in Figure 9 shows that the 90th percentile of lambdas across firms is orders of magnitude larger than the 10th percentile and many times larger than the median, with low time-series correlation to each other. This evidence is consistent with the TAQ data producing very noisy estimates of trading costs.

Table XI reports summary statistics of the distribution of trading costs from our model versus the TAQ effective spread and price impact measure lambda. The table reports the time-series average and standard deviation of the 10th, median, and 90th percentiles of each variable. Since our model and the TAQ lambda capture price impact, which depends on trade size, we report results for a trade size of 1% of DTV in each stock.

The 10th percentile has estimated trading costs of 8.68 bps on average according to our model, with a time-series standard deviation of 2.3 bps. The TAQ effective spread produces an average 6.56 bps for the 10th percentile of firms with a standard deviation of 5.4 bps. At the median, our model generates 14.55 bps (standard deviation of 3.2 bps), while the TAQ effective spread produces 19.74 bps with a standard deviation of 15.7 bps. At the 90th percentile, the differences are very large. Our model indicates an average cost of 38.1 bps with a standard deviation of 6.5 bps, whereas the TAQ effective spread is 78.9 bps at the 90th percentile, with a standard deviation of 39.9. The extreme values for the TAQ lambda and their volatility through time are also apparent in Table XI. These numbers indicate that the TAQ data not only produce higher estimates of trading costs, but also that they generate noisy estimates of trading costs that vary significantly through time, and results in extreme cross-sectional and time-series variation in estimated costs that seems unreasonably large.[17] Finally, the last row of Table XI reports the time-series correlations between our price impact model and the two TAQ measures at the 10th, median, and 90th percentiles. The correlation is positive but small at the 10th and 50th percentiles for the effective spread, but is negative at the 90th percentile. For the TAQ lambda, the 10th percentile and median are positively correlated to our model's 10th percentile and median cost estimate, but the 90th percentiles close to zero correlation, consistent with the TAQ data providing less reliable estimates of trading costs for the most extreme stocks.

E.3 Correlations across trading cost measures

Table XII reports time-series correlations for median trading cost estimates from various measures of liquidity and trading costs that include our price impact model, Amihud (2002), modified Roll (1984), proportion of zero return days, and the TAQ effective spread and lambda. We also report time-series correlations of all measures with firm size, idiosyncratic risk, market volatility (VIX), and total dollar trading volume. Importantly, these are *time-series* correlations for the median cost estimate with average firm size, idiosyncratic volatility, and aggregate daily trading volume over

[17] The fact that the TAQ models provide such extreme and unreliable estimates of trading costs, especially at the extremes, suggests that these models may be ill-suited to analyze the trading costs of various strategies based on empirical anomalies at any reasonable portfolio size, as in Lesmond, Schill, and Zhou (2003) and Novy-Marx and Velikov (2017), or to use in calculating break-even capacity fund sizes of various anomalies to assess how many dollars could be traded in an anomaly before trading costs eliminated its excess returns, as in Korajczyk and Sadka (2004). This explains why Frazzini, Israel, and Moskowitz (2017) get such different results than these papers on the trading costs and break-even fund sizes of various trading strategies.

time. Panel A reports the correlations for the U.S. time-series and Panel B for the international time-series (all countries ex-U.S.). Lacking TAQ data internationally the effective spread and lambda variables are not present for the international sample.

As Panel A of Table XII shows, the median market impact cost we measure varies positively through time with the other liquidity variables, ranging from 0.35 for the TAQ effective spread to 0.75 for the proportion of zero trading days. This result contrasts with the lack of correlation we found cross-sectionally in Table VII, where the cross-sectional relation between our price impact measure and the other liquidity and trading cost measures from the literature was weak and subsumed by the other variables in our model. The contrast between these two sets of results indicates that the liquidity and cost measures of individual stocks do not match up well with real trading costs *across* firms, likely due to the substantial noise in these estimates at the firm level. However, in *aggregate* all of these measures are picking up general trends in aggregate liquidity, where averaging across stocks mitigates the noise. The last four columns of Panel A also show that the time-series variation in these aggregate cost measures line up intuitively with other variables likely related to market liquidity, such as average firm size, firm volatility, and market volatility. Panel B of Table XII shows similar patterns internationally (with the exception of proportion of zero return days).

Figure 10 shows, however, that while the time-series variation in the median cost across firms behaves relatively well across different cost measures, the extreme parts of the distribution do not. Figure 10 plots the correlations of our market impact measure for the 10^{th}, 50^{th}, and 90^{th} percentiles relative to the same percentiles for the other measures. The time-series correlations of the 10^{th} and 90^{th} percentiles across measures are much weaker and inconsistent across measures (sometimes switching signs). These results further show that other measures of liquidity and trading costs have less reliability at both ends of the distribution across firms and do not match the live trading costs we find for these stocks. These other measures therefore seem to be able to capture some of the aggregate movements in market liquidity, but have a much more difficult time capturing cross-sectional variation in liquidity and costs across securities.

E.4 Correlations within trading cost measures

Table XIII provides a final test of the reliability of various trading cost measures by looking at correlations *within* a given measure. Specifically, for a given trading cost measure we calculate the time-series correlation of the 10^{th}, median, and 90^{th} percentile firm, in the U.S. and internationally, to test whether changes in costs/illiquidity for the median firm are associated with similar changes for the 10^{th} and 90^{th} percentile firms, and whether those changes are correlated internationally.

Panel A of Table XIII reports the results for our market impact model. The left side of the panel reports results covering the period over which we have live trade data (1998 to 2016), though includes all common stocks from CRSP and Compustat whether they were traded by our manager or not. The right hand side of the table reports results over the 1984 to 2016 time period that includes the out of sample period 14 years pre-dating our sample of executed trades. As the table shows, time-series variation in estimated costs at the 10th, 50th, and 90th percentile of the firm distribution are highly correlated. The shaded blue numbers show that in the U.S. the cost of the 10th percentile of firms is 0.99 correlated to the cost of the median firm over time, and 0.69 correlated to the cost of the 90th percentile firm over time, in the 1998 to 2016 sample period. Intuitively, this makes sense – if the median firm becomes more expensive to trade due to market liquidity, it seems reasonable that the 10th and 90th percentile firms are also more expensive to trade. This suggests there is some commonality to liquidity as hypothesized in the literature (Pastor and Stambaugh (2003), Acharya and Pedersen (2005), Korajczyk and Sadka (2008)). The numbers internationally are even stronger, where the time-series correlation of the 10th percentile and median firm to the 90th percentile firm is 0.87 and 0.91, respectively. The right hand side of Panel A reports the same correlations over the longer sample period 1984 to 2016, which finds similarly strong positive time-series correlations across different parts of the cross-sectional distribution of firms.

The shaded pink numbers in Panel A highlight the correlations *across* markets – U.S. and international. For our price impact cost estimates, we find that the 10th percentile of costs in the U.S. is 0.73 correlated to the 10th percentile of costs internationally. Likewise, medians are 0.82 correlated and the 90th percentiles are 0.59 correlated. Hence, there appears to be commonality across markets in terms of our trading cost measure over time.

Panels B through F of Table XIII repeat the same exercise for the Amihud (2002), modified Roll (1984), proportion of zero return days, TAQ effective spread, and TAQ lambda measures, respectively. For the Amihud measure in Panel B, the time-series correlation of the 10th, 50th, and 90th percentiles are zero or negative. In addition, the international correlations are positive for the 10th percentile and median, but are zero for the 90th percentile. This is more evidence that the Amihud measure is very noisy, especially for stocks estimated with the highest illiquidity.

Panel C shows similar results for the modified Roll measure. Here, we cannot even measure the correlation at the 10th percentile because the Roll measure is zeroed out for all of these firms, and hence does not provide a lot of cross-sectional firm heterogeneity. For the 90th percentile, the correlations are weak and vary substantially. The correlations for the prop zero measure in Panel D are weak as well, though a bit stronger internationally. For the effective TAQ spread in Panel E, we get consistent positive correlations over time from firms in different parts of the cross-sectional distribution, and for the TAQ lambda in Panel F we get similarly positive but lower correlations.

Adding up all of the evidence in this section, our trading cost model, which is calibrated to live executed trades from a large institutional investor, seems to provide more stable and reliable trading cost estimates out of sample that: 1) capture the variation in costs implied by other measures across firms; 2) match the real-world costs of live portfolios, which other models and measures fail to do; 3) match the estimated costs from other institutional sources of costs from brokers and consulting firms better than other models; 4) deliver less extreme and more stable estimates of costs; and 5) vary over time and across the distribution of firms over time in an intuitive way that other models and measures often fail to do. These results suggest our trading cost model provides a much more reasonable and reliable estimate of costs than previous models and measures.

V. Conclusion

We examine the trading costs of a large institutional trader using a dataset of live trade orders and executions in 21 international equity markets over a 19-year period. The trading costs represent those of a large arbitrageur, who is likely the marginal investor in markets, and are many times smaller than those claimed in the literature. Building a trading cost model calibrated to our live trade data, we find that the model provides more reliable trading cost estimates out of sample that subsume other measures in the literature, match the costs of live portfolios and other institutional sources of costs, and deliver less extreme and more stable estimates over time and across firms in an intuitive way that other models fail to accomplish.

The out of sample performance of our model and its ability to capture real trading costs may make it useful for assessing costs in other settings. The model is based on observable characteristics about the firm and market and can be applied in the U.S. and international markets – trade size being the only unobservable choice variable that must be input into the model, which requires specifying a portfolio size. Our model can be useful for future asset pricing research. For example, Frazzini, Israel, and Moskowitz (2017) use the model to examine how various trading strategies based on asset pricing anomalies survive trading costs at different fund sizes. They also compute a break-even capacity for each anomaly where fund size is calibrated so that trading costs would be large enough to fully erode the expected return of the strategy. Consistent with our results here, they find that trading costs (and break-even capacities) for strategies based on a host of anomalies are substantially lower (higher) than those suggested in the literature (Lesmond, Schill, and Zhou (2003), Korajczyk and Sadka (2004), Novy-Marx and Velikov (2017)).

Our model can also be useful for settings where a measure of trading costs for a large arbitrageur is desired, such as event studies, mutual fund or hedge fund trading, and possibly for evaluating the impact of correlated trading effects from institutions (e.g., Lo (2007), Lou and Polk (2014), Alquist, Jiang, and Moskowitz (2018)).

References

Acharya, V. and L.H. Pedersen (2005), "Asset Pricing with Liquidity Risk, *Journal of Financial Economics* 77, 375–410.

Amihud Y. (2002), "Illiquidity and Stock Returns: Cross-Section and Time Series Effects, *Journal of Financial Markets* 5, 31–56.

Almgren, R. (2003), "Optimal execution with nonlinear impact functions and trading-enhanced risk," *Applied Mathematical Finance,* 10, 1-18.

Almgren R., and N. Chriss (2000), "Optimal execution of portfolio transactions," *Journal of Risk*, 3, 5-40.

Almgren R., C. Thum, E. Hauptmann, and H. Li. Equity (2005), "Market impact." *Risk*, 18, 57–62.

Alquist, R., S. Jiang, and T.J. Moskowitz (2018), "Crowding," Working paper Yale University and AQR Capital.

Asness, C., T.J. Moskowitz, and L.H. Pedersen (2013), "Value and Momentum Everywhere," *Journal of Finance*, Vol. 68 (3), 929-985.

Barra (1997), "Market Impact Model Handbook."

Bekaert, G., C. Harvey, and C. Lundblad (2007), "Liquidity and Expected Returns: Lessons from Emerging Markets," *Review of Financial Studies,* 20(6), 2007, 1783-1831.

Blocher, J. and R. E. Whaley (2015), "Passive Investing: The Role of Securities Lending," working paper Vanderbilt University.

Breen, W. J., L. Hodrick, and R. Korajczyk (2002), "Predicting equity liquidity," *Management Science* 48, 470–483.

Chordia, T., Roll, R., Subrahmanyam, A. (2001), "Market liquidity and trading activity." *Journal of Finance* 56, 501-530.

Chordia, T., Roll, R., Subrahmanyam, A. (2002). "Order imbalance, liquidity and market returns." *Journal of Financial Economics* 65, 111-130.

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

Engle, R., R. Ferstenberg, and J. Russell (2012), “Measuring and Modeling Execution Cost and Risk,” *The Journal of Portfolio Management*, Winter.

Fama, Eugene F. and Kenneth R. French (1993), “Common Risk Factors in the Returns on Stocks and Bonds,” *The Journal of Financial Economics*, vol. 33, pp. 3-56.

Fama, E., MacBeth, J., (1973), “Risk, return, and equilibrium: empirical tests.” *Journal of Political Economy* 607-636.

Fujimoto, A. (2003), “Macroeconomic sources of systematic liquidity.” Working Paper, Yale University.

Garleanu, N. and Pedersen L.H. (2012), “Dynamic Trading with Predictable Returns and Transaction Costs,” *Journal of Finance*.

Glosten, L. R., Harris, L. E. (1988), “Estimating the components of the bid/ask spread.” *Journal of Financial Economics* 21, 123-142.

Goyenko, R. (2006), “Stock and bond pricing with liquidity risk.” Working Paper, Indiana University.

Goyenko, Holden, and Trzcinka (2009), “Do Liquidity Measures Measure Liquidity?” *Journal of Financial Economics* 92 (2009) 153–181.

Grinold and Kahn (1999), “Active Portfolio Management,” McGraw Hill.

Hasbrouck, J., (1991a), “Measuring the information content of stock trades.” *Journal of Finance* 46, 179-207.

Hasbrouck, J., (1991b) “The summary informativeness of stock trades: an econometric analysis.” *Review of Financial Studies* 4, 571-595.

Hasbrouk, J. (2009), “Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data,” *The Journal of Finance,* Vol. 54(3), 1445-1477.

Hausman, A., A. W. Lo, and A. C. MacKinlay (1992), “An ordered probit analysis of transaction stock prices,” *Journal of Financial Economics* 31, 319–379.

Holden, C. (2009), “New low-frequency liquidity measures,” Working Paper, Indiana University.

Huang, R. D. and H. R. Stoll (1997). “The components of the bid-ask spread: A general approach.” *Review of Financial Studies* 10 (4), 995–1034.

Huberman, G. and W. Stanzl (2001). “Optimal liquidity trading,” *Review of Finance*, Vol. 9 (2), 165-200.

Jegadeesh, N. and S. Titman (1993), “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency,” *The Journal of Finance*, vol. 48, pp. 65-91.

Jones, C. (2001), “A century of stock market liquidity and trading costs.” Working Paper, Columbia University.

Jones, C. M. and M. L. Lipson (1999) “Execution costs of institutional equity orders.” *Journal of Financial Intermediation* 8, 123–140.

Kahn, R. N. (1993). “How the execution of trades is best operationalized.” In K. F. Sherrerd (Ed.), Execution Techniques, True Trading Costs, and the Microstructure of Markets. AIMR.

Keim, D.B. (1999), “An analysis of mutual fund design: the case of investing in small-cap stocks,” *The Journal of Financial Economics*, vol. 51, pp. 173-194.

Keim, D. B. and A. Madhavan (1995). “Anatomy of the trading process: Empirical evidence on the behavior of institutional traders.” *Journal of Financial Economics* 37, 371– 398.

Keim, D.B., and A. Madhavan (1997), “Execution costs and investment style: an interexchange analysis of institutional equity trades,”*Journal of Financial Economics*, vol. 46, pp. 265-292.

Knez, P. and M. Ready (1996), “Estimating the profits from trading strategies,” *Review of Financial Studies* 9, 1121–1163.

Korajczyk, R.A., and R. Sadka (2004), “Are momentum profits robust to trading costs?” *Journal of Finance*, vol. 59, pp. 1039-1082.

Kyle, A. S. (1985). “Continuous auctions and insider trading.” *Econometrica* 53, 1315–1336.

Lee, C. and M. Ready (1991), “Inferring trade direction from intraday data,” *Journal of Finance* 46, 733–754.

Lehmann, B. (2003), “What We Measure in Execution Cost Measurement”, *Journal of Financial Markets*, vol. 6, pp. 227-231.

Lesmond, D.A., M.J. Schill, and C. Zhou (2003), “The illusory nature of momentum profits,” *Journal of Financial Economics*, vol. 71, pp. 349–380.

Lesmond, D. (2005). “Liquidity of emerging markets,” *Journal of Financial Economics* 77, 411–452.

Lesmond, D. A., J. Ogden, and C. Trzcinka (1999), “A new estimate of transaction costs.” *Review of Financial Studies* 12, 1113–1141.

Lo, A. (2007), “What Happened to the Quants in 2007?” *Journal of Investment Management 5 (2007), 29-78.*

Lou, D. and C. Polk (2014), “Comomentum: Inferring Arbitrage Activity from Return Correlations,” Working paper, London School of Economics.

Loeb, T. F. (1983). “Trading costs: The critical link between investment information and results.” *Financial Analysts Journal* 39, 39–44.

Novy-Marx, R. and M. Velikov (2017), “A Taxonomy of Anomalies and their Trading Costs,” *forthcoming Review of Financial Studies*.

Pástor, L., Stambaugh, R. F. (2003) “Liquidity risk and expected stock returns.” *Journal of Political Economy* 111, 642-685.

Patton, J. and B. Weller (2018), “What You See is Not What You Get: The Costs of Trading Market Anomalies,” Working paper, Duke University.

Perold, A. F. (1988), “The implementation shortfall: paper vs. reality,” *Journal of Portfolio Management*, vol. 14, pp. 4–9.

Roll, R. (1984), “A simple implicit measure of the effective bid-ask spread in an efficient market,” *Journal of Finance* 39, 1127–1139.

Sadka, R., (2006). “Momentum and post-earnings announcement drift anomalies: The role of liquidity risk.” *Journal of Financial Economics* 80, 309-349.

Watanabe, A. and M. Watanabe (2006), “Time-varying liquidity risk and the cross section of stock returns.” *Review of Financial Studies*, Volume 21, Issue 6, 1, 2449–2486.

Werner, Ingrid (2003), “NYSE Order Flow, Spreads and Information”, *Journal of Financial Markets*, vol. 6, pp. 309-335.

Figure 1: Anatomy of a Trade Execution

The two figures plot an example of a live execution of a buy order for Microsoft, executed on May 8, 2014 from a long-only account. The trade is for 1.4% of Microsoft's average one-year rolling daily trading volume (DTV), and the trade was executed over a time interval of 6 minutes and 3 seconds, from 9:40:01 am to 9:46:04 am. The total market impact of the trade is 4.3 bps. The first graph plots the limit orders submitted by the trading algorithm at various prices, the TAQ best bid offered across all trading venues throughout the time interval, and the executed buying price of the limit orders, all time-stamped to the millisecond. The second graph illustrates what happens at the execution times by plotting the TAQ best bid and ask prices at the time of order submission, the executed prices, and the TAQ best bid and ask prices at the time of execution.

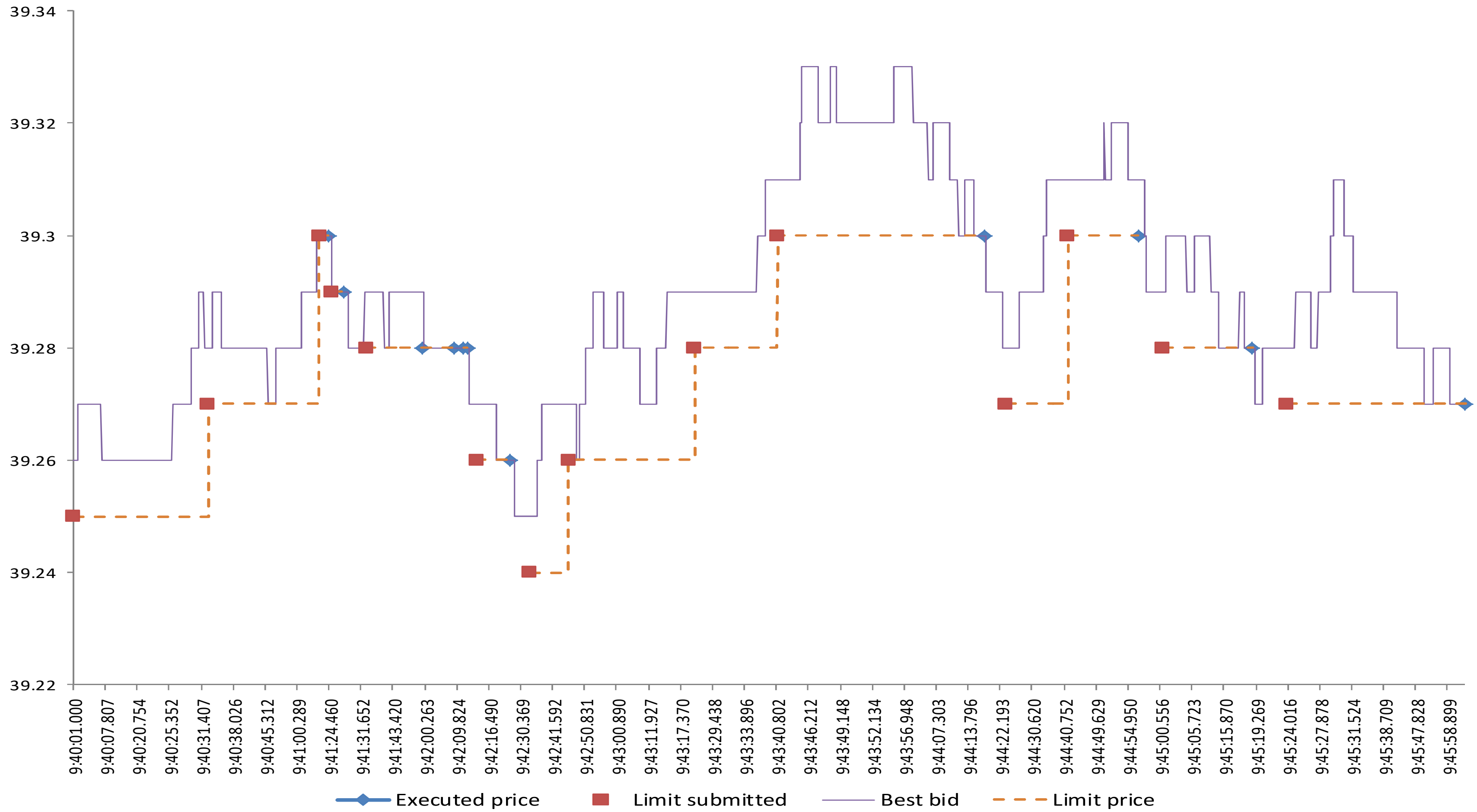


Figure 1: Anatomy of a Trade Execution (continued)

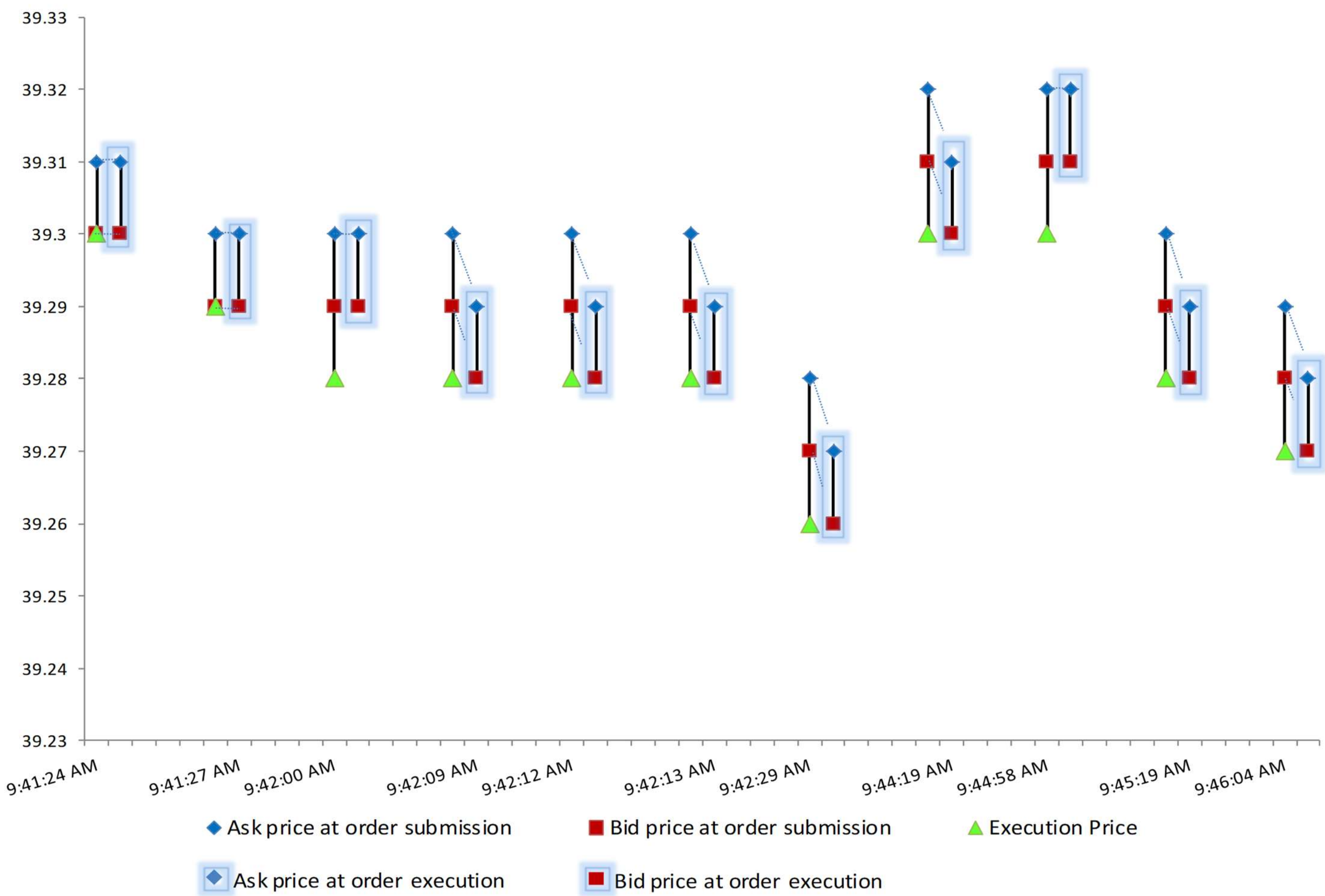


Table I
Trade Execution Data Summary Statistics (1998 to 2016)

This table reports summary statistics of the trade execution database. Panel A reports the total amount traded by year. Amounts are in billion USD. "U.S." are trades executed in the United States. "International" are trades executed outside of the United States (in 20 other countries). "Large Cap" are trades in large capitalization securities, "Small Cap" are trades in small capitalization securities, determined based on the portfolio's benchmark (e.g. for the US large cap is the Russell 1000 and small cap is below the Russell 1000 in market cap, typically within the Russell 2000 universe). "Long Short" are trades executed in long-short accounts and "Long only" are trades executed in long-only accounts. Relaxed constraint portfolios (130-30 and 140-40) are classified as "Long only" (these portfolios are small, representing 0.19% of total trades). Panel B reports the market coverage of our execution database of stocks across seven regions: US, Europe, Japan, UK, Canada, Australia, and Pacific. Panel C reports times-series averages of yearly cross-sectional summary statistics on trading, including the number of stocks traded per year, number of trade baskets per year, average trade size, fraction of (daily) trading volume traded on average, which is the trade's dollar size divided by the stock's average prior one-year dollar volume, trade horizon (how long the trade takes to complete), as well the number of limit orders and fill rates used to execute the trades. Finally, the last two rows report the fraction of orders executed aggressively, which means paying the spread at the time of execution, and the average bid-ask spread in markets at the time the trade is initiated ("arrival price spread"). Statistics are calculated across all developed market equity transactions (cash equities and equity swaps) between August 1998 and June 2016.

Panel A: Amount Traded (Billion USD)

		By region		By size		By portfolio type	
Year	Total	U.S.	International	Large Cap	Small Cap	Long short	Long only
1998*	2.96	1.29	1.67	2.96		2.96	
1999	5.29	1.99	3.30	5.29		5.29	
2000	1.99	0.76	1.23	1.99		1.86	0.13
2001	1.08	0.55	0.53	1.08		1.00	0.08
2002	4.21	0.71	3.50	4.21	0.00	1.40	2.81
2003	5.43	2.69	2.75	5.43	0.00	4.17	1.26
2004	10.00	2.95	7.05	9.99	0.01	6.38	3.62
2005	16.16	8.06	8.10	15.75	0.41	11.45	4.71
2006	67.01	34.79	32.22	64.23	2.78	44.69	22.31
2007	129.46	50.70	78.76	125.21	4.25	96.65	32.81
2008	108.29	25.06	83.24	104.27	4.02	69.30	38.99
2009	111.12	18.58	92.54	108.12	2.99	85.50	25.62
2010	117.17	29.15	88.02	113.78	3.38	91.94	25.23
2011	146.50	56.62	89.88	141.93	4.58	115.69	30.81
2012	179.09	121.39	57.70	173.41	5.68	141.97	37.13
2013	173.94	112.75	61.18	167.11	6.82	117.25	56.69
2014	223.34	153.72	69.62	217.41	5.93	169.99	53.35
2015	263.26	167.39	95.87	256.04	7.22	185.30	77.96
2016*	135.10	82.85	52.25	130.87	4.23	93.33	41.77
Total	1,701.39	871.99	829.40	1,649.07	52.32	1,246.11	455.28

*The data begins in August 1998 and ends in June of 2016, so only a partial year of trading for 1998 and 2016.

Table I (continued)
Trade Execution Data Summary Statistics (1998 to 2016)

Panel B: Market Coverage

	Numer of	Firm size (Billion USD) Mean	Median	Fraction of Market	Fraction of Firms
All	9,543	15.42	4.94	0.715	0.174
United States	4,984	16.16	4.26	0.718	0.209
Europe	1,252	17.56	7.28	0.703	0.140
Japan	1,302	9.04	3.98	0.712	0.245
United Kingdom	556	19.60	6.18	0.808	0.106
Canada	415	14.71	7.48	0.692	0.077
Australia	354	14.72	4.51	0.761	0.116
Pacific	680	15.62	6.39	0.592	0.215

Panel C: Summary Statistics on Trading

	Mean	Median	Std	Min	Max
Number of stocks per year	3,463.0	4,485.0	1,649.6	311.0	5,041.0
Number of markets per year	18.0	20.0	4.0	8.0	21.0
Number of exchanges per year	25.6	26.0	4.9	12.0	32.0
Number of trade baskets per year	1,352.7	1,469.0	1,334.1	9.0	3,615.0
Average trade size (1,000$)	607.2	356.9	917.6	53.4	5,964.4
Fraction of trading volume (%)	0.9	0.4	1.7	0.1	13.1
Trade horizon (days)	2.7	1.7	2.0	1.0	9.8
Number of stocks per trade basket	135.0	72.0	203.8	1.0	2,496.0
Number of parent orders per trade basket	153.7	78.0	265.3	1.0	7,848.0
Number of parent orders per year (x 1,000)	581.3	523.7	146.6	445.2	809.6
Number of child orders per year (x 1,000,000)	222.9	207.4	35.2	190.3	270.8
Number of executions per year (x 1,000,000)	36.4	26.1	22.4	17.7	75.7
Ratio of fills to cancelled orders per year	0.13	0.08	0.11	0.04	0.32
Fraction of orders executed aggressively	0.15	0.04	0.23	0.00	1.00
Arrival price spread (bps)	21.33	14.30	21.89	0.00	145.89

Figure 2. Event-Time Average Market Impact

This figure plots the event-time average market impact (MI) in our data for all US trades made within one day duration, where averages are reported for 30 minute intervals during the day and two-hour intervals overnight and for the next trading day. The data includes all available developed market equity transactions (cash equities and equity swaps) in our database between August 1998 and June 2016. Market impact is in basis points and is calculated as the weighted average total change in price from the time trading starts in a given security to the last trade completed, weighted by the amount traded. Plotted are the averages of all market impact calculations across all stocks during an average trading day, where we take the average of all of the trades within each 30 minute interval throughout the trading day. Also plotted are the average market impact for each two-hour interval overnight and over the following trading day (*t+1*). The shaded areas indicate 95% confidence bands.

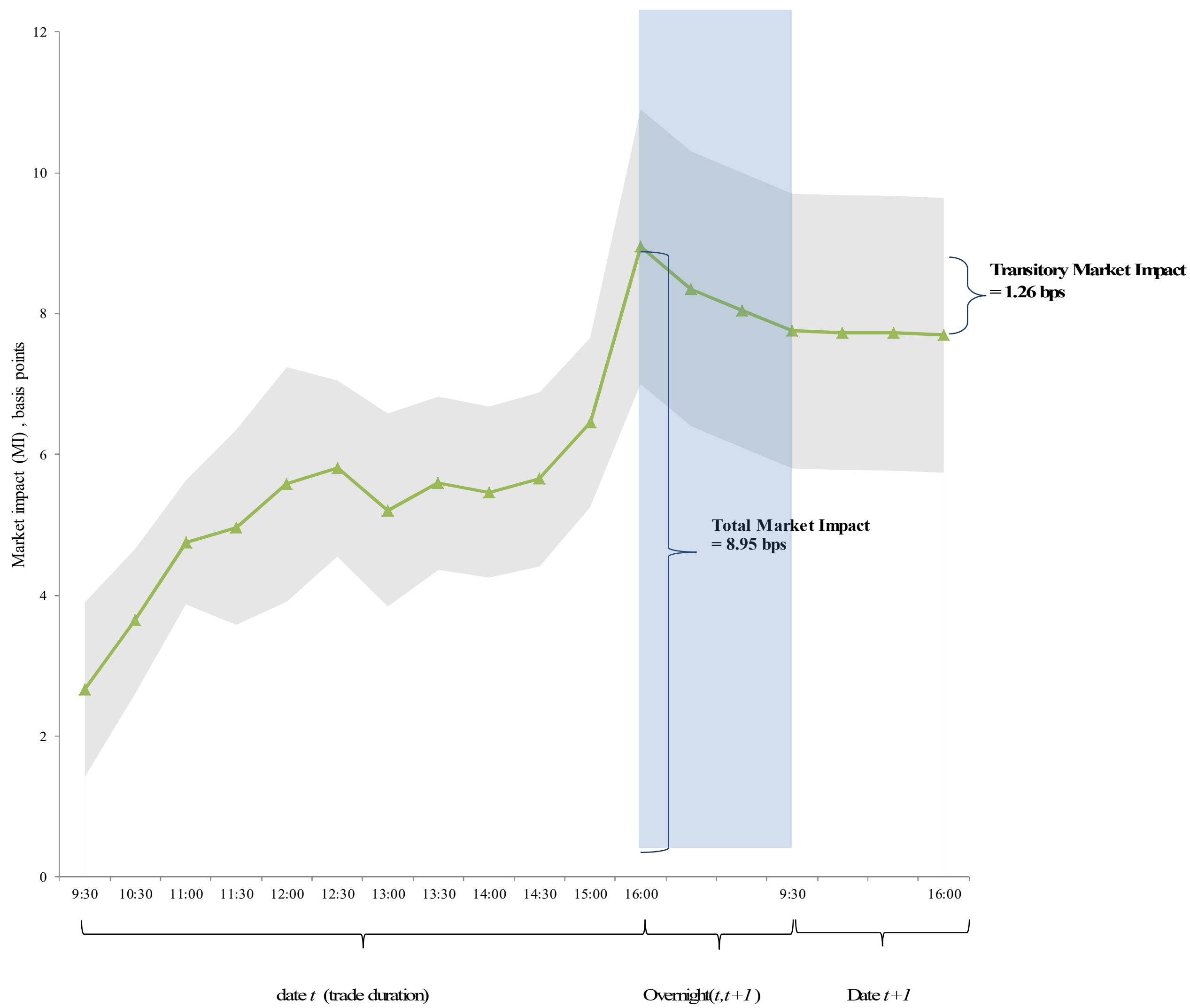


Table II
Realized Trading Costs

This table shows average market impact (MI) and implementation shortfall (IS) for all $1.7 trillion worth of trades in the trade execution database. Each month, we compute average, median and weighted average (by dollars traded) cost ("vw mean") of all trades. Time-series averages of the cross sectional estimates are reported, where we weight each monthly observation by the number of trades (stock-order type-date triplet, e.g., IBM-buy to cover-20010112) executed during the month. Statistics pertain to all available developed market equity transactions (cash equities and equity swaps) in our data between August 1998 and June 2016. The distinction between large and small cap is based on the portfolio's benchmark (e.g. for the US large cap is the Russell 1000 and small cap is below the Russell 1000 in market cap, typically within the Russell 2000 universe). "Long Short" are trades executed in long-short accounts and "Long only" are trades executed in long-only accounts. Relaxed constraint portfolios (130-30 and 140-40) are classified as "Long only" (these portfolios are small, representing 0.19% of total trades). Panel A reports statistics over the full sample period from August 1998 to June 2016 and Panel B from January 2006 to June 2016. MI and IS are in basis points and standard errors are reported at the bottom of each panel. Panel C reports MI (vw mean) for different types of orders: buy long, buy to cover, sell long, and sell short, by region and size.

Panel A: Full sample 199808 to 201606

	All	By region			By size		By portfolio type	
		US Nyse-Amex	US Nasdaq	Int.	Large cap	Small cap	Long short	Long only
MI mean	9.97	7.86	10.21	11.30	8.90	18.95	8.37	13.66
MI median	6.18	5.06	5.03	7.72	5.54	13.53	5.13	8.97
MI vw mean	15.14	13.03	14.88	15.97	14.68	23.04	14.72	15.11
IS mean	11.02	8.73	10.85	12.61	9.93	20.18	9.61	14.27
IS median	8.63	7.10	7.11	10.54	7.89	16.43	7.53	11.60
IS vw mean	16.06	13.78	17.02	16.95	15.57	25.30	15.58	15.88
Standard errors								
MI mean	0.51	0.60	0.96	0.57	0.54	0.91	0.59	0.67
MI median	0.33	0.44	0.52	0.42	0.34	0.81	0.37	0.58
MI vw mean	0.68	0.86	1.28	0.71	0.70	1.25	0.78	0.79
IS mean	0.70	0.79	1.21	0.77	0.73	1.24	0.81	0.86
IS median	0.46	0.60	0.73	0.60	0.48	1.16	0.52	0.71
IS vw mean	0.88	1.17	1.62	0.84	0.90	1.49	0.97	0.98

Panel B: Recent sample 200601 to 201606

	All	By Region			By Size		By Portfolio type	
		US Nyse-Amex	US Nasdaq	Int.	Large cap	Small cap	Long short	Long only
MI mean	9.54	7.40	9.80	10.90	8.45	18.86	7.98	13.18
MI median	6.15	4.99	4.96	7.73	5.52	13.43	5.11	8.90
MI vw mean	12.95	10.86	12.57	13.89	12.59	20.16	13.11	13.09
IS mean	10.90	8.52	10.66	12.57	9.84	19.93	9.52	14.11
IS median	8.55	6.99	7.00	10.55	7.82	16.32	7.46	11.55
IS vw mean	15.97	13.71	16.85	16.86	15.48	25.21	15.49	15.80
Standard errors								
MI mean	0.62	0.77	1.35	0.73	0.60	1.90	0.58	1.15
MI median	0.33	0.43	0.52	0.43	0.34	0.81	0.36	0.59
MI vw mean	0.82	1.03	1.59	0.93	0.85	1.77	0.88	1.33
IS mean	0.70	0.78	1.19	0.78	0.73	1.23	0.80	0.88
IS median	0.46	0.59	0.73	0.61	0.48	1.16	0.51	0.74
IS vw mean	0.88	1.19	1.60	0.85	0.90	1.49	0.96	1.02

Table II (continued)
Realized Trading Costs

Panel C: Market impact by trade type								
		Fraction of sample		All	By region		By size	
		Dollars	Trades		US	Int.	Large cap	Small cap
MI (VW-mean)	Buy Long	0.36	0.34	12.53	14.70	10.36	11.72	22.68
	Buy Cover	0.15	0.16	15.53	16.75	13.62	15.33	21.79
	Sell Long	0.31	0.31	15.93	10.88	19.79	15.35	25.26
	Sell Short	0.18	0.19	16.55	8.79	22.28	16.42	26.71
Differences	Buy Cover - Buy Long			3.00	2.05	3.26	3.61	-0.89
	Sell Short - Sell Cover			0.62	-2.09	2.48	1.06	1.45
t-statistics	Buy Cover - Buy Long			1.33	0.62	1.29	1.51	-0.10
	Sell Short - Sell Cover			0.33	-0.68	0.92	0.53	0.15

Figure 3. Average Market Impact by Exchange, Size, and Trade Type

This table shows average market impact (MI) by exchange, size, and trade type. Each month, we compute the average cost of all trade baskets executed during the month. Time-series averages of the cross sectional estimates are plotted, where we weight each monthly observation by the number of trades (stock-order type-date triplet, e.g., IBM-buy to cover-20010112) executed during the month. All available developed market equity transactions (cash equities and equity swaps) are included from our data between August 1998 and June 2016. Market impact is in basis points.

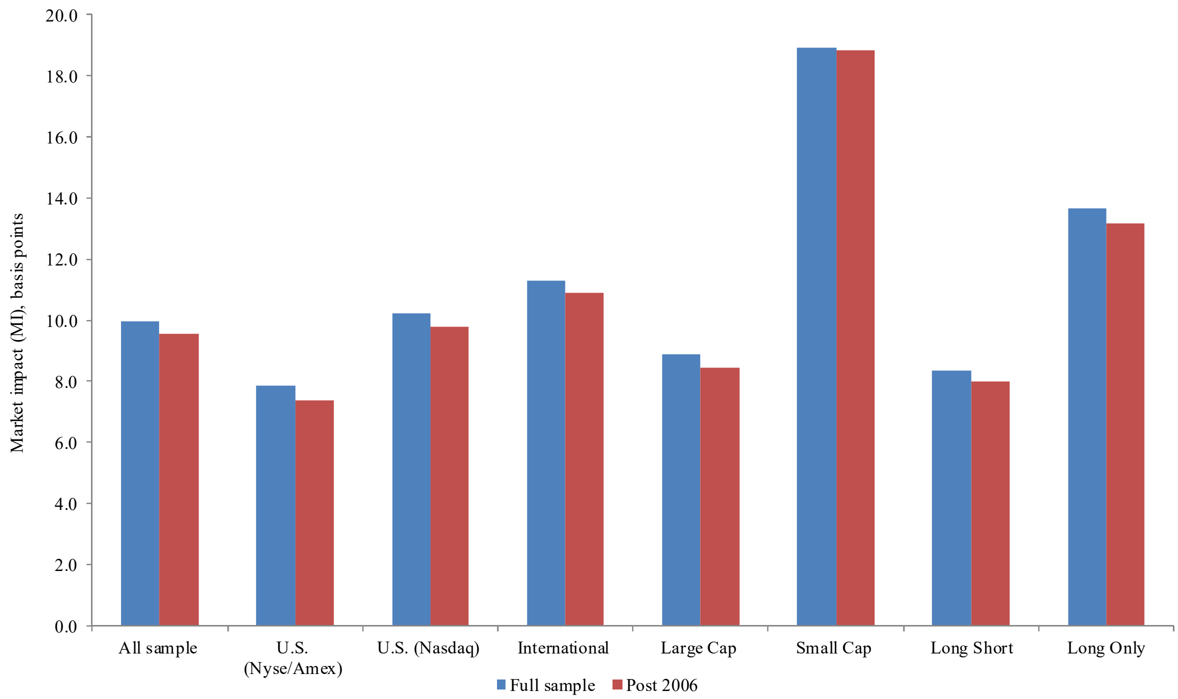


Table III
Permanent and Temporary Components of Realized Trading Costs

The table reports the value-weighted (by dollar trades) average market impact (MI) decomposed into transitory and permanent components. The transitory component of a trade's price impact is measured as the component of a trade's price impact that reverses the following trading day as defined in equation (4) and the permanent component is the difference between the total price impact and the transitory component, as defined in equation (5). Time-series averages of the cross sectional estimates are reported, where we weight each monthly observation by the number of trades (stock-order type-date triplet, e.g., IBM-buy to cover-20010112) executed during the month. Statistics pertain to all available developed market equity transactions (cash equities and equity swaps) in our data between August 1998 and June 2016. The distinction between large and small cap is based on the portfolio's benchmark (e.g. for the US large cap is the Russell 1000 and small cap is below the Russell 1000 in market cap, typically within the Russell 2000 universe). "Long Short" are trades executed in long-short accounts and "Long only" are trades executed in long-only accounts. Relaxed constraint portfolios (130-30 and 140-40) are classified as "Long only" (these portfolios are small, representing 0.19% of total trades). Panel A reports statistics over the full sample period from August 1998 to June 2016 and Panel B from January 2006 to June 2016.

Panel A: Full sample 199808 to 201606

	All	By region			By size		By portfolio type	
		US Nyse-Amex	US Nasdaq	Int.	Large cap	Small cap	Long short	Long only
MI vw mean	15.14	13.03	14.88	15.97	14.68	23.04	14.72	15.11
MI vw mean transitory	2.08	2.18	2.16	1.93	2.00	2.66	1.53	1.92
MI vw mean permanent	13.06	10.84	12.72	14.04	12.69	20.38	13.19	13.19
Temporary (%)	14%	17%	15%	12%	14%	12%	10%	13%
Permanent (%)	86%	83%	85%	88%	86%	88%	90%	87%
Standard errors								
MI vw mean	0.69	0.88	1.26	0.71	0.70	1.24	0.78	0.80
MI vw mean transitory	0.52	0.80	1.28	0.54	0.54	1.31	0.54	0.92
MI vw mean permanent	0.82	1.03	1.59	0.93	0.85	1.77	0.88	1.33

Panel B: Recent sample 200601 to 201606

	All	By region			By size		By portfolio type	
		US Nyse-Amex	US Nasdaq	Int.	Large cap	Small cap	Long short	Long only
MI vw mean	12.95	10.86	12.57	13.89	12.59	20.16	13.11	13.09
MI vw mean transitory	2.11	2.33	2.70	1.86	2.03	3.17	1.55	2.07
MI vw mean permanent	10.84	8.54	9.87	12.03	10.56	16.99	11.56	11.02
Temporary (%)	16%	21%	21%	13%	16%	16%	12%	16%
Permanent (%)	84%	79%	79%	87%	84%	84%	88%	84%
Standard errors								
MI vw mean	0.78	0.97	1.16	0.81	0.79	1.30	0.88	0.83
MI vw mean transitory	0.49	0.86	1.19	0.56	0.51	1.27	0.57	0.83
MI vw mean permanent	0.91	1.17	1.54	1.05	0.93	1.88	1.02	1.22

Table IV
Realized Trading Costs Relative to VWAP

This table shows average market impact (MI) relative to the TAQ stock's value-weighted average price (VWAP) during the execution interval. Each calendar month, we compute average, median and weighted (by dollars traded) average cost ("vw_mean") of all trades during the month. Reported are the time-series averages of the cross sectional estimates, where we weight each monthly observation by the number of trades (stock-order type-date triplet, e.g., IBM-buy to cover-20010112) executed during the month. The table reports results for the U.S., where TAQ data is available and includes all US equity transactions (cash equities and equity swaps) in our data between August 1998 and June 2016. The distinction between large cap and small cap is based on the portfolio's benchmark. The distinction between "long-short" and "long only" depends on which type of portfolio the trade originated, where relaxed constraint portfolios (130-30 and 140-40) are classified as "long only." Market impact and implementation shortfall are measured relative to the VWAP during the trading interval following equation (6) and are reported in basis points with standard errors reported in the panel below the point estimates. The top panel of the table reports results for the full sample from 1998 to 2016 and the bottom panel reports results over the more recent time period from 2006 to 2016.

Full sample: 199808 - 201606	All sample	By Region		By size		By portfolio type	
		Nyse-Amex	Nasdaq	Lage cap	Small cap	Long-short	Long only
MI mean	4.50	3.95	5.61	3.41	9.72	3.63	6.96
MI median	2.62	2.62	2.62	2.10	6.38	2.45	3.44
MI vw mean	6.61	6.49	6.91	5.80	11.71	5.70	8.61
IS mean	4.50	3.95	5.61	3.41	9.72	3.63	6.96
IS median	2.62	2.62	2.62	2.10	6.38	2.45	3.44
IS vw-mean	6.61	6.49	6.91	5.80	11.71	5.70	8.61
Standard errors							
MI mean	0.59	0.42	0.96	0.65	0.83	0.66	0.65
MI median	0.25	0.24	0.26	0.27	0.65	0.30	0.39
MI vw mean	0.83	0.74	1.30	0.87	1.15	0.86	1.23
IS mean	0.59	0.42	0.96	0.65	0.83	0.66	0.65
IS median	0.25	0.24	0.26	0.27	0.65	0.30	0.39
IS vw-mean	0.83	0.74	1.30	0.87	1.15	0.86	1.23

Recent sample: 200601 - 201606	All sample	By Region		By size		By portfolio type	
		Nyse-Amex	Nasdaq	Lage cap	Small cap	Long-short	Long only
MI mean	4.47	3.90	5.63	3.29	10.19	3.60	6.95
MI median	2.49	2.50	2.47	1.90	6.83	2.25	3.52
MI vw mean	6.59	6.45	7.03	5.84	11.73	5.80	8.34
IS mean	4.47	3.90	5.63	3.29	10.19	3.60	6.95
IS median	2.49	2.50	2.47	1.90	6.83	2.25	3.52
IS vw-mean	6.59	6.45	7.03	5.84	11.73	5.80	8.34
Standard errors							
MI mean	0.32	0.32	0.43	0.27	0.85	0.27	0.65
MI median	0.14	0.15	0.14	0.14	0.59	0.16	0.40
MI vw mean	0.79	0.84	0.94	0.76	1.22	0.68	1.28
IS mean	0.32	0.32	0.43	0.27	0.85	0.27	0.65
IS median	0.14	0.15	0.14	0.14	0.59	0.16	0.40
IS vw-mean	0.79	0.84	0.94	0.76	1.22	0.68	1.28

Figure 4. Market Impact by Country

The graph plots the median, average, and value-weighted (by trade size) average market impact (MI) by country. Each calendar month, we compute the average cost of all trade baskets executed during the month and plot the time-series averages of the cross sectional estimates, where we weight each monthly observation by the number of trades (stock-order type-date triplet, e.g., IBM-buy to cover-20010112) executed during the month. All developed market equity transactions (cash equities and equity swaps) in our data are included between August 1998 and June 2016.

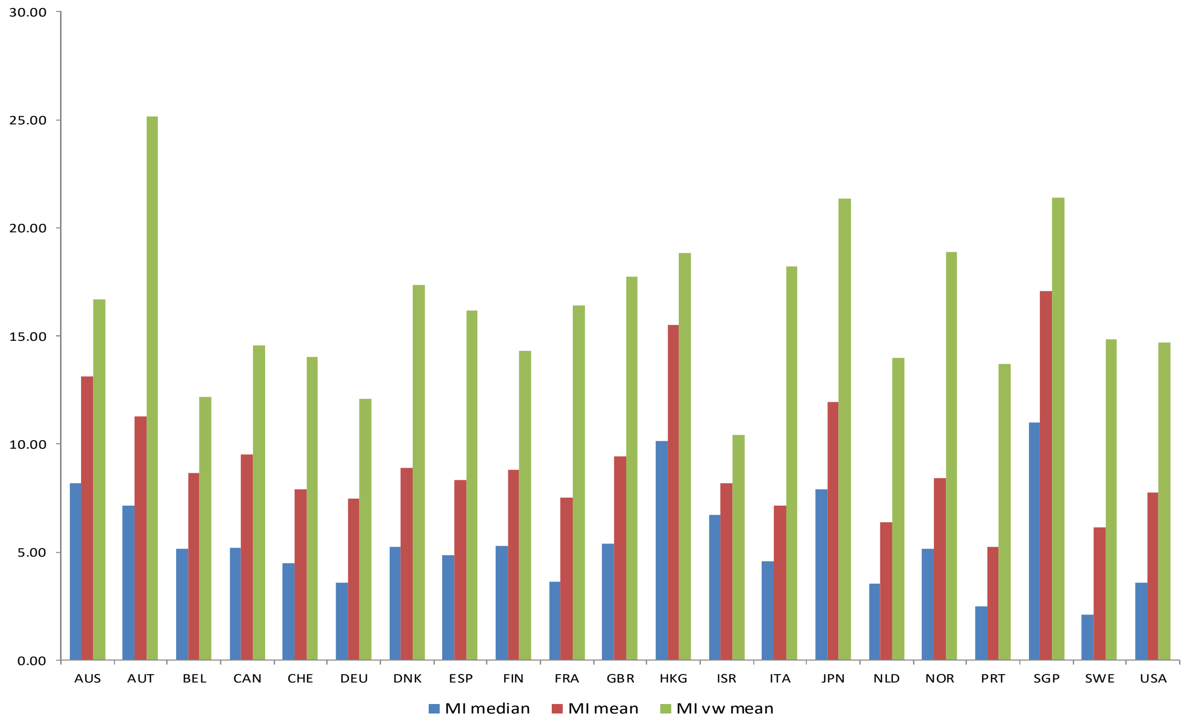


Table V
Realized Trading Costs from New Inflows

The table reports the average market impact (MI) estimates from new inflows of funds into long-only accounts compared to all *other* long-only trades (excluding new inflows). Each calendar month, we compute average, median and weighted average (by dollar trades) trading costs ("vw mean") of all trades during the month. Time-series averages of the cross sectional estimates are reported, where we weight each monthly observation by the number of trades (stock-order type-date triplet, e.g., IBM-buy to cover-20010112) executed during the month. Statistics are reported for all available developed market equity transactions executed in long-only accounts (cash equities and equity swaps) in our data between August 1998 and June 2016 coming solely from new inflows, defined as the first trade for a given long-only account. A graph highlighting the market impact from new inflows versus other trades is also shown. Market impact is in basis points.

Market impact of trades from new flows					
Long-only trades, 199808 - 201606	Trade type	Inflows only	All other trades	Difference	*t*-statistic
MI mean	All trades	14.99	13.57	1.42	0.36
MI median	All trades	11.77	8.92	2.85	0.77
MI vw mean	All trades	11.40	15.24	-3.84	-1.08
MI mean	Large cap	14.16	11.24	2.92	0.62
MI median	Large cap	11.29	7.43	3.86	0.88
MI vw mean	Large cap	11.30	14.63	-3.34	-0.84
MI mean	Small cap	17.62	18.90	-1.27	-0.28
MI median	Small cap	13.37	13.45	-0.08	-0.02
MI vw mean	Small cap	24.08	22.78	1.30	0.22

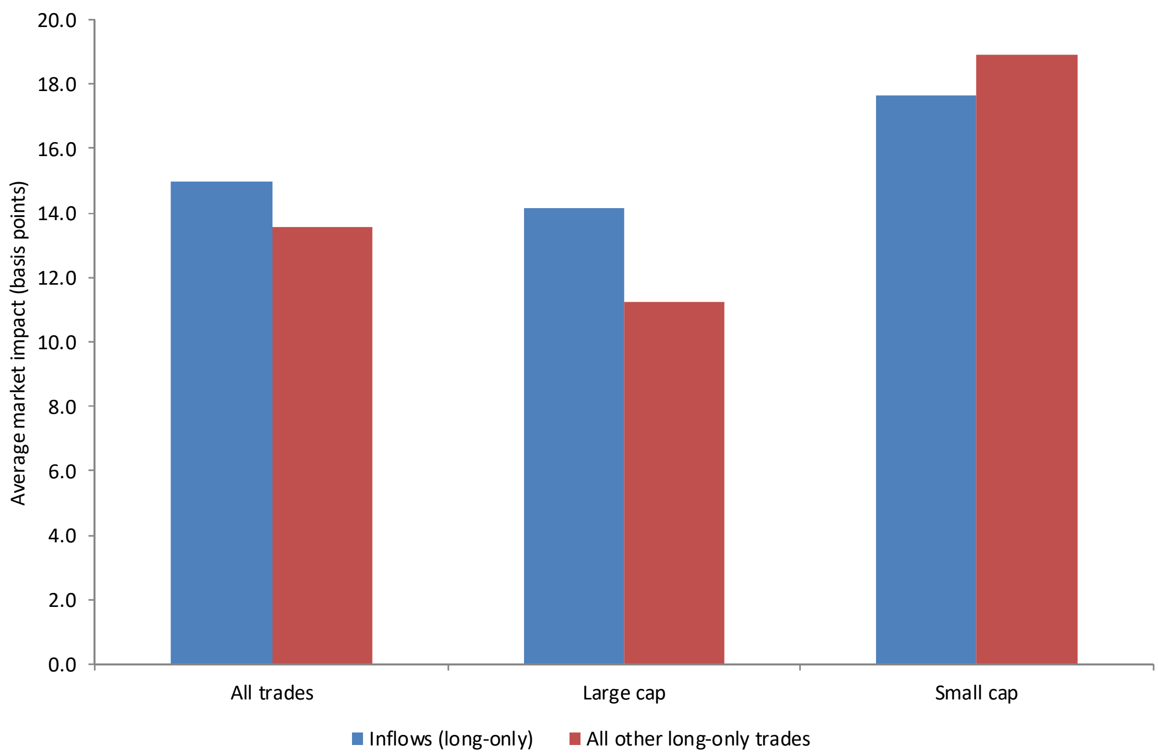


Table VI
The Effect of Market Structure on Trading Costs

Reported are regression results for the median, mean, and value-weighted mean market impact costs across the 21 countries in our sample on dummy variable for various country stock exchange features that govern how an investor trades equities in each market. The market features are detailed in Table A4 in the Appendix. Regressions are run on each market feature separately in univariate regressions, as well as on all market features simultaneously in a single multivariate regression. Regression coefficient estimates are reported along with their associated *t*-statistics in parentheses. **Bold** indicates significance at the 5% level or lower.

Dependent variable =	Median MI		Mean MI		VW mean MI	
	Univariate	Multivariate	Univariate	Multivariate	Univariate	Multivariate
Separate Oddlot Handling	**2.45**	-1.26	**3.82**	-0.33	**2.37**	-0.06
	(2.60)	(-1.07)	(3.43)	(-0.23)	(1.50)	(-0.02)
Intraday Triggered Auctions	**2.61**	1.89	**3.92**	3.04	**3.48**	3.36
	(2.96)	(2.47)	(3.75)	(3.25)	(2.46)	(1.62)
OTC Trading	**-3.83**	-2.12	**-5.55**	-4.12	**-2.24**	-5.92
	(-3.22)	(-1.43)	(-3.89)	(-2.28)	(-1.02)	(-1.48)
Order Protection Rule	**-1.22**	7.81	**-0.69**	3.81	**-1.88**	-9.53
	(-0.70)	(1.68)	(-0.31)	(0.67)	(-0.71)	(-0.76)
Number of LIT Venues	**-0.49**	-1.48	**-0.50**	-0.99	**-0.27**	0.53
	(-2.63)	(-2.24)	(-1.95)	(-1.23)	(-0.83)	(0.30)
Fragmentation	**-4.43**	1.62	**-5.02**	1.64	**-0.63**	3.52
	(-4.12)	(0.81)	(-3.29)	(0.67)	(-0.28)	(0.65)
Transaction Tax	**0.00**	0.00	**0.00**	-0.01	**0.01**	-0.01
	(-0.02)	(0.20)	(0.10)	(-0.39)	(0.50)	(-0.17)
Short Sale Uptick Rule	**2.00**	1.08	**2.87**	0.63	**2.28**	1.15
	(1.42)	(0.94)	(1.60)	(0.45)	(1.04)	(0.37)

Figure 5. Market Impact by Fraction of Trading Volume

This figure plots the average market impact (MI) for actual live trades from our execution database. We sort all trades into 30 bins based on their fraction of daily volume and compute average market impact for each bucket. This table includes all available developed market equity transactions (cash equities and equity swaps) in our data between August 1998 and June 2016. Market impact is in basis points.

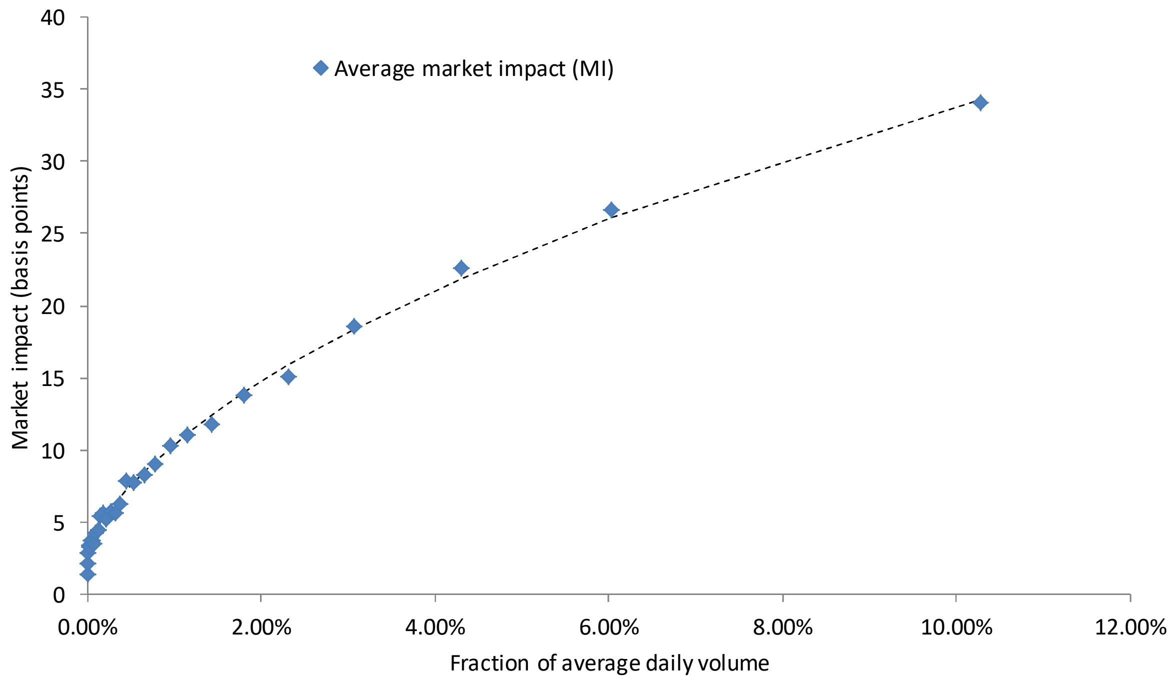


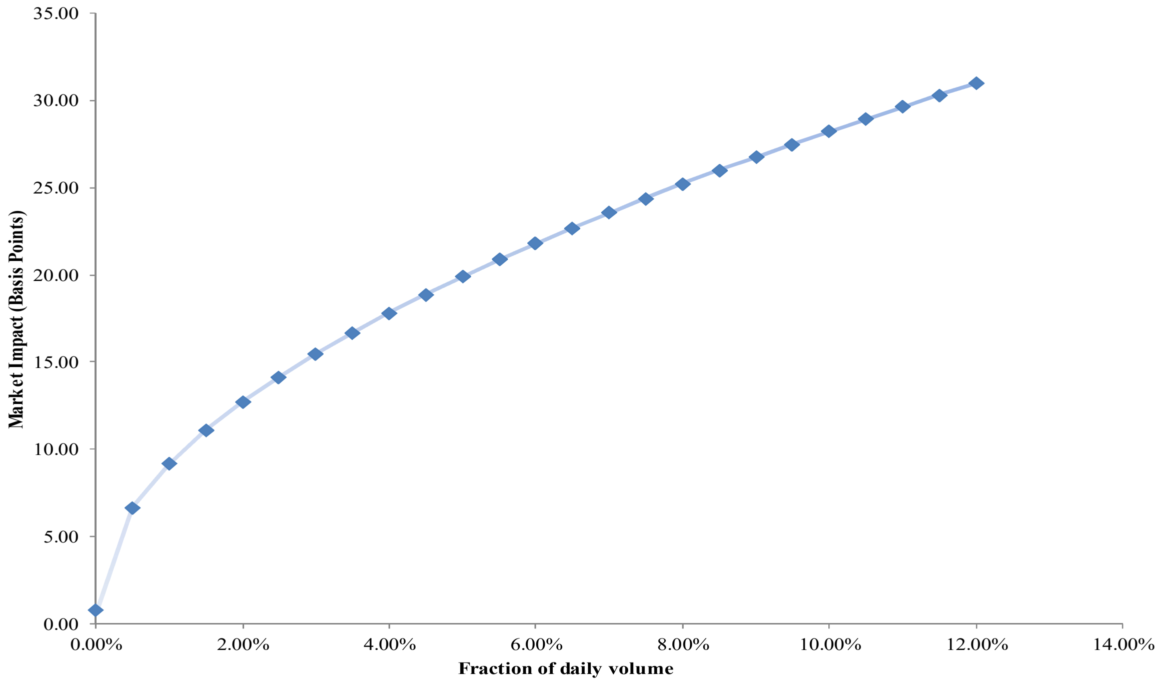
Table VII
Trading Cost Model Calibrated with Live Trade Data

We estimate a trading cost model calibrated to our live trading data. Reported are estimates of parameters for a trading cost model that is a function of three sets of variables: market-wide variables, stock characteristics, and trade characteristics. The estimates are results from pooled regressions of a trade's market impact (MI), in basis points, on the following explanatory variables: "Beta*IndexRet*buysell" is the contemporaneous (beta-adjusted) market return that controls for market movements occurring at the time of the trade, where Beta is the stock's predicted beta at the time of order submission, "IndexRet" is the corresponding index return over the life of the trade (the movement of the market), and "buysell" is a dummy equal to 1 for buy orders and -1 for sell orders. "Time trend" is a linear time trend. Log of ME is equal to the log of 1 plus the market value of equity, where ME is in billions of USD. "Fraction of daily volume" is equal to the trade's dollar size divided by the stock's average one-year dollar volume (in %). Both a linear and square root function of fraction of daily volume are included in the regression. "Idiosyncratic Volatility" is the volatility of the residuals of a regression of one-year daily stock returns on the corresponding value-weighted benchmark (annualized %), "VIX" is the monthly variance of the CRSP-value weighted index, computed using daily returns (annualized %). The DGTW-adjusted return is the return on the stock minus the return on a portfolio of similar stocks matched on size, book-to-market, and momentum (past one year return) from Daniel, et al. (1997), which is interacted with the "buysell" dummy. Country fixed effects are included where indicated, *t*-statistics are shown below the coefficient estimates and 5% statistical significance is indicated in bold. Standard errors are clustered by calendar month. Regression estimates are provided over the full sample, the US only, and outside of the US ("International"), covering all equity trades in 21 countries from August 1998 to June 2016.

	All sample					United States					International				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Beta*IndexRet*buysell	**0.25**	**0.25**	**0.25**	**0.25**	**0.23**	**0.30**	**0.30**	**0.30**	**0.30**	**0.28**	**0.22**	**0.22**	**0.22**	**0.22**	**0.14**
	(25.76)	(25.78)	(25.79)	(25.81)	(11.77)	(13.96)	(13.96)	(13.96)	(13.95)	(11.07)	(21.22)	(21.21)	(21.19)	(21.31)	(15.02)
Time trend (Jun 1926 = 1)	**-0.04**	-0.03	**-0.03**	0.00	-0.01	-0.02	0.00	-0.01	0.02	0.01	**-0.06**	**-0.05**	**-0.06**	**-0.03**	**-0.03**
	(-2.72)	(-1.96)	(-2.29)	(-0.31)	(-0.82)	(-0.82)	(-0.13)	(-0.46)	(1.00)	(0.54)	(-4.55)	(-3.67)	(-3.96)	(-2.14)	(-3.50)
Log of ME (Billion USD)	**-3.66**	**-2.61**	**-1.90**	**-0.62**	**-0.62**	**-3.28**	**-2.23**	**-1.56**	-0.20	-0.14	**-4.39**	**-3.17**	**-2.47**	**-1.18**	**-1.40**
	(-18.04)	(-13.90)	(-10.00)	(-5.14)	(-4.60)	(-14.17)	(-10.83)	(-6.91)	(-1.10)	(-0.77)	(-17.18)	(-12.70)	(-10.00)	(-8.09)	(-9.45)
Fraction of daily volume		**1.97**	**0.36**	0.22	-0.13		**2.56**	0.58	0.35	-0.53		**1.69**	**0.34**	0.25	**0.29**
		(15.29)	(2.30)	(1.55)	(-0.72)		(10.34)	(1.67)	(1.06)	(-1.37)		(12.43)	(2.12)	(1.72)	(2.05)
Sqrt(Fraction of daily volume)			**7.33**	**8.27**	**8.89**			**7.88**	**9.32**	**11.21**			**6.57**	**7.22**	**5.97**
			(11.26)	(13.23)	(10.39)			(7.11)	(8.56)	(8.54)			(11.00)	(13.18)	(12.72)
Idiosyncratic Volatility				**0.30**	**0.28**				**0.32**	**0.31**				**0.29**	**0.25**
				(10.67)	(9.50)				(7.87)	(7.49)				(9.76)	(8.94)
Vix				**0.17**	**0.15**				**0.13**	0.12				**0.21**	**0.20**
				(2.74)	(2.91)				(2.06)	(1.95)				(2.61)	(2.83)
DGTW-adjusted return*buysell				.	0.04					0.03					**0.13**
					(1.54)					(1.33)					(14.51)
Observations (1,000s)	3,470	3,470	3,470	3,470	3,470	1,722	1,722	1,722	1,722	1,722	1,748	1,748	1,748	1,748	1,748
Adjusted R^2	0.103	0.105	0.105	0.106	0.149	0.117	0.118	0.119	0.119	0.152	0.094	0.095	0.096	0.096	0.212
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Figure 6. Calibrated Trading Cost Model

The first graph plots the price impact function – market impact as a function of the fraction of daily volume traded (%DTV) – from our calibrated trading cost model in Table VII (column (5)) over the entire sample of trades from August 1998 to June 2016. All the other right-hand side variables are evaluated at the sample mean. The second graph plots the price impact functions using our calibrated model on live trading data estimated separately from 1998 to 2005, 2006 to 2016, 2007 to 2009, and 2010 to 2016.



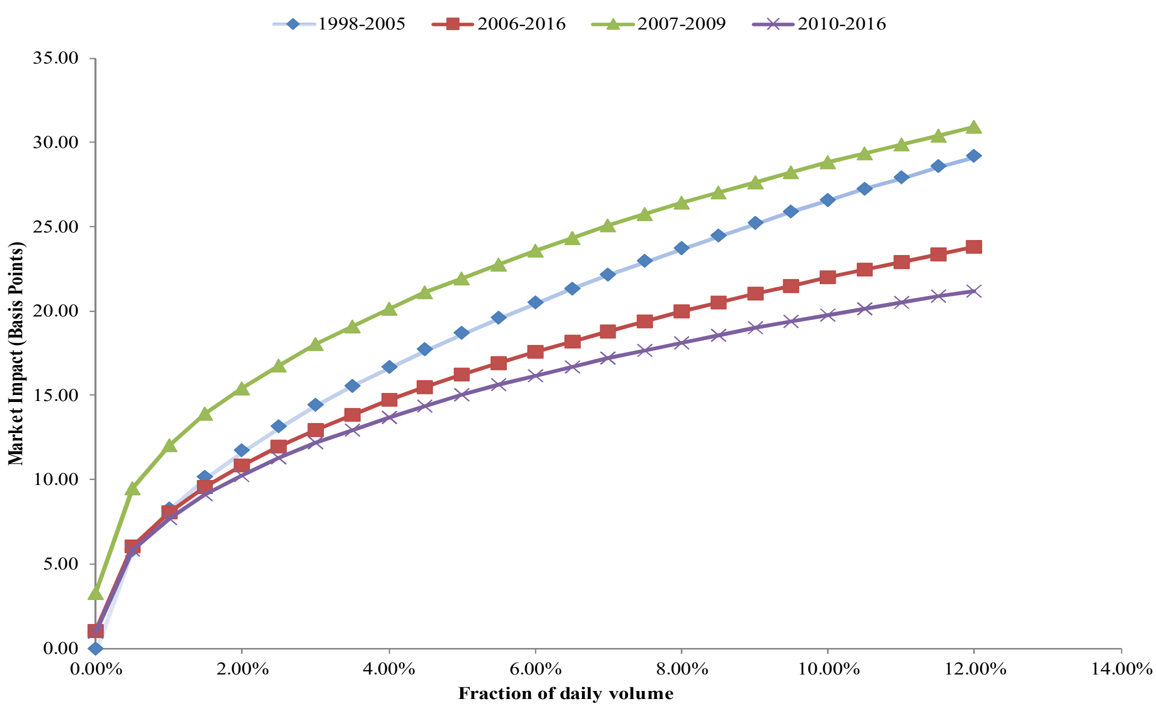


Table VIII
Comparison to Other Trading Cost and Liquidity Measures from the Literature

This table reports estimates of the regressions in Table VII of the live executed trades' market impact (MI), in basis points, on the explanatory variables from Table VII plus additional trading cost and liquidity variables from the literature that include the modified measure of Roll (1984), the liquidity measure of Amihud (2002), the proportion of trading days with zero returns, the effective bid-ask spread from the TAQ database, and an estimate of Kyle's (1985) lambda from the TAQ data. Coefficient estimates and their *t*-statistics (in parentheses) are reported, with bold indicating significance at the 5% level, along with the adjusted *R*-square from the regression and the marginal *R*-square after accounting for beta times the index return and a time trend. Descriptions of each variable are found in the appendix and in Hasbrouck (2009). Panel A reports results for U.S. market impact measures and Panel B for international exchanges. Due to lack of data availability, the estimates involving TAQ data pertain only to the U.S.

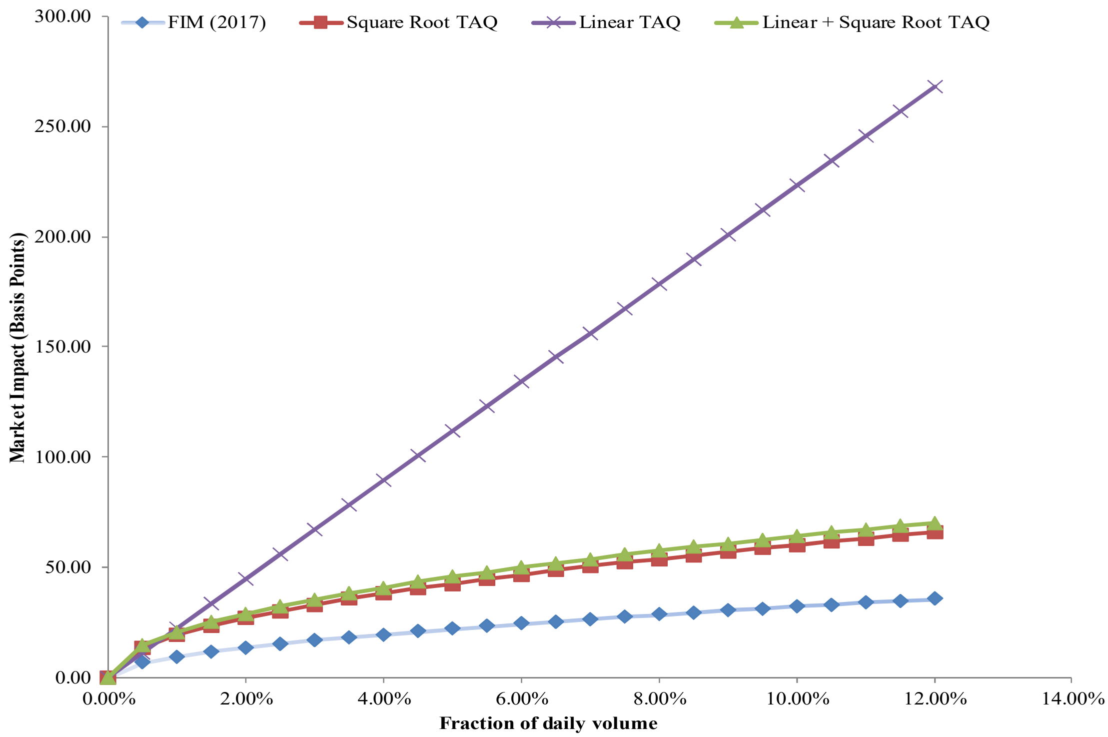
	Panel A: United States											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Modified Roll	**0.03**	0.01									-0.01	0.01
	(3.01)	(1.39)									(-0.80)	(0.82)
Amihud			0.04	-0.03							-0.08	-0.06
			(0.98)	(-0.87)							(-1.89)	(-1.59)
PropZero					102.33	-53.98					15.94	-34.45
					(1.32)	(-0.75)					(0.37)	(-0.83)
TAQ Effective Spread							**0.30**	-0.04			0.05	0.00
							(2.21)	(-0.33)			(0.74)	(-0.06)
TAQ Lambda									**118.53**	0.38	**179.96**	**60.13**
									(4.33)	(0.01)	(5.53)	(2.19)
Beta*IndexRet*buysell	**0.30**	0.01	**0.30**	0.01	**0.30**	0.01	**0.30**	0.01	**0.30**	0.01	**0.30**	0.01
	(13.96)	(0.53)	(13.96)	(0.54)	(13.96)	(0.53)	(13.95)	(0.55)	(13.95)	(0.55)	(13.95)	(0.55)
Time trend	-0.04	0.02	**-0.04**	0.01	**-0.04**	0.02	**-0.04**	0.02	-0.02	0.02	-0.02	0.02
	(-1.83)	(1.03)	(-2.31)	(0.93)	(-2.31)	(0.95)	(-1.98)	(0.99)	(-0.97)	(1.01)	(-0.87)	(1.00)
Log of ME (Billion USD)		-0.13		-0.42		-0.25		-0.22		-0.16		0.05
		(-0.78)		(-1.33)		(-1.31)		(-1.01)		(-0.36)		(0.15)
Fraction of daily volume		0.35		0.41		0.37		0.43		0.43		0.50
		(0.99)		(1.27)		(1.07)		(1.24)		(1.21)		(1.55)
Sqrt(Fraction of daily volume)		**9.04**		**9.07**		**9.03**		**8.86**		**8.86**		**8.88**
		(8.40)		(8.38)		(8.32)		(8.01)		(7.87)		(7.88)
Idiosyncratic Volatility		**0.32**		**0.33**		**0.32**		**0.33**		**0.32**		**0.29**
		(7.85)		(8.39)		(8.05)		(7.83)		(7.41)		(7.05)
Vix		0.12		**0.12**		**0.12**		**0.13**		**0.13**		0.10
		(1.94)		(2.03)		(2.00)		(2.05)		(2.00)		(1.82)
DGTW Ret*buysell		**0.27**		**0.27**		**0.27**		**0.27**		**0.27**		**0.27**
		(22.20)		(22.27)		(22.23)		(22.11)		(22.12)		(22.17)
Adj. R^2	0.1154	0.1560	0.1155	0.1561	0.1154	0.1561	0.1155	0.1559	0.1161	0.1559	0.1163	0.1561
Adj. R^2 after beta and trend	0.0001	0.0407	0.0002	0.0408	0.0001	0.0408	0.0002	0.0406	0.0008	0.0406	0.0010	0.0408

Table VIII (continued)
Comparison to Other Trading Cost and Liquidity Measures from the Literature

	Panel B: International							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Modified Roll	**0.03**	0.01					**0.02**	0.01
	(3.10)	(1.66)					(2.17)	(1.38)
Amihud			**0.21**	**0.06**			**0.20**	**0.06**
			(11.17)	(4.82)			(10.89)	(4.77)
PropZero					**38.65**	16.07	3.46	12.22
					(2.70)	(1.36)	(0.25)	(1.05)
Beta*IndexRet*buysell	**0.22**	**-0.05**	**0.22**	**-0.05**	**0.22**	**-0.05**	**0.22**	**-0.05**
	(21.20)	(-6.63)	(21.20)	(-6.63)	(21.20)	(-6.63)	(21.21)	(-6.63)
Time trend	**-0.07**	**-0.03**	**-0.07**	**-0.03**	**-0.08**	**-0.03**	**-0.06**	**-0.03**
	(-4.93)	(-2.36)	(-4.44)	(-2.35)	(-5.64)	(-2.65)	(-4.81)	(-2.59)
Log of ME (Billion USD)		**-1.42**		**-0.90**		**-1.35**		**-0.89**
		(-10.08)		(-5.02)		(-8.98)		(-4.87)
Fraction of daily volume		0.19		0.18		0.19		0.18
		(1.38)		(1.30)		(1.38)		(1.30)
Sqrt(Fraction of daily volume)		**6.81**		**6.68**		**6.83**		**6.67**
		(13.81)		(13.58)		(13.86)		(13.55)
Idiosyncratic Volatility		**0.27**		**0.26**		**0.28**		**0.26**
		(10.06)		(9.29)		(10.08)		(9.34)
Vix		**0.18**		**0.18**		**0.18**		**0.18**
		(2.98)		(2.98)		(3.01)		(3.03)
DGTW Ret*buysell		**0.27**		**0.27**		**0.27**		**0.27**
		(47.63)		(47.61)		(47.63)		(47.62)
Adj. R^2	0.0921	0.1532	0.0933	0.1533	0.0920	0.1532	0.0933	0.1533
Adj. R^2 after beta and trend	0.0000	0.0612	0.0012	0.0613	0.0000	0.0612	0.0013	0.0613

Figure 7. Comparison to Other Trading Cost Models

The first graph plots the price impact function – market impact as a function of the fraction of daily volume traded (%DTV) – from our calibrated trading cost model in Table VII (column (5)) using live trades from the execution database versus the price impact functions estimated from TAQ data over the same sample period using the linear, square root, and linear + square root functions of %DTV. All of the calibrated parameters are held constant except those on trade size (%DTV). The second graph removes the linear model from TAQ data for scale.



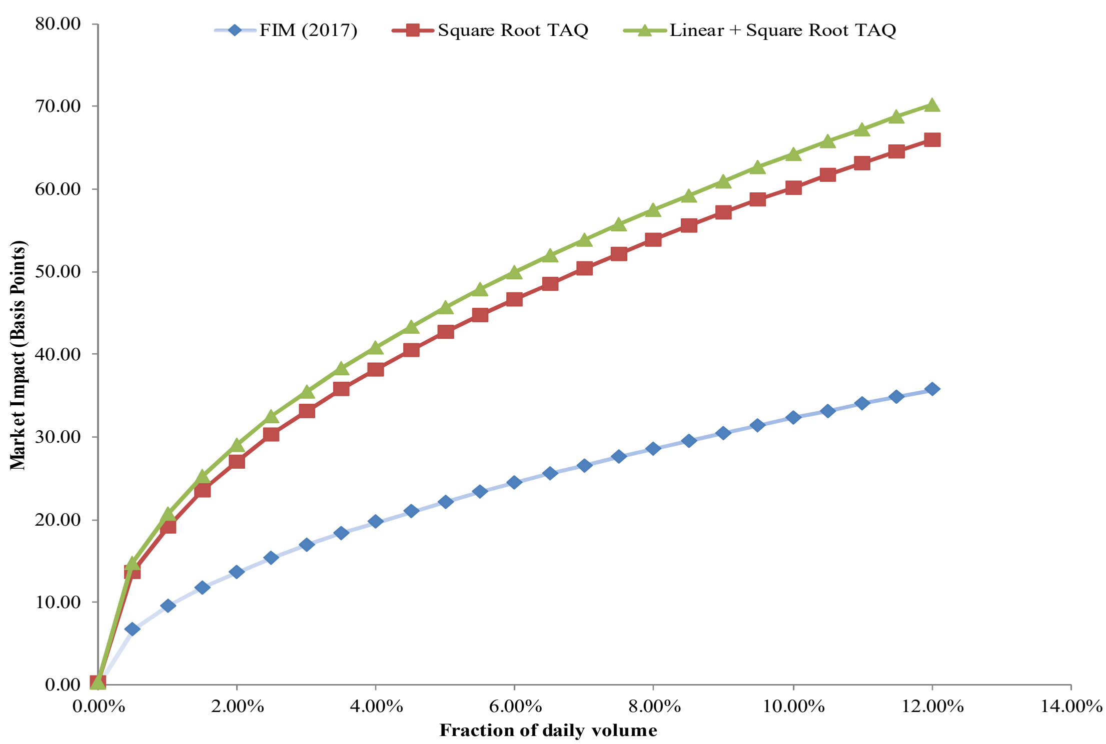


Table IX
Out of Sample Test of the Trading Cost Model on S&P 500 and Russell 2000 Index Funds

The table reports an out-of-sample test of our calibrated trading cost model using live trades from our execution database ("FIM, trade data") to forecast the expected trading costs of two well-known indices – the S&P 500 and the Russell 2000 – where we can compare them to an approximation of the realized costs of trading these two indices by looking at Vanguard and iShares ETFs that trade these indices in live funds. The table also reports results from the same exercise for a model that is linear in %DTV and is estimated using TAQ data ("linear, TAQ"), which is a model used in the literature (see Korajczyk and Sadka (2004)), as well a model that is a square root function of %DTV (following Kyle (1985)) using the TAQ data (Square root, TAQ), and a model that combines both linear and square root terms for %DTV using TAQ data. Since a key input into both models is the trade size, which is a function of the size of the fund being traded, we use a total fund size (NAV) equal to the approximate total dollars being benchmarked against each index (supplied for the S&P 500 by Dow Jones and for the Russell 2000 by Russell). Panel A reports the calibration data on the indices and Panel B reports the estimated trading costs under each model, including an approximation of the actual annualized trading costs of the indices using the Vanguard S&P 500 Index Admiral Fund and the iShares Russell 2000 ETF, where trading costs are estimated from expense ratios and differences in returns between the funds and the index as detailed in the appendix.

	S&P 500 Index	Russell 2000 Index
Panel A: Calibration data		
Fund size ($ Million)	7,800,000	856,000
Turnover (annual)	4.0%	17.0%
Number of stocks	500	2000
Number of rebalances	12	12
Average Trade (Million USD)	52.00	6.06
Average daily volume (Million USD)	662.83	14.76
Average fraction of daily volumne traded	7.85%	41.1%
Panel B: Estimated total cost (annualized bps)		
ETF (live fund data) actual costs	4.72	12.87
Trade data, FIM	4.81	12.36
Linear TAQ	29.79	155.79
Square root TAQ	9.07	20.70
Linear + square root TAQ	9.69	21.78

Table X
Comparison of Trading Costs to Other Sources – By Trade Size, Time, and Country

Panel A reports average realized price impact trading costs across different trade sizes (% of daily trading volume) from three different brokers: ITG, Deutsche Bank (DB), and JP Morgan (JPM). Six different levels of trade size are reported: 0.25-0.5% DTV, 0.5-1% DTV, 1-1.5% DTV, 1.5-2% DTV, 2-5% DTV, and 5-10% DTV. The common sample period for these trading costs is 2008 to 2011. The last four columns report estimated rather than actual trading costs at the various trade sizes from the regression specification of Table VII column (5), calibrated using the live trade database ("Trade data, FIM model"), as well as estimated trading costs at every trade size from using TAQ data, with a linear function, square root function, and linear plus square root function of %DTV. Panel B reports a comparison of average trading costs over time, measured from the execution database, versus average costs from ANcerno (formerly Abel Noser), a consulting firm that monitors institutional investor execution costs, whose database contains equity trades from 750 institutions executing 104 million trades. We obtain these cost estimates from two studies that use the ANcerno data: 1) Anand, Irvine, Puckett, and Venkataraman (2012) who examine average realized costs for the period 1999 to 2008, and separately from 1999 to 2006 and 2007 to 2008 for U.S. stocks. The average trade size as a percentage of daily trading volume (%DTV) in their study is 2.4%, hence we report average market impact costs from our execution database for comparison for trades of 2.4% of DTV. 2) Di Maggio, Franzoni, Kermani, and Sommavilla (2017), who report price impact and trading fees for the top 30 brokers in the ANcerno database from 1999 to 2014. The average trade size in their sample is 0.5% of DTV, so we report average market impact costs from our execution database for comparison for trades of 0.5% of DTV.

Panel A: Comparison of Trading Costs Across Trade Size (%DTV)

Average costs from 2008 -2011	Realized average trading costs				Estimated trading costs			
%DTV	**ITG**	**DB**	**JPM**	**Average**	**Trade data, FIM model**	**Linear TAQ**	**Square Root TAQ**	**Linear + Square Root TAQ**
0.25-0.50%	4.00	4.50	8.00	5.50	5.85	8.65	11.72	12.61
0.50-1.0%	8.00	10.00	14.00	10.67	8.24	17.01	16.46	17.70
1.0-1.5%	10.00	13.00	16.00	13.00	10.75	28.16	21.37	22.96
1.5-2.0%	10.00	13.00	16.00	13.00	12.79	39.31	25.29	27.16
2.0-5.0%	17.00	17.50		17.25	17.97	78.34	34.88	37.36
5.0-10.0%	22.00			22.00	27.28	167.55	51.45	54.93

Panel B: Comparison of Trading Costs Over Time

	Execution database average costs	**ANcerno database average costs**
	MI	MI
Avg. trade size = 2.4% DTV		
1999-2008	15.4	24.1
2007-2008	18.2	24.5
1999-2006	28.7	24.0
Avg. trade size = 0.5% DTV		
1999-2014	7.10	10.52

Figure 8. Time-Series of the Distribution of Trading Costs from the Market Impact Model

Each month we use the market impact model from Table VII, calibrated with live trading data, to estimate costs across stocks and plot the 10th, median, and 90th percentile of costs over time for two levels of trade size: 1% of daily trading volume (DTV) and 5% of DTV. The graph plots the time series of costs estimated from the model over both the period for which we have live trading data from 1998 to 2016 as well as back to 1984 using the model coefficients and the characteristics of firms and market conditions (such as the VIX and a time trend) to extrapolate costs prior to the live trading period (shaded region).

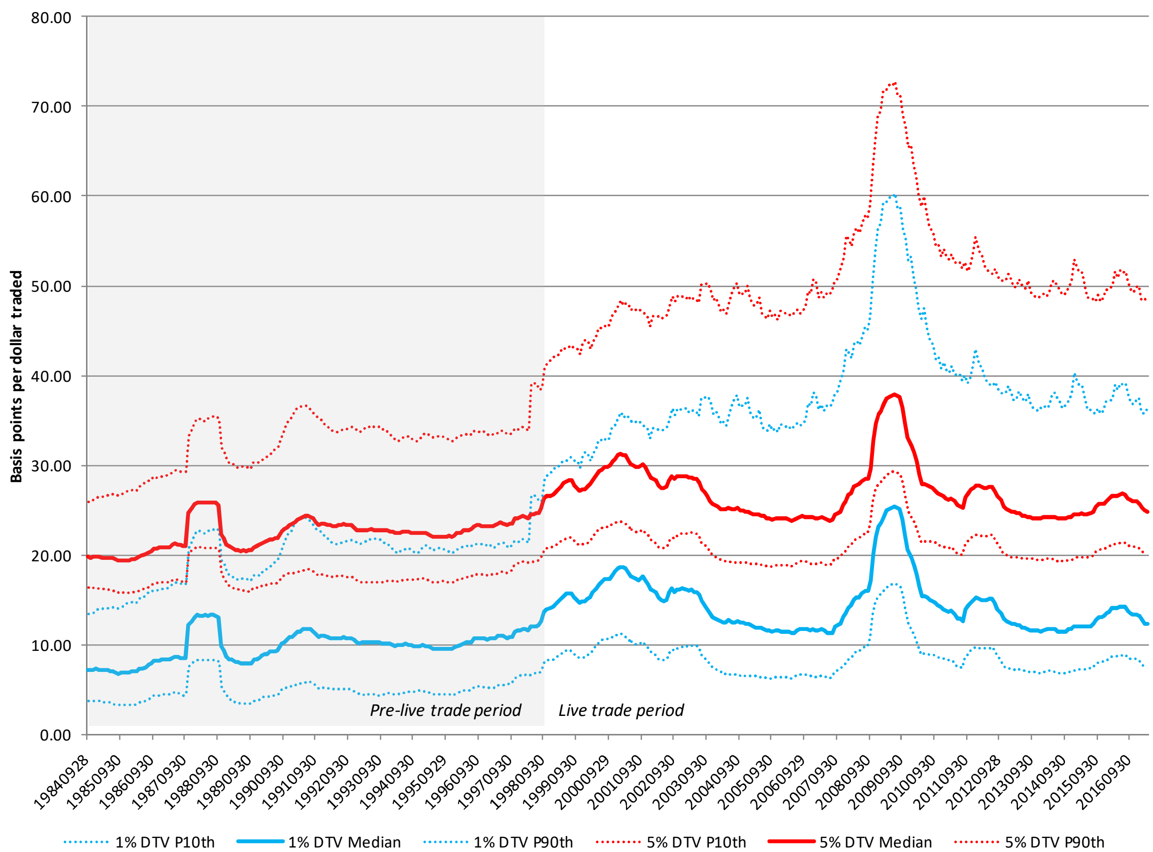
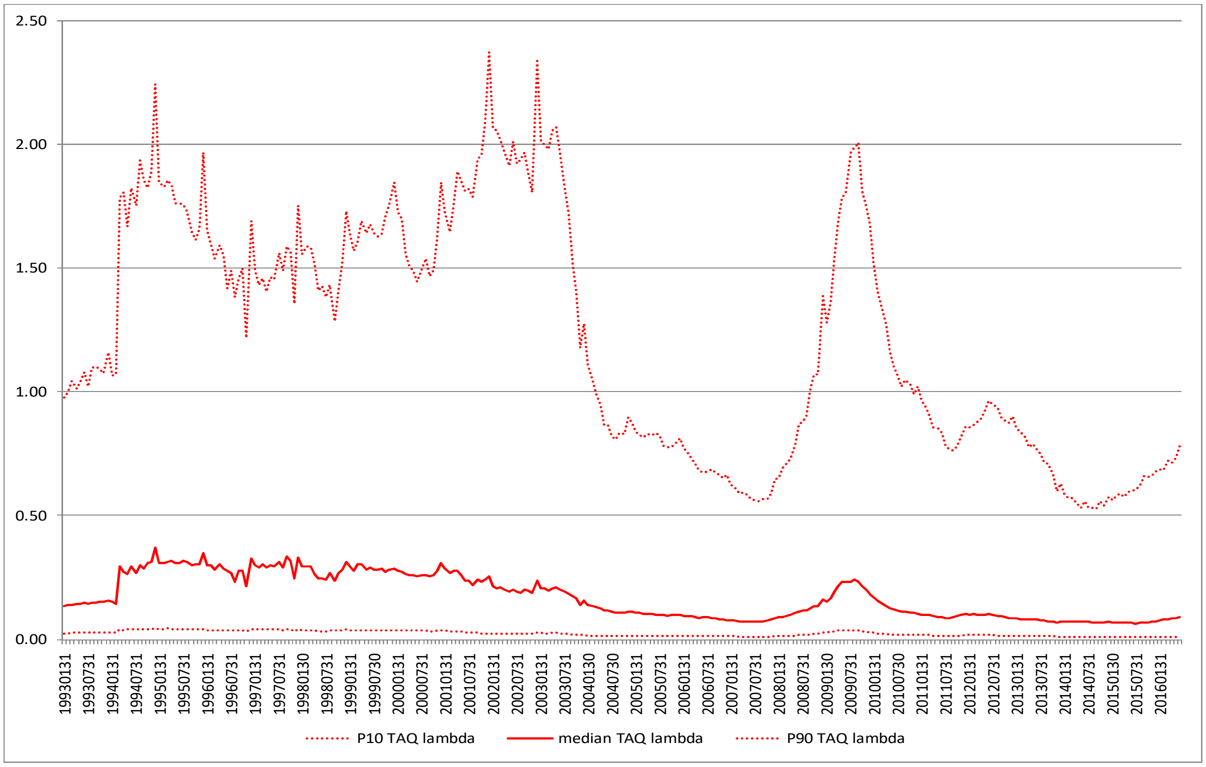


Figure 9. Time-Series of the Distribution of Effective Spreads and Lambda from the TAQ Data

Each month we use the TAQ data to estimate effective spreads and Kyle's (1985) lambda across stocks following the equations in the Appendix. We plot the 10th, median, and 90th percentile of lambda in the first plot and of the effective spread in the second plot over time from March 1993 to June 2016.



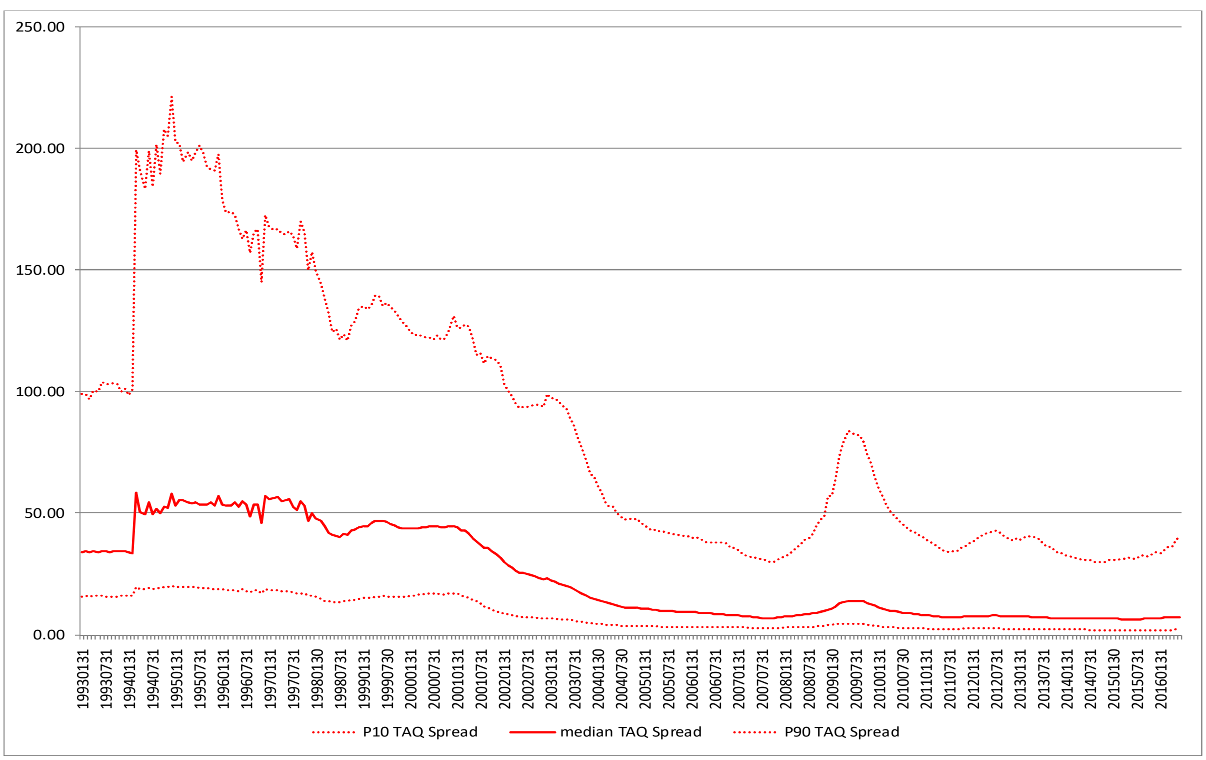


Table XI

Summary Statistics of the Distribution of Estimated Trading Costs: Live Trade Data vs. TAQ

The table reports the time-series average and time-series standard deviation of the 10^{th}, median, and 90^{th} percentiles of the cross-sectional distribution of estimated trading costs across stocks from our market impact model in Table VII calibrated using live trading data (and assuming a trade size of 1% of DTV), the TAQ effective spread used by Novy-Marx and Vehlikov (2017), and Kyle's (1985) lambda estimated from TAQ data to measure price impact, used by Korajczyk and Sadka (2004). Statistics are calculated over the live trading period (1998 to 2016) from which our trading cost model is calibrated and pertain only to the U.S., where TAQ data is available. The final two rows of the table report the time-series correlations between the estimated costs from our calibrated model and the TAQ effective spread and lambda for each percentile of the cross-sectional distribution.

Estimated Trading Costs			
	10th %	Median	90th %
FIM market impact model (1% DTV)			
Time-series average	8.68	14.55	38.09
Time-series stdev	2.27	3.15	6.48
TAQ spread only			
Time-series average	6.56	19.74	78.94
Time-series stdev	5.40	15.65	39.92
TAQ lambda			
Time-series average	0.02	0.21	1.55
Time-series stdev	0.01	0.18	0.73
Time-series correlation (MI model, TAQ Spread)	0.27	0.35	-0.25
Time-series correlation (MI model, TAQ Lambda)	0.68	0.59	0.12

Table XII
Time-Series Correlations Across Trading Cost Measures

The table reports time-series correlations of various trading cost measures. We examine the correlation between the estimated trading cost from our market impact model in Table VII calibrated using live trading data with those from the literature such as Amihud's (2002) measure, the modified Roll measure, the proportion of zero return trading days, the effective bid-ask spread from TAQ, and Kyle's (1985) lambda estimated from TAQ data, which are used to estimate trading costs in Hasbrouck (2009), Korajczyk and Sadka (2004), and Novy-Marx and Vehlikov (2017). Trading costs are estimated for a trade size of 1% daily trading volume (DTV), which is the average daily dollar volume over the prior 30 days. Each month, each trading cost measures is calculated across all stocks and the median is recorded for each measure in each month. The time-series correlation in the medians for each measure are then calculated over the entire common sample period from August 1998 to June 2016. Panel A reports results for the U.S., where TAQ data is available, and Panel B for all non-U.S. countries, averaged across the 20 international countries. The last four columns of both panels report the time-series correlations of each median trading cost measure with average firm size, average idiosyncratic volatility, the VIX level, and average dollar trading volume each month.

MI	Amihud	Roll	Prop zero	TAQ spread	TAQ lambda	log(Size)	Idio. Vol.	Vix	DTV
Panel A: U.S. time-series correlations for median trading costs									
1.00	0.51	0.74	0.75	0.35	0.58	-0.59	0.65	0.97	0.08
	1.00	0.10	0.49	-0.40	-0.18	-0.09	0.59	0.54	0.51
		1.00	0.58	0.53	0.72	-0.79	0.50	0.69	-0.51
			1.00	0.28	0.46	-0.57	0.49	0.73	-0.03
				1.00	0.95	-0.18	-0.26	0.28	-0.63
					1.00	-0.37	-0.04	0.52	-0.61
Panel B: International time-series correlations for median trading costs									
1.00	0.61	0.42	-0.57			-0.57	0.89	0.78	-0.24
	1.00	0.57	0.07			-0.63	0.82	0.76	-0.17
		1.00	-0.12			-0.64	0.57	0.39	0.11
			1.00			0.34	-0.33	-0.11	-0.03

Figure 10. Time-Series Correlation of Median, 10th and 90th Percentiles of Market Impact with Other Trading Cost Measures

Reported are the time-series correlation of the median, 10th, and 90th percentiles of the cross-sectional distribution of market impact measures each month with the median, 10th, and 90th percentiles of the cross-sectional distribution of other trading cost measures each month. The market impact measure is the estimated trading cost from our market impact model in Table VII calibrated using live trading data. The other trading cost measures are those from the literature such as Amihud's (2002) measure, the modified Roll measure, the proportion of zero return trading days, the effective bid-ask spread from TAQ, and Kyle's (1985) lambda estimated from TAQ data. Trading costs are estimated for a trade size of 1% daily trading volume (DTV), which is the average daily dollar volume over the prior 30 days. Each month, each trading cost measure is calculated across all stocks and the median, 10th and 90th percentile of the cross-sectional distribution is recorded each month. The time-series correlations of the medians, 10th and 90th percentiles for each measure are then calculated over the entire common sample period from August 1998 to June 2016 and plotted below.

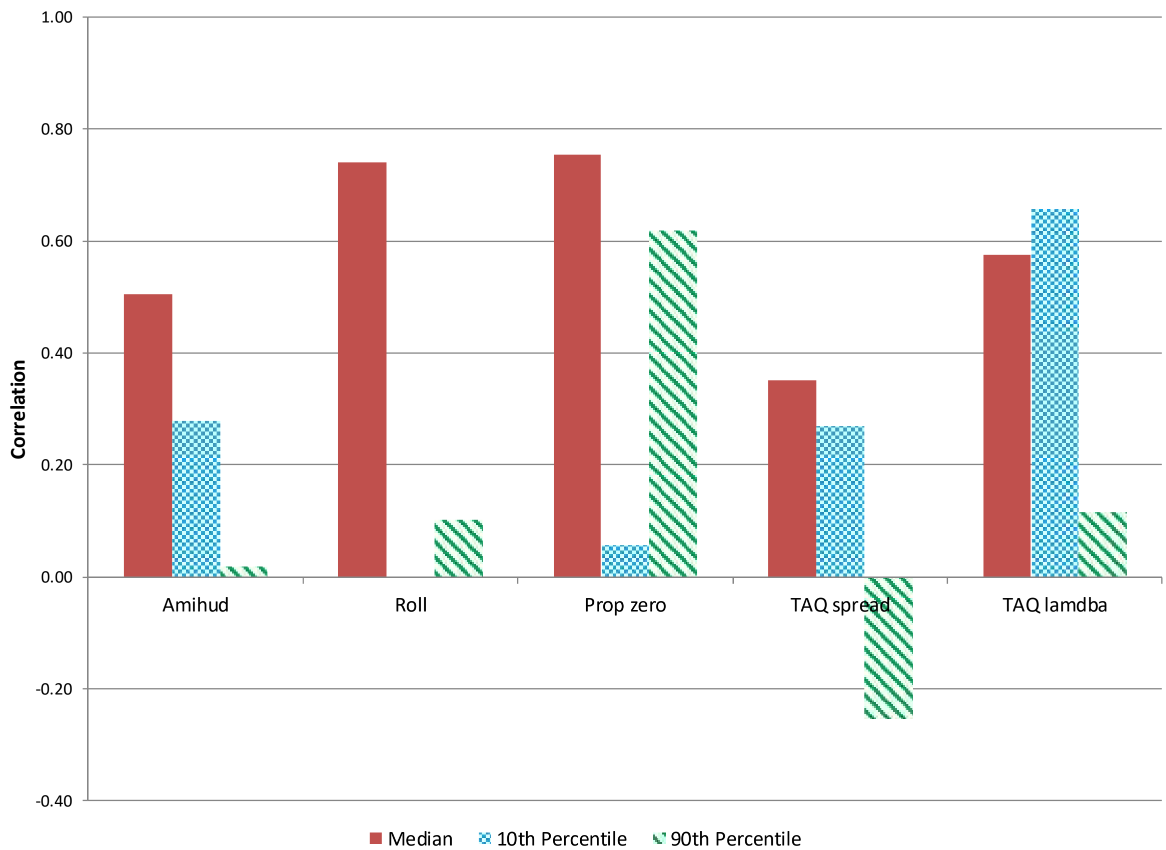


Table XIII
Time-Series Correlations Within Trading Cost Measures

Reported are time-series correlations between the 10th, median, and 90th percentiles of the cross-sectional distribution of each trading cost measure separately. Correlations between the percentiles within a given trading cost measure are reported for the U.S. and for the other 20 countries (internationally, excluding the U.S., averaged across the other countries) highlighted in blue, and correlations between the U.S. and international markets for the same measure are highlighted in red. Panel A reports correlations within our market impact measure based on the model in Table VII calibrated using live trading data. Panels B, C, D, E, and F report correlations for Amihud's (2002) measure, modified Roll, the proportion of zero return trading days, TAQ effective spread, and TAQ Kyle's (1985) lambda, respectively. Results are reported over the live trading period (1998 to 2016) from which our trading cost model is calibrated as well as over the longest common sample period available for each measure in the U.S. and internationally, which is 1984 to 2016 for every measure, except those using the TAQ data, in which case the sample period is 1993 to 2016, if we project the market impact model back in time prior to 1998. TAQ data pertains only to the U.S.

Time-series correlations of percentiles for each measure											
1998 to 2016						1984 to 2016					
U.S.			International			U.S.			International		
10th %	Median	90th %	10th %	Median	90th %	10th %	Median	90th %	10th %	Median	90th %
Panel A: Market impact											
1.00	0.99	0.69	0.73	0.77	0.86	1.00	0.99	0.86	0.37	0.59	0.82
	1.00	0.65	0.80	0.82	0.88		1.00	0.83	0.49	0.68	0.84
		1.00	0.19	0.26	0.59			1.00	-0.07	0.20	0.70
			1.00	0.99	0.87				1.00	0.95	0.56
				1.00	0.91					1.00	0.76
					1.00						1.00
Panel B: Amihud											
1.00	-0.35	-0.14	0.83	0.16	-0.25	1.00	-0.43	-0.32	0.81	0.40	-0.10
	1.00	0.02	-0.18	0.74	0.44		1.00	0.21	-0.28	0.19	0.35
		1.00	-0.12	-0.03	-0.09			1.00	-0.24	-0.11	0.03
			1.00	0.48	0.02				1.00	0.78	0.36
				1.00	0.36					1.00	0.68
					1.00						1.00
Panel C: Modified Roll											
1.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	1.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
	1.00	0.22	*0.00*	0.56	0.77		1.00	0.58	*0.00*	0.30	0.63
		1.00	*0.00*	0.01	0.27			1.00	*0.00*	0.49	0.76
			0.00	*0.00*	*0.00*				*0.00*	*0.00*	*0.00*
				1.00	0.69					1.00	0.74
					1.00						1.00
Panel D: Proportion of zero return days											
1.00	0.21	-0.81	-0.45	-0.46	-0.59	1.00	0.85	-0.45	-0.28	-0.43	-0.66
	1.00	0.07	-0.18	-0.05	-0.13		1.00	-0.18	-0.20	-0.32	-0.53
		1.00	0.60	0.64	0.73			1.00	0.36	0.37	0.63
			1.00	0.96	0.93				1.00	0.91	0.66
				1.00	0.95					1.00	0.81
					1.00						1.00
Panel E: TAQ effective spread											
1.00	0.99	0.93				1.00	0.98	0.95			
	1.00	0.96					1.00	0.97			
		1.00						1.00			
Panel F: TAQ lambda											
1.00	0.94	0.81				1.00	0.95	0.74			
	1.00	0.86					1.00	0.82			
		1.00						1.00			

Trading Costs

Appendix

ANDREA FRAZZINI, RONEN ISRAEL, AND TOBIAS J. MOSKOWITZ

This Appendix contains additional analysis and results.

- Table A1 reports additional summary statistics of our merged CRSP/Xpressfeed/Compustat global data.
- Table A2 reports coverage of our Trade Execution Data.
- Table A3 reports average market impact of our Trade Execution data based on averages pooled across stocks and time.
- Table A4 reports a list of market structure feature across the 21 country stock exchanges for which we have live trading data.
- Table A5 reports additional regression results for our price impact model that includes a pre-decimalization dummy variable.
- Table A6 reports price impact trading cost model estimates for different time periods.
- Table A7 reports details of how other trading cost and liquidity measures in the literature are calculated.
- Table A8 reports details of how we approximate actual trading costs from live ETF returns for the S&P 500 and Russell 2000 indices.
- Figure A1 plots log-log graph of market impact on trade size (percentage of daily volume traded).

Table A1
Summary Statistics on Coverage of the CRSP-Compustat/XpressFeed Global data

The table reports summary statistics on the equities for each country covered by the trading execution database. The sample includes all common stocks on merge CRSP/ Compustat North America and Compustat Global datasets. Reported are the number of firms in the CRSP and Compustat data as well as their median firm size (in USD billions) in comparison to the number of firms covered by the trade execution database and their median size. The last two columns report the fraction of the market and fraction of the number of firms covered by the trading execution database relative to the CRSP/ Compustat North America and Compustat Global common stock universes.

Code	Country	Market		Trade Execution data			
		Number of firms	Median Firm size (Billion-USD)	Number of firms	Median Firm size (Billion-USD)	Fraction of market	Fraction of firms
AUS	Australia	3,040	0.02	354	4.51	0.761	0.116
AUT	Austria	176	0.13	31	5.60	0.427	0.176
BEL	Belgium	308	0.14	52	7.94	0.685	0.169
CAN	Canada	5,380	0.01	415	7.48	0.692	0.077
CHE	Switzerland	445	0.32	123	7.59	0.766	0.276
DEU	Germany	1,674	0.05	186	8.26	0.684	0.111
DNK	Denmark	354	0.07	52	5.78	0.752	0.147
ESP	Spain	360	0.41	91	8.15	0.805	0.253
FIN	Finland	227	0.16	45	4.97	0.729	0.198
FRA	France	1,675	0.08	198	9.84	0.756	0.118
GBR	United Kingdom	5,270	0.06	556	6.18	0.808	0.106
HKG	Hong Kong	2,100	0.12	527	7.11	0.575	0.251
ISR	Israel	713	0.06	39	1.43	0.176	0.055
ITA	Italy	560	0.23	159	5.61	0.736	0.284
JPN	Japan	5,322	0.11	1,302	3.98	0.712	0.245
NLD	Netherlands	349	0.30	86	7.74	0.767	0.246
NOR	Norway	456	0.12	51	6.09	0.578	0.112
PRT	Portugal	123	0.14	21	8.04	0.422	0.171
SGP	Singapore	1,067	0.06	153	4.93	0.664	0.143
SWE	Sweden	1,026	0.06	118	6.82	0.758	0.115
USA	United States	23,860	0.03	4,984	4.26	0.718	0.209

Table A2

Fraction of Market Cap Covered by Trading Database By Region Over Time

The table reports the fraction of market cap covered by the trade execution database by region over time from 1998 to 2016, reported separately for the entire database across all developed regions, the U.S., Europe, Japan, UK, Canada, Australia, and the Pacific.

Fraction of market cap by region over time								
	All	USA	Europe	Japan	UK	Canada	Australia	Pacific
1998	0.17	0.18	0.14		0.49			
1999	0.31	0.36	0.35		0.65			
2000	0.47	0.40	0.54	0.56	0.74		0.26	0.43
2001	0.50	0.45	0.58	0.52	0.79		0.43	0.32
2002	0.56	0.49	0.59	0.64	0.83	0.56	0.69	0.46
2003	0.69	0.71	0.65	0.68	0.80	0.61	0.70	0.54
2004	0.71	0.72	0.67	0.70	0.82	0.63	0.69	0.54
2005	0.75	0.81	0.67	0.72	0.80	0.66	0.73	0.54
2006	0.77	0.81	0.72	0.81	0.78	0.74	0.74	0.54
2007	0.78	0.80	0.76	0.85	0.82	0.77	0.79	0.61
2008	0.81	0.83	0.76	0.86	0.83	0.81	0.87	0.67
2009	0.82	0.83	0.81	0.87	0.89	0.79	0.88	0.68
2010	0.81	0.82	0.80	0.85	0.86	0.73	0.83	0.68
2011	0.81	0.82	0.81	0.86	0.87	0.79	0.85	0.49
2012	0.82	0.83	0.82	0.86	0.88	0.82	0.87	0.58
2013	0.82	0.84	0.82	0.88	0.87	0.85	0.88	0.42
2014	0.83	0.83	0.81	0.86	0.85	0.84	0.87	0.78
2015	0.81	0.80	0.81	0.83	0.84	0.83	0.82	0.74
2016	0.81	0.81	0.79	0.85	0.85	0.80	0.79	0.67
Mean	0.69	0.69	0.68	0.78	0.80	0.75	0.75	0.57
Median	0.78	0.81	0.76	0.85	0.83	0.79	0.79	0.54

Table A3
Trade Execution Data, Market Impact – Pooled Means

The table reports average market impact (MI) and implementation shortfall (IS) using pooled means across the entire database across all traded stocks and time. We compute average, median and weighted (by dollar trades) average cost ("vw_mean") of all trades in the database. All available developed market equity transactions (cash equities and equity swaps) in our data are included between August 1998 and June 2016. The distinction between large cap and small cap is based on the portfolio's benchmark. The distinction between "long-short" and "long only" trades is based on the portfolio in which those trades resided, where relaxed constraint portfolios (130-30 and 140-40) are classified as "long only". Market impact and implementation shortfall are in basis points (annualized). The top panel reports results for the full sample of trades from 1998 to 2016 and the bottom panel for the more recent sample from 2006 to 2016.

Pooled means	All	By region			By size		By portfolio type	
		Nyse-Amex	Nasdaq	International	Large cap	Small cap	Long-short	Long only
Full sample: 199808 - 200616								
MI mean	8.90	7.11	9.10	10.04	8.00	17.97	7.62	11.96
MI median	4.69	3.15	4.63	6.04	4.13	11.64	3.79	7.03
MI vw mean	16.11	13.87	16.80	17.51	15.83	23.05	17.59	13.42
IS mean	9.82	7.80	9.67	11.21	8.89	19.14	8.82	12.21
IS median	6.79	4.83	6.32	8.45	6.14	14.59	5.77	9.56
IS vw mean	17.35	14.55	18.68	18.98	17.04	25.10	19.30	13.80
Recent sample: 200606 - 201606								
MI mean	8.78	7.04	9.01	9.88	7.86	18.14	7.52	11.89
MI median	4.61	3.10	4.57	5.94	4.04	11.71	3.72	6.98
MI vw mean	15.50	13.66	15.56	16.80	15.19	23.02	16.68	13.40
IS mean	9.70	7.72	9.54	11.08	8.76	19.19	8.73	12.10
IS median	6.70	4.76	6.21	8.37	6.03	14.67	5.68	9.53
IS vw mean	16.74	14.30	17.34	18.32	16.40	25.07	18.40	13.78

Table A4
Market Structure Features Across Country Stock Exchanges

The table lists the different market structure features across stock exchanges for the 21 countries we examine in our trading sample. Fragmentation represents whether securities are traded on multiple venues within a country, and number of LIT venues indicates the number of LIT pools or markets available to trade within a country, where a LIT market displays bid and ask prices at which market participants are willing to trade (the opposite of a "dark" pool). Order protection rules are provisions that investors receive an execution price equivalent to what is being quoted on any other exchanges within the country where the security is traded, including OTC markets potentially. Separate odd lot handling indicates that odd lot shares are traded separately on a different venue from the exchange (e.g., OTC), where an odd lot is the residual shares from the standardized lot size for the exchange (e.g., for 100 share lots, a trade of 115 shares would have a 15 odd lot trade executed separately in a country that separates odd lot handling). *Additional history on Japan uptick rule: Prior to 3/6/2002: Short sell orders prohibited at a price below the last traded price. 3/6/2002 to 11/5/2013: the price of short sale must be at a price above the last traded price of the security if that price was lower than the price in the previous trade, or at the last traded price if that price was higher than the price in the previous trade.11/5/2013 to present: Same as previous uptick rule, although the price restrictions are not always applied. It becomes active once the traded price drops by 10 percent from the base price of each issue (trigger rule). At the time the price drops by 10 percent from the base price of the day, the price restriction becomes active and it lasts until the end of the next trading day. Please note if the price drops by 10 percent for consecutive days, the price restriction is applied continuously.

Country	Short Sale Uptick Rule	Transaction Tax	Fragmentation (Y/N)	Number of LIT Venues	Order Protection Rule	OTC Trading (Y/N)	Daily Auctions	Intraday Triggered Auctions	Separate Oddlot Handling
Australia	N	N	Y	2	n/a	Y	open/close	news based	n/a everything trades in 1 share
Austria	N	N	Y	4	N	Y	open/intraday/close	Yes, scheduled	N
Belgium	N	N	Y	4	N	Y	open/close	No	N
Canada	N	N	Y	9	Yes - OPR (full depth of book price protection across protected venues, no locking/crossing market)	N	open/close	N	Y
Switzerland	N	Few IFTT names	Y	4	N	Y	open/close	No	N
Germany	N	N	Y	4	N	Y	open/intraday/close	Yes, scheduled	N
Denmark	N	N	Y	4	N	Y	open/close	No	N
Spain	N	N	Y	4	N	Y	open/close	No	N
Finland	N	N	Y	4	N	Y	open/close	No	N
France	N	30 bps (Buys)	Y	4	N	Y	open/close	No	N
United Kingdom	N	50 bps (Buys) + £1 (>10k)	Y	4	N	Y	open/close (UK) open/intraday/close (UK Sets)	Yes, scheduled for UK Sets	MQAT rule (Minimum quantity at touch applies for some LSE segments)
Hong Kong	Yes - Can only join the best offer (or above)	10.77bps	N	1	n/a	Y (stamp still applies for change in beneficial ownership)	open/close (not all stocks have a close auction)	n/a but volatility control mechanism exists	manually traded
Israel	N	1.95 bps (Exch. Fee - all)	N	1	N	Y	open/close	No	N
Italy	N	10 bps (Buys) / OTC 20 bps	Y	4	N	Y	open/close	No	N
Japan*	Uptick rule kicks in when stock moves down more than 10% from previous close. Uptick will then apply for the next trading day as well.	N	Y	3	n/a	Y	am open, am close, pm open, pm close	special quotes (volatility auctions) exist	manually traded
Netherlands	N	N	Y	4	N	Y	open/close	No	N
Norway	N	N	Y	4	N	Y	open/close	No	N
Portugal	N	N	Y	4	N	Y	open/close	No	N
Singapore	N	3.25bps but there is a trading fee of 0.75 bps that the industry is starting to pass back	N	1	n/a	Y (additional charges apply for OTC)	open/close (but looking to introduce lunch break again)	n/a but circuit breaker and cool off period exists	manual routing into odd lot board (illiquid)
Sweden	N	N	Y	4	N	Y	open/close	No	n/a everything trades in 1 share
United States	Rule 201: circuit breaker triggered for a security in which the price declines by 10% or more from the prior day's close. Short sales must trade at a higher price than the bid	N	Y	12	Yes - Reg NMS	Y	open/close	triggered as a result of a trading halt upon resumption	N

Table A5
Price Impact Regression Results with a Pre-Decimalization Dummy

This table shows results from pooled regressions of a trade's market impact (MI), in basis points, on the explanatory variables that include the contemporaneous market return, firm size, volatility and trade size (all measured at order submission), as well as "Beta*IndexRet*buysell," which is the contemporaneous (beta-adjusted) market return using the stock's predicted beta at time of order submission, where "indexRet" is the corresponding index return over the duration of the trade. "BuySell" is a dummy equal to 1 for buy orders and -1 for sell orders. "Size" is equal to the log of 1 plus the market value of equity log(1+ME) in Billion USD. "Fraction of daily volume" is equal to the trade's dollar size divided by the stock's average one-year dollar volume (in %). "Idiosyncratic Volatility" is the volatility of the residuals of a regression of one-year daily stock returns on the corresponding value-weighted benchmark (annualized, in %), VIX is the monthly variance of the CRSP-value weighted index, computed using daily returns (annualized, in %). A linear time trend is also included along with a dummy variable equal to one before U.S. markets instituted decimalization ("pre-decimalization dummy"). Coefficient estimates and their associated *t*-statistics are reported below with 5% statistical significance indicated in bold. Standard errors are clustered by calendar month.

	United States				
	(1)	(2)	(3)	(4)	(5)
Beta*IndexRet*buysell	**0.30**	**0.30**	**0.30**	**0.30**	**0.28**
	(13.96)	(13.96)	(13.96)	(13.95)	(11.07)
Time trend (Jun 1926 = 1)	-0.01	0.00	0.00	0.03	0.02
	(-0.31)	(0.20)	(-0.10)	(1.23)	(0.81)
Pre -decimalization Dummy	**40.05**	**27.52**	**28.39**	**24.24**	**24.67**
	(5.46)	(4.29)	(4.32)	(3.65)	(3.59)
Log of ME (Billion USD)	**-3.32**	**-2.28**	**-1.61**	-0.29	-0.23
	(-14.36)	(-11.09)	(-7.08)	(-1.55)	(-1.24)
Fraction of daily volume		**2.50**	0.51	0.30	-0.58
		(9.96)	(1.49)	(0.90)	(-1.54)
Sqrt(Fraction of daily volume)			**7.94**	**9.35**	**11.25**
			(7.22)	(8.65)	(8.62)
Idiosyncratic Volatility				**0.31**	**0.29**
				(7.37)	(7.01)
Vix				**0.14**	**0.13**
				(2.16)	(2.07)
DGTW-adjusted return*buysell					0.03
					(1.33)
Observations (1,000s)	1,722	1,722	1,722	1,722	1,722
Adjusted R^2	0.117	0.118	0.119	0.119	0.152
Country Fixed Effects	No	No	No	No	No

Table A6
Price Impact Regression for Different Sample Periods

The trading cost model calibrated to our live trading data estimated in Table VII is estimated over different sample periods: 1998 to 2005, 2006 to 2016, 2007-2009, and 2010 to 2016. Reported are estimates of parameters for a trading cost model following the specification of column (5) of Table VII. The estimates are results from pooled regressions of a trade's market impact (MI), in basis points, on the following explanatory variables: "Beta*IndexRet*buysell" is the contemporaneous (beta-adjusted) market return that controls for market movements occurring at the time of the trade, where Beta is the stock's predicted beta at the time of order submission, "IndexRet" is the corresponding index return over the life of the trade (the movement of the market), and "buysell" is a dummy equal to 1 for buy orders and -1 for sell orders. "Time trend" is a linear time trend. Log of ME is equal to the log of 1 plus the market value of equity, where ME is in billions of USD. "Fraction of daily volume" is equal to the trade's dollar size divided by the stock's average one-year dollar volume (in %). Both a linear and square root function of fraction of daily volume are included in the regression. "Idiosyncratic Volatility" is the volatility of the residuals of a regression of one-year daily stock returns on the corresponding value-weighted benchmark (annualized %), "VIX" is the monthly variance of the CRSP-value weighted index, computed using daily returns (annualized %). The DGTW-adjusted return is the return on the stock minus the return on a portfolio of similar stocks matched on size, book-to-market, and momentum (past one year return) from Daniel, et al. (1997), which is interacted with the "buysell" dummy. Country fixed effects are included where indicated, *t*-statistics are shown below the coefficient estimates and 5% statistical significance is indicated in bold. Standard errors are clustered by calendar month. Regression estimates are provided over the full sample, the US only, and outside of the US ("International"), covering all equity trades in 21 countries from August 1998 to June 2016.

	All Sample				United States				International			
Period	1998-2005	2006-2016	2007 - 2009	2010-2016	1998-2005	2006-2016	2007 - 2009	2010-2016	1998-2005	2006-2016	2007 - 2009	2010-2016
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Beta*IndexRet*buysell	0.02	**-0.03**	**-0.02**	**-0.03**	0.03	0.01	0.02	0.01	0.02	**-0.05**	**-0.05**	**-0.06**
	(0.39)	(-3.70)	(-2.46)	(-2.59)	(0.28)	(0.72)	(0.55)	(0.73)	(0.50)	(-6.56)	(-3.55)	(-6.06)
Time trend	0.14	-0.01	-0.20	0.01	0.68	0.00	-0.10	0.01	**-0.58**	-0.02	**-0.33**	0.01
	(0.37)	(-0.52)	(-1.88)	(0.97)	(1.23)	(0.34)	(-0.99)	(0.80)	(-3.55)	(-1.34)	(-2.40)	(0.60)
Log of ME (billion USD)	4.57	**-0.76**	**-1.62**	**-0.84**	8.66	-0.29	-0.60	**-0.35**	-2.27	**-1.42**	**-2.77**	**-1.43**
	(1.36)	(-6.70)	(-3.98)	(-7.82)	(1.87)	(-1.69)	(-1.30)	(-2.34)	(-1.54)	(-10.26)	(-4.51)	(-11.92)
Fraction of daily volume	0.28	0.10	-0.27	0.03	-0.16	0.56	0.27	0.57	0.45	0.08	-0.22	0.02
	(0.79)	(0.74)	(-1.07)	(0.16)	(-0.14)	(1.66)	(0.32)	(1.90)	(1.26)	(0.56)	(-1.13)	(0.09)
Sqrt(fraction of daily volume)	**9.23**	**8.23**	**11.71**	**7.39**	12.84	**8.61**	**12.41**	**7.55**	**5.65**	**7.10**	**9.81**	**6.42**
	(3.01)	(13.93)	(8.34)	(16.75)	(1.67)	(7.90)	(3.52)	(9.63)	(2.02)	(14.18)	(8.80)	(11.45)
Idiosyncratic volatility	**0.70**	**0.28**	**0.27**	**0.24**	**0.86**	**0.32**	**0.30**	**0.30**	0.48	**0.26**	**0.31**	**0.18**
	(3.07)	(10.88)	(3.55)	(11.51)	(3.29)	(7.97)	(2.72)	(10.10)	(1.93)	(9.57)	(3.79)	(7.78)
VIX	0.32	**0.15**	**0.22**	**0.04**	1.37	0.11	0.13	0.05	**-0.88**	**0.19**	**0.31**	0.03
	(0.45)	(2.97)	(2.18)	(2.31)	(1.18)	(1.83)	(0.95)	(1.77)	(-2.31)	(3.07)	(3.67)	(1.18)
DGTW Ret*buysell	**0.42**	**0.27**	**0.27**	**0.26**	**0.44**	**0.27**	**0.27**	**0.25**	**0.39**	**0.27**	**0.27**	**0.26**
	(10.64)	(33.51)	(20.81)	(40.27)	(5.89)	(22.09)	(15.71)	(28.30)	(19.33)	(47.05)	(29.17)	(40.76)
Observations (1,000s)	3,078	3,078	3,078	3,078	1,652	1,652	1,652	1,652	1,426	1,426	1,426	1,426
Adjusted R^2	0.153	0.155	0.167	0.139	0.163	0.157	0.175	0.137	0.146	0.155	0.159	0.145
Country Fixed Effects	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes

Table A7
Definitions of Other Trading Cost and Liquidity Measures from the Literature

Detailed descriptions of other trading cost and liquidity measures from the literature are provided, including the modified measure of Roll (1984), the illiquidity measure of Amihud (2002), the proportion of trading days with zero returns (examined by Hasbrouck (2009) and Goyenko, Holden, and Trzcinka (2009)), the effective bid-ask spread from the TAQ database, and an estimate of price impact from the Kyle (1985) model (Kyle's lambda), estimated from signed trades from the TAQ database. The calculations of these measures follow Hasbrouck (2009).

Measure	Description	Calculation	
TAQ effective spread	The difference between the log transaction price and the log midpoint quote, averaged over all trades that month, weighted by the dollar value of the trade using high frequency TAQ data for each stock in each month over the period January 1993 to December 2016.	$\lvert \log(P_k) - \log(P_k^{mid}) \rvert$	where P_k is the price at trade k and P_k^{mid} is the midpoint quote at the time of trade k.
Modified Roll	The square root of the negative of the autocovariance of the log price change for each trading day of a stock over the last month, where positive autocovariances are set to zero.	$\sqrt{-Cov(\Delta p_t, \Delta p_{t-1})}$	where Δp_t is the log price change at time t.
Amihud (2002)	The absolute value of the daily return on a stock divided by its daily dollar trading volume, averaged over the last month.	$\frac{\lvert ret_t \rvert}{\$volume_t}$	where ret_t is the daily return on a stock at time t and \$volume is the daily dollar volume on day t.
Prop zero	Proportion of trading days in the year that had a zero price change from the previous day.	$\frac{\#days\ ret_t = 0}{\#days}$	where $\#days$ is the number of trading days over the last year and $\#days\ ret_t = 0$ are the number of trading days with zero return.
TAQ lambda	Price impact measure from the Kyle (1985) model based on high frequency TAQ data by regressing the change in price over five minute trading intervals on aggregated signed dollar volume over the same trading interval, where the coefficient from the regression is Kyle's lambda representing the elasticity of price changes with respect to dollar volume traded.	$\Delta p_\tau = \lambda\left(signed\sqrt{\$volume_\tau}\right) + \varepsilon_\tau$	where Δp_t is the log price change over time interval τ (5 minutes) and signed dollar volume is aggregated volume over the same trading interval signed by the order imbalance measure of Lee and Ready (1991), which uses the traded price relative to the midquote price to sign the trade. The coefficient estimate, λ, is the average price impact measure over the prior month of all five-minute trading intervals for the stock from the TAQ database.

Figure A1. Log-Log plot of Price Impact on Fraction of Daily Trading Volume Traded.
Plotted is the log-log plot of actual market impact from live trades against the percentage of daily trading volume traded from live trading data. A best-fit line is also plotted along with the estimated linear regression equation, including the R^2.

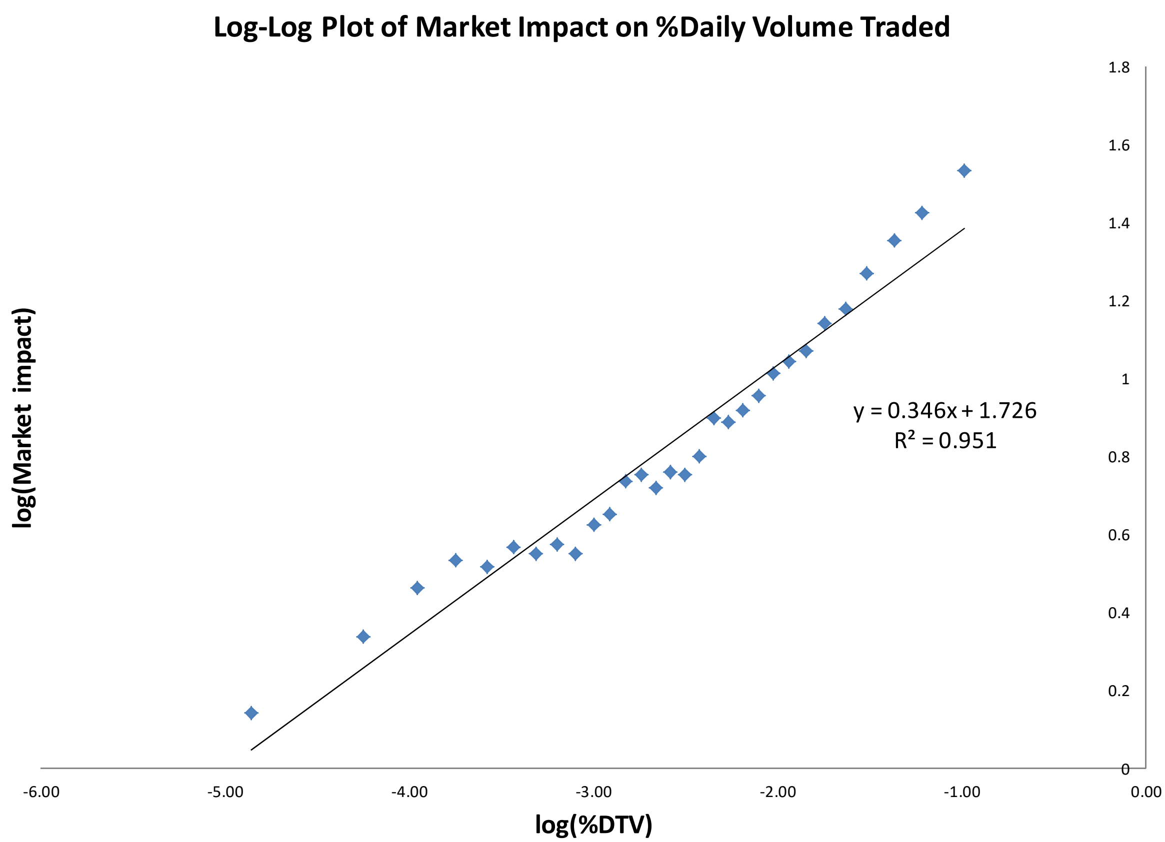


Table A8

Computing Transactions Costs from Passive Index ETFs

Reported are the inputs from exchange traded funds that track passive indices – the S&P 500 and the Russell 2000 – which we use to back out actual transactions costs of trading these funds and compare them to estimates from our model calibrated to live trading data and other models calibrated to TAQ data. The table reports the average return, most recent expense ratio, beta to the index, and annual tracking error for the Vanguard 500 Index Admiral Fund and the iShares Russell 2000 ETF. Also reported are the average returns the S&P 500 and Russell 2000 indices themselves. The sample period is from January 2001 to January 2016, where we have a full year of ETF return data. Finally, to calculate transactions costs an estimate of other revenue generated by the fund from trading, the largest candidate being lending fees from the fund lending its shares, must be added back. To gauge the amount of lending fee revenue we use the estimates from Blocher and Whaley (2015), who calculate lending fee returns from a host of ETFs and we apply their average estimates for large cap and small cap funds to the S&P 500 and Russell 2000 ETFs, respectively. Below the table are the details of the calculations that deliver the transaction cost estimates for both index funds.

Name	**Vanguard 500 Index Admiral**	**Russell 2000 ETF**	**S&P 500 Index**	**R2000 Index**
Ticker	VFIAX	IWM	XIUSA04G92	XIUSA000O5
Expense Ratio	0.04%	0.20%		
Time period	1/01 - 6/16	1/01 - 6/16	1/01 - 6/16	1/01 - 6/16
Mean	6.08%	8.85%	6.10%	8.87%
Beta to index	1.00	1.00		
Tracking Error	0.03%	1.34%		
Lending fee from Blocher and Whaley (2015) Table VIII	0.064%	0.310%		
Total cost estimate (bps)	**4.72**	**12.87**		

To measure actual transactions costs from ETFs, we examine the difference in returns between the ETF and the underlying index. Since the ETFs are designed to track the indices, if the funds had zero tracking error, any deviations in returns between the fund and the index will be due to transactions costs and transactions revenue, since the index itself does not have to transact. Of course, as Table A8 shows above, the tracking errors of the funds are not zero, but are quite small.[i] Hence, our calculations should be a close approximation to the actual transactions costs and benefits facing the fund. Transactions costs include the costs of trading, such as price impact and commissions, while the largest benefit from transacting would be the fees funds receive from lending their shares to other investors. The index, of course, does not incur any of these costs nor receive any benefit from transacting.

To approximate the transactions cost of trading the index from ETF returns, we decompose returns as follows.

Decomposition of ETF returns:

$$\begin{aligned} r &= r^h - ER + LF - TC \\ &= r^b + r^\varepsilon - ER + LF - TC \end{aligned}$$

where r is the ETF net return, ER is the expense ratio (expenses incurred by the fund divided by the fund's total assets, and hence expressed as a return), LF are the proceeds from security lending (as a percentage of the fund's assets, also expressed in returns) and r^h is the gross return of the ETF holdings, which, as the second equation shows, can be decomposed into the gross benchmark return r^b and the return from active deviation from the benchmark, $r^\varepsilon = r^b - r^h$. Since index funds seek to track the index, the return from active deviations should average to zero, $E(r^\varepsilon) = 0$, hence:

$$TC = E(r^b) - E(r) - ER + E(LF).$$

We collect data for two ETFs: The Vanguard 500 Admiral Index Fund tracking the S&P 500 (ticker: VFIAX) and the iShares Russell 2000 ETF tracking the Russell 2000 (ticker: IWM). From their most recent annual report we obtain the expense ratio of each fund, as well their reported annualized turnover, which allows us to calculate a transaction cost per year. To estimate the average return deviation from the benchmark we obtain monthly ETF and underlying index returns from Morningstar over the sample period January 2001 (the full calendar year available for both ETFs) to June 2016 (the end of our trading sample). In order to estimate expected lending fee revenue per dollar of assets under management, we use the average lending fee returns for large cap and small cap stocks, respectively for the S&P 500 and Russell 2000, obtained from Blocher and Whaley (2015), Table VIII, who calculate lending fees for all U.S. ETFs and estimate them separately for large and small cap funds.

Calibration model implied total costs

We then compare the costs estimated for the ETFs to those implied by our market impact model calibrated by live trading data, as well as models used in the literature calibrated by TAQ data. To compute the implied trading costs from these calibrated models for the S&P 500 and Russell 2000 indices, we use the following procedure.

For a given fund size (NAV) we compute the average dollar volume per stock at each rebalance as a fraction of the stock's daily volume:

$$x = m/dtv$$

where $m = \left(\frac{turn \cdot NAV}{n}\right)/h$, $turn$ is the fund's annualized turnover, $turn \cdot NAV$ is the fund total trading, n is the number of stocks in the portfolio and h is the number of rebalances per year. We assume monthly rebalancing ($h = 12$) . We compute dtv as the portfolio weighted average daily volume of each stock:

$$dtv = \sum_i w_i v_i$$

where v_i is stock i's one-year median daily volume. For a given x, total costs are computed as

$$TC = turn \cdot MI(x)$$

where $MI(\cdot)$ is the calibrated market impact function from whichever model and data is chosen.

[i] Non-zero tracking error can occur due to flows in and out of the funds that create differences in turnover, slight deviations in the timing of executed trades, special corporate events, and even potentially some active choices by the fund manager to reduce trading costs or increase lending revenue that causes some additional tracking error to the index.