

Market Microstructure Invariants *

Albert S. Kyle

Robert H. Smith School of Business
University of Maryland
akyle@rhsmith.umd.edu

Anna A. Obizhaeva

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

September 10, 2010

Abstract

A model of market microstructure invariance is presented based on the intuition that stocks with high and low levels of trading activity differ in the rate at which the time clock generating trading activity ticks. This model as well as two alternative benchmark models are consistent with traditional adverse selection and inventory models. The model is tested using a database of over 400,000 portfolio transition trades provided by a leading vendor of portfolio transition services. In a pooled cross-section of stocks, holding volatility constant, the model of market microstructure invariance predicts that a one percent increase in trading activity leads to a decrease of two-thirds of one percent in mean portfolio transition order size as a fraction of expected daily volume, leads to no change in the higher moments of the distribution of order size, leads to an increase in market impact costs (measured in basis points) of one-third of one percent for trading a given percentage of expected daily trading volume, and leads to a decrease in bid-ask spreads of one-third of one percent. Compared with the predictions of alternative models, empirical results match closely the predictions of the model of market microstructure invariance, with order size conforming closely to a log-normal distribution. The result is a simple empirical transaction cost formula for stocks, which estimates market impact and bid-ask spread costs as a function of dollar trade size, average daily dollar volume, and volatility.

*We are grateful to Georgios Skoulakis, Mark Loewenstein, and Vish Viswanathan for helpful comments. Obizhaeva is also grateful to the Paul Woolley Center the at London School of Economics for its hospitality as well as Ross McLellan, Simon Myrgren, Sébastien Page, and especially Mark Kritzman for their help.

1 Introduction

When portfolio managers trade stocks, they can be modeled as playing trading games. Trading game models typically make specific assumptions about the consistency of beliefs across traders, the flow of public and private information which informed traders use to trade, the flow of orders from liquidity traders, and auction mechanisms in the context of which market makers compete to take the other side of trades. Some models emphasize adverse selection, such as Treynor (1971) and Kyle (1985); some models emphasize inventory dynamics, such as Grossman and Miller (1988) and Campbell and Kyle (1993); and some models emphasize both, such as Grossman and Stiglitz (1980) and Jiang Wang (1993). The purpose of this paper is to develop and test a model of order flow dynamics, market depth, and bid-ask spreads based on simple invariance principles, without making specific assumptions about how informed traders, liquidity traders, and market makers interact in a specific game-theoretic context.

Our model is based on the intuition that trading time is different from calendar time. For actively traded stocks, a “trading day” corresponds to a much shorter period of clock time than for inactively traded stocks. The idea that financial markets operate in trading time that is different from calendar time is not new. Rather, we continue a long-standing discussion by Mandelbrot and Taylor (1967), Clark (1973), Hasbrouck (1999, 2007), and Dufour and Engle (2000) concerning whether “time” in financial markets should be measured relative to “calendar time,” “volume time,” or “transaction time.”

The theory of market microstructure invariance developed in this paper generates empirical predictions concerning how quantities traded, market impact, and bid-ask spreads vary across stocks with different levels of trading activity. These predictions are tested using data on portfolio transitions provided by a major vendor of portfolio transition services.

1.1 Theories of Market Microstructure Invariants

The theoretical model is based on the simple reduced-form assumption that innovations in the order flow, which we call “bets,” follow a Poisson process whose arrival rate varies across stocks. “Bet size” is a random variable defined as the product of dollar bet value (dollar share price times share quantity) and volatility (percentage standard deviation of returns per day). In a similar spirit, “trading activity” is defined as the product of the arrival rate of bets and the expected absolute value of bet size. While bet size is a theoretical concept which is not directly observable, trading activity corresponds to the empirically observable product of average dollar volume and volatility.

Modigliani-Miller Irrelevance and Time Clock Irrelevance We obtain a trading game invariant by applying two irrelevance principles to the idea that bets follow a Poisson process: “Modigliani-Miller irrelevance” implies that bet size and bet arrival rate are unaffected by stock splits or changes in leverage because prices, quantities, and volatility adjust to offset splits or changes in leverage. “Time-clock irrelevance” implies that speeding up or slowing down the Poisson arrival rate of bets does not change the trading game in a fundamental manner, much like the fundamentals of the game of chess are not changed if the adjustments are made in the amount of time players are given to make moves. Although

speeding up the time clock increases trading activity, both by increasing the number of bets per calendar day and by increasing bet size through increasing volatility, when bet size is adjusted for the speed of the time clock by being divided by the square root of bet arrival rate, the time-clock-adjusted bet size has a constant probability distribution during Modigliani-Miller and time-clock irrelevance transformations. We call this variable the trading game invariant, an example of a market microstructure invariant.

Market Microstructure Invariance as an Empirical Hypothesis. It is important to make a clear distinction between theoretical irrelevancies and empirical hypotheses. Modigliani-Miller irrelevance and time clock irrelevance are theoretical concepts which apply to hypothetical splits, leverage changes, and time clock changes for a given stock at a given time. These irrelevance principles do not themselves imply that the invariant distribution above is the same for all stocks. We convert these irrelevance principles into empirically testable hypotheses by making three empirical hypotheses:

- Trading Game Invariance: The probability distribution of the trading game invariant (defined as ratio of bet size to the square root of bet arrival rate) is the same across stocks and across time for the same stock.
- Market Impact Invariance: For all stocks, the same constant fraction ψ of price volatility results from the linear price impact of bets.
- Bid-Ask Spread Invariance: For all stocks, the expected bid-ask spread cost of a bet is the same fraction ϕ of market impact costs.

We call these three empirical hypotheses “market microstructure invariance.” We call the three variables that remain constant “market microstructure invariants.”

Market microstructure invariance captures the intuition that all trading games are fundamentally the same when compared across stocks and across time, up to some irrelevant Modigliani-Miller and time-clock transformations, none of which change the rules of trading games. Market microstructure invariance implies that the expected price impact cost of a bet and the expected bid-ask spread cost of a bet is the same across different stocks.

Our idea of time clock irrelevance is similar to the approach of Mandelbrot and Taylor (1967), who modeled trading in transaction time. Note, however, that since the concept of a bet is different from a concept of transaction, the time clocks implied by these two processes may be quite different.

The model of market microstructure invariance leads to testable implications concerning how bet size, price impact, and bid-ask spreads vary with the level of trading activity across stocks and across time for the same stock. Note that implications concerning bet size are derived based solely on the hypothesis of trading game invariance, whereas implications concerning transaction costs require additional assumptions:

- Trading game invariance implies that if trading activity is observed to increase by one percent, then the increase in trading activity resulted from an increase in the arrival rate of bets by two-thirds of one percent and an increase in bet size of one-third of one percent. Trading game invariance implies that the shape of the distribution of bet size does not change as the level of trading activity changes.

- Market impact invariance together with trading game invariance implies that increasing trading activity by one percent increases the market impact cost of trading one percent of average daily volume by one-third of one percent, when trading costs are measured in basis points per dollar traded, holding volatility constant.
- Bid-ask spread invariance together with trading game invariance implies that increasing trading activity by one percent decreases the bid-ask spread by one-third of one percent, when costs are measured in basis points per dollar traded, holding volatility constant.

The predictions for bet size result from the following intuition. Speeding up the time clock increases trading activity both by increasing the arrival rate of bets and by increasing the variance of returns proportionally. Since price volatility is defined as the standard deviation of returns (square root of variance), volatility increases half as fast as variance. Therefore, as trading activity increases, the arrival rate of bets increases twice as fast as volatility. Since trading activity is equal to the product of the bet arrival rate of bets and their size, this implies that if trading activity increases by one percent, the arrival rate of bets must increase by two-thirds of one percent and the size of bets must increase by one-third of one percent.

The model of market microstructure invariance suggests that there is a link between the arrival rate of bets and what finance researchers call “market efficiency.” When arrival rate of bets is high and a “trading day” is short, the flow of information is fast, data is processed quickly, new ideas are generated more often, i.e., the market is “more efficient.” The arrival rate of bets is not directly observable, but the model implies that it is inversely proportional to the level of trading activity raised to the two-thirds power. For a security with a price P dollars, trading volume V shares per calendar day, and daily returns volatility σ , the speed of a trading process can be calculated as $h \cdot (\sigma \cdot P \cdot V)^{-2/3}$ with a constant of proportionality h being unidentifiable in our data.

Two Alternative Models. We compare the predictions of our proposed model of market microstructure invariance with the predictions of two alternative models.

- The “model of invariant bet frequency” assumes that as trading activity increases, the size of bets increases proportionally, but the arrival rate of bets remains constant. This model implies that order size as a fraction of daily volume is the same across stocks. This model predicts that as trading activity increases, the cost of executing a trade of one percent of average daily volume remains constant in basis points per dollar traded (holding returns volatility constant). We believe that this model is the “default model” that implicitly but incorrectly guides the intuition of many asset managers and researchers. This model underlies some conventional illiquidity measures, e.g., the illiquidity ratio suggested in Amihud (2002) and the Amivest measure of liquidity proposed by Cooper, Groth, and Avera (1985).
- The “model of invariant bet size” assumes that as trading activity increases, the size of bets remains the same but the number of bets increases proportionately with trading volume. This model predicts that bet size as a fraction of average daily volume decreases at a rate proportionate to the rate at which trading activity increases. This model predicts that if trading activity increases one per cent, the market impact cost

of executing an order for a given percentage of expected daily volume increases by one-half of one percent, and the bid-ask spread cost decreases by one-half of one percent. This model is in some ways similar to the approach of Hasbrouck (2009) and Gabaix et al. (2006).

We show that all three models are consistent with adverse selection and inventory models. The models differ in how they convert the theoretical implications of adverse selection and inventory models into predictions about bet size and market impact. A key issue concerns the manner in which order imbalances are related to trading volume. If we define “order imbalances” as the sum of all bets in a given day, then for a given level of trading volume, the standard deviation of order imbalances is proportional to $\gamma^{-1/2}$, where γ is the arrival rate of bets. If γ is large, the bets tend to cancel out by the law of large numbers and order imbalances are small. If price volatility results from bets, a large value of γ implies large impact from each bet. The three alternative models make different assumptions concerning how γ varies with trading activity and therefore lead to different empirical predictions about cross-sectional variation in bet size, market impact costs, and spread costs.

1.2 Empirical Results

For all three models, the empirical implications concerning bet size, market impact, and bid-ask spread are tested using a database of more than 400,000 portfolio transition trades provided by a major vendor of portfolio transition services. In a portfolio transition, an incumbent portfolio manager is replaced by a newly hired one. A skillful transition manager replaces the incumbent’s legacy portfolio with a new portfolio by selling a portfolio held by the incumbent manager and buying a portfolio chosen by the new manager in a cost efficient manner.

Order Size. We exploit the data on portfolio transition orders to test implications concerning the size of bets. To justify our test, we invoke the assumption that transition orders are if not equal, then at least proportional to sizes of bets in the cross-section. All three models then can be conveniently nested into the same regression specification that relates order size to trading activity, with different power coefficients distinguishing the different models. A log-linear OLS regression of portfolio transition order size on the level of trading activity provides strong support for the model of trading game invariance. The estimate of -0.63 is remarkably close to the predicted value of $-2/3$ and quite different from the predictions of alternative models of 0 and -1 . Several robustness checks for different sub-samples and alternative regression specifications provide additional evidence in favor of our model.

As predicted by the model, we find that shape of the distribution of orders sizes, adjusted to conform to the predictions of the model of trading game invariance, is remarkably stable across volume and volatility groups. The shape of the distribution of unsigned bet size is close to a log-normal, not a normal one as assumed in Kyle (1985).

Implementation Shortfall and Portfolio Transitions We use data on the implementation shortfall of portfolio transitions to test implications concerning transaction costs. Perold (1988) defines implementation shortfall as the difference between a “paper trading”

benchmark and actual trading results, marking to market unexecuted shares at post-trade prices. For our purposes, the paper trading benchmark for a given transition is defined to be the price which would have been obtained if all shares were executed at the market closing price the day before any trades implementing a given transition began to take place. This benchmark is compared against the actual prices at which the transition trades are later executed. The difference is measured in basis points. Implementation shortfall includes the effect of both market impact and bid-ask spread, as well as random changes in the stock price between the benchmark date and the time when the trades are executed. We make the identifying assumption that the returns on the stock would otherwise have had a mean return of zero, which implies that the mean of the implementation shortfall is an unbiased estimate of expected transaction costs.

There are two major problems usually associated with using implementation shortfall to estimate transaction costs. The database of portfolio transition data avoids both of them.

- **Selection Bias.** When transaction costs are estimated based on data sets containing prices and quantities of executed orders, these estimates are most likely biased downwards. The reason is that high cost orders are often canceled before execution and thus not observed in the data. Obizhaeva (2009b) finds that the selection bias associated with unexecuted trades can be substantial. The data on portfolio transitions does not suffer from a selection bias problem. The transition manager's job is to sell the entire legacy portfolio and replace it with the entire new portfolio. The lists of securities in both portfolios are specified in a transition mandate before trading. Assuming the transition manager executes each portfolio fully, the problem of selection bias due to unexecuted orders goes away.
- **Low Statistical Power:** Black (1986) alerts that noise makes it difficult to test theories about how financial markets work. Indeed, the market impact costs and spread costs of a trade are usually small relative to the noise introduced by other random changes in stock prices. The dataset of portfolio transitions avoids the problem of low statistical power in two ways. First, there are more than 400,000 orders executed over the period 2001-2005, providing many degrees of freedom. Second, many of the trades are for a substantial fraction of average daily volume. This magnifies price impact relative to other trading noise, making it easier to detect linear market impact in the data.

Market Impact and Bid-Ask Spread Costs. Our theoretical model as well as two alternative models imply that market impact and bid-ask spread can be estimated from a non-linear regression in which the left-hand side is implementation shortfall of portfolio transition trades measured in basis points per dollar traded. As before, the models differ only in their predictions for the power coefficient of trading activity for market depth and spread. Our tests show that the model of market microstructure invariance predicts transaction costs from market impact and spread much better than the other two alternatives. Its empirical prediction that a one percent increase in trading activity increases the market impact (in units of daily standard deviation) by one-third of one percent is almost exactly the point estimate from non-linear regressions based on implementation shortfall.

A Transaction Cost Formula. The empirical estimates of implementation shortfall imply a simple operational formula for market impact and bid-ask spread as functions of observable dollar trading volume and volatility. The formula scales transaction costs so the market impact $\bar{\lambda}$ and bid-ask spread $\bar{\kappa}$ are measured in basis points for trading one percent of average daily volume for a benchmark stock defined arbitrarily to have a price of \$40 per share, trading volume of one million shares per day, and volatility of 2% per day. The expected implementation shortfall $C(X)$ of executing X shares of a stock with a stock price P dollars, trading volume V shares per calendar day, and daily returns volatility σ , can be calculated as,

$$C(X) = \frac{1}{2}\bar{\lambda} \cdot \left(\frac{\sigma \cdot P \cdot V}{(0.02)(40)(10^6)} \right)^{1/3} \cdot \frac{\sigma}{0.02} \cdot \frac{X}{(0.01)V} + \frac{1}{2}\bar{\kappa} \cdot \left(\frac{\sigma \cdot P \cdot V}{(0.02)(40)(10^6)} \right)^{-1/3} \cdot \frac{\sigma}{0.02},$$

where the estimated constant for half-market-impact is $\bar{\lambda}/2 = 2.89$ basis points and for half-spread is $\bar{\kappa}/2 = 7.91$ basis points. For a trade of one percent of average daily volume for the benchmark stock, i.e. $\sigma = 0.02$, $P = 40$, $V = 10^6$, $X = 0.01 \cdot 10^6 = 10,000$, this formula reduces to $C(X) = \bar{\lambda}/2 + \bar{\kappa}/2 = 2.89 + 7.91 = 10.80$ basis points. The formula $C(X)$ shows how to extrapolate the estimates $\bar{\lambda}/2$ and $\bar{\kappa}/2$ for the benchmark stock to trades of different size X in securities with different dollar volume $P \cdot V$ and different volatility σ .

The remainder of this paper describes the theory of market microstructure invariance, the nature of its empirical predictions, and the results of some initial tests based on portfolio transitions data.

2 Time Clock Irrelevance and Modigliani-Miller Irrelevance

Trading in a stock can be thought of as a game in which innovations in order flow influence prices over time. Instead of modeling this game by making explicit assumptions about the structure of information, the motivations of traders, and the consistency of their beliefs, we take a reduced form approach. We assume that innovations in the order flow follow a compound Poisson process with arrival rate of γ innovations per day, and we let \tilde{Q} denote a random variable with probability distribution representing a typical innovation in this order flow. In this paper, we informally refer to \tilde{Q} as an innovation in the order flow, instead of formally making reference to the probability distribution of the independently distributed bets themselves. A positive value of \tilde{Q} represents buying and a negative value of \tilde{Q} represents selling. We think of \tilde{Q} as having a mean of zero.

2.1 Trading Activity as a Sequence of Bets

We call the innovations in the order flow \tilde{Q} “bets” instead of “orders” or “trades” because we want to make a distinction between statistically independent decisions to take risky positions (bets) and the specific sequences of potentially correlated orders and trades that implement the decisions to take on risk. For example, after performing weeks of research on a target stock, a trader might make a decision to purchase 100,000 shares. This decision represents a

bet with $\tilde{Q} = 100,000$. The trader might implement this bet by placing a sequence of orders to purchase 20,000 shares of stock per day for five days in a row. Each of these orders may be broken into smaller pieces for execution. For example, on day one there may be trades of 2,000, 3,000, 5,000, and 10,000 shares executed at different prices. Each of these smaller trades may each show up in TAQ data as multiple “prints.” In the context of a compound Poisson process, which requires independent innovations, it would not be appropriate to model each of the five orders to purchase 20,000 shares as independent bets.

Similarly, if an analyst makes a recommendation to ten different customers that they buy a stock, and all of the customers together buy 100,000 shares as a result of the recommendation, it is reasonable to think of this purchase of 100,000 shares as one bet with $\tilde{Q} = 100,000$, not as a sequence of ten different bets. Since the ten purchases are all based on the same information—the recommendation of the analyst—the decisions are probably also not statistically independent like the compound Poisson assumption requires.

Economically, the size of the risk transferred by a bet is more meaningful than the number of shares \tilde{Q} or the dollar volume represented by \tilde{Q} shares. Let P denote the price of the stock in dollars per share, and let σ denote the percentage standard deviation of returns on the stock per day. A good measure of the risk transferred by the bet is the product of \tilde{Q} , P , and σ . Therefore, we define “bet risk” or “bet size” \tilde{B} by

$$\tilde{B} = \tilde{Q} \cdot P \cdot \sigma. \quad (1)$$

Let V denote expected daily trading volume, measured in shares per day. Since expected trading volume is the product of bet arrival rate and expected unsigned bet size, we have

$$V = \gamma \cdot E\{|\tilde{Q}|\}. \quad (2)$$

Economically, the risk that is transferred by these traded shares is more meaningful than share trading volume. Therefore, analogously to the definition of bet risk, we define “trading activity,” denoted W , as the product of expected trading volume, stock price, and volatility:

$$W = V \cdot P \cdot \sigma. \quad (3)$$

Note that since trading volume, stock price, and volatility can all be observed or estimated empirically, trading activity is empirically observable. Observable trading activity can be equivalently defined as the product of two unobservable quantities: the expected number of bets per day and expected (unsigned) bet size:

$$W = \gamma \cdot E\{|\tilde{B}|\}. \quad (4)$$

One of the main contributions of this paper is to provide an empirical approach to decompose observable trading activity into unobservable bet arrival rate and bet size.

2.2 Trading Game Invariant

We obtain a theoretical principle of trading game invariance by applying two different irrelevance principles to the idea that trading is generated by bets which follow a compound

Poisson process. We refer to these irrelevance principles as “Modigliani-Miller irrelevance” and “time-clock irrelevance.”

For our purposes, Modigliani-Miller irrelevance refers to the idea that relevant economic magnitudes—bet size and bet rates—are not affected by share splits or changes in leverage. If shares are split by a factor η , then the price will be divided by the split factor η and traders will multiply shares traded by the split factor η , leaving dollar volume and other relevant economic variables unaffected. If a company pays a debt-financed dividend equal to a fraction δ of the price of the firm’s shares, then the ex-dividend price of the shares will decrease by δP and the volatility of the ex-dividend shares will increase by a factor $1/(1 - \delta)$. Assuming no bankruptcy, the risk of trading \tilde{Q} shares will be unaffected by the dividend. A negative dividend can be implemented as a rights issue (combined with a reverse stock split). From a microstructure perspective, both stock splits and debt-finance dividends are irrelevant because traders can easily undo these effects by changing quantities traded and changing the leverage of their own portfolios.

Time-clock irrelevance is the key concept in this paper. Imagine slowing down trading or speeding up trading by slowing down or speeding up the arrival of new information and other innovations in the stochastic process which drives trading. As a result of such a change in the speed of the trading game’s time clock, traders will make economically equivalent trades at economically equivalent prices, but the trades and price changes will unfold proportionately faster or more slowly depending on the factor by which the time clock has been speeded up or slowed down.

Intuitively, time clock irrelevance refers to the idea that speeding up or slowing down a trading game does not have an effect on the “fundamentals” of the game. For example, if clock speed doubles, the rate of flow of information contained in news articles, analysts’ reports, and corporate disclosures doubles; the amount of effort it takes traders to process relevant information doubles; and the frequency of bets based on such information doubles; all this implies that returns variance also doubles, assuming market prices reflect the arrival of new information. In order to spend the same amount of effort formulating and executing each trade, without working harder, the trader might either halve the number of stocks he trades or halve his “market share” of bets in the stocks he does trade.

Our intuition is that trading games are similar in this respect to chess. In chess, the game can be played quickly with a fast time clock which allows each trader only a few seconds to make a move. Alternatively, chess can be played slowly, with unlimited time between moves. With a fast time clock, moves occur very quickly compared to chess played slowly, but the rules of the game are otherwise the same. The same moves could be made in a slow game as in a fast game. And the same intellectual effort could lie behind each move in fast and slow games if the chess player varies the number of fast games played in parallel proportionally with the speed of the clock.

Consider speeding up or slowing down the trading game’s time clock by changing one hour to H hours. If $H > 1$, the time clock slows down; if $H < 1$, the time clock speeds up. Speeding up the time clock speeds up trading activity for two reasons. First, the number of bets per day—and therefore dollar volume—increases proportionately with $1/H$. Second, speeding up the time clock increases returns variance proportionately with $1/H$, so volatility (the square root of variance) increases proportionately with $1/H^{1/2}$; as a result, each bet becomes riskier. Combining both the volume effect and the volatility effect implies that

trading activity increases proportionately with $1/H^{3/2}$.

We can easily express the effects of Modigliani-Miller irrelevance and time-clock irrelevance algebraically. Let γ^* , \tilde{Q}^* , P^* , and σ^* denote benchmark levels of bet arrival rate, bet random variable, stock price, and volatility. Now change these benchmark levels to new levels γ , \tilde{Q} , P , and σ by applying a split factor η , a debt-finance proportionate dividend δ , and change in time clock speed H . The old and new levels are related by the following equations:

$$\gamma = \frac{1}{H} \cdot \gamma^*, \quad (5)$$

$$\tilde{Q} = \eta \cdot \tilde{Q}^*, \quad (6)$$

$$P = \frac{(1 - \delta)}{\eta} \cdot P^*, \quad (7)$$

$$\sigma = \frac{1}{H^{1/2} \cdot (1 - \delta)} \cdot \sigma^* \quad (8)$$

Modigliani-Miller irrelevant transformations η and δ do not affect bet risk $\tilde{B} = \tilde{Q}P\sigma$ or trading activity $W = \gamma E\{|\tilde{B}|\}$, but the time-clock change H affects both:

$$\tilde{B} = \tilde{B}^* \cdot H^{-1/2} \quad (9)$$

$$W = W^* \cdot H^{-3/2} \quad (10)$$

The change in time clock speed $1/H$ is not directly observable, but we can easily solve the above equation for H to express the change in time clock as a function of the change in observable trading activity:

$$H = \left(\frac{W}{W^*} \right)^{-2/3} \quad (11)$$

This equation allows us to express both changes in bet risk and changes in bet arrival rate—neither directly observable—as functions of observable trading activity:

$$\tilde{B} = \tilde{B}^* \cdot \left(\frac{W}{W^*} \right)^{1/3}, \quad (12)$$

$$\gamma = \gamma^* \cdot \left(\frac{W}{W^*} \right)^{2/3}. \quad (13)$$

The above two equations (12) and (13) incorporate the important point. As trading activity W increases, both arrival rate of bets and their variance increase because information arrives faster. Arrival rate of bets increases, however, twice as fast as bet size. The two-to-one ratio emerges because converting bet variance to a standard deviation requires taking a square root. The coefficients are $2/3$ and $1/3$ because the exponent associated with bet arrival rate is twice the exponent associated with bet size and these exponents must sum to one, since trading activity is the product of expected bet size and bet arrival rate.

Let \tilde{I} denote the “trading game invariant,” which we define as the ratio of bet size to the square root of the bet arrival rate:

$$\tilde{I} := \frac{\tilde{B}}{\gamma^{1/2}}. \quad (14)$$

Since bet risk is defined by $\tilde{B} = \tilde{Q}P\sigma$, we can represent \tilde{I} by dividing the product of equations (6),(7), and (8) by the square root of equation (5). The coefficients involving η , $1 - \delta$, and H all cancel, resulting in

$$\tilde{I} = \frac{\tilde{Q} \cdot P \cdot \sigma}{\gamma^{1/2}} = \frac{\tilde{Q}^* \cdot P^* \cdot \sigma^*}{\gamma^{*1/2}}. \quad (15)$$

We refer to \tilde{I} as a trading game invariant because its distribution is unaffected by splits, leverage, and time clock changes. The invariance of \tilde{I} can also be seen by dividing equation (12) by the square root of equation (13). The fact that the coefficients involving trading activity W cancel shows that \tilde{I} is invariant to changes in trading activity and therefore to changes in the speed of the time clock.

The invariance of \tilde{I} implies a precise relationship between bet size \tilde{B} and bet frequency γ . In equations (14) and (15), the numerator is bet size and the denominator is the square root of the arrival rate of bets. The invariance of \tilde{I} implies that as trading volume increases, arrival rate of bets increases twice as fast as bet size. Speeding up the time clock makes bets riskier by increasing bet variance proportionately with the speed of the clock. By dividing bet risk by the square root of the bet arrival rate, the standard deviation of bet risk is re-scaled so that bet risk is invariant not only to splits and debt-financed dividends, but also to time clock speed.

If we think about the invariant \tilde{I} as a product of dollar trade $P \cdot \tilde{Q}$ and volatility $\sigma\gamma^{-1/2}$, scaled so that there is one bet per clock tick, we can see that \tilde{I} represents bet size in transaction time. Our concept of trading game invariance is similar in spirit to the idea of Mandelbrot and Taylor (1967) that markets operate in the transaction time, with the clock ticking once per transaction. It also has something in common with Dufour and Engle (2000), who examine price risk between trades. Of course, “bet time” and “transaction time” are different concepts because one bet may generate many transactions. Our trading game invariant captures the idea that the risk of bets can be measured in a manner invariant to the rate at which the time clock ticks by re-scaling time so that one new bet is expected to arrive for each tick on the re-scaled time clock.

The distributions of \tilde{I} , \tilde{Q} , and \tilde{B} all have the same shape, but are scaled differently. We have

$$\frac{\tilde{I}}{E\{|\tilde{I}|\}} = \frac{\tilde{Q}}{E\{|\tilde{Q}|\}} = \frac{\tilde{B}}{E\{|\tilde{B}|\}}. \quad (16)$$

By solving the equations (4) and (14) for γ and $E\{|\tilde{B}|\}$, we obtain

$$\gamma = E\{|\tilde{I}|\}^{-2/3} \cdot W^{2/3}, \quad (17)$$

$$E\{|\tilde{B}|\} = E\{|\tilde{I}|\}^{2/3} \cdot W^{1/3}. \quad (18)$$

These equations both capture the idea that as trading activity W increases due to speeding up the time clock, bet frequency increases twice as fast as bet size. These equations are similar to equations (12) and (13). The difference is that equations (17) and (18) use the invariant to express bet rate and expected bet size as a function of W , while equations (12) and (13) use benchmark levels of variables to express bet rate and expected bet size as functions of W .

Equation (18) implies that the expected size of bets increases one third as fast as trading activity. It follows from this that bet size as a fraction of average daily volume should decrease at a rate two-thirds as fast as trading activity increases. This can be shown precisely by combining equations (2), (16), and (18) to obtain

$$\frac{\tilde{Q}}{\bar{V}} = \frac{\tilde{I}}{E\{|\tilde{I}|\}^{1/3}} \cdot W^{-2/3}. \quad (19)$$

2.3 Market Impact and Bid-Ask Spread Costs

The concepts of Modigliani-Miller irrelevance and time clock irrelevance can be also applied to transaction costs. Assume that the trading costs associated with executing a bet consist of two components, a permanent linear price impact and a transitory bid-ask spread. Let C_L denote the expected price impact cost of executing a bet, and let C_S denote the expected round-trip bid-ask spread cost of executing a bet.

Following the notation of Kyle (1985), let λ denote the price impact of trading one share of stock, measured in units of dollars per share per share, i.e., dollars per share-squared. Assume that the expected price impact cost C_L is obtained by walking up or down a linear demand schedule with slope λ . We therefore have

$$C_L = \frac{1}{2} \cdot \lambda \cdot E\{\tilde{Q}^2\}. \quad (20)$$

Let κ denote the bid-ask spread for one share of stock, measured in dollars per share. The expected round-trip bid-ask spread cost is given by

$$C_K = \kappa \cdot E\{|\tilde{Q}|\}. \quad (21)$$

We assume that the expected impact cost C_L and spread cost C_K of executing a given bet are not affected by stock splits, leveraging dividends, and time clock changes. This assumption captures the intuition that splits, dividends, and time clock changes do not affect the essential fundamentals of the trading game.

While stock splits have a proportionate effect on the number of shares traded in a bet, changes in leverage and time clock do not affect the number of shares traded in a bet. Thus, if λ^* and κ^* denote benchmark levels of price impact and bid-ask spread parameter, then the values of these parameters after applying a stock split η , a leveraging dividend δ , and a time clock change H are given by the following formulas, which do not depend on δ or H :

$$\lambda = \lambda^*/\eta^2 \quad (22)$$

and

$$\kappa = \kappa^*/\eta. \quad (23)$$

In order to make price impact λ and price impact costs C_L endogenous, we make the following assumption:

- Price impact leads to a fraction ψ^2 of price variance.

This assumption allows price volatility to result from both “announcement effects” and “trading effects.” This assumption implies that the product of the variance of the price impact of a bet and the expected number of bets per day γ is equal to the product of ψ^2 and price variance $\sigma^2 P^2$. Since the price impact of a bet is $\lambda \cdot \tilde{Q}$, the variance of the price impact of a bet is $\lambda^2 \cdot E\{\tilde{Q}^2\}$. Thus, we obtain

$$\psi^2 \cdot \sigma^2 \cdot P^2 = \gamma \cdot \lambda^2 \cdot E\{\tilde{Q}^2\}. \quad (24)$$

Solving for the market depth parameter λ , we obtain

$$\lambda = \frac{\psi \cdot \sigma \cdot P}{\gamma^{1/2} \cdot E\{\tilde{Q}^2\}^{1/2}}. \quad (25)$$

Plugging equation (25) into equation (20) and using the definition of \tilde{B} from equation (1) yields the following solution for C_L :

$$C_L = \frac{1}{2} \cdot \psi \cdot \frac{E\{\tilde{B}^2\}^{1/2}}{\gamma^{1/2}}. \quad (26)$$

This equation provides a remarkably simple expression for linear market impact costs as a function of bet size and bet rate, consistent with Modigliani-Miller irrelevance. While the derivation depends upon the assumption of linear price impact of bets, the derivation of this equation does not depend on the assumption of time clock irrelevance.

In order to create a stable relationship between the bid-ask spread κ and expected bid-ask spread costs C_K endogenous, we make the following reduced form assumption:

- The expected bid-ask spread cost of a bet is a constant fraction ϕ of the expected price impact cost of a bet.

This assumption states

$$C_K = \phi \cdot C_L. \quad (27)$$

Using equations (27) and (21), the bid-ask spread κ can be written

$$\kappa = \frac{\phi \cdot C_L}{E\{|\tilde{Q}|\}}. \quad (28)$$

2.4 Time Clock Irrelevance and Trading Costs.

The square of equation (14) can be plugged into equation (26) to express expected market impact costs C_L as a function of moments of the invariant \tilde{I} :

$$C_L = \frac{1}{2} \cdot \psi \cdot E\{\tilde{I}^2\}^{1/2}. \quad (29)$$

Under the assumption of time clock irrelevance, the expected market impact cost C_L does not depend on trading activity W . This is not a surprising result, since the intuition for time clock irrelevance is that speeding up or slowing down the trading game does not affect the “fundamentals” of the game such as market impact costs.

From equation (25), the market impact of trading one entire day's expected trading volume $V = \gamma E|\tilde{Q}|$, expressed as a fraction of one day's price volatility σP , is given by

$$\frac{\lambda \cdot V}{\sigma \cdot P} = \psi \cdot \frac{E\{|\tilde{Q}|\}}{E\{\tilde{Q}^2\}^{1/2}} \cdot \gamma^{1/2}. \quad (30)$$

Using equations (16) and (17), the market impact of trading one day's expected trading volume can be expressed as a function of moments of the invariant \tilde{I} and W as follows:

$$\frac{\lambda \cdot V}{\sigma \cdot P} = \psi \cdot \frac{E\{|\tilde{I}|^{2/3}\}}{E\{\tilde{I}^2\}^{1/2}} \cdot W^{1/3}. \quad (31)$$

Combining the definition of bet size in equation (1) with equation (28), the bid-ask spread as a fraction of one day's dollar volatility $\sigma \cdot P$ can be written

$$\frac{\kappa}{\sigma \cdot P} = \frac{\phi \cdot C_L}{E\{|\tilde{B}|\}}. \quad (32)$$

Plugging the solution for C_L in equation (29) and the solution for $E\{|\tilde{B}|\}$ from equation (18) into equation (32) yields a right-hand-side which only depends on moments of the trading game invariant \tilde{I} and trading activity W :

$$\frac{\kappa}{\sigma \cdot P} = \frac{1}{2} \cdot \phi \cdot \psi \cdot \frac{E\{\tilde{I}^2\}^{1/2}}{E\{|\tilde{I}|^{2/3}\}} \cdot W^{-1/3}. \quad (33)$$

Consider a trade which represents a constant fraction of average daily trading volume, say one percent. Equations (31) and (33) imply that, holding price P and volatility σ constant, a one percent increase in trading activity leads to a one-third of one percent increase in impact costs and a one-third of one percent decrease in spread costs. As trading activity increases, the proportion of total trading costs associated with price impact increases and the proportion of total trading costs associated with bid-ask spread falls.

What is the intuition behind these results? In our framework, trading activity increases when the time clock speeds up, but this also increases volatility. As the time clock speeds up and increases trading activity W on the right-hand side of the above equations, the volatility parameter σ and the price P on the left-hand side of the above equations can be held constant by making appropriate offsetting changes in the split factor η and dividend factor $1 - \delta$. For this reason, it is useful to consider Modigliani-Miller irrelevant changes together with time clock changes and express all formulas in terms of variable such as bets \tilde{B} , trading activity W , scaled price impact $\lambda V/\sigma P$ and scaled spread $\kappa/\sigma P$, unaffected by Modigliani-Miller irrelevant changes, as we do in this paper.

Consider what happens to price impact after an increase in trading activity of one percent. From equation (17), the number of bets represented by one day's trading volume increases by two-thirds of one percent. Since impact cost is quadratic in bet size, the impact cost of buying or selling one day's trading volume increases by four-thirds of one percent. This increase is one-third of one percent more than the increase in trading volume, consistent with the power one-third applied to W on the right-hand side of equation (31).

Consider what happens to the bid-ask spread after an increase in trading activity of one percent. A bet becomes riskier because volatility increases by one-third of one percent. Scaling the unchanged spread by volatility therefore reduces the spread by one-third of one percent, consistent with equation (33).

3 Market Microstructure Invariance as an Empirical Hypothesis

Market Microstructure Invariants. So far, we have presented theoretical results about hypothetical splits, dividends, and time clock changes for one individual stock at a particular time. These results do not assume that the invariant distribution \tilde{I} for different stocks (or the same stock at different times) is the same, that the parameter ψ for different stocks is the same, or the parameter ϕ for different stocks is the same. Therefore, these results do not lead to hypotheses about bet sizes, price impact, or bid-ask spreads which can be tested empirically based on trading data for a cross-section or time series of stocks. To generate empirically testable hypotheses, it is necessary to make empirically testable conjectures about the cross sectional and time series properties of different stocks. We therefore make the following three empirical hypotheses:

- Trading Game Invariance: For all stocks, the distribution of the trading game invariant \tilde{I} is the same.
- Market Impact Invariance: For all stocks, the linear price impact of bets explains a fraction of price volatility ψ which is constant across stocks.
- Bid-Ask Spread Invariance: For all stocks, the bid-ask spread is the same constant fraction ϕ of the expected market impact costs of a bet.

We refer to all three of these empirical hypotheses as “market microstructure invariance.” We refer to the distribution \tilde{I} as well as the two parameters ψ and ϕ as “market microstructure invariants.”

The intuitive idea behind trading game invariance is that—except for Modigliani-Miller irrelevant transformations and time clock irrelevant changes—the trading game for different stocks is the same. For stocks with high levels of trading activity, the trading game is played faster than for stocks with low levels of trading activity. For stocks with high stock prices, the game is the same as for stocks with low stock prices. For stocks with high volatility, the game is that same as for stocks with low volatility.

To summarize, we apply three theoretical irrelevancy assumptions—two Modigliani-Miller irrelevancies (split irrelevance and leverage irrelevance) and time clock irrelevance—to construct theoretical invariants for an individual stock. We then introduce three empirical hypotheses—trading game invariance, market impact invariance, and bid-ask spread invariance—to get empirical market microstructure invariants. The intuition of invariants is that a reduced form approach to trading games leads to empirically testable implications for bet size, market impact, and bid-ask spreads which might be satisfied for many different market microstructure models. Most importantly, time clock irrelevance—in the context of the

other assumptions—makes it possible to derive empirically testable implications concerning how these variables vary across stocks with different levels of trading activity.

Empirical Implications for Bet Size.

To understand the cross-sectional empirical implications concerning bet size, consider two different stocks. Let the first stock be a “benchmark stock” with price P^* , volatility σ^* , bet arrival rate γ^* , and bet quantity \tilde{Q}^* . Let analogous quantities for the second stock be P , σ , γ , and \tilde{Q} . Now use equation (5) to change the time clock parameter H for the benchmark stock so that its bet arrival rate is the same as the second stock. Next use equation (8) to change the leverage factor δ so that the volatilities are the same. Then use equation (7) to change the split factor η so that the stock prices are the same. From equation (6), an additional restriction is needed to insure that the split factor η also makes the distribution of \tilde{Q}^* the same as \tilde{Q} . The needed restriction is precisely captured by trading game invariance. This can be seen from equation (16), where the assumption that the distribution of the trading game invariant \tilde{I} is the same for both stocks guarantees that \tilde{Q}^* and \tilde{Q} have the same distribution. From this, it follows that the other equations above can be interpreted as empirical predictions about different stocks.

For example, equations (12) and (13) can be interpreted as describing how bet size and bet frequency vary as functions of relative trading activity across different stocks or across the same stock at different times.

The hypothesis of trading game invariance implies that if two stocks have the same level of trading activity, this level of expected trading activity results from the same number of expected bets and the same expected bet size. Moreover, if trading activity differs across different stocks or if trading activity for the same stock differs at different times, trading game invariance implies that the relationship between trading activity, bet size, and bet frequency is given by equations (17) and (18), or equivalently by equations (12) and (13). For example, if one stock’s trading activity is 8 times higher than another, then the more active stock will have 4 times as many bets, and bets will on average be 2 times larger than the less active stock. Moreover, from equation (16), the distribution of the shape of bet size will be same across stocks with different levels of trading activity.

To test this hypothesis, an empirical proxy for bets is needed. In the second half of this paper, we use portfolio transition orders as an empirical proxy for bets.

It is intuitive for traders to measure the size of transactions as a fraction of average daily volume. Let \tilde{Q}^*/V^* and \tilde{Q}/V denote the distribution of bet shares as a fraction of average daily volume for the benchmark stock and another stock. Equation (19) shows precisely how this distribution depends on trading activity and the distribution of the invariant \tilde{I} . From equation (19), it follows that

$$\frac{\tilde{Q}}{V} = \frac{\tilde{Q}^*}{V^*} \cdot \left(\frac{W}{W^*}\right)^{-2/3}. \tag{34}$$

This equation states that as trading activity increases cross-sectionally across stocks, the size of bets as a fraction of average daily volume decreases two-thirds as fast.

Empirical Implications for Market Impact and Bid-Ask Spread.

From equation (29), the combination of trading game invariance and market impact invariance implies that the expected market impact cost of a bet is the same across all stocks. Adding the assumption

of bid-ask spread invariance implies that the expected bid-ask spread cost of a bet is the same across all stocks. Instead of testing these invariance implications for market impact costs and bid-ask spread costs for bets, we test equivalent implications for implications about how market impact costs of trading a given fraction of average daily volume and bid-ask spread costs vary with changes in levels of trading activity.

The market impact hypothesis and bid-ask spread hypothesis imply that if two stocks have the same level of trading activity, then the impact costs and spread costs of trading the same percentage of average daily volume will be the same for both stocks. These hypotheses implies that if trading activity varies across different stocks, impact costs and spread costs will vary as a function of trading activity W as shown in equations (31) and (33). In particular, combining two versions of equation (31)—one for the benchmark stock and one for a second stock—implies that the market impact of trading an entire day’s trading volume, scaled by stock price and volatility, satisfies

$$\frac{\lambda \cdot V}{\sigma \cdot P} = \frac{\lambda^* \cdot V^*}{\sigma^* \cdot P^*} \cdot \left(\frac{W}{W^*} \right)^{1/3} . \quad (35)$$

This equation states that as trading activity increases cross-sectionally across stocks, the market impact of trading an entire day’s trading volume, scaled appropriately for volatility and stock price, increases one-third as fast as trading activity increases. Since price impact is linear in trade size, the same relationship holds for trading any constant fraction of average daily volume. Equation (35) is based on the hypothesis of market impact invariance because the proportionality factor ψ is the same in both versions of equation (31).

Similarly, equation (33) implies that the bid-ask spread, scaled by volatility and stock price, satisfies

$$\frac{\kappa}{\sigma \cdot P} = \frac{\kappa^*}{\sigma^* \cdot P^*} \cdot \left(\frac{W}{W^*} \right)^{-1/3} . \quad (36)$$

This equation states that as trading activity increases cross-sectionally across stocks, the bid-ask spread, scaled appropriately for volatility and stock price, decreases one-third as fast as trading activity increases. This equation is based on the hypotheses of bid-ask spread invariance and market impact invariance because the proportionality factors ϕ and ψ are the same in versions of equation (33) for the benchmark stock and the second stock. It also requires the trading game invariance.

Consider a bet that is some constant fraction of average daily volume, say 1%. The above two equations imply that as the level of trading activity increases, the market impact portion of the total trading costs will increase and the bid-ask spread portion of total trading costs will decrease.

To test the above two equations, data on transaction costs is needed. In the second half of this paper, we use data on portfolio transitions to estimate market impact and spreads using the concept of implementation shortfall.

Formulated implications for transaction costs require additional assumptions of market impact invariance and bid-ask spread invariance on top of the assumption of trading game invariance. We expect that the empirical evidence on the implications for bet size will be stronger than that on implications for transaction costs, as the former requires only the assumption of trading game invariance.

An Illustrative Example. The bet size hypothesis and the bet cost hypothesis are based on the intuition that when investors trade stocks with different levels of trading activity, they are playing the same game, but at different speeds. Using the chess analogy, trading an active stock is like playing chess with a fast time clock. Trading an inactive stock is like playing chess with a slow time clock. Except for the speed of the time clock, the trading games look the same.

Figure 1 provides a numerical example using two stocks to illustrate intuitively how the model works. In this example, we assume that $\psi = 1$, i.e., all price volatility results from the linear impact of trading. One stock is a benchmark stock with expected daily trading volume V^* of one million shares, price P^* of \$40 per share, and volatility σ^* of 200 basis points per day. Suppose the trading volume is expected to consist of four bets per day ($\gamma = 4$). Each bet has an expected volume of one-fourth of average daily volume, or 250,000 shares. Each bet is expected to contribute one fourth of a day's price variance, or one half of a day's price volatility (standard deviation), which equal 100 basis points per order.

Consider another stock with the same share price and same volatility, but with eight times the trading volume, or 8 million shares per day. How do microstructure variables change in response to increasing trading volume eight-fold, while holding stock price and volatility constant?

Trading game invariance implies that bet frequency increases twice as fast as bet size (in logs). To increase trading volume by a factor of eight, this implies that bet frequency increases by a factor of four and bet size increases by a factor of two. Thus, the daily trading volume of the second stock is expected to consist of 16 independent bets ($4 \cdot 4$), each with expected size of 500,000 shares ($250,000 \cdot 2$).

As we have discussed, there exists a transformation of the time clock and the capital structure of the less active stock so that its trading game becomes identical to the more active stock. Indeed, start with the inactive stock with 4 orders of 250,000 shares placed over one calendar day and the daily volatility of returns being equal to 200 basis points. Speed up the time clock by a factor of 4 ($H = 1/4$), i.e., think about the trading games as being played in transaction time with a trading day of the active stock being 4 times shorter ($8^{-2/3}$) than a trading day of inactive stock. Since speeding up the time clock by a factor of four doubles the volatility to 400 basis points per day, de-lever the stock with a rights offering where one new share is offered at the current market price (minus epsilon, to insure full subscription). This is equivalent to $\delta = -1$. No stock split is needed ($\eta = 1$). After these changes, we obtain a \$40 stock with sixteen orders of 500,000 shares per calendar day and daily volatility of 200 basis points, exactly the parameters of the active stock. According to our model of market microstructure invariance, it should not be a surprise that any stock can be transformed into another one through certain irrelevance transformations.

How different are the market depth and spread between these two stocks? In the inactive benchmark stock, each bet contributes one fourth of daily variance, or one half of daily volatility ($(1/4)^{1/2}$). Therefore, the market impact of one-fourth of trading volume is equal to 100 basis points ($(200/4)^{1/2}$ basis points). In the more active stock, each bet contributes one sixteenth of daily volume, one sixteenth of daily variance, or one fourth of daily standard deviation ($(1/16)^{1/2}$), which is 50 basis points of volatility. For this more active stock, since the market impact of one sixteenth of daily volume is 50 basis points, the market impact of one fourth of trading volume then would be equal to 200 basis points ($200/16^{1/2} \cdot 4$). As the

model implies, this is twice the impact of one fourth of a day's trading volume in the less active stock.

Concerning the bid-ask spread, if the bid-ask spread in the less active stock is equal to κ dollars per share, then the bid-ask spread in the more active stock will be equal to $\kappa/2$ (given by $\kappa \cdot 8^{-1/3}$).

An Equilibrium Story of a Trading Game. These hypotheses are consistent with the following story about the “equilibrium” trading game in markets. Trading activity and price dynamics in financial markets are dominated by institutional investors. Institutional investors are highly trained professionals, with skill levels similar to those of professionals with MBA degrees or higher. These investment professional search for investment opportunities in both active and inactive markets. The search process involves sifting through many different ideas. When a good enough idea is discovered, the investment professional places a bet, with the size of the bet increasing in the quality of the idea. In equilibrium, the cost of finding a good idea is equal to the profit to be made trading on the idea, factoring in the value of the institutional investor's time and the costs of executing the necessary trades.

Our two hypotheses are consistent with the idea that when markets with different levels of trading activity are compared, we will find institutional investors with similar skills in both markets. These investors will generate about the same numbers of bets in the same amount of time, i.e., the costs of discovering a good bet in an inactive market will be the same as the cost of discovering a good bet in an active market. Over the lifetime of the bets, the risks and returns on the bets will be the same in active and inactive markets. In this sense, the economics of trading active and inactive stocks are the same. In other respects, trading in active markets will look different. Bets in the more active markets will be riskier (in dollar standard deviation per day), but these riskier bets will be held for shorter periods of time.

For example, consider five markets. Suppose that one of the markets has the same parameters as the more active market in figure 1. In particular, it has trading volume eight times higher than each of the other four markets, which are like the inactive market in figure 1. The model of microstructure invariance implies that bet frequency in the active market is four times bet frequency in an inactive market, and bet size in the active market is twice bet size in the inactive market. Suppose that one investment professional specializes in the active market and another investment professional specializes in the four inactive markets. Both investors engage in a costly search for trading ideas, which requires substantial inputs in the form of intellectual effort processing market information. Both investors generate the same share of volume in their respective markets, they have the same Sharpe ratios associated with their overall trading, they make the same expected profits, and they have the same expected pay. Their market impact costs and spread costs per bet are the same. The difference between their expected trading profits and expected trading costs covers their salaries, and leaves enough profit for their employer to justify hiring them to search for trading ideas. Over some period of time, the institutional investor specializing in the active market is expected to make four bets, all in the same stock. Over the same period of time, the investor specializing in the inactive markets is expected to make four bets also, one bet in each of the four different stocks. The individual bets made by the investor in the active market are expected to be twice as large as the individual bets made by the investor in the

inactive markets (500,000 shares versus 250,000 shares), but the bets are expected to be held for a time period only one-fourth as long. Thus, the investor in the active stock trades twice the dollar volume as the investor in the inactive market, but since he holds each bet for a shorter time, they are no more or less risky over their lives than the smaller bets in the less active stocks.

Returning to the chess analogy, the investor in the active market is similar to a chess player playing numerous games of chess sequentially, with a fast time clock for each game. The investor in the inactive stocks is similar to a chess player participating four games in parallel, with a slower clock for each game. The basic rules of chess are the same for both games. Furthermore, the level of play in the fast games is the same as the level of play in the slow games because the players in both the fast and slow games spend the same amount of time contemplating each move. The player of the slow games has more time between moves, but he spends the same amount of time contemplating each move because he is playing many games in parallel.

4 Alternative Models

We examine two alternatives to our model proposed above. These alternative models serve three purposes. First, they provide empirical alternatives which can be easily tested against our proposed model because they nest easily into a common specification. Second, the two alternative models are similar in spirit to the models of Amihud (2002) and Hasbrouck (2009) respectively, i.e., we believe the alternative models capture the thinking behind the current literature. Third, we believe the alternative models conform to the way in which traders think about trading costs.

Our first alternative, which we call the “Model of Invariant Bet Frequency,” is based on the assumption that the number of bets per day is invariant across stocks with different levels of trading activity. This model implies that as trading activity increases, bet size increases proportionally.

Our second alternative model, which we call the “Model of Invariant Bet Size,” is based on the assumption that the size of bets is invariant across stocks with different levels of trading activity. This model implies that as trading activity increases, the number of bets increases proportionally.

Both of our proposed alternative models satisfy Modigliani-Miller irrelevance. Both of our proposed models satisfy the empirical hypotheses of market impact invariance and bid-ask spread invariance. Neither is consistent with time clock irrelevance and, consequently, trading game invariance. Instead, each alternative model relies on an invariance hypothesis which is different from the hypothesis of trading game invariance, which we have in our model of microstructure invariance. As a result, neither model implies that the expected market impact cost and expected bid-ask spread cost of a bet is constant across different stocks. Trading game invariance is crucial for distinguishing our model from alternatives. The trading game invariant is a key concept in our paper.

Theoretical models of market microstructure, such as Kyle (1985), often deal with an individual stock, casting the theory in terms of an abstract “trading day” which does not map in a natural manner into clock time. For the purpose of using models based on an

abstract trading day to measure variables empirically, there is no *a priori* reason to assume that this trading day is one calendar day, the same for all stocks. According to the invariance hypothesis, the trading day is an endogenously determined period of time which is shorter for more actively traded stocks. A naive assumption that the trading day is one calendar day, the same for all stocks, may seem to be consistent with one of our proposed alternative models.

4.1 Model of Invariant Bet Frequency

The Model of Invariant Bet Frequency proposes that variation in trading activity W comes entirely from variation in the distribution of bet sizes \tilde{B} , while bet frequency γ remains constant across stocks. In effect, this model assumes a common time clock for all stocks, with bets generated proportionally to the rate at which this time clock ticks. The assumption of constant bet frequency implies that average bet size, as a fraction of average daily volume, is a constant. Using notation similar to equation (34), the implication of this model for trade size as a fraction of average daily volume can be written

$$\frac{\tilde{Q}}{V} = \frac{\tilde{Q}^*}{V^*} \cdot \left(\frac{W}{W^*}\right)^0. \quad (37)$$

The exponent of zero in equation (37) emphasizes how equations (37) and (34) are nested into a common specification. Furthermore, equation (37) also implies that the shape of the distribution of bets is constant across stocks; only the size of the bets changes as trading activity changes.

Market impact invariance implies that the price impact of trading an entire day's trading volume, expressed as a fraction of daily volatility, does not depend on trading activity W . We can express this result in a manner which nests equation (35) and the following equation in a common specification:

$$\frac{\lambda \cdot V}{\sigma P} = \frac{\lambda^* \cdot V^*}{\sigma^* \cdot P^*} \cdot \left(\frac{W}{W^*}\right)^0. \quad (38)$$

Similarly, bid-ask spread invariance implies that the spread can vary in a manner nested into a common specification with equation (36):

$$\frac{\kappa}{\sigma \cdot P} = \frac{\kappa^*}{\sigma^* \cdot P^*} \cdot \left(\frac{W}{W^*}\right)^0. \quad (39)$$

In equations (38) and (39), the exponent of zero captures the model predictions that price impact as a fraction of daily volume and spread—controlling for price and volatility—are both constant.

This theory is intuitively plausible. It assumes that the average bet size is a constant fraction of trading volume. It predicts that when market impact is measured in units of price standard deviation σP , then the impact of trading a given percentage of average daily volume V is constant across stocks of different trading activity W . Since the spread is proportional to market impact and average trade size, the spread measured in units of price standard deviation σP remains constant as well.

An Illustrative Example. Figure 1 illustrates the model of invariant bet frequency using a specific example. The figure shows the difference between a less active benchmark stock and a more active stock with eight times larger trading volume. The larger volume of the more active stock is explained by larger bet size (keeping price and volatility constant). Both stocks have four bets per day, but the two million share bet quantity of the more active stock is eight times larger than the 250,000 share bet quantity of the benchmark stock. For both stocks, each bet contributes one fourth of average daily volume and therefore one-half of average volatility. The market impact of one fourth of trading volume is 100 basis points for both stocks ($= 200/4^{1/2}$), consistent with equation (38). The spread remains equal to 100 basis points ($= 200/4^{1/2}$), consistent with equation (39).

We believe that the model of invariant bet frequency is the “default model” that implicitly but incorrectly guides the intuition of many asset managers. For example, this model justifies trading no more than some constant percentage of average daily volume—say 5%—uniformly for all stocks, regardless of their level of trading activity. When a basket containing both active and inactive stocks is traded in such a way that each stock in the basket represents the same proportion of average daily volume, this model also justifies attributing the same number of basis points in transaction costs for individual stocks.

This model also captures the spirit of some models in the finance literature. For example, the Amihud (2002) measure of illiquidity fits this framework. In the Amihud model, the price impact of trading a dollar’s worth of stock is proportional to the absolute value of daily returns $E\{|\tilde{r}|\}$ and inversely proportional to average daily dollar volume $V \cdot P$. To convert Amihud’s units of price impact per dollar of volume into our units of price impact per share of stock, we can divide our units of price impact per share squared by P^2 . Thus, Amihud’s model can be written

$$\frac{\lambda_a}{P^2} = \text{const} \cdot \frac{E\{|\tilde{r}|\}}{V \cdot P}. \quad (40)$$

This can be equivalently expressed as

$$\frac{\lambda_a V}{\sigma \cdot P} = \text{const} \cdot \frac{E\{|\tilde{r}|\}}{\sigma}. \quad (41)$$

This equation is equivalent to equation (38) if the ratio of $E\{|\tilde{r}|\}$ to σ is constant. Thus, Amihud’s model is the same as the model of invariant bet frequency except for his use of absolute first moment instead of standard deviation to measure returns volatility.

4.2 Model of Invariant Bet Size

The model of invariant bet size proposes that variation in trading activity W comes entirely from variation in the frequency of bets γ , while the distribution of bet sizes \tilde{B} remains constant across stocks. In effect, this model assumes that as the time clock speeds up, the number of bets increases proportionally, but the size of the bets does not change. This model therefore assumes that average bet quantity, as a fraction of average daily volume, shrinks proportionally as average volume increases.

Using notation similar to equation (34), the implication of this model for trade size as a fraction of average daily volume can be written

$$\frac{\tilde{Q}}{V} = \frac{\tilde{Q}^*}{V^*} \cdot \left(\frac{W}{W^*}\right)^{-1}. \quad (42)$$

The exponent of minus-one in equation (42) emphasizes how equations (42) and (34) are nested into a common specification. Furthermore, equation (42) also implies that the shape of the distribution of bets is constant across stocks; only the number of bets changes as trading activity changes.

Market impact invariance implies that the price impact of trading an entire day's trading volume, expressed as a fraction of daily volatility, increases with trading activity W . We can express this result in a manner which nests equation (35) and the following equation into a common specification:

$$\frac{\lambda \cdot V}{\sigma \cdot P} = \frac{\lambda^* \cdot V^*}{\sigma^* \cdot P^*} \cdot \left(\frac{W}{W^*}\right)^{1/2}. \quad (43)$$

Similarly, the bid-ask spread invariance implies that the spread varies in a manner nested into a common specification with equation (36):

$$\frac{\kappa}{\sigma \cdot P} = \frac{\kappa^*}{\sigma^* \cdot P^*} \cdot \left(\frac{W}{W^*}\right)^{-1/2}. \quad (44)$$

This theory is also intuitively plausible. It assumes that bet size is the same across stocks with different levels of trading volume. It predicts that when market impact is measured in units of price standard deviation σ , then the impact of trading a given percentage of average daily volume V increases across stocks half as fast as trading activity W increases, while the bid-ask spread (controlling for volatility) decreases half as fast as trading activity increases.

An Illustrative Example. Figure 1 illustrates the model of invariant bet size using a specific example. The figure shows the difference between a less active benchmark stock and a more active stock with eight times larger trading volume. The larger volume of the more active stock is explained by a larger number of bets (keeping price, volatility, and bet size constant). Both stocks have an average bet size of 250,000 shares, but the more active stock has 32 bets per day while the less active stock has only four. For the less active stock, each of the four bets generates one fourth of a day's variance, or equivalently one half of a day's volatility, which equals 100 basis points. For the more active stock, each of the 32 bets generates 1/32 of a day's variance, or equivalently a fraction $(1/32)^{1/2}$ of a day's standard deviation, which equals $100/8^{1/2}$ basis points. Therefore, trading one fourth of a day's volume in the more active stock, equivalent to eight perfectly correlated bets in the same direction, generates price impact of $100 \cdot 8^{1/2}$ basis points, as implied by the model. The bid-ask spread in the more active stock is lower than the less active stock by a factor of $8^{-1/2}$.

This model also captures the spirit of some models in the finance literature. For example, Hasbrouck (2009) suggests measuring the price impact of trading one dollar of stock as being proportional to $E\{|\tilde{r}|\}$ and inversely proportional to the square root of average dollar volume. Measuring price impact in dollars per share squared, Hasbrouck's model can be written

$$\frac{\lambda_h}{P^2} = \text{const} \cdot \frac{E\{|\tilde{r}|\}}{(V \cdot P)^{1/2}}. \quad (45)$$

This can be equivalently expressed as

$$\frac{\lambda_h V}{\sigma \cdot P} = \text{const} \cdot W^{1/2} \cdot \frac{E\{|\tilde{r}|\}}{\sigma^{1/2}}. \quad (46)$$

This equation is equivalent to equation (38) if both $E\{|\tilde{r}|\}$ and σ do not vary across stocks. Note that if returns \tilde{r} are multiplied by some constant factor, both $E\{|\tilde{r}|\}$ and σ change proportionally but $E\{|\tilde{r}|\}$ and $\sigma^{1/2}$ do not change proportionally. Thus, while Hasbrouck’s model is similar to the model of invariant bet size, he controls for volatility in a manner that is not consistent with Modigliani-Miller equivalence. Hasbrouck’s approach is similar to the model of invariance bet size, but his assumptions do not match the assumptions of the model exactly.

4.3 Comparison of Models

Trading activity is the product of the number and size of bets. The three proposed models make different assumptions about how the number and size of bets changes in response to an increase in trading activity:

- Our model of market microstructure assumes that as trading activity increases, the number of bets increases two-thirds as fast as trading activity and the size of bets increases one-third as fast as trading activity.
- The model of invariance bet frequency assumes that as trading activity increases, the number of bets does not increase at all but the size of bets increases proportionally with trading activity.
- The model of invariant bet size assumes that as trading activity increases, the number of bets increases proportionally with trading activity but the size of bets does not increase at all.

The three models make different assumptions about the relationship between “trading volume” and “order imbalances,” which further lead to different predictions concerning transaction costs. In theory, these variables represent a link between volatility and standard deviation of order imbalances. When they are expressed in terms of volatility and observable trading volume for purposes of empirical tests, the assumption made about the relationship between trading volume and order imbalances becomes crucial.

- Our model of market microstructure invariance assumes that standard deviation of order imbalances is proportional to trading volume raised to the two-thirds power, keeping price volatility constant. This choice is made naturally by assuming that trading games are the same across stocks, except for the speed with which they are being played.
- The model of invariant bet frequency takes the approach that standard deviation of order imbalances is proportional to trading volume. This suggests measuring illiquidity as the ratio of daily returns to daily volume.
- The model of invariant bet size takes the approach that standard deviation of order imbalances is proportional to a square root of trading volume. This suggests measuring illiquidity as the ratio of daily returns to the square root of daily volume.

Equilibrium Stories. The three models lead to different stories about what equilibrium trading in stock markets looks like. In equilibrium, traders search for ideas. When a good idea is found, a trader places a bet, expecting to profit from the good idea. The three models imply differences in markets with different levels of trading activity:

- In comparing active markets with inactive markets, our model is consistent with the idea that traders in all markets have the same level of skills and earn the same wages, traders discover ideas and place bets at the same rate, bets have the same expected profitability and execution costs. Bets in active markets are larger, but since they are held for shorter periods of time, the risk of the bets over their lifetimes does not vary across markets. In order to place bets at the same rate, traders specializing in inactive markets cover more stocks than traders specializing in active markets.
- In comparing active markets with inactive markets, the model of invariant bet rate is consistent with the idea that big traders trade in active markets and little traders trade in inactive markets. Traders in active markets find big ideas and place big bets, which they believe to be more profitable than bets in smaller markets. If traders place bets at similar rates, traders in active markets expect to make higher profits than traders in less active markets.
- In comparing active markets with inactive markets, the model of invariant bet size is consistent with the idea that a small number of traders cover inactive markets and a large number of traders cover active markets. If traders place bets at similar rates, the traders expect to earn similar levels of profits.

4.4 Nested Model Formulation for Testing

The model of microstructure invariance leads to empirical predictions about how bet size, market impact, and bid-ask spreads vary with the level of trading activity. The two alternative models—the model of invariance bet size and the model of invariant bet frequency—also make predictions about how bet size, market impact, and bid-ask spreads vary with the level of trading activity. These implications all have a convenient nested form that we use below in empirical tests.

For all three models, the implications about bet size in equations (34), (37), and (42) can be nested as

$$\frac{\tilde{Q}}{V} = \frac{\tilde{Q}^*}{V^*} \cdot \left(\frac{W}{W^*}\right)^{\alpha_0}. \quad (47)$$

For all three models, the implications about market impact from equations (35), (38), and (43) can be nested as

$$\frac{\lambda \cdot V}{\sigma \cdot P} = \frac{\lambda^* \cdot V^*}{\sigma^* \cdot P^*} \cdot \left(\frac{W}{W^*}\right)^{\alpha_1}. \quad (48)$$

Similarly, the implications concerning bid-ask spreads from equations (36), (39), and (44) can be nested as

$$\frac{\kappa}{\sigma \cdot P} = \frac{\kappa^*}{\sigma^* \cdot P^*} \cdot \left(\frac{W}{W^*}\right)^{\alpha_2}. \quad (49)$$

Our proposed model of market microstructure invariance implies

$$\alpha_0 = -2/3, \alpha_1 = 1/3, \alpha_2 = -1/3. \quad (50)$$

The model of invariant bet frequency implies

$$\alpha_0 = 0, \alpha_1 = 0, \alpha_2 = 0. \quad (51)$$

The model of invariant bet size implies

$$\alpha_0 = -1, \alpha_1 = 1/2, \alpha_2 = -1/2. \quad (52)$$

Below, we test these models in a slightly modified form.

For order size, we also examine the additional hypothesis that \tilde{Q} is a symmetric distribution such that unsigned bet quantity for the benchmark stock $|\tilde{Q}^*|$ is log-normally distributed and therefore unsigned bet quantity for all stocks $|\tilde{Q}|$ is also log-normally distributed. Letting \bar{q} denote the median of the log-normal distribution of unsigned bet quantity for the benchmark stock $|\tilde{Q}^*|$, then equation (47) can be written in a log-linear form

$$\ln \left[\frac{|\tilde{Q}|}{V} \right] = \ln(\bar{q}) + \alpha_0 \cdot \ln \left[\frac{W}{W^*} \right] + \tilde{\epsilon}, \quad (53)$$

where $\tilde{\epsilon}$ has a zero-mean normal distribution. This log-linear formulation is convenient for econometric testing.

For market impact and bid ask spread, we combine equations (48) and (49) into one measure of transaction costs which includes both market impact and bid-ask spread. Let X denote the unsigned quantity traded. Let $C(X)$ denote the transaction costs associated with trading the quantity X , measured in basis points per share. For the benchmark stock with trading activity W^* and volatility σ^* , let $\bar{\lambda}$ denote the market impact costs, measured in basis points, associated with trading one percent of average daily volume V . Similarly, let $\bar{\kappa}$ denote the bid-ask spread cost, also measured in basis points. Note that $\bar{\lambda}$ is different from λ in that the units have been changed to basis points (cents per hundred dollars) rather than dollars per share per share. Similarly, $\bar{\kappa}$ is different from κ in that units have been changed to basis points from dollars per share. Then $C(X)$ is given by

$$C(X) = \frac{1}{2} \bar{\lambda} \cdot \frac{\sigma}{\sigma^*} \cdot \left(\frac{W}{W^*} \right)^{\alpha_1} \cdot \frac{X}{(0.01)V} + \frac{1}{2} \bar{\kappa} \cdot \frac{\sigma}{\sigma^*} \cdot \left(\frac{W}{W^*} \right)^{\alpha_2}. \quad (54)$$

The first term on the right-hand side measures the component of transaction cost due to market impact, and the second term measures the component of transaction costs due to bid-ask spread. If $X = 0.01 \cdot V^*$ represents a trade equal to one percent of average volume in the benchmark stock, we have $\sigma/\sigma^* = W/W^* = X/(0.01 \cdot V) = 1$, in which case $C(X) = \bar{\lambda}/2 + \bar{\kappa}/2$. For example, if $\bar{\lambda}/2 = 8$ and $\bar{\kappa}/2 = 3$, then the expected cost of executing an order equal to one percent of average daily volume in the benchmark stock is 11 basis points. For the empirical work in this paper, the benchmark stock is defined as having a price of \$40 per share, expected trading volume of one million shares per day, and volatility of two percent per day.

In calculating the transaction cost in equation (54), both the price impact parameter $\bar{\lambda}$ and the bid-ask spread parameter $\bar{\kappa}$ are divided by two. The fraction one-half multiplies the market impact parameter $\bar{\lambda}$ because we assume that the order is executed by “walking the book” in a manner which averages the prices of the early quantities executed with an impact near zero and the last shares with impact near $\bar{\lambda}$. This is consistent with $\bar{\lambda}$ measuring the market impact of a trade of one percent of average daily volume in the benchmark stock. In equation (54), the fraction one-half multiplies $\bar{\kappa}$ because $\bar{\kappa}$ measures the two-sided bid-ask spread cost of a buy and a sell, while $C(X)$ measures a one-way transaction cost, which is either a buy or a sell, but not both. This is consistent with $\bar{\kappa}$ being a two-sided bid-ask spread for the benchmark stock.

It is challenging to design a reasonable test of these predictions. For empirical implementation, it is important to remember that they are formulated for independent orders \tilde{Q} , arriving to the market. More precisely, one must think about \tilde{Q} as the innovation to the target holdings of all investors who got the same “trading idea.” Of course, both independent orders \tilde{Q} and their arrival rates γ are unobservable. In actual trading, one independent trading decision may be shared between a number of investors who submit perfectly correlated orders. Each of these orders may be broken down into smaller pieces for execution, and an execution of a trade may have several different counter-parties and prices. Thus, independent trading decisions usually generate multiple reports in the data. This complicates the testing of trading game invariance.

For example, the TAQ database gives a time-stamped record of trades printed for NYSE and NASDAQ stocks. It is probably not a good idea to estimate γ as the average number of prints in TAQ data and also not a good idea to estimate the average \tilde{Q} as the average print size in TAQ data. Suppose that an independent trade generates on average μ prints. This may happen because several investors share the same trading idea or because their orders are shredded into several trades. Then the number of trade prints in TAQ data is $\gamma_{TAQ} = \mu \cdot \gamma$ per day, and the average trade size is $E\{\tilde{Q}_{TAQ}\} = E\{\tilde{Q}\}/\mu$. If the number of TAQ prints and the average TAQ print size are used to estimate the mean of the invariant $E\{\tilde{I}\}$ equal to $E\{\tilde{Q}\} \cdot P \cdot \sigma/\gamma^{1/2}$, the result is $E\{\tilde{Q}_{TAQ}\} \cdot P \cdot \sigma/\gamma_{TAQ}^{1/2} = E\{\tilde{I}\} \cdot \mu^{-3/2}$. This estimate is biased by a factor $\mu^{-3/2}$.

The parameter μ is usually not observable; moreover, it may vary across stocks. Since μ is unobservable, using average trade frequency and average trade size from TAQ data does not make it possible to calibrate the average level of price impact. If μ may vary across stocks in an unknown manner, it is not possible to use average trade frequency and average trade size from TAQ data to explain how price impact varies cross-sectionally across stocks. Whether μ is constant or varies across stocks, as a function of say stock price (based on tick size), is an interesting issue for further research. We expect that in TAQ data, the minimum tick size of one penny and the minimum lot size of one hundred shares may influence the distribution of print sizes as a function of stock price. Order shredding and algorithmic trading have evolved over time, as computer technology has reduced the costs of submitting and canceling orders.

In our empirical tests, we use a sample of portfolio transition orders representing entire orders. While it is possible that our portfolio transition orders may be on average larger or smaller than typical “bets,” we believe it is reasonable to assume that the factor by which

they may be larger or smaller than average bets does not vary across stocks with higher and lower levels of trading activity.

5 Order Imbalances in Adverse Selection and Inventory Models

So far, we have discussed our model and two alternatives using a reduced form approach which is not tied to a specific model of trading. All three models are, however, consistent with the adverse selection and inventory models in the literature. We illustrate this claim using the adverse selection model of Kyle (1985) and the inventory model of Grossman and Miller (1988).

5.1 Adverse Selection Models

In the model of Kyle (1985), competitive risk neutral market makers trade continuously against an informed trader and noise traders over one trading day. The informed trader privately observes the liquidation value of the risky asset, which has a normal distribution with variance σ_V^2 . The noise traders' inventory follows a Brownian motion process with variance σ_U^2 . The optimal strategy of the informed trader is to trade gradually so that the private information is revealed at the end of the trading day. In equilibrium, prices follow a Brownian motion with volatility σ_P , measured in dollars per share per square-root-of-a-day. Market depth is a constant given by

$$\lambda = \frac{\sigma_V}{\sigma_U}, \quad (55)$$

and this implies $\sigma_P = \sigma_V$. Assuming prices are far enough from zero, a reasonable approximation to σ_P in our notation is $\sigma_P = \sigma \cdot P$. Letting V denote daily trading volume as before, the price impact of trading a day's trading volume, scaled by volatility, is given by

$$\frac{\lambda \cdot V}{\sigma \cdot P} = \frac{V}{\sigma_U}. \quad (56)$$

This reduces the problem of calculating the level of market depth to the problem of calculating trading volume and calculating the standard deviation of noise trading. Neither of these problems is straightforward. In our opinion, the market microstructure literature has struggled with how to do these calculations in a manner which creates a graceful bridge from theory to empirical work.

Consider the problem of measuring trading volume. In the model, noise traders continuously place orders so that their inventory follows a Brownian motion process. In theory, the Brownian motion assumption implies that the trading volume of noise traders is infinite. For empirical implementation, it is reasonable to approximate the Brownian motion process as a compound process with trade arrival rate γ and trade size \tilde{Q} . In other words, we can think of noise trading as independent bets generated by a Poisson process, as in Back and Baruch (2004). With this approximation, our assumption that bets are statistically independent trading decisions is consistent with the independent increments implied by Brownian motion. An issue concerns the distinction between noise trading and informed trading. In

the model of Kyle (1985), the noise trades have no information content, but the trading volume is dominated by noise trading. In the three models of this paper, it is intuitively appropriate to think of bets as being based on information. To bridge this difference, we imagine that the traders placing the bets believe that their bets have information content, but the truth is that the vast majority of bets are based on noise. Occasionally, however, a bet is generated by a truly informed trader. With this intuition, we can think of trading volume as being defined by $V = \gamma \cdot E\{|\tilde{Q}|\}$, as we have assumed in our discussion of all three models in this paper so far.

Consider next the problem of measuring σ_U , the standard deviation of the inventory of noise traders. Empirically, σ_U is often approximated by examining order imbalance based on differences between uptick and downtick volume, using for example the Lee and Ready (1991) algorithm. Theoretically, since the bets are independently distributed, we have $\sigma_U = \gamma^{1/2} \cdot E\{\tilde{Q}^2\}^{1/2}$.

Plugging the expressions for V and σ_U into equation (56) yields

$$\frac{\lambda \cdot V}{\sigma \cdot P} = \frac{\gamma \cdot E\{|\tilde{Q}|\}}{\gamma^{1/2} \cdot E\{\tilde{Q}^2\}^{1/2}} = \frac{E\{|\tilde{Q}|\}}{E\{\tilde{Q}^2\}^{1/2}} \cdot \gamma^{1/2} \quad (57)$$

While the number of shares represented by bets may vary across stocks, it is reasonable to believe that the shape of the distribution is the same across stocks. For example, the model of Kyle (1985) suggests a normal distribution. As discussed later, in our empirical work below, we find evidence that $|\tilde{Q}|$ has a log-normal distribution. Both the log-normal and the normal distributions imply that the ratio of $E\{|\tilde{Q}|\}$ to $E\{\tilde{Q}^2\}^{1/2}$ is a constant. This reduces the problem of measuring market depth to the problem of measuring the square root of the arrival rate of bets $\gamma^{1/2}$.

Both our proposed model and the two alternative models are consistent with equation (57). The three models differ in how the bet rate γ is related to trading activity W . In our model of market microstructure invariance, $\gamma^{1/2}$ is proportional to $W^{1/3}$; in the model of invariant bet rate, $\gamma^{1/2}$ is a constant; in the model of invariance bet size, $\gamma^{1/2}$ is proportional to $W^{1/2}$.

The model of Kyle (1985) is consistent, at least in spirit, with all three of our proposed models. Although we prefer the model of market microstructure invariance on aesthetic grounds, which of the three models is better is ultimately an empirical question, which we address below.

5.2 Inventory Models

All three of the models proposed in this paper are also consistent with inventory models in which investors who hold inventories, such as market makers, respond to incentives to hold inventories based on *ex ante* expected profits. The following example, which has some of the flavor of the inventory model of Grossman and Miller (1988), illustrates this point.

Consider an asset whose payoff \tilde{v} is normally distributed with mean μ and variance σ_v^2 . Suppose noise traders buy a quantity \tilde{u} , which is normally distributed with mean zero and variance σ_u^2 . Let there be M competitive atomistic market makers who hold positive or negative inventories in order to profit from expected increases or decreases in price. Each

market maker has exponential utility with risk aversion parameter equal to 1. Market makers must pay a cost to buy a ticket to become market makers. The number of market makers M responds to financial incentives, equating the cost of a ticket with the certainty equivalent of the trading profits that can be obtained by holding inventories.

It is well-known that the demand function by market makers is $x = M \cdot (\mu - p) / \sigma_V^2$. Market clearing implies $x + \tilde{u} = 0$. As a result, the price is given by the formula $p = (\sigma_V^2 / M) \cdot u - \mu$. It is left as an exercise for the reader to show that the certainty equivalent of market maker trading profits is a function of the expected squared Sharpe ratio available from trading. The number of market makers M will therefore drive the expected squared Sharpe ratio to some level A^2 , which equates the certainty equivalent of trading profits to the cost of becoming a market maker. Conditional on \tilde{u} , trading profits are normally distributed with mean $(\sigma_V^2 / M) \cdot \tilde{u}$ and variance σ_V^2 . The expected squared Sharpe ratio (unconditional on \tilde{u}) is therefore $\sigma_V^2 \sigma_U^2 / M^2$. Equating this to A and solving for M yields $M = \sigma_V^2 \sigma_U^2 / A$. Allowing M to adjust endogenously, the price is given by $p = (A \cdot \sigma_V / \sigma_U) \cdot \tilde{u} - \mu$. The market impact of increasing \tilde{u} by one share of stock is therefore

$$\lambda = A \cdot \frac{\sigma_V}{\sigma_U}. \quad (58)$$

This calculation of λ is the same as the calculation in the model of Kyle (1985) in equation (55), up to a proportionality constant A . Therefore, the entire discussion of adverse selection models also applies to inventory models.

In particular, all three of the models proposed in this paper are consistent with inventory models. The market impact of trading one day's trading volume is proportional to the square root of the order arrival rate $\gamma^{1/2}$. The three different models give different implications for the relationship between trading activity and market impact because the three different models imply a different relationship between the bet rate γ and trading activity.

6 Data

6.1 Portfolio Transitions Data

The empirical implications of all of the three theoretical models are tested using a proprietary database of portfolio transitions from a leading vendor of portfolio transition services. During the evaluation period, this portfolio transition vendor supervised more than 30 percent of outsourced U.S. portfolio transitions. The sample includes about 2,680 portfolio transitions executed over the period from 2001 to 2005. This database is derived from the post-transition reports prepared by transition managers for their U.S. clients. This is the same database used by Obizhaeva (2009a, 2009b).

The portfolio transitions database contains data on individual transactions. Each observation has the following fields: a trade date, an identifier of a portfolio transition, its starting and ending dates, the name of the stock traded, the number of shares traded, buy or sell indicator, the average execution price, the pre-transition benchmark price, commissions, and fees. The data is given on separate lines for three trading venues: internal crossing networks, external crossing networks, and open market transactions. It is also given separately for

each trading day in a trading package executed over several days. Old “legacy” and new “target” portfolios usually overlap. For example, both portfolios may have positions in some large and therefore widely held securities. Instead of first selling overlapping holdings from legacy portfolios and then acquiring them into target portfolios by buying them back, these positions are transferred from one account to another as “in-kind” transfers which do not incur transaction costs. Thus, if the legacy portfolio holds 10,000 shares of IBM stock and the new portfolio holds 4,000 shares of IBM in a portfolio transition, then 4,000 shares are transferred in-kind and recorded as in-kind transfers. The balance of 6,000 shares is sold. If the transition manager sells these shares in two days with open market trades on the first day and both external crosses and open market trades on the second day, then there will be four lines in the database corresponding to IBM stock for this portfolio transition: a 4,000 share in-kind transaction, an open market trade the first day, an open market trade the second day, and an external cross the second day. Our empirical results do not depend at all on in-kind transfers. Instead, our empirical results are based on open market trades, external crosses, and internal crosses.

The original data is further grouped at order level. For example, aforementioned transactions are combined into one line corresponding to the order for IBM stock in portfolio transition A. This observation contains the name of the stock, the pre-transition benchmark price, buy or sell indicator, the number of shares executed over different trading venues, the average execution price for each of them, as well as the data on portfolio transition such as its beginning and ending dates.

The portfolio transition data are then merged with the CRSP data to add data on stock prices, returns, and volume. Only common stocks (CRSP share codes of 10 and 11) listed on the New York Stock Exchange (NYSE), the American Stock Exchange (Amex), and NASDAQ in the period of January 2001 through December 2005 are included in the sample. ADRs, REITS, and closed-end funds were excluded. Also excluded were stocks with missing CRSP information necessary to construct variables used for empirical tests, low-priced stocks defined as stocks with prices less than 5 dollars, and transition observations which appeared to contain typographical errors and obvious inaccuracies. Since it was unclear from the data whether adjustments for dividends and stock splits were made in a consistent manner across all transitions, all observations with non-zero payouts during the first week following the starting date of portfolio transitions were excluded from statistical tests.

After exclusions, the number of daily observations was 441,865 orders (204,780 buy orders and 237,085 sell orders).

Portfolio Transitions as Bets. The three proposed models make very different assumptions about how the sizes of bets vary across stocks with different levels of trading activity. To test these different assumptions empirically, it is necessary to identify the theoretical concept of a bet Q with actual data. Who are the traders placing bets in the stock market? One partial answer to this question is that professional equity managers are representative of such traders. Although these asset managers may try to trade on the basis of private information they work hard to collect, the difficulty professional asset managers have in beating the market suggests that many of their trades do not contain much private information, and thus may be considered as liquidity trades in the context of models like Kyle (1985).

If the portfolios put together by professional asset managers result from bets, then the differences in these portfolios represent the results of numerous bets in many different stocks. Therefore, we make the identifying assumption that the differences in professionally managed portfolios, while not exactly a random sample of bets themselves, vary in a manner proportional to the size of bets.

Our notation makes a distinction between the theoretical concept of a bet, denoted \tilde{Q} , and the individual trades made by transition managers. To emphasize the distinction, we use the notation X_i to represent the unsigned number of shares transacted in a given security during a given portfolio transition. The notation X_i represents the actual buy orders for target portfolios and the actual sell orders for legacy portfolios, excluding shares transferred in-kind. The index i ranges across 441,685 stock-transition pairs.

Portfolio transitions represent transactions in the differences between portfolios of two different professional asset managers. Note that the quantities traded often do not match the levels of positions in legacy and target portfolios, but rather the quantities traded match the differences in positions. When legacy and target portfolios overlap, the overlapping positions are transferred from one account to another one as “in-kind” transactions. These in-kind transactions are transfers of positions, not trades. Therefore, these in-kind transfers are excluded from the empirical tests in this paper. As a result, the trades used in the empirical tests below represent differences in portfolio across two different asset managers. We focus on transactions rather than positions because our models are designed to explain the cross-sectional differences in the execution data. The three models establish a link between trading activity (the product of volume, price, and volatility) and trading costs with trade sizes. The models are not meant to explain the absolute levels of holdings.

Portfolio Transitions and Implementation Shortfall. To estimate transaction costs $C(X)$, we use the concept of implementation shortfall as developed by Perold (1988). Specifically, we estimate costs by comparing the average execution prices of portfolio transition trades with closing prices the evening before the transition trades begin executing. In many applications, the use of implementation shortfall to measure transaction costs is problematic. Portfolio transitions avoid the usual problems associated with implementation shortfall and therefore provide one of the rare situations where implementation shortfall works well as a measure of transaction costs.

The fundamental problem with using implementation shortfall to measure transaction costs is that the actual quantities traded may not be known at the start date due to order cancellations or changes in trading intentions which occur after the start date and affect actual quantities traded. Statistically, the resulting selection bias problem can lead to significant underestimation of transaction costs if orders tend to be either canceled when prices move in an unfavorable direction or increased when prices move in a favorable direction. Implementation shortfall can also lead to biased estimates of transaction costs if the trading decisions are based on short-lived private information which is incorporated into prices during the period when the trades occur. Portfolio transition data has several important properties which make it particularly advantageous for estimating transaction costs using implementation shortfall.

For each stock in a portfolio transition, the quantities to be traded are known precisely at

a specific time before the trades are actually executed. The composition of legacy and target portfolios is fixed in the mandates that transition managers receive the night before portfolio transitions begin. These managers then execute orders regardless of the unfolding price dynamics. This makes it reasonable to assume that the initial orders or trading intentions are exactly equal to the quantities subsequently traded. Thus, portfolio transition data tend not to be affected by the selection bias problem that would affect databases of trades where the quantities traded change in a manner correlated with price changes between the time orders are placed and the time they are executed, canceled, or increased.

The timing of portfolio transitions is likely determined by a schedule of investment committee meetings of institutional sponsors, who make decisions to undertake transitions. The investment committee meets regularly on schedules set well in advance of the meetings. Among the issues boards discuss are the replacement of fund managers and the changes of asset mix. If a decision is made to replace a portfolio manager, then a portfolio transition is arranged shortly after the meeting. These decisions are unlikely to be correlated with short-term price dynamics of individual securities during the period of the transition. This makes it possible to obtain estimates of price impact and spread that are not affected by short-lived information likely to be incorporated into prices during the period the transition trades are executed.

These properties of portfolio transitions are not often shared by other data. Consider a database built up from trades by a mutual fund, a hedge fund, or a proprietary trading desk at an investment bank. In such samples, the trading intentions of traders may not be recorded in the database. For example, a database might time stamp a record of a trader placing an order to buy 100,000 shares of stock but not time stamp a record of the trader's secret intention to buy another 200,000 shares after the first 100,000 shares is bought. Furthermore, trading intentions before traders begin trading may not coincide with realized trades because the trader changes his mind as market conditions change. Traders often condition their trading strategies on prices by using limit orders or by canceling parts of their orders, thus hard-wiring into their strategies a selection bias problem for using such data to estimate transaction costs. The trading intentions themselves can be significantly affected by overall price dynamics, e.g., traders may be following trends or playing contrarian strategies. This dependence of actually traded quantities on prices, consequently, makes it impossible to use implementation shortfall in a meaningful way to estimate market depth and bid-ask spreads from data on trades only.

6.2 CRSP Data: Prices, Volume, and Volatility

Our three models use trading activity to explain how transaction costs and expected trade size vary across stocks. Trading activity is the product of trading volume (in shares), share price (in dollars), and volatility (percentage standard deviation of daily returns). To measure implementation shortfall, a pre-trade benchmark price is needed. The components of trading activity and the pre-trade benchmark are calculated from CRSP data.

As a pre-trade price, denoted $P_{0,i}$ for i th order, we use the closing price of the corresponding security on the evening before the portfolio transition trades begin. More precisely, a portfolio transition involves trades in numerous stocks. Typically, many of the stocks are traded on the first day of the transition. For each stock in the transition, the benchmark

price $P_{0,i}$ is the closing price the evening before the first trade is made in any of the stocks in the portfolio transition, even if a particular stock itself is not traded on the first day.

As expected trading volume during portfolio transitions, denoted V_i for i th order, we use the average daily trading volume (in the number of shares) of the corresponding security in the pre-transition month.

We estimate the expected volatility of daily returns, denoted σ_i for i th order, using past daily CRSP returns for the stock involved in the i th trade. We use two different estimates of volatility, a simple estimate equal to average daily volatility from the past month and a more complicated estimate from an ARIMA model.

For each security, we first calculate the monthly standard deviation of returns from daily CRSP returns data. Let $r_{i,t,k}$ denote the CRSP return for the k th day of month t for stock involved in the i th trade. Letting $N_{i,t}$ denote the number of CRSP trading days in month t , then the standard deviation for month t for the stock in i th trade, denoted $\sigma_{i,t}^m$, is

$$\sigma_{i,t}^m = \left[\sum_{k=1}^{N_{i,t}} r_{i,t,k}^2 \right]^{1/2} \quad (59)$$

We do not de-mean the returns data since the mean return in a month is very small relative to the standard deviation. We also do not adjust the estimates for autocorrelation of returns by adding a cross-product of adjacent returns, since this might result in negative estimates of volatility for some stocks.

One simple estimate of daily volatility for the stock in trade i for month t , denoted $\sigma_{i,t}^h$, is the monthly standard deviation converted to daily units:

$$\sigma_{i,t}^h = \frac{1}{N_{i,t}^{1/2}} \sigma_{i,t}^m. \quad (60)$$

We also estimate an ARIMA model to obtain another forecast of the daily return standard deviations for each stock j and month t . To reduce effects from the positive skewness of the standard deviation estimates, we use a logarithmic transformation for the volatility. We estimate a third-order moving average process for the changes in $\ln \sigma_{i,t}^m$ over the whole sample from 2001 to 2005:

$$(1 - L) \ln \sigma_{i,t}^m = \Theta_0 + (1 - \Theta_1 L - \Theta_2 L^2 - \Theta_3 L^3) u_t \quad (61)$$

The conditional forecast for the volatility of daily returns is

$$\sigma_{i,t}^e = \frac{1}{N_{i,t}^{1/2}} \exp \left[\ln \sigma_{i,t}^m + \frac{1}{2} \hat{V}(u) \right] \quad (62)$$

where $\hat{V}(u)$ is the variance of the prediction errors of the ARIMA model.

In the empirical tests below, both $\sigma_{i,t-1}^e$ and $\sigma_{i,t-1}^h$ are used as proxies for σ_i in the i th transition trade. It is possible that using these proxies in our regressions may introduce an error-in-variables problem due to the volatility estimates themselves having errors. The empirical results are quantitatively similar for both proxies. Thus, only results for the estimates based on $\sigma_{i,t-1}^e$ are reported. We use the pre-transition variables known before portfolio transition trades in order to avoid any spurious effects from using contemporaneous variables, except to the extent that the ARIMA model uses in-sample data to estimate model parameters.

6.3 Descriptive Statistics

Table 1 reports statistical characteristics of both securities traded and individual transition trades. Statistics are calculated for all securities in aggregate as well as separately for ten groups of stocks sorted by average daily dollar volume. Instead of dividing the securities into ten deciles with the same number of securities, volume break points are set at the 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of trading volume for the universe of stocks listed on the NYSE with CRSP share codes of 10 and 11. Group 1 contains stocks in the bottom 30th percentile by dollar trading volume. Group 10 contains stocks in the top 5th percentile. Smaller percentiles for the more active stocks make it possible to focus on the stocks which are most important economically. For each month, the thresholds are recalculated and the stocks are reshuffled across bins.

Panel A of Table 1 reports statistical properties of the securities in the sample. There is a column for each of the ten groupings as well as a column which reports aggregate statistics. For the entire sample of stocks, the median trading volume is \$19.99 million per day, ranging from \$1.22 million for the lowest volume decile to \$212.55 million for the highest volume decile. Since the average dollar volume ranges over more than two orders of magnitude, this variation in the data should create statistical power helpful in determining how transaction costs and trade size vary with dollar volume. Panel A reports that the median volatility for all stocks is a standard deviation in returns of 1.85 percent per day. Volatility tends to be slightly higher in the lower volume deciles than the higher ones. The volatility for the lowest volume decile is 2.04 percent, and it is 1.76 percent for the highest volume group. This implies that the amount of variation in volatility across stocks is somewhat small.

Panel A reports that the median bid-ask spread, a quoted spread obtained from the transition database, is 11.54 basis points. Its mean is 23.67 basis points. From lowest volume grouping to highest volume grouping, the median bid-ask spread declines monotonically across groups from 38.16 basis points in the lowest volume group to 4.83 basis points in the highest volume group. This monotonic decline of almost one order of magnitude in reported bid-ask spreads is so large that significant statistical power should be generated to differentiate the predictions of the three models for bid-ask spreads. This, of course, assumes that the spreads reported in Panel A, which are quoted spreads not estimated from implementation shortfall, also show up in statistical estimates based on implementation shortfall.

For example, our proposed model of market microstructure invariance predicts that spreads should decrease one-third of one percent for each increase of one percent in trading volume, holding volatility constant. From lowest to highest quintile, volume increases by a factor of $212.55/1.22 = 174.22$. A back-of-the-envelope prediction for the decrease in spreads across these groups is the one-third power of the increase in volume, i.e., $174.22^{1/3} = 5.58$. The actual decrease in spreads is a factor of $38.16/4.83 = 7.90$. While this back-of-the-envelope calculation suggests that spreads decrease more than our model predicts, the difference between 5.58 and 7.90 is small enough to warrant further statistical investigation. It is possible that the estimates of effective spreads obtained from implementation shortfall are different from quoted spreads, and thus results based on implementation shortfall will be different.

Panel B of Table 1 reports properties of order sizes in the portfolio transition data. The

mean portfolio transition order is 3.90 percent of average daily volume of the stock traded. The means decline monotonically across the ten volume groups from 15.64 percent in the smallest group to 0.49 percent in the largest. The median portfolio transitions order is 0.56 percent of average daily volume. The median also declines monotonically across the ten volume groups, from 3.48 percent in the smallest to 0.14 percent in the largest. The fact that the medians are much smaller than the means indicates that the order size is skewed to the right. This is to be expected, since the order size is a non-negative number, and there may be some very small trades from highly diversified portfolios involving smaller transitions as well as very large trades from less diversified portfolios involving larger transitions. Below we show that the distribution of order size is very close to a log-normal distribution. For a log-normal distribution, the log of the median is the same as the mean of the log. Thus, we consider both the mean and the median in the following discussion.

The significant variation in mean and median order size as a fraction of average daily volume across dollar volume deciles is expected to have several important effects on statistical estimates.

On the one hand, the larger order size in the lower deciles generates more statistical power for using implementation shortfall to estimate market impact in the lower deciles than in the higher ones (holding constant market impact of an order with the size being a constant percentage of trading volume). On the other hand, our proposed model predicts that price impact increases as volume increases, holding order size as a percentage of volume constant. Of the other two models, one predicts no change in impact while the other predicts an even larger increase. Since each of the three models makes very different predictions concerning how market impact varies with trading activity, all three models will try to extrapolate the statistical power from one volume group to another, but the extrapolation will operate differently for each model.

Second, the variation in the order size across volume groups makes it possible to test the assumptions of the three models concerning how the size of liquidity trades varies across stock with different trading activity levels. From highest to lowest group, median daily volume increases by a factor of $212.55/1.22 = 174.22$. According to our proposed model of trading game invariance, average trade size as a percent of volume should decrease by two-thirds of one percent for every one percent increase in volume, holding volatility constant. As a back-of-the-envelope calculation, this implies that the decrease in trade size from lowest to highest quintile should be the two-thirds power of the increase in dollar volume, i.e., $174^{2/3} = 31.2$. The actual median trade size decreases as a fraction of average volume by a factor of $3.48/0.14 = 24.86$, from lowest group to highest group. While the back-of-the-envelope calculation of 31.2 does not exactly match the factor of 24.86, the numbers are close enough to suggest that further statistical investigation is warranted.

7 Empirical Results: Order Size

All three proposed models have distinctly different implications concerning the cross-sectional variation of quantities traded, market impact, and bid-ask spread. Portfolio transitions data are used to test these implications.

Order size data for portfolio transitions is used to test model implications for liquidity

order sizes. These tests provide evidence for our hypothesis of trading game invariance is. They do not provide any evidence on hypotheses of market impact invariance and bid-ask spread invariance, which are parts of our model of market microstructure invariance.

Implementation shortfall data for portfolio transitions is used to test model implications for market impact and bid-ask spread. These tests provide evidence for joint hypotheses of trading game invariance with market impact invariance or bid-ask invariance, respectively. For consistency, we refer to all tests as tests of market microstructure invariance.

7.1 Nested Log-Linear Specification

The three theoretical models make distinctly different assumptions concerning how the sizes of bets vary with the level of trading activity. The predictions (53) can be expressed as a simple linear regression of the form

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + \tilde{\epsilon}. \quad (63)$$

This equation relates the size of the trade X_i as a fraction of average volume V_i to the level of trading activity W_i , defined as the product of benchmark price $P_{0,i}$, last month's trading activity V_i , and estimated volatility σ_i :

$$W_i = P_{0,i} \cdot V_i \cdot \sigma_i. \quad (64)$$

The scaling constant $W^* = (40)(10^6)(0.02)$ corresponds to W_i for the hypothetical benchmark stock with price \$40 per share, trading volume of one million shares per day, and volatility of 0.02. In this regression, the model of market microstructure invariance predicts $\alpha_0 = -2/3$, the model of invariant bet frequency predicts $\alpha_0 = 0$, and the model of invariant bet size predicts $\alpha_0 = -1$.

A potential econometric difficulty with the specification in equation (63) is that taking the log of trade size as a fraction of average daily volume has the potential to create large negative outliers out of tiny, economically meaningless trades. Below, we show that the shape of the distribution of trade size (conditional of the level of trading activity W) closely matches a log-normal distribution. This implies that the error term in the regression above is normally distributed. As a result, we can conclude that tiny orders are not skewing the distribution to the left, potentially distorting the estimates.

OLS Estimates of Order Size. Table 2 presents estimates for the coefficients in equation (63). The first column of the table reports the results of a regression pooling all the data. The four other columns in the table report results for four separate regressions in which the four parameters are estimated separately for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells.

The estimate for α_0 is $\hat{\alpha}_0 = -0.63$ with standard error of 0.008. Economically, the point estimate for α_0 is close to the value predicted by the model of market microstructure invariance $\alpha_0 = -2/3$, but this model is strongly rejected ($F = 17.03, p < 0.0001$) because the standard error is very small. This point estimate is so different from the predictions of the two other models that they are rejected by overwhelming margins.

When sample is broken down into NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells, the point estimates -0.63 , -0.60 , -0.71 , and -0.61 are consistently close to the predicted value of $-2/3$, but the model of market microstructure invariance is rejected in all cases due to the low standard error.

Separate Coefficients for Price, Volume, and Volatility. The model of market microstructure invariance predicts that order size is explained by trading activity. Time clock irrelevance and Modigliani-Miller irrelevance suggest that the individual components of trading activity—price, volume, and volatility—should have no additional explanatory power beyond the explanatory power provided by trading activity W .

To test this idea, table 3 presents results for the OLS regression

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] - \frac{2}{3} \ln \left[\frac{W_i}{W_*} \right] + b_1 \cdot \ln \left[\frac{\sigma_i}{(0.02)} \right] + b_2 \cdot \ln \left[\frac{P_{0,i}}{(40)} \right] + b_3 \cdot \ln \left[\frac{V_i}{(10^6)} \right] + \tilde{\epsilon}_i. \quad (65)$$

This regression imposes on $\ln(W_i/W_*)$ the coefficient $\alpha_0 = -2/3$ predicted by the model of market microstructure invariance. It then allows the coefficient on the three components of W_i to vary freely. Thus, the model of trading game invariance predicts $b_1 = b_2 = b_3 = 0$. The model of invariant bet frequency predicts $b_1 = b_2 = b_3 = 2/3$, and the model of invariant bet size predicts $b_1 = b_2 = b_3 = -1/3$.

A comparison of table 3 with table 2 shows that increasing the number of parameters estimated from two to four increases the adjusted R^2 of the regression from 0.3188 to 0.3211.

The table reports point estimates for the coefficient on volatility of $\hat{b}_1 = 0.25$, the coefficient on price of $\hat{b}_2 = 0.16$, and the coefficient on share volume of $\hat{b}_3 = 0.01$, with corresponding standard errors of 0.031, 0.014, and 0.009, respectively (t-values of 8.17, 11.05, and 0.86). The regression fails to reject the hypothesis $b_3 = 0$, supporting the model of market microstructure invariance. The coefficients on volatility and price are significantly positive, indicating that trade size, as a fraction of average daily volume, does not decrease with increasing volatility and volume as fast as predicted by the model of market microstructure invariance.

Comparison of Three Models. Table 4 estimates the constant term in the regression under the assumption that the coefficient α_0 of $\ln(W_i)$ in equation (63) is fixed at the values implied by the three models. For the model of market microstructure invariance, fixing the coefficient at $\alpha_0 = -2/3$ results in a constant term estimate of log order size as a fraction of average daily volume, $\ln[\bar{q}]$, equal to -5.69 . We will see later that distribution of order sizes is close to be log-normal. With a log-normal distribution, the log of the median is the mean of the log. This implies a median order size of 33.75 basis points of volume, or 0.3375% of average daily volume, for the benchmark stock.

For each sample, the reported maximum likelihood function is much greater for the model of market microstructure invariance than for the two alternative models. For prior probabilities assigning a non-trivial probability to the model of market microstructure invariance, a Bayesian statistician will assign posteriors of about 100% to the proposed model and conclude that alternative models are highly unlikely. His conclusion will be almost independent of his priors, since the extensive amount of data overwhelms the priors.

Economic Interpretation. From the perspective of meaningful economic magnitudes, the R^2 results in tables 2, 3, and 4 provide strong support for the model of market microstructure invariance. When the coefficient on $\ln W$ is fixed at the model-implied value of $\alpha_0 = -2/3$ and only one parameter (constant term) is estimated, we obtain an adjusted R^2 of $R^2 = 0.3177$. By contrast, when the values of α_0 implied by the two alternative models are fixed at either $\alpha_0 = 0$ or $\alpha_0 = -1$, then we obtain adjusted R^2 values of $R^2 = 0.0000$ (by definition) and $R^2 = 0.2105$, respectively. Clearly, the model of market microstructure invariance has superior explanatory power. The results reflect the fact that portfolio transition orders for less active stocks tend to be for a much greater percentage of average daily volume than portfolio transition orders for more active stocks.

When the parameter α_0 is estimated rather than held fixed, changing α_0 from the model value of $\alpha_0 = -2/3$ to the estimated value of $\hat{\alpha}_0 = -0.63$ increases the adjusted R^2 only modestly, from $R^2 = 0.3177$ reported in table 4 to $R^2 = 0.3188$ reported in table 2. Although statistically significant, the addition of the extra variable does not add much explanatory power. When the specification is further relaxed to estimate coefficients on price, volume, and volatility separately, the addition of two extra parameters increases the adjusted R^2 from $R^2 = 0.3188$ reported in table 2 to $R^2 = 0.3211$ reported in table 3. Although statistically significant, the improvement in R^2 is again modest. We interpret these results as implying that portfolio transition order size as a fraction of average daily volume varies across stock with different levels of trading activity in a manner consistent with our model of market microstructure invariance.

Dummy Variables as a Robustness Check. Figure 2 uses ten dummy variables for volume groups to estimate the three versions of the regression

$$\ln \left[\frac{X_i}{V_i} \right] = \left[\sum_{j=1}^{10} \mathbb{I}_{j,i} \cdot \ln [\bar{q}_j] \right] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + \tilde{\epsilon}, \quad (66)$$

in which the value of α_0 is fixed at the value predicted by the three theoretical models, and the ten dummy variables are estimated. Dummy variables $\bar{q}_j, j = 1, \dots, 10$ quantify the average size of liquidity order for the benchmark stock based on data for j th volume group. If the regression is well-specified, then the values of the dummy variables resulting from fixing α_0 at the corresponding level should be constant across volume groups. In figure 2, the ten dummy variables are plotted, along with their 95% confidence bounds. The value of the constant term from the one-parameter regression is shown as a horizontal line. If the regression is well-specified, then the values of the dummy variables \bar{q}_j should line up along the horizontal line.

In the first graph in the figure, the ten dummy variables resulting from fixing $\alpha_0 = -2/3$, as implied by the model of market microstructure invariance, are plotted. It can be seen that these dummy variables line up nicely along the horizontal line. Upon close inspection however, it is possible to notice that the 95% confidence bounds are so narrow that some of the points lie outside the 95% confidence bound, consistent with the previous rejection of the model.

In the second graph in the figure, the ten dummy variables resulting from fixing $\alpha_0 = 0$, as implied by the model of invariant trade frequency, are plotted. Instead of lining up on the

horizontal line, the dummy variables decline monotonically from a level very far above the line to a level very far below it, far outside the 95% confidence bounds. This is consistent with strong rejection of this model.

In the third graph in the figure, the ten dummy variable resulting from fixing $\alpha_0 = -1$, as implied by the model of invariant trade frequency, are plotted. Instead of lining up on the horizontal line, the dummy variables increase monotonically from a level far below the line to a level far above it, far outside the 95% confidence bounds. This is consistent with strong rejection of this model as well.

What is the intuition underlying these patterns? Why does the model of invariant bet frequency significantly overestimate trade size for high-volume stocks? Why does the model of invariant bet size significantly underestimate it? The reason is that the first model incorrectly attributes large trading volume entirely to large bet size whereas the second model mistakenly explains this volume entirely by high bet frequency. If our model of market microstructure invariance, however, is true, then two-thirds of this volume comes from higher frequency of bets placed and one-third from their larger size. Note that consistent with the two-to-one ratio between bet frequency and bet size, the deviation from a horizontal line is twice as large for the first alternative model as for the second one.

The data on the average size of portfolio transition orders strongly support assumptions made in the model of market microstructure invariance and soundly reject assumptions made in alternative models. The reason is that variations in trading activity are associated with variations in both frequency and size of liquidity trades; neither remains constant.

7.2 Invariant Order Size Distribution as Log-Normal

The model of market microstructure invariance predicts that the entire distribution of order sizes is invariant across stocks when order size X is adjusted for existing differences in trading activity W . Specifically, the distribution of $(X_i/V_i) \cdot W_i^{2/3}$ should be invariant to differences in the level of trading activity. So far, our tests have provided evidence that this invariance holds reasonably well for the means of distributions, but we have not examined whether it holds for the higher moments.

Figure 3 illustrates whether it is also true for entire distributions. We calculate empirical distribution for logs of W -adjusted transition order sizes given by $\ln[(X_i/V_i) \cdot W_i^{2/3}]$ for stocks sorted into ten volume groups and five standard deviation groups. As before, volume groups are based on average dollar trading volume with thresholds corresponding to 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of dollar volume. Standard deviation groups are quintiles based on thresholds corresponding to 20th, 40th, 60th, 80th percentiles of standard deviations. These thresholds are obtained for the sample of common NYSE-listed stocks. On each plot, we superimpose a bell-shaped density function of a normal distribution with the common mean and variance across subplots, corresponding to a mean of -5.69 and a variance of 2.50 estimated for the entire sample of W -adjusted order sizes. If the model of market microstructure invariance describes data well and order sizes are distributed as a log-normal, then these density functions should be identical.

Figure 3 shows a plot of the distributions for volume groups 1, 4, 7, 9, and 10 and for volatility groups 1, 3 and 5. These 15 distributions of W -adjusted order sizes are strikingly similar. Results for the remaining 35 subgroups also look very similar and therefore are not

presented in this paper. The visual similarity of distributions is consistent with the similarity of their first four moments. The means range from -6.06 to -5.42, close to the mean of -5.69 for the entire sample. The variances range from 2.19 to 2.82, also close to the variance of 2.50 for the entire sample. The skewness ranges from -0.21 to 0.05, close to skewness of zero for the normal distribution. The kurtosis ranges from 2.66 to 3.48, also close to the kurtosis of 3 for a normal random variable. These results suggest that the assumption that order sizes are distributed as log-normal random variables is a reasonable one. Moreover, order sizes adjusted for differences in trading activity W indeed have very similar distributions.

These results are consistent with the interpretation that the error terms in the regressions reported in tables 2, 3, and 4 are approximately IID normal. If the distribution of order sizes were not log-normal, the taking of logs in the regressions reported in tables 2, 3, and 4 could have created problems related to heteroscedasticity. For example, taking logs of very small trade sizes could potentially have created outliers in the distribution, giving tiny, economically meaningless trades an inordinately large influence on reported results. The fact that the distribution of order sizes fits the invariant log-normal supports the use of OLS regressions as statistically appropriate.

8 Empirical Results: Market Impact and Bid-Ask Spread

The three theoretical models make distinctly different predictions concerning how transaction costs vary with the level of activity. These predictions can be expressed in terms of a simple equation (54) that relates the transaction cost $C(X)$ to four parameters. For a trade in the benchmark stock equal to one percent of average daily volume, the two parameters $\bar{\lambda}$ and $\bar{\kappa}$ represent the market impact and bid-ask spread in basis points. The two remaining parameters, the exponents α_1 and α_2 , describe how the models extrapolate market impact and spread costs across stocks with different levels of activity. Since the three models make dramatically different predictions concerning α_1 and α_2 , it should be possible to test the models by estimating all four parameters.

Identifying Assumptions and Estimation Strategy. We make the identifying assumption that, in a correctly specified model, the implementation shortfall from the portfolio transition database is an unbiased estimate of the transaction cost $C(X)$. We can think of implementation shortfall as representing the sum of two components: (1) the transaction costs incurred as a result of market impact and bid-ask spread and (2) the effect of random price changes between the time the benchmark price is set and the time the trades are executed. Since implementation shortfall is an unbiased estimate of transaction costs, we can think of the random price changes as an error in a regression. This suggests an estimation strategy of adding an error term to $C(X)$, then estimating the four parameters using a non-linear regression. The regression is non-linear because the exponent parameters α_1 and α_2 appear in $C(X)$ in a non-linear manner. Implementation shortfall data for portfolio transitions is used to test model predictions for market impact and bid-ask spread.

To implement this strategy, two adjustments are made, one based on statistics and one

based on economics.

First, since the errors in the regression are likely to be proportional in size to the return volatility of the stock, both the right-hand-side and left-hand-side variables are divided by return volatility σ_i . This has the effect of making a crude correction for a heteroscedasticity problem which would otherwise occur. Furthermore, the imperfectly observed volatility σ_i is replaced by its estimate $\sigma_{i,t-1}^e$. To the extent that $\sigma_{i,t-1}^e$ is an imperfect estimate of σ_i , the problem of bias associated with errors in variables is reduced by placing this variable on the right-hand-side.

Second, we adjust our estimation procedure for the fact that transition managers have access to different pools of liquidity. Transition orders can be executed through internal crossing networks, through external crossing networks, or in open market transactions. Market impact and bid-ask spread may be different across trading venues. Some of the portfolio transitions are the result of internal crosses. In an internal cross, one of the transition manager's customers buys from the other at some price. In fact, it is possible that both the buyer and the seller represent different portfolio transitions being implemented simultaneously. Internal crosses with other types of customers also occur, for example, crosses against flows from a passive investment management unit affiliated with the same firm as the transition management unit. Since the buyer and the seller pay the same price, it seems reasonable to assume that there is no effective spread incurred for internal crosses but there is spread for external crosses and open market transactions.

Concerning market impact for crosses and open market transactions, it is assumed that the transition manager optimally chooses the percentages of the orders to execute via these trading venues. To the extent that crosses are cheaper than open market transactions, this is expected to show up as a larger percentage of the orders being crossed than executed in open markets, not as lower market impact and spread costs on crosses. The fact that both crosses and open market transactions are used in a significant proportion of orders suggests that there are significant pools of liquidity in both crossing networks and open markets, i.e., neither dominates the other. We therefore assume that there is market impact associated with internal crosses that is equal in magnitude to the impact of external crosses and open market trades. We will later relax the assumptions of our benchmark specification and allow transaction costs vary across trading platforms.

Impact and Spread Estimates in Nested Non-Linear Regression. Let X_i denote the number of shares in the i th order. Let $X_{omt,i}$ and $X_{ec,i}$ denote the number of these shares executed in open market transactions and external crosses, respectively. Then the number shares crossed internally, denoted $X_{ic,i}$ is by definition given by $X_{ic,i} = X_i - X_{omt,i} - X_{ec,i}$.

With the two aforementioned adjustments, the four parameters $\bar{\lambda}$, $\bar{\kappa}$, α_1 , α_2 are estimated in the following non-linear regression:

$$\frac{\mathbb{I}_{BS,i}(P_{ex,i} - P_{0,i})}{P_{0,i}} \cdot 10^4 \cdot \frac{(0.02)}{\sigma_i} = \frac{1}{2} \cdot \bar{\lambda} \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_1} \cdot \frac{X_i}{(0.01)V_i} + \frac{1}{2} \cdot \bar{\kappa} \cdot \frac{(X_{omt,i} + X_{ec,i})}{X_i} \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} + \tilde{\epsilon}. \quad (67)$$

In this non-linear regression, the observed data items have subscript i : $P_{ex,i}$, $P_{0,i}$, $\mathbb{I}_{BS,i}$, W_i , X_i , V_i , σ_i . Variable $\mathbb{I}_{BS,i}$ denotes trading direction being equal to 1 for buy orders and -1 for sell orders. Since $P_{0,i}$ denotes the benchmark price established the night before the transition

begins and $P_{ex,i}$ denotes the average execution price, the expression $\mathbb{I}_{BS,i}(P_{ex,i} - P_{0,i})/P_{0,i} \cdot 10^4$ is the implementation shortfall measured in basis points. The term $(0.02)/\sigma_i$ adjusts for heteroscedasticity. The trading activity variable W_i is defined as the product of benchmark price $P_{0,i}$, last month's volume V_i , and estimated volatility σ_i . The scaling constant $W^* = (40)(10^6)(0.02)$ corresponds to the trading activity for the hypothetical benchmark stock with price \$40 per share, trading volume of one million shares per day, and volatility of 0.02. The term $X_i/(0.01)V_i$ is the size of the trade relative to average volume, scaled so that the size has a value of one for a trade of one percent of average daily volume. The variables are scaled so that $\bar{\lambda}/2$ estimates in basis points the market impact costs of a trade of one percent of average daily volume, and $\bar{\kappa}/2$ estimates in basis points the effective spread cost.

To adjust standard errors for positive contemporaneous correlation in returns, the 441,865 observations are pooled by week over the 2001-2005 period into 4,389 clusters across 17 industry categories using the pooling option on Stata.

Recall that the three models make very different predictions concerning α_1 and α_2 . The model of market microstructure invariance predicts $\alpha_1 = 1/3$ and $\alpha_2 = -1/3$. The model of invariant bet frequency predicts $\alpha_1 = 0$ and $\alpha_2 = 0$. The model of invariant bet size predicts $\alpha_1 = 1/2$ and $\alpha_2 = -1/2$.

The results of the non-linear regression are reported in Table 5. The estimates for the parameters α_1 and α_2 are strongly supportive of the model of market microstructure invariance over the alternatives. The estimate for α_1 is $\hat{\alpha}_1 = 0.33$ with standard error of 0.024. This point estimate is almost exactly equal to the value of $1/3$ predicted by the model of market microstructure invariance. Furthermore, the standard error is sufficiently small that predictions of the other two models, $\alpha_1 = 0.50$ and $\alpha_1 = 0$, are soundly rejected.

The estimate for α_2 is $\hat{\alpha}_2 = -0.39$ with standard error of 0.025. This estimate is somewhat more negative than the value $\alpha_2 = -1/3$ predicted by the model of market microstructure invariance, by a margin of slightly more than two standard errors. The result suggests that effective bid-ask spreads decrease faster than the model predicts as trading activity increases. This is consistent with the back-of-the-envelope calculation from Table 1 suggesting that quoted bid-ask spreads decline faster than the model predicts as activity increases.

A Stata F-test for the joint hypothesis $\alpha_1 = 1/3$, $\alpha_2 = -1/3$ is rejected with a borderline p-value of 0.0742. Similar F-tests soundly reject the other two models with p-values less than 0.0001.

The estimate for half-price-impact is $\hat{\lambda}/2 = 2.85$ basis points with standard error of 0.245 ($t = 11.60$), and the estimate for half-spread is $\hat{\kappa}/2 = 6.30$ basis points with standard error of 1.131 ($t = 6.31$). These estimates imply that a hypothetical trade in the benchmark stock equal to one percent of daily volume incurs a market impact cost of 2.85 basis points and a spread cost of 6.30 basis points. The total cost of 9.15 basis points represents 3.66 cents per share for a \$40 stock, or \$366 for the hypothetical 10,000 share benchmark block.

The estimate for the bid-ask spread $\hat{\kappa}$ is double the point estimate for the half-spread $\hat{\kappa}/2$, i.e. 12.60 basis points. This estimate is somewhat higher than the median spread of 8.09 basis points reported in Table 1 for volume group 7, to which the hypothetical benchmark stock would belong. It is, however, similar to its mean value of 12.14 basis points.

Similarly, the estimate for $\hat{\lambda}$ is double the estimate of 2.85 basis points for $\hat{\lambda}/2$, i.e., it is 5.70 basis points. This implies that a trade of 10,000 shares, one percent of average daily volume in the benchmark stock, increases the \$40 price by 5.70 basis points, or 2.28 cents

per share.

When the four parameters are estimated separately for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells, the results are also supportive of the model of market microstructure invariance. In three of the four regressions (with the exception of NYSE Buys), the estimated coefficient for α_1 is close to the predicted value of $1/3$, but α_2 is more negative than predicted. In these three cases, F-tests either fail to reject or narrowly reject the predictions of market microstructure invariance that $\alpha_1 = 1/3, \alpha_2 = -1/3$, with p-values of 0.1057, 0.9114, and 0.0443.

The disaggregated results for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells also suggest that buying is more expensive than selling. For NYSE and NASDAQ, both estimated market impact costs and estimated spread costs are larger for buy orders than for sell orders by margins that are economically meaningful if not statistically significant. For example, the effective spread for NASDAQ Buys is estimated to be more than twice as large as the effective spread for NASDAQ Sells. This is consistent with the idea that the market believes that buy orders contain more information than sell orders. See Obizhaeva (2009a) for further discussion of this idea. It is also consistent with the possibility that closing benchmark prices are biased towards the bid side of the market.

A More General Specification.

Table 6 reports the results of a non-linear regression with a more general specification than Table 5. Three separate market impact parameters and three separate spread parameters are estimated for open market trades, external crosses, and internal crosses. In addition, the exponents on the three components of market activity (volume, price, volatility) are allowed to differ. The regression estimated is

$$\begin{aligned} & \frac{\mathbb{I}_{BS,i}(P_{ex,i} - P_{0,i})}{P_{0,i}} \cdot 10^4 \cdot \frac{(0.02)}{\sigma_i} \\ &= \frac{1}{2} \cdot \frac{\bar{\lambda}_{omt,i} X_{omt,i} + \bar{\lambda}_{ec,i} X_{ec,i} + \bar{\lambda}_{ic,i} X_{ic,i}}{(0.01)V_i} \cdot \left[\frac{W_i}{W^*} \right]^{1/3} \cdot \frac{\sigma_i^{\beta_1} \cdot P_{0,i}^{\beta_2} \cdot V_i^{\beta_3}}{(0.02)(40)(10^6)} \\ &+ \frac{1}{2} \cdot \frac{\bar{\kappa}_{omt,i} X_{omt,i} + \bar{\kappa}_{ec,i} X_{ec,i} + \bar{\kappa}_{ic,i} X_{ic,i}}{X_i} \cdot \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot \frac{\sigma_i^{\beta_4} \cdot P_{0,i}^{\beta_5} \cdot V_i^{\beta_6}}{(0.02)(40)(10^6)} + \tilde{\epsilon}. \end{aligned} \quad (68)$$

Because the exponents on the W -terms are set to be $1/3$ and $-1/3$, the model of market microstructure invariance predicts

$$\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0. \quad (69)$$

The model of invariant bet frequency predicts

$$\beta_1 = \beta_2 = \beta_3 = -1/3, \quad \beta_4 = \beta_5 = \beta_6 = 1/3. \quad (70)$$

The model of invariant bet size predicts

$$\beta_1 = \beta_2 = \beta_3 = 1/6, \quad \beta_4 = \beta_5 = \beta_6 = -1/6. \quad (71)$$

The first column of the table presents the results for all buys and sells. The remaining four columns present results for separate regressions for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells.

F-tests of the above restrictions for the model of invariant bet frequency ($F = 78.25$) and the model of invariant bet size ($F = 13.71$) are rejected very strongly ($p \ll 0.0001$). An F-test of the restrictions of our model in equation (69) is rejected less strongly, with $F = 4.55$, $p = 0.0001$. From the table, it appears that one reason for this rejection is that bid-ask spreads decrease faster than predicted as trading volume increases. The estimate of β_6 is -0.09 with standard error of 0.025 ($t = 3.46$). The rapid decrease in spreads as trading volume increases is consistent with the results from Table 1. The bid-ask spread, however, does not decrease as fast as predicted when stock price increases since β_5 is estimated as 0.18 with standard error of 0.062 ($t = 1.83$). Our tests do not show any significant deviation of bid-ask spread from predicted values with respect to volatility. Another reason for the rejection is that the estimates of β_1 and β_2 are quite negative. The estimate of β_1 is -0.31 with standard error of 0.191 ($t = -1.61$), and the estimate of β_2 is -0.22 with standard error of 0.191 ($t = -2.26$). The estimate of β_3 is close to zero. These results say that market impact behaves as predicted by the model of market microstructure invariance when the number of shares traded increases, but market impact decreases relative to what is predicted when volatility and stock price increase.

The rejection of the model of market microstructure invariance seems to be related to the fact that the exponents for volatility and price behave differently from the coefficients for share volume; the coefficients for volatility behave similarly to the coefficients for price. This suggests that the rejection might depend in a subtle manner on tick effects. When volatility is high and stock price is high, the tick size is small relative to a typical day's trading range.

Despite increasing the number of estimated parameters from four to twelve, the adjusted R^2 in the aggregate regression increases only from 0.0123 to 0.0129 .

The estimates for the three half spread parameters are $\hat{k}_{omt}/2 = 6.56$, $\hat{k}_{ec}/2 = 6.26$, and $\hat{k}_{ic}/2 = 0.25$. These results support the assumption that there is no spread associated with internal crosses, and the spread associated with external crosses is the same as the spread associated with open market trades.

The point estimates for market impact parameters are $\hat{\lambda}_{omt}/2 = 4.49$, $\hat{\lambda}_{ec}/2 = 2.17$, and $\hat{\lambda}_{ic}/2 = 2.41$. These results support the assumption that internal crosses do have market impact. The results, however, suggest that the impact for open market trades may be greater than the impact for internal and external crosses. While interpreting these results, we have to keep in mind, however, that the separate estimates of transaction costs parameters for different trading venues may suffer from selection bias since transition managers optimally chooses trading venues to minimize the total costs.

Comparison of Three Models. Table 7 presents estimates for equation (67) with the parameters α_1 and α_2 restricted to be as predicted in the model of market microstructure invariance, the model of invariant bet frequency, and the model of invariant bet size, respectively. For each of the three models, only two parameters are estimated: half-price-impact $\bar{\lambda}/2$ and half spread $\bar{\kappa}/2$.

For the model of market microstructure invariance, the reduction from four parameters to two parameters reduces the adjusted R^2 from 0.0123 to 0.0122 , consistent with very mild rejection of the model reported in Table 5. Furthermore, the parameter estimates for half-market impact $\hat{\lambda}/2$ and half bid-ask spread $\hat{\kappa}/2$ do not change much.

For the model of invariant bet frequency, the reduction from four parameters to two parameters reduces the R^2 greatly, from 0.0123 to 0.0075, consistent with very strong rejection of the model. Furthermore, the point estimate for half market impact drops enormously, from $\hat{\lambda}/2 = 2.85$ to $\hat{\lambda}/2 = 0.3788$. This is offset by a large increase in the estimated half spread, from $\hat{\kappa}/2 = 6.31$ to $\hat{\kappa}/2 = 15.29$. The model of invariant bet frequency is intuitively appealing since it suggests that the market impact of trading a given percentage of average daily volume is constant as a fraction of daily returns standard deviation, regardless of the level of trading activity in the stock. The strong rejection of this model, combined with the large changes in estimated coefficients, suggests that this model leads to the misleading empirical result that market impact is less important than it really is, and bid-ask spread is more important than it really is. Therefore, one of the justifications for the model of market microstructure invariance is that it allows for the importance of market impact to be estimated more accurately from a better specified model.

For the model of invariant bet size, the reduction from four parameters to two parameters reduces the adjusted R^2 from 0.0123 to 0.0110, consistent with a strong rejection of this model. The point estimates for half market impact and half bid-ask spread change in the opposite direction, with the point estimate for market impact increasing from 2.85 to 3.92 and the point estimate for half spread decreasing from 6.31 to 3.46.

In all three specifications, separate regressions for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells continue to suggest that buying is more expensive than selling or that benchmark prices are biased towards the bid side of the market.

For each subsample, the model of market microstructure invariance has the largest maximum log-likelihood function among three models under consideration. This evidence implies that, according to various likelihood-based comparison methods, our model delivers predictions that are more consistent with portfolio transition data than predictions of alternative models. Almost regardless of priors, a Bayesian statistician will calculate the posterior probability of our proposed model as being close to one and the posterior probabilities of alternative models as being close to zero.

Dummy Variables as Robustness Check. Figure 4 presents the results of three linear regressions, one for each of the three proposed models. The regression represents a modification of equation (67) in two ways. First, similarly to Table 7, for each of the three models, the values of α_1 and α_2 are fixed at the levels predicted by the models. Second, a dummy variable for each of the ten volume groups is associated with a half-market impact parameter and a half spread parameter for each group. The result is a regression with twenty coefficients, two coefficients for each volume bin, with one coefficient $\bar{\lambda}_j$ for half market impact and one coefficient $\bar{\kappa}_j$ for half spread. The regression equation can be written

$$\frac{\mathbb{I}_{BS,i}(P_{ex,i} - P_{0,i})}{P_{0,i}} \cdot 10^4 \cdot \frac{(0.02)}{\sigma_i} = \left(\sum_{j=1}^{10} \mathbb{I}_{j,i} \cdot \frac{1}{2} \cdot \bar{\lambda}_j \right) \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_1} \cdot \frac{X_i}{(0.01)V_i} + \left(\sum_{j=1}^{10} \mathbb{I}_{j,i} \cdot \frac{1}{2} \cdot \bar{\kappa}_j \right) \cdot \frac{(X_{omt,i} + X_{ec,i})}{X_i} \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} + \tilde{\epsilon}. \quad (72)$$

In the figure, for each of the three models, there is a graph of the estimates of the ten dummy variables for half market impact $\bar{\lambda}/2$ and a graph of the estimates of the ten dummy

variables for half-spread $\bar{\kappa}/2$. Each graph also shows the 95% confidence intervals around the point estimates, as well as a horizontal line showing the point estimate from Table 7. If the model is well specified, then the ten dummy variables should be the same and should equal the point estimates from Table 7.

For the model of market microstructure invariance, all of the point estimates lie either within the 95% confidence bands or slightly outside the 95% confidence bands, consistent with the mild rejection of the model discussed above. For the smallest volume group, the estimate for the half-spread has a very small confidence band which anchors the point estimate close to the two-parameter model. For the smallest volume group, the estimate for half-market impact also has a relatively small confidence band which anchors it close to the two-parameter model as well. For the two largest groups, the half-spread estimates are somewhat larger than the unconstrained estimate and the half-market impact estimates are somewhat smaller. The data seem to be saying that for the very largest stocks, there is a somewhat bigger spread and somewhat less market impact than implied by the model of market microstructure invariance. For trade groups from 2 to 6, the data are saying the opposite, i.e., that the half spread should be smaller and the half market impact larger than in the two parameter model. These patterns suggest that the transition manager might be using basket trades and then splitting the total costs among individual stocks in a manner consistent with the model of invariant bet frequency, i.e., assigning smaller than needed costs to large stocks and larger than needed costs to small stocks.

Similar graphs for the dummy variables in the model of invariant bet frequency are presented in the middle of the figure. This model predicts that effective bid-ask spreads do not decline as trading activity increases. It is clear from the figure, however, that the estimated effective bid-ask spreads for the smallest volume group are far greater than the estimated bid-ask spreads for the other nine groups. This places the effective spread for the smallest group far above the point estimate from the two parameter model and very far outside the 95% confidence bands. For the market impact parameters, the model generates a great deal of power from the smallest volume group because the trade sizes are large relative to volume for this group. The point estimate of half-market impact for the smallest volume group is therefore very close to the point estimate from the two parameter model. But this forces the dummy variables for half-market impact for the nine large volume groups to lie far above the point estimate from the two parameter model. If the smallest volume group were eliminated from consideration, it appears from the figure that the model of invariant bet frequency would perform almost as well as the model of market microstructure invariance. It is possible that the transition manager trades baskets of stocks and then marks the prices according to the model of invariant bet frequency. Perhaps basket trades tend to occur in the largest stocks; the model might look good for spurious reasons especially among the largest stocks.

Graphs of the dummy variables for the model of invariant bet size are presented on the right-hand side of the figure. The model generates very precise estimates of spreads for the smallest size group. For the larger size groups, the predicted spreads are much larger than the point estimates from the two parameter model. For half market impact, the dummy variables decrease almost monotonically, indicating that the rapid increase in market impact implied by the model value of $\alpha_1 = 1/2$ is greater than what is consistent with the data.

9 Practical Implications

This paper provides a simple formula for calculation of expected transaction costs as a function of observable dollar trading volume and volatility. Using portfolio transition data, we estimate the level of transaction costs for a benchmark stock with price of \$40 per share, trading volume of one million shares per day, and volatility of 2% per day. In particular, Table 7 shows that, if the exponent parameters are set to the values implied by the model of market microstructure invariance, then the estimated values of half market impact $\bar{\lambda}^*/2 = 2.89$ basis points and half bid-ask spread $\bar{\kappa}^*/2 = 7.91$ basis points. These estimates imply that a trade of one percent of average daily volume in the benchmark stock incurs a market impact cost of 2.89 basis points and a bid-ask spread cost of 7.91 basis points.

The model of market microstructure invariance describes how transaction costs vary across stocks with different levels of trading activity. A simple formula for expected costs $C(X)$ shows how to extrapolate the estimated transaction costs for the benchmark stock to any other security. The expected trading costs for an order of X shares, denoted $C(X)$, can be calculated as,

$$C(X) = \frac{1}{2}\bar{\lambda}^* \cdot \left(\frac{W}{(0.02)(40)(10^6)} \right)^{1/3} \cdot \frac{\sigma}{0.02} \cdot \frac{X}{(0.01)V} + \frac{1}{2}\bar{\kappa}^* \cdot \left(\frac{W}{(0.02)(40)(10^6)} \right)^{-1/3} \cdot \frac{\sigma}{0.02},$$

where W is the trading activity equal to the product of share price P , daily trading volume V , and daily returns volatility σ . The values for $\bar{\lambda}^*/2$ and $\bar{\kappa}^*/2$ are fixed at the estimated levels of $\bar{\lambda}^*/2 = 2.89$ and $\bar{\kappa}^*/2 = 7.91$ basis points, respectively.

We believe this formula may help to determine transaction costs not only for the U.S. stocks but also for other securities, since markets for most of them are likely to share the same underlying principles. Given the estimates $\bar{\lambda}^*/2$ and $\bar{\kappa}^*/2$ for the benchmark stock, the suggested extrapolation may allow us to find expected transaction costs, for example, for fixed income and foreign exchange securities as well as for securities in other markets across the world, after adjusting our extrapolation method for differences in exchange rates. Whether the formula for transaction costs is indeed of such a general nature is an interesting issue for future research.

Our paper provides answers to some important questions concerning management of transaction costs. Using the formula for expected costs and their components, asset managers can better forecast what expenses they will incur during implementation of their investment strategies as well as what amount of funds can be allocated to a strategy before it becomes unprofitable. Understanding cross-sectional variation in transaction costs has other practical implications. For example, when comparing execution quality across brokers specializing in stocks with different trading activity, performance metrics should take account of non-linearities documented in our paper. When executing basket trades, it may not be appropriate to assign the same number of basis points of transactions to each stock in a basket. Instead, transaction cost attribution to individual securities in a basket should take account of the dollar trade size, volatility, and level of trading activity of the stock, in the context of an appropriate model. Finally, our analysis has implications for trading strategies. If it is reasonable to restrict trading of the benchmark stock to say 1% of average daily volume, then a smaller percentage would be appropriate for more active stocks and a larger percentage would be appropriate for less active stocks.

10 Conclusion

This paper proposes three theoretical models that differ in their assumptions about what features of trading games remain invariant as games themselves vary across securities with different levels of trading activity. Our preferred model of trading game invariance is based on the intuition that deep parameters of the trading game itself are invariant, but the length of the trading games varies across stocks because the time clock in different markets ticks at different rates. All three models make dramatically different predictions concerning how market impact and bid-ask spreads vary cross-sectionally across stocks.

Data on portfolio transitions are used to test the models in two ways. First, under the identifying assumption that portfolio transitions are proportional in size to independent bets made by market participants, the size of portfolio transition orders is used to test implications of the models concerning how order sizes vary across stock with different levels of trading activity. Second, their implications for market impact and spreads are tested using estimates derived from implementation shortfall.

The empirical results are supportive of the model of market microstructure invariance, but with some caveats. The model assumes that if trading activity increases by one percent, trade size as a fraction of daily volume falls by two-thirds of one percent. The trade size regressions provide strong support for this assumption. The coefficient estimate of -0.63 is remarkably close to the predicted value of $-2/3$. The predictions of the model of market microstructure invariance for market impact and bid-ask spread are also supported by data. The empirical prediction that a one percent increase in trading activity increases the market impact (in units of daily standard deviation) by one-third of one percent is almost exactly the point estimate from non-linear regressions based on implementation shortfall. The empirical predictions that a one percent increase in trading activity decreases the bid-ask spread (in units of standard deviation) by one-third of one percent is matched by data reasonably closely.

There are, however, several issues which need further investigation. First, the statistical power behind implementation shortfall results come mostly from the 30 percent of stocks in the lowest dollar volume group. For the top 70 percent of stocks by dollar volume, it may be difficult to distinguish the model of market microstructure invariance from the model of invariant bet frequency. Second, in the implementation shortfall regressions, the bid-ask spreads decrease with increased activity somewhat faster than the model of market microstructure invariance predicts. Third, our measure of trading activity can be thought of as the product of share volume and price volatility in dollars per share. Although the model predicts that these two components of trading activity should behave similarly, both the implementation shortfall regressions and the trade size regressions suggest that they behave differently. Trading volume (measured in shares) seems to be more consistent with the model of trading game invariance than dollar price volatility. It is possible that these issues have something to do with the interaction between tick size effects and trading volume.

Interesting issues for further research include testing the three proposed models on different databases.

- The model's predictions concerning spreads can be tested using quoted spreads from TAQ data. Although it is difficult to measure the level of market depth from TAQ data

using, for example, the approach of Lee and Ready (1991), the model’s cross-sectional implications concerning market impact might be testable using this approach. The predictions concerning noise trading quantities can be tested using changes in holdings of mutual funds or other reporting institutional traders.

- If we make the assumption that the flow of information per trading game is constant across stocks and the flow of news articles is proportional to the flow of information, then the hypothesis of trading game invariance implies a particular relationship between the number of news articles for individual stocks and the measure of trading activity. If we make the assumption that the flow of analyst forecasts is proportional to the flow of information, then the hypothesis of trading game invariance also implies a particular relationship between the number of analyst forecasts for individual stocks and the measure of trading activity. These predictions can be tested as well.
- It is also possible that the model tested on stock data in this paper can be generalized to other markets. For example, market impact and spreads in bond markets, currency markets, or futures markets may be consistent with the regressions estimated for stocks in this paper.

Our insights about the structure of a trading process may also help to improve some existing measures of liquidity, such as the probability of informed trading (PIN) introduced by Easley et al. (1996).

The predictions of our model are broadly consistent with established empirical regularities in financial markets summarized in Bouchaud, Farmer, and Lillo (2008). They report, for example, that trading frequencies are usually found to be proportional to stock size raised to a power ranging between 0.44 and 0.86 (predicted to be two-third), market impact is proportional to stock size raised to the power of -0.30 (predicted to be minus one-third), and the percentage bid-ask spread is proportional to the volatility per trade (as predicted). Although these similarities provide complimentary evidence in favor of our model, they have to be taken with caution. First, this comparison is not a straightforward one, because we have to rely on the additional assumptions that volatility does not vary significantly across stocks and that stock size is proportional to volume. Second, since these empirical regularities are documented using *ex post* trading data—data on realized trading costs of executed trades—they may be distorted by various selection biases.

Numerous power-law relations have been empirically established in physics, linguistics, geophysics, sociology, and other fields. Examples include relationships between the strength of gravity and the distance between objects, the number of earthquakes and their magnitude, the frequency of words and their rank in frequency tables, the weight of fish and their length, as well as many others. Financial data also exhibit power-law regularities. Finding a theoretical explanation for power laws remains a challenge in social sciences. In our model of market microstructure invariance, power laws between financial variables appear naturally when the time is scaled so that these variables are measured in the same “units” across stocks. As in physics, the existence of power laws in financial data may have a deep origin in the universal processes generating these relations, and diverse systems may turned out to share the same fundamental principles, if studied at appropriate scale.

References

- Amihud, Yakov, 2002, "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets* 5, 31-56.
- Back, Kerry, and Shmuel Baruch, 2004, "Information in Securities Markets: Kyle meets Glosten and Milgrom," *Econometrica* 72, 433-465.
- Black, Fischer, 1986, "Noise," *Journal of Finance* 41, 529-543.
- Black, Fischer, 1995, "Equilibrium Exchanges," *Financial Analysts Journal* 51, 23-29.
- Bouchaud, Jean-Philippe, J. Dooyne Farmer, Fabrizio Lillo, 2009, "How Markets Slowly Digest Changes in Supply and Demand," in *Handbook of Financial Markets: Dynamics and Evolution*, North Holland, 47-160.
- Campbell, John Y., and Albert S. Kyle, 1993, "Smart Money, Noise Trading, and Stock Price Behavior," *Review of Economic Studies* 60, 1-34.
- Clark, Peter, 1973, "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135-155.
- Cooper, Kerry S., John C. Groth, and William E. Avera, 1985, "Liquidity, Exchange Listing, and Stock Return Performance," *Journal of Economics and Business*, 37, 19-33.
- Dufour, Alfonso, and Robert F. Engle, 2000, Time and the Price Impact of a Trade, *Journal of Finance*, 55(6), 2467-2598.
- Easley, David, Nicholas Kiefer, Maureen O'Hara, and Joseph Paperman, 1996, "Liquidity, Information, and Infrequently Traded Stocks," *Journal of Finance*, 51, 1405-1436.
- Foster, F. Douglas and S. Viswanathan, 1993, "The Effect of Public Information and Competition on Trading Volume and Price Volatility," *Review of Financial Studies* 6, 23-56.
- Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley, 2006, "Institutional Investors and Stock Market Volatility," *Quarterly Journal of Economics*, 121 (2), 461-504.
- Grossman, Sanford J., and Merton H. Miller, 1988, "Liquidity and Market Structure," *The Journal of Finance* 43(3), 617-633.
- Grossman, Sanford J., and Joseph E. Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets," *American Economic Review* 70(3), 393-408.
- Hasbrouck, Joel, 1999, "Trading Fast and Slow: Security Market Events in Real Time," *Working Paper*.
- Hasbrouck, Joel, 2007, *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading*, Oxford University Press.

- Hasbrouck, Joel, 2009, "Trading Costs and Returns for US Equities: Estimating Effective Costs from Daily Data," *Journal of Finance* 64(3), 1445–1477.
- Kyle, Albert S., 1985, "Continuous Auctions and Insider Trading", *Econometrica* 53(6), 1315–1335.
- Lee, Charles, and Mark Ready, 1991, "Inferring Trade Direction from Intraday Data," *Journal of Finance* 46, 733–746.
- Mandelbrot, Benoît, and Howard M. Taylor, 1967, "On the Distribution of Stock Price Differences," *Operations Research* 15(6), 1057–1062.
- Treynor, Jack, 1995, "The Only Game in Town," *The Financial Analysts Journal* 51(1), 81-83 (reprinted from *The Financial Analysts Journal* 22, 1971, 12–14, 22).
- Obizhaeva, Anna A., 2009a, "Portfolio Transitions and Price Dynamics," *University of Maryland, Working paper*.
- Obizhaeva, Anna A., 2009b, "Selection Bias in Liquidity Estimates," *University of Maryland, Working paper*.
- Perold, André, 1988, "The Implementation Shortfall: Paper vs. Reality," *Journal of Portfolio Management* 14, 4–9.
- Wang, Jiang, 1993, "A Model of Intertemporal Asset Prices Under Asymmetric Information," *The Review of Economic Studies* 60(2), 249–282.

Table 1: Descriptive Statistics.

	All	1	2	3	4	5	6	7	8	9	10
Med(V) (m \$)	19.99	1.22	5.14	9.97	15.92	23.92	31.45	42.11	60.16	101.51	212.55
Med(σ)	1.89	2.04	2.00	1.92	1.95	1.88	1.85	1.79	1.78	1.76	1.76
Med($Sprd$) (bps)	11.54	38.16	18.34	13.53	11.81	10.12	9.34	8.09	7.16	5.92	4.83
Mean($Sprd$) (bps)	23.67	64.05	31.27	21.83	18.40	15.65	13.86	12.14	11.00	9.02	7.46
<i>Panel B: Properties of Orders</i>											
	All	1	2	3	4	5	6	7	8	9	10
Avg(X/V) (%)	3.90	15.64	4.58	2.63	1.82	1.36	1.18	1.07	0.88	0.69	0.49
Med(X/V) (%)	0.56	3.48	1.39	0.80	0.54	0.40	0.35	0.30	0.25	0.20	0.14
Avg OMT Share	0.31	0.38	0.33	0.32	0.32	0.31	0.31	0.30	0.29	0.27	0.24
Avg EC Share	0.40	0.42	0.42	0.41	0.41	0.41	0.40	0.40	0.39	0.38	0.36
Avg IC Share	0.29	0.20	0.25	0.26	0.27	0.28	0.29	0.30	0.32	0.35	0.40
# Obs	441,865	65,081	68,545	41,559	49,532	28,621	30,087	30,710	35,733	42,331	49,666

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

Table 2: OLS Estimates of Order Size.

	NYSE			NASDAQ	
	All	Buy	Sell	Buy	Sell
$\ln [\bar{q}]$	-5.67*** (0.017)	-5.68*** (0.022)	-5.63*** (0.018)	-5.75*** (0.033)	-5.65*** (0.031)
α_0	-0.63*** (0.008)	-0.63*** (0.010)	-0.60*** (0.008)	-0.71*** (0.019)	-0.61*** (0.012)
<i>Model of Market Microstructure Invariance: $\alpha_0 = -2/3$</i>					
F-test	17.01	13.74	72.00	6.53	18.56
p-val	0.0000	0.0002	0.0000	0.0107	0.0000
<i>Model of Invariant Bet Frequency: $\alpha_0 = 0$</i>					
F-test	5664.91	3740.45	5667.60	1440.32	2427.51
p-val	0.0000	0.0000	0.0000	0.0000	0.0000
<i>Model of Invariant Bet Size: $\alpha_0 = -1$</i>					
F-test	1920.13	1306.11	2537.08	229.30	966.99
p-val	0.0000	0.0000	0.0000	0.0000	0.0000
Q^*/V^*	34.68	34.08	35.96	31.85	35.27
$d/g/n$	2/1/4389	2/1/4018	2/1/4198	2/1/2855	2/1/2977
#Obs	441,865	135,006	152,701	69,774	84,384
Adj. R^2	0.3188	0.2588	0.2643	0.4364	0.3648
R^2	0.3188	0.2588	0.2643	0.4364	0.3648

Table presents the estimates $\ln \bar{q}$ and α_0 for the regression:

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + \tilde{\epsilon}.$$

Each observation corresponds to order i . The left-hand side variable is the logarithm of the order size X_i as a fraction of expected daily volume V_i . The trading activity W_i is the product of expected daily volatility σ_i , benchmark price $P_{0,i}$, and expected daily volume V_i , measured as the last month's average daily volume. The scaling constant $W^* = (0.02)(40)(10^6)$ corresponds to the trading activity for the benchmark stock with volatility of 0.02, price \$40 per share, and trading volume of one million shares per day. \bar{q} is the measure of order size such that the median order size Q^*/V^* for a benchmark stock is calculated as $\exp(\bar{q}) \cdot 10^4$ in basis points. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. F-statistics and p-values are reported for three models with d parameters, g restrictions, and n clusters in the regression. The sample ranges from January 2001 to December 2005. ***, **, * denotes significance at 1%, 5% and 10% levels, respectively.

Table 3: OLS Estimates for Order Size: A More General Specification.

	NYSE			NASDAQ	
	All	Buy	Sell	Buy	Sell
$\ln [\bar{q}]$	-5.65*** (0.016)	-5.66*** (0.022)	-5.58*** (0.018)	-5.80*** (0.034)	-5.63*** (0.029)
b_1	0.25*** (0.031)	0.30*** (0.041)	0.36** (0.036)	0.17* (0.073)	0.24** (0.073)
b_2	0.16*** (0.014)	0.11*** (0.018)	0.21*** (0.014)	0.02 (0.039)	0.22*** (0.028)
b_3	0.01 (0.009)	0.02 (0.011)	0.03** (0.009)	-0.07*** (0.018)	0.02 (0.015)
<i>Model of Market Microstructure Invariance: $b_1 = b_2 = b_3 = 0$</i>					
F-test	47.57	19.92	79.29	8.77	20.97
p-val	0.0000	0.0000	0.0000	0.0000	0.0000
<i>Model of Invariant Bet Frequency: $b_1 = b_2 = b_3 = 2/3$</i>					
F-test	2044.60	1286.85	1939.03	567.85	808.05
p-val	0.0000	0.0000	0.0000	0.0000	0.0000
<i>Model of Invariant Bet Size: $b_1 = b_2 = b_3 = -1/3$</i>					
F-test	747.74	465.71	947.01	78.19	325.63
p-val	0.0000	0.0000	0.0000	0.0000	0.0000
Q^*/V^*	35.31	34.98	32.99	30.35	35.95
$d/g/n$	4/3/4389	4/3/4018	4/3/4198	4/3/2855	4/3/2977
#Obs	441,865	135,006	152,701	69,774	84,384
Adj. R^2	0.3211	0.2614	0.2682	0.4382	0.3674
R^2	0.3213	0.2616	0.2684	0.4384	0.3676

Table presents the estimates $\ln \bar{q}$, and b_1, b_2, b_3 for the regression:

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] - \frac{2}{3} \cdot \ln \left[\frac{W_i}{W^*} \right] + b_1 \cdot \ln \left[\frac{\sigma_i}{(0.02)} \right] + b_2 \cdot \ln \left[\frac{P_{0,i}}{(40)} \right] + b_3 \cdot \ln \left[\frac{V_i}{(10^6)} \right] + \tilde{\epsilon}.$$

Each observation corresponds to order i . The left-hand side variable is the logarithm of the order size X_i as a fraction of expected daily volume V_i . The trading activity W_i is the product of expected daily volatility σ_i , benchmark price $P_{0,i}$, and expected daily volume V_i , measured as the last month's average daily volume. The scaling constant $W^* = (0.02)(40)(10^6)$ corresponds to the trading activity for the benchmark stock with volatility of 0.02, price \$40 per share, and trading volume of one million shares per day. \bar{q} is the measure of order size such that the median order size Q^*/V^* for a benchmark stock is calculated as $\exp(\bar{q}) \cdot 10^4$ in basis points. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. F-statistics and p-values are reported for three models with d parameters, g restrictions, and n clusters in the regression. The sample ranges from January 2001 to December 2005. ***, **, * denotes significance at 1%, 5% and 10% levels, respectively.

Table 4: OLS Estimates of Order Size: Comparison of Three Models.

	NYSE			NASDAQ	
	All	Buy	Sell	Buy	Sell
<i>Model of Market Microstructure Invariance: $\alpha_0 = -2/3$</i>					
$\ln [\bar{q}]$	-5.69*** (0.018)	-5.70*** (0.022)	-5.67*** (0.019)	-5.70*** (0.039)	-5.71*** (0.035)
Q^*/V^*	33.75	33.35	34.61	33.60	32.99
Adj. R^2	0.3177	0.2577	0.2608	0.4343	0.3618
R^2	0.3178	0.2579	0.2609	0.4345	0.3620
$\log(L)$	-828,757	-255,637	-287,149	-127,591	-158,149
<i>Model of Invariant Bet Frequency: $\alpha_0 = 0$</i>					
$\ln [\bar{q}]$	-5.17*** (0.021)	-5.33*** (0.025)	-5.29*** (0.021)	-4.95*** (0.047)	-4.88*** (0.043)
Q^*/V^*	56.85	48.44	50.42	70.83	75.97
Adj. R^2	-0.0002	-0.0002	-0.0002	-0.0004	-0.0003
R^2	0.0000	0.0000	0.0000	0.0000	0.0000
$\log(L)$	-913,255	-275,771	-310,235	-147,479	-177,111
<i>Model of Invariant Bet Size: $\alpha_0 = -1$</i>					
$\ln [\bar{q}]$	-5.95*** (0.021)	-5.89*** (0.024)	-5.86*** (0.022)	-6.07*** (0.045)	-6.13*** (0.045)
Q^*/V^*	26.05	27.67	28.51	23.11	21.77
Adj. R^2	0.2105	0.1683	0.1458	0.3669	0.2192
R^2	0.2107	0.1685	0.1460	0.3671	0.2195
$\log(L)$	-860,973	-263,318	-298,186	-131,520	-166,656
#Obs	441,865	135,006	152,701	69,774	84,384

Table presents the estimates $\ln [\bar{q}]$ for the regression:

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + \tilde{\epsilon},$$

with α_0 restricted to be as predicted in proposed models. Each observation corresponds to order i . The left-hand side variable is the logarithm of the order size X_i as a fraction of expected daily volume V_i . The trading activity W_i is the product of expected daily volatility σ_i , benchmark price $P_{0,i}$, and expected daily volume V_i , measured as the last month's average daily volume. The scaling constant $W^* = (0.02)(40)(10^6)$ corresponds to the trading activity for the benchmark stock with volatility of 0.02, price \$40 per share, and trading volume of one million shares per day. \bar{q} is the measure of order size such that the median order size Q^*/V^* for a benchmark stock is calculated as $\exp(\bar{q}) \cdot 10^4$ in basis points. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The log-likelihoods and R^2 are shown for each model. The sample ranges from January 2001 to December 2005. ***, **, *denotes significance at 1%, 5% and 10% levels, respectively.

Table 5: Impact and Spread Estimates in Nested Non-Linear Regression.

	NYSE			NASDAQ	
	All	Buy	Sell	Buy	Sell
$1/2\bar{\lambda}$	2.85*** (0.245)	2.50*** (0.515)	2.33*** (0.365)	4.20*** (0.753)	2.99*** (0.662)
α_1	0.33*** (0.024)	0.18*** (0.045)	0.33*** (0.054)	0.33*** (0.053)	0.35*** (0.045)
$1/2\bar{\kappa}$	6.30*** (1.131)	14.94*** (2.529)	2.82* (1.394)	8.38* (3.328)	3.94** (1.498)
α_2	-0.39*** (0.025)	-0.19*** (0.045)	-0.46*** (0.061)	-0.36*** (0.061)	-0.45*** (0.047)
<i>Model of Market Microstructure Invariance: $\alpha_1 = 1/3, \alpha_2 = -1/3$</i>					
F-test	2.62	8.51	2.25	0.09	3.12
p-val	0.0731	0.0002	0.1057	0.9114	0.0443
<i>Model of Invariant Bet Frequency: $\alpha_1 = 0, \alpha_2 = 0$</i>					
F-test	176.14	14.79	47.03	33.11	71.06
p-val	0.0000	0.0000	0.0000	0.0000	0.0000
<i>Model of Invariant Bet Size: $\alpha_1 = 1/2, \alpha_2 = -1/2$</i>					
F-test	30.30	39.63	5.23	7.21	5.92
p-val	0.0000	0.0000	0.0054	0.0007	0.0027
$d/g/n$	4/2/4389	4/2/4018	4/2/4198	4/2/2855	4/2/2977
#Obs	441,865	135,006	152,701	69,774	84,384
R^2	0.0126	0.0136	0.0067	0.0211	0.0195
Adj. R^2	0.0123	0.0134	0.0064	0.0208	0.0192

Table presents the estimates for $\bar{\lambda}$, $\bar{\kappa}$, α_1 , and α_2 in the regression:

$$\frac{\mathbb{I}_{BS,i}(P_{ex,i} - P_{0,i})}{P_{0,i}} \cdot 10^4 \cdot \frac{(0.02)}{\sigma_i} = \frac{1}{2} \cdot \bar{\lambda} \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_1} \cdot \frac{X_i}{(0.01)V_i} + \frac{1}{2} \cdot \bar{\kappa} \cdot \frac{(X_{omt,i} + X_{ec,i})}{X_i} \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} + \tilde{\epsilon}$$

Each observation corresponds to order i . The left-hand side variable is the implementation shortfall in basis points, where $\mathbb{I}_{BS,i}$ is a buy/sell indicator, $P_{ex,i}$ and $P_{0,i}$ are the execution and benchmark prices. The term $(0.02)/\sigma_i$ adjusts for heteroscedasticity. The trading activity W_i is the product of expected volatility σ_i , benchmark price $P_{0,i}$, and expected volume V_i . The scaling constant $W^* = (0.02)(40)(10^6)$ is the trading activity for the benchmark stock with volatility of 0.02, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order with $X_{omt,i}$ and $X_{ec,i}$ shares executed in open market and external crossing networks, respectively. $\bar{\lambda}/2$ is the market impact costs of executing a trade of one percent of daily volume in a benchmark stock, and $\bar{\kappa}/2$ is the effective spread cost. Both $\bar{\lambda}/2$ and $\bar{\kappa}/2$ are in basis points. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. F-statistics and p-values are reported for three models with d parameters, g restrictions, and n clusters in the regression. The sample ranges from January 2001 to December 2005. ***, **, * denotes significance at 1%, 5% and 10% levels, respectively.

Table 6: Impact and Spread Estimates: A More General Specification.

	All	NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
$\frac{1}{2}\bar{\lambda}_{omt}$	4.49*** (0.601)	2.41*** (0.433)	4.35*** (0.798)	5.39*** (1.449)	4.33*** (1.058)
$\frac{1}{2}\bar{\lambda}_{ec}$	2.17*** (0.386)	3.02*** (0.471)	1.77** (0.542)	3.64*** (0.968)	1.34 (0.705)
$\frac{1}{2}\bar{\lambda}_{ic}$	2.40*** (0.334)	2.20*** (0.622)	1.77*** (0.420)	2.07* (0.891)	1.51** (0.504)
β_1	-0.31 (0.191)	-0.86*** (0.141)	-0.37 (0.327)	-0.10 (0.347)	-1.05*** (0.284)
β_2	-0.22* (0.191)	-0.01 (0.141)	-0.43* (0.327)	-0.17 (0.347)	-0.32 (0.284)
β_3	0.04 (0.043)	-0.19*** (0.031)	0.13 (0.077)	0.01 (0.053)	-0.04 (0.037)
$\frac{1}{2}\bar{\kappa}_{omt}$	6.56*** (1.227)	18.55*** (3.539)	3.05* (1.279)	14.43*** (3.967)	4.69* (1.819)
$\frac{1}{2}\bar{\kappa}_{ec}$	6.26*** (1.124)	8.99*** (2.355)	4.98** (1.520)	11.13** (3.690)	5.08** (1.841)
$\frac{1}{2}\bar{\kappa}_{ic}$	0.26 (1.846)	5.31 (3.987)	-4.38** (1.429)	7.78 (7.699)	0.70 (1.317)
β_4	0.10 (0.160)	-0.06 (0.249)	0.60* (0.260)	-0.29 (0.297)	0.99*** (0.258)
β_5	0.18** (0.062)	-0.22 (0.172)	0.06 (0.123)	0.26 (0.139)	0.36*** (0.104)
β_6	-0.09*** (0.025)	0.26*** (0.055)	-0.12* (0.054)	0.05 (0.050)	-0.11* (0.049)
<i>Model of Market Microstructure Invariance: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$.</i>					
F-test	4.59	20.96	2.80	1.38	11.26
p-val	0.0001	0.0000	0.0102	0.2168	0.0000
<i>Model of Invariant Bet Frequency: $\beta_1 = \beta_2 = \beta_3 = -1/3, \beta_4 = \beta_5 = \beta_6 = 1/3$.</i>					
F-test	78.26	12.06	26.46	10.71	23.27
p-val	0.0000	0.0000	0.0000	0.0000	0.0000
<i>Model of Invariant Bet Size: $\beta_1 = \beta_2 = \beta_3 = 1/6, \beta_4 = \beta_5 = \beta_6 = -1/6$.</i>					
F-test	13.77	44.25	5.94	6.85	28.79
p-val	0.0000	0.0000	0.0000	0.0000	0.0000
$d/g/n$	12/6/4389	12/6/4018	12/6/4198	12/6/2855	12/6/2977
#Obs	441,865	135,006	152,701	69,774	84,384
Adj. R^2	0.0129	0.0147	0.0076	0.0222	0.0214
R^2	0.0131	0.0150	0.0079	0.0225	0.0217

Table presents the estimates for $\bar{\lambda}_{omt}, \bar{\lambda}_{ec}, \bar{\lambda}_{ic}, \bar{\kappa}_{omt}, \bar{\kappa}_{ec}, \bar{\kappa}_{ic}, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ in the regression:

$$\begin{aligned} & \frac{\mathbb{I}_{BS,i}(P_{ex,i} - P_{0,i})}{P_{0,i}} \cdot 10^4 \cdot \frac{(0.02)}{\sigma_i} \\ &= \frac{1}{2} \cdot \frac{\bar{\lambda}_{omt,i}X_{omt,i} + \bar{\lambda}_{ec,i}X_{ec,i} + \bar{\lambda}_{ic,i}X_{ic,i}}{(0.01)V_i} \cdot \left[\frac{W_i}{W^*} \right]^{1/3} \cdot \frac{\sigma_i^{\beta_1} \cdot P_{0,i}^{\beta_2} \cdot V_i^{\beta_3}}{(0.02)(40)(10^6)} \\ &+ \frac{1}{2} \cdot \frac{\bar{\kappa}_{omt,i}X_{omt,i} + \bar{\kappa}_{ec,i}X_{ec,i} + \bar{\kappa}_{ic,i}X_{ic,i}}{X_i} \cdot \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot \frac{\sigma_i^{\beta_4} \cdot P_{0,i}^{\beta_5} \cdot V_i^{\beta_6}}{(0.02)(40)(10^6)} + \tilde{\epsilon}. \end{aligned}$$

Each observation corresponds to order i . The left-hand side variable is the implementation shortfall in basis points, where $\mathbb{I}_{BS,i}$ is a buy/sell indicator, $P_{ex,i}$ and $P_{0,i}$ are the execution and benchmark prices. The term $(0.02)/\sigma_i$ adjusts for heteroscedasticity. The trading activity W_i is the product of expected volatility σ_i , benchmark price $P_{0,i}$, and expected volume V_i . The scaling constant $W^* = (0.02)(40)(10^6)$ is the trading activity for the benchmark stock with volatility of 0.02, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order with $X_{omt,i}$, $X_{ec,i}$ and $X_{ic,i}$ shares executed in open market, external crossing networks, and internal crossing networks, respectively. $\bar{\lambda}_{omt}/2, \bar{\lambda}_{ec}/2, \bar{\lambda}_{ic}/2$ are the market impact costs of executing a trade of one percent of daily volume in a benchmark stock using open market trades, external crosses, and internal crosses, respectively. $\bar{\kappa}_{omt}/2, \bar{\kappa}_{ec}/2, \bar{\kappa}_{ic}/2$ are the effective spread costs for corresponding trading venues. Both market impact and spread costs are in basis points. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. F-statistics and p-values are reported for three models with d parameters, g restrictions, and n clusters in the regression. The sample ranges from January 2001 to December 2005. ***, **, *denotes significance at 1%, 5% and 10% levels, respectively.

Table 7: Market Impact and Spread Estimates: Comparison of Three Models.

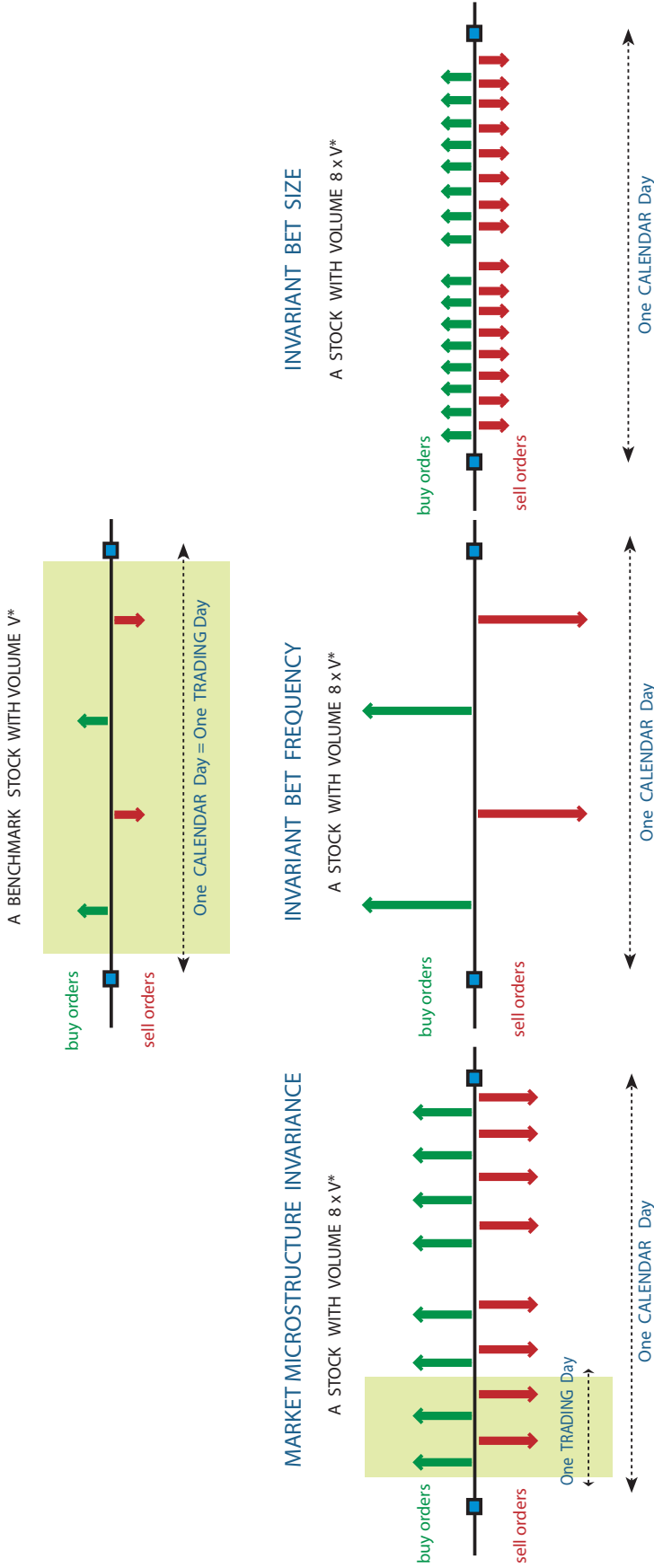
	All	NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
<i>Model of Market Microstructure Invariance: $\alpha_1 = 1/3, \alpha_2 = -1/3$</i>					
$1/2 \hat{\lambda}$	2.8898*** (0.1948)	3.4199*** (0.4421)	2.3411*** (0.3231)	4.2361*** (0.4342)	2.7800*** (0.3300)
$1/2 \hat{\kappa}$	7.9036*** (0.6889)	10.9695*** (1.3724)	4.7553*** (1.1065)	9.2870*** (1.5153)	7.3187*** (1.0067)
Adj. R^2	0.0122	0.0129	0.0062	0.0207	0.0184
R^2	0.0125	0.0131	0.0064	0.0211	0.0188
$\log(L)$	1061,808	324,590	366,995	167,793	202,588
<i>Model of Invariant Bet Frequency: $\alpha_1 = 0, \alpha_2 = 0$</i>					
$1/2 \hat{\lambda}$	0.3788*** (0.0884)	1.3291*** (0.1130)	0.4712*** (0.1041)	1.1857*** (0.1377)	0.2694*** (0.0457)
$1/2 \hat{\kappa}$	15.2686*** (1.5658)	19.1143*** (2.5247)	6.4113** (2.3157)	19.2409*** (4.4536)	13.1085*** (3.5423)
Adj. R^2	0.0075	0.0126	0.0038	0.0178	0.0105
R^2	0.0077	0.0128	0.0040	0.0182	0.0108
$\log(L)$	1060,752	324,568	366,811	167,689	202,247
<i>Model of Invariant Bet Size: $\alpha_1 = 1/2, \alpha_2 = -1/2$</i>					
$1/2 \hat{\lambda}$	3.9202*** (0.3036)	3.9578*** (0.6274)	2.8730*** (0.4339)	6.5057*** (0.6467)	4.6229*** (0.6037)
$1/2 \hat{\kappa}$	3.4648*** (0.2917)	5.7598*** (0.6487)	2.5020*** (0.4105)	4.2232*** (0.5461)	3.0040*** (0.3587)
Adj. R^2	0.0110	0.0108	0.0056	0.0195	0.0182
R^2	0.0112	0.0110	0.0058	0.0198	0.0185
$\log(L)$	1061,533	324,446	366,950	167,748	202,578
#Obs	441,865	135,006	152,701	69,774	84,384

Table presents the estimates $\bar{\lambda}, \bar{\kappa}$ for the regression:

$$\frac{\mathbb{I}_{BS,i}(P_{ex,i} - P_{0,i})}{P_{0,i}} \cdot 10^4 \cdot \frac{(0.02)}{\sigma_i} = \frac{1}{2} \cdot \bar{\lambda} \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_1} \cdot \frac{X_i}{(0.01)V_i} + \frac{1}{2} \cdot \bar{\kappa} \cdot \frac{(X_{omt,i} + X_{ec,i})}{X_i} \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} + \tilde{\epsilon}_i,$$

with α_1 and α_2 restricted to be as predicted in proposed models. Each observation corresponds to order i . The left-hand side variable is the implementation shortfall in basis points, where $\mathbb{I}_{BS,i}$ is a buy/sell indicator, $P_{ex,i}$ and $P_{0,i}$ are the execution and benchmark prices. The term $(0.02)/\sigma_i$ adjusts for heteroscedasticity. The trading activity W_i is the product of expected volatility σ_i , benchmark price $P_{0,i}$, and expected volume V_i . The scaling constant $W^* = (0.02)(40)(10^6)$ is the trading activity for the benchmark stock with volatility of 0.02, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order with $X_{omt,i}$ and $X_{ec,i}$ shares executed in open market and external crossing networks, respectively. $\bar{\lambda}/2$ is the market impact costs of executing a trade of one percent of daily volume in a benchmark stock, and $\bar{\kappa}/2$ is the effective spread cost. Both $\bar{\lambda}/2$ and $\bar{\kappa}/2$ are in basis points. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. F-statistics and p-values are reported for three models with d parameters, g restrictions, and n clusters in the regression. The sample ranges from January 2001 to December 2005. ***, **, *denotes significance at 1%, 5% and 10% levels, respectively. The log-likelihoods and R^2 are shown for each model. The sample ranges from January 2001 to December 2005. ***, **, *denotes significance at 1%, 5% and 10% levels, respectively.

Figure 1: Intuition Behind Models.



The figure illustrates the assumptions of our models. The trading game for the benchmark stock (top chart) is compared to the trading game for a stock with trading volume eight times larger than that of the benchmark stock (bottom chart). The benchmark stock has four bets expected per calendar day; buy orders are marked in green color, and sell orders are marked in red color. Prices and volatilities are assumed to be the same for both stocks. Thus, the length of the bars is proportional to dollar bet size. The model of market microstructure invariance assumes that the time clock operates four times faster for the more active stock. This generates four times as many bets per day. To prevent speeding up the time clock from increasing volatility in the model of “market microstructure invariance,” the stock has been delevered by a factor of two to keep volatility constant. This delevering multiplies dollar bet size by a factor of two. The model of “invariant bet frequency” assumes that the difference in volume comes from the difference in bet sizes. The model of “invariant bet size” assumes that the difference in volume comes from the difference in bet frequencies.

Figure 2: Dummy Variables as a Robustness Check in OLS Regression for Order Size.

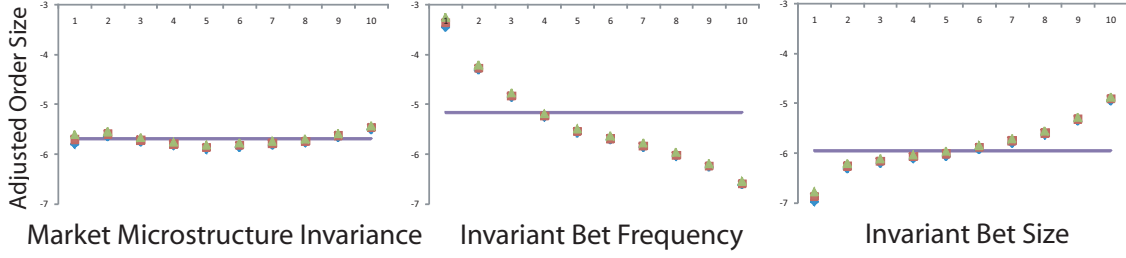


Figure shows the dummy variables $\ln [\bar{q}_j]$ with the 95%-confidence intervals for the ten volume groups and three proposed models from regression

$$\ln \left[\frac{X_i}{V_i} \right] = \left[\sum_{j=1}^{10} \mathbb{I}_{j,i} \cdot \ln [\bar{q}_j] \right] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + \tilde{\epsilon},$$

where j th dummy variable corresponds to the average logarithm of the order size X_i as a fraction of expected daily volume V_i for j th volume group. Each observation corresponds to order i . $\mathbb{I}_{j,i}$ is an indicator equal to one if order i is executed in a stock from volume group j . In the model of market microstructure invariance, $\alpha_0 = -2/3$; in the model of invariant bet frequency, $\alpha_0 = 0$; in the model of invariant bet size, $\alpha_0 = -1$. The trading activity W_i is the product of expected daily volatility σ_i , benchmark price $P_{0,i}$, and expected daily volume V_i , measured as the last month's average daily volume. The scaling constant $W^* = (0.02)(40)(10^6)$ corresponds to the trading activity for the benchmark stock with volatility of 0.02, price \$40 per share, and trading volume of one million shares per day. The ten volume groups are based on the pre-transition dollar trading volume with thresholds corresponding to 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of the dollar volume for common NYSE-listed stocks. Group 1 (Group 10) contains stocks with the lowest (highest) trading volume. The standard errors are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005. On each subplot, the point estimate of $\ln [\bar{q}]$ from Table 4 is superimposed for each model.

Figure 3: Invariant Order Size Distribution as Log-Normal.

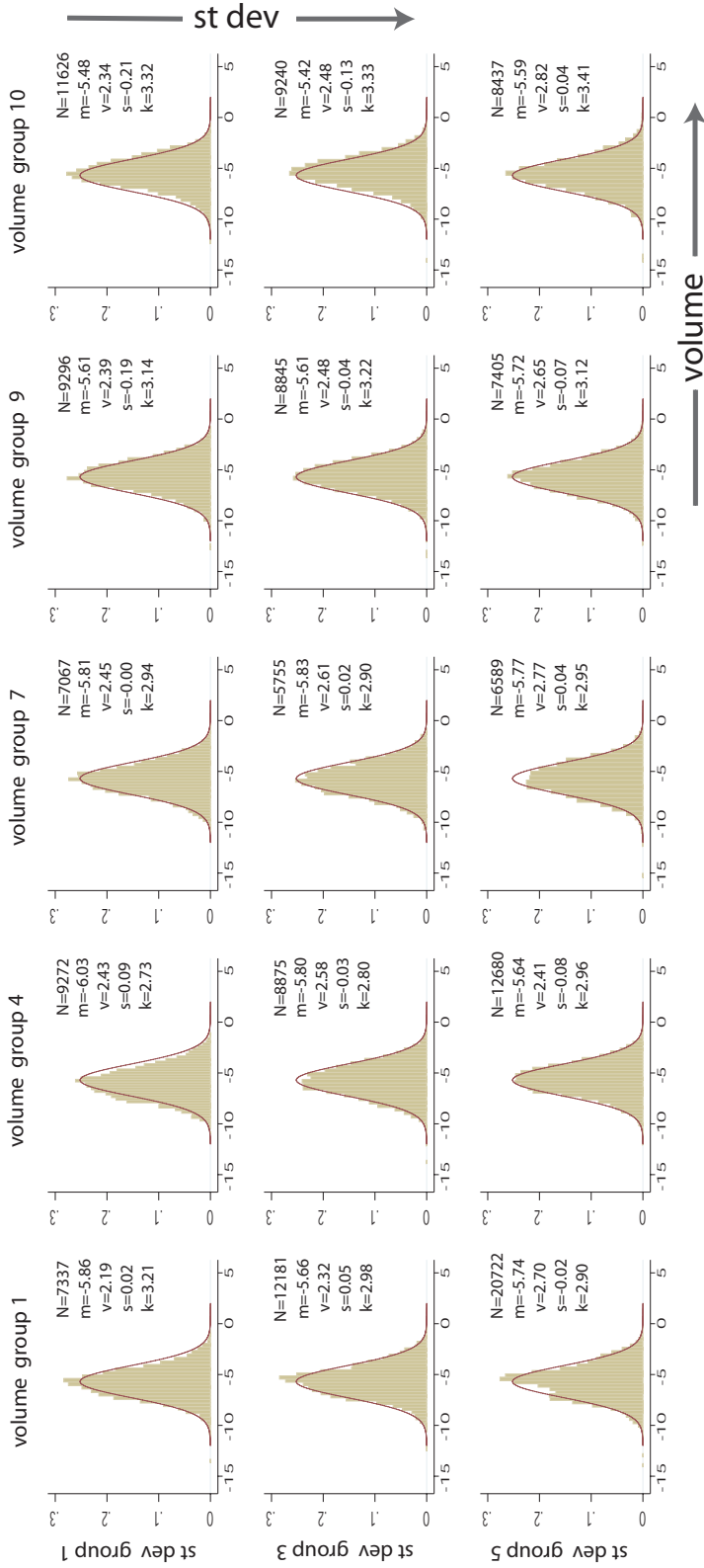


Figure shows a distribution of transition order sizes, adjusted for differences in the trading activity, for stocks sorted into 10 volume groups and 5 volatility groups (only volume groups 1, 4, 7, 9, 10 and volatility groups 1, 3, 5 are reported). The adjustment is done according to the model of market microstructure invariance, i.e. $\frac{X_i}{V_i} \cdot W_i^{2/3}$, where X_i is an order size in shares, V_i is the average daily volume in shares, and W_i is the measure of trading activity. The ten volume groups are based on average dollar trading volume with thresholds corresponding to 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of the dollar volume for common NYSE-listed stocks. Volume group 1 (group 10) has stocks with the lowest (highest) trading volume. The five volatility groups are based on thresholds corresponding to 20th, 40th, 60th, and 80th percentiles for common NYSE-listed stocks. Volatility group 1 (group 5) has stocks with the lowest (highest) volatility. Each subplot also shows the number of observations (N), the mean (m), the variance (v), the skewness (s), and the kurtosis (k) for depicted distribution. The normal distribution with the common mean of -5.69 and variance of 2.50 is imposed on each subplot. The common mean and variance are calculated as the mean and variance of adjusted order sizes for the entire sample. The sample ranges from January 2001 to December 2005.

Figure 4: Market Impact and Spread: Dummy Variables as Robustness Check.

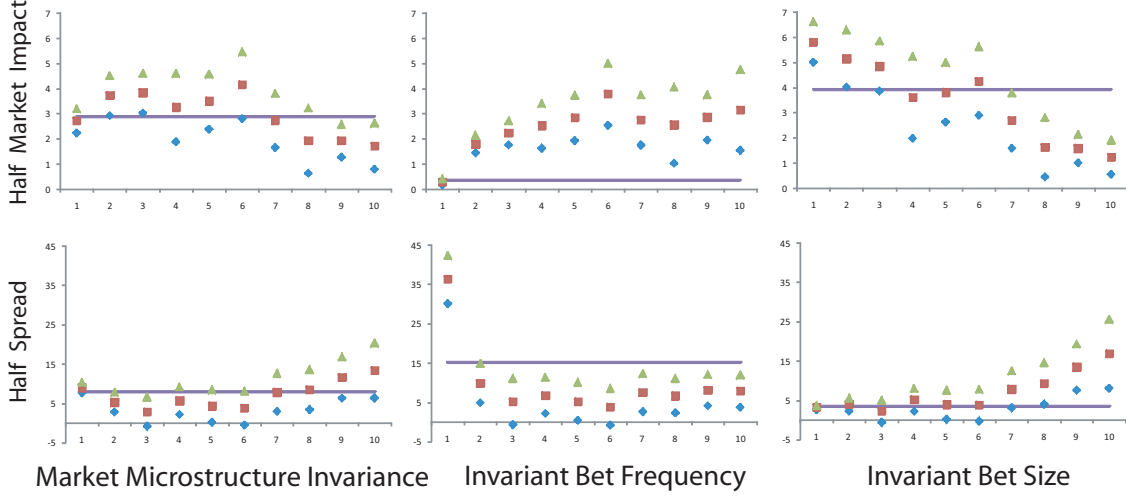


Figure shows the estimates of half market impact $\lambda_j/2$ (top) and half spread $\bar{\kappa}_j/2$ (bottom) with the 95%-confidence intervals for the ten volume groups and three proposed models from the regression:

$$\frac{\mathbb{I}_{BS,i}(P_{ex,i} - P_{0,i})}{P_{0,i}} \cdot 10^4 \cdot \frac{(0.02)}{\sigma_i} = \left(\sum_{j=1}^{10} \mathbb{I}_{j,i} \cdot \frac{1}{2} \bar{\lambda}_j \right) \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_1} \cdot \frac{X_i}{(0.01)V_i} + \left(\sum_{j=1}^{10} \mathbb{I}_{j,i} \cdot \frac{1}{2} \bar{\kappa}_j \right) \cdot \frac{(X_{omt,i} + X_{ec,i})}{X_i} \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} + \tilde{\epsilon},$$

where $\bar{\lambda}_j/2$ and $\bar{\kappa}_j/2$ correspond to the market impact costs of a trade of one percent of daily volume and the spread cost for j th volume group; both in basis points. Each observation corresponds to order i . $\mathbb{I}_{j,i}$ is an indicator equal to one if order i is executed in a stock from volume group j . In the model of market microstructure invariance, $\alpha_1 = 1/3, \alpha_2 = -1/3$. In the model of invariant bet frequency, $\alpha_1 = 0, \alpha_2 = 0$. In the model of invariant bet size, $\alpha_1 = 1/2, \alpha_2 = -1/2$. The left-hand side variable is the implementation shortfall in basis points, where $\mathbb{I}_{BS,i}$ is a buy/sell indicator, $P_{ex,i}$ and $P_{0,i}$ are the execution and benchmark prices. The term $(0.02)/\sigma_i$ adjusts for heteroscedasticity. The trading activity W_i is the product of expected volatility σ_i , benchmark price $P_{0,i}$, and expected volume V_i . The scaling constant $W^* = (0.02)(40)(10^6)$ is the trading activity for the benchmark stock with volatility of 0.02, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order with $X_{omt,i}$ and $X_{ec,i}$ shares executed in open market and external crossing networks, respectively. The ten volume groups are based on the pre-transition dollar volume with thresholds corresponding to 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of the dollar volume for common NYSE-listed stocks. Group 1 (Group 10) contains stocks with the lowest (highest) volume. The standard errors are clustered at weekly levels for 17 industries. The sample ranges from January 2001 to December 2005. On each subplot, the point estimates of $\bar{\lambda}/2$ and $\bar{\kappa}/2$ from Table 7 are superimposed as horizontal lines for each model.