

Market Microstructure Invariance: Theory and Empirical Tests*

Albert S. Kyle
University of Maryland
akyle@rhsmith.umd.edu

Anna A. Obizhaeva
New Economic School
aobizhaeva@nes.ru

October 17, 2014[†]

Abstract

Using the intuition that financial markets transfer risks in business time, we define “market microstructure invariance” as the hypothesis that the distribution of risk transfers (“bets”), transactions costs, market resiliency, and pricing accuracy are constant across assets when measured per unit of business time. A structural model of risk-neutral informed trading and noise trading with linear price impact shows that invariance relationships arise when the costs of informative signals are constant. The invariance hypotheses imply that microstructure variables like bet size and transactions costs have empirically testable relationships to observable dollar volume and volatility. Since portfolio transitions can be viewed as natural experiments for measuring transactions costs and individual orders can be treated as proxies for bets, we test these empirical predictions using a dataset of 400,000+ portfolio transition orders and find that the predictions closely match the data.

Keywords: market microstructure, liquidity, bid-ask spread, market impact, transactions costs, market efficiency, resiliency, order size, number of orders, portfolio transitions, efficient markets hypothesis, econophysics, invariance, scaling, structural estimation.

*We are grateful to Elena Asparouhova, Peter Bossaerts, Xavier Gabaix, Lawrence Glosten, Pankaj Jain, Mark Loewenstein, Natalie Popovic, Sergey N. Smirnov, Georgios Skoulakis, Vish Viswanathan, and Wenyuan Xu for helpful comments. Obizhaeva is also grateful to the Paul Woolley Center at the London School of Economics for its hospitality as well as Ross McLellan, Simon Myrgren, Sébastien Page, and especially Mark Kritzman for their help.

[†]This paper is a revised version of a previous manuscript (June 7, 2013) which combined and superseded two earlier papers: the theoretical paper “Market Microstructure Invariants: Theory and Implications of Calibration” (December 2011) and the empirical paper “Market Microstructure Invariants: Empirical Evidence from Portfolio Transitions” (December 2011). These two papers superseded an older combined manuscript “Market Microstructure Invariants” (May 2011).

This paper proposes a modeling principle for financial markets that we call “market microstructure invariance.” When portfolio managers trade financial assets, they can be modeled as playing trading games in which risks are transferred. Market microstructure invariance begins with the intuition that these risk transfers, which we call “bets,” take place in business time. The rate at which business time passes—market “velocity”—is the rate at which new bets arrive into the market. For actively traded assets, business time passes quickly; for inactively traded assets, business time passes slowly. Market microstructure invariance hypothesizes that microstructure characteristics, which vary when measured in units of calendar time, become constants—“microstructure invariants”—when measured in units of business time.

In section 1, we formulate the three invariance principles as empirical hypotheses, conjectured to apply for all securities and across time:

- The distribution of the dollar risk transferred by a bet is the same when the dollar risk it transfers is measured in units of business time.
- The dollar expected transactions cost of executing a bet is the same function of the size of the bet when its size is measured as the dollar risk it transfers in units of business time.
- Pricing accuracy is the same if its reciprocal is scaled by returns volatility per unit of business time, and market resiliency is the same if it is measured in units of business time.

When measured in calendar time, invariance implies that the size distribution of risk transfers, the number of bets, illiquidity, bid-ask spreads, market impact, pricing accuracy, and resiliency become proportional to powers of market velocity. Market velocity itself is proportional to the two-thirds power of calendar-time “trading activity,” which we define as the product of empirically observable dollar volume and returns volatility. This gives specific testable empirical content to the invariance hypotheses. For example, the size distribution of bets, as a fraction of trading volume, is inversely proportional to the two-thirds power of trading activity. The transactions cost function is the product of an invariant cost function of bet size and an asset-specific measure of illiquidity, which is proportional to the cube root of the ratio of returns variance to dollar volume. Other market microstructure characteristics are also proportional to powers of observable dollar volume and returns volatility.

In section 2, we develop a structural model showing that all three microstructure invariance hypotheses are consistent with a dynamic infinite-horizon model of market microstructure with risk-neutral informed trading, noise trading, market making, and endogenous production of information. The invariance relationships are derived under the assumption that the effort required to generate one discrete bet does not vary across stocks and time. Since bets are based on the arrival of discrete chunks of information, the structural model describes how the invariance relationships reflect differences in the granularity of information flows across markets. The invariance

of pricing accuracy and market resiliency requires the additional assumption that private information has the same signal-to-noise ratio across markets.

In section 3, we show how invariance can be used to derive testable implications from theoretical models of market microstructure. Invariance provides guidance on how to construct good empirical proxies for some difficult-to-observe microstructure concepts such as “order imbalances.” It imposes a discipline on empirical tests by showing how to specify regressions and scale explanatory variables so that estimated regression coefficients can be assumed to be constant across observations.

In section 4, we describe the portfolio transitions data used to test invariance relationships concerning bet size and transactions costs. The dataset consists of more than 400,000 portfolio transition orders executed over the period 2001 through 2005 by a leading vendor of portfolio transition services. Portfolio transitions are used by institutional sponsors to transfer funds from legacy portfolio managers to new managers in order to replace fund managers, change asset allocations, or accommodate cash inflows and outflows. Portfolio transitions provide a good natural experiment for identifying bets and measuring transactions costs.

In section 5, we examine whether bet sizes are consistent with the invariance hypothesis under the identifying assumption that portfolio transition orders are proportional to bets. As implied by invariance, the size distribution of the product of the ratio of order size as a fraction of average daily volume and the two-thirds power of trading activity indeed resembles an invariant distribution. Regression analysis also confirms this finding.

The portfolio transition orders have a log-normal distribution with estimated log-variance of 2.53 (figure 2). The log-normal empirical distribution of order size (a bi-modal “signed” log-normal distribution for buy and sell orders) has much more kurtosis than the normal distribution often assumed for analytical convenience in the theoretical literature. The fat tails of the estimated log-normal distribution suggest that very large bets represent a large fraction of trading volume and generate an even larger fraction of returns variance. Execution of large bets may trigger noticeable market dislocations.

In section 6, we use implementation shortfall to examine whether transactions costs are consistent with the invariance hypothesis. Even though statistical tests usually reject invariance hypotheses, the results are economically close to those implied by invariance. Consistent with invariance, transactions cost functions can be closely approximated by the product of an asset-specific illiquidity measure (proportional to the cube root of the ratio of returns variance to dollar volume) and an invariant function of bet size (figure 5). Invariance itself does not impose a particular form on the transactions cost function. Empirically, both a linear model and a square root model explain transactions costs well. A square root model explains transactions costs for orders in the 90th to 99th percentiles better than a linear model; a linear model explains transaction costs for the largest 1% of orders slightly better than the square root model. Quoted spreads are also consistent with the predictions of invariance.

In section 7, we present estimates of the distribution of bet size, the number of bets, and linear and square root transactions cost functions based on calibration of a handful of parameters.

The potential benefits of invariance principles for empirical market microstructure are enormous. In the area of transactions cost measurement, for example, controlled experiments are costly and natural experiments are rare; even well-specified tests of transactions cost models tend to have low statistical power. Market microstructure invariance defines parsimonious structural relationships leading to precise predictions about how various microstructure characteristics, including transactions costs, vary across stocks with different dollar volume and volatility. These predictions can be tested with structural estimates of a handful of parameters using limited data from many different stocks.

The idea of using invariance principles in finance and economics, at least implicitly, is not new. The theory of Modigliani and Miller (1958) is an example of an invariance principle. The idea of measuring trading in financial markets in business time or transaction time is not new either. The “time-change” literature has a long history, beginning with Mandelbrot and Taylor (1967), who link business time to transactions, and Clark (1973), who links business time to volume. Allais (1956) and Allais (1966) are other early examples of models with time deformation. More recent papers include Hasbrouck (1999), Ané and Geman (2000), Dufour and Engle (2000), Plerou et al. (2000), and Derman (2002). Invariance is not based on the idea that returns volatility is constant in transaction time, like some of these papers; instead, invariance is based on the different idea that the dollar risk transferred by the average bet is the same in bet time.

Given values of a tiny number of proportionality constants, the invariance relationships allow microscopic features of the market for a financial asset to be inferred from macroscopic market characteristics such as dollar volume and returns volatility. The units in which these proportionality factors are measured are consistent with their intended economic content. Making empirical predictions on the basis of invariance principles is well established in physics. Our analysis is similar in spirit to inferring the size and number of molecules in a mole of gas from measurable large-scale physical quantities. It is also similar to methods physicists use to model turbulence.

In both physics and market microstructure, application of invariance principles requires that certain assumptions be met. For example, the laws of physics hold in simplest form for objects traveling in a vacuum but have to be modified when resistance from air generates friction. Similarly, in market microstructure, invariance relationships may hold only under idealized conditions. For example, the predictions of invariance may hold most closely when tick size is small, market makers are competitive, and transactions fees and taxes are minimal. Invariance principles provide a benchmark from which the importance of frictions such as a large tick size, non-competitive market access, or high fees and taxes can be measured.

1 Market Microstructure Invariance as an Empirical Hypothesis

Market microstructure characteristics such as order size, order arrival rate, price impact, bid-ask spread, pricing accuracy, and market resiliency vary across assets and across time. We define “market microstructure invariance” as the empirical hypothesis that this variation almost disappears when these characteristics are examined at an asset-specific “business-time” scale which measures the rate at which risk transfer takes place.

Although the discussion below is based on cross-sectional implications of invariance for equity markets for individual stocks, we believe that invariance principles generalize to markets for commodities, bonds, currencies, and aggregate indices such as exchange-traded funds and stock index futures contracts. For simplicity, we assume that a bet transfers only idiosyncratic risk about a single security, not market risk; modeling both idiosyncratic and market risks simultaneously takes us beyond the scope of this paper.

Notation. In the market for an individual stock, institutional asset managers buy and sell shares to implement “bets.” We think of a bet as a decision to acquire a long-term position of a specific size in a stock, distributed approximately independently from other such decisions. Intermediaries with short-term trading strategies—market makers, high frequency traders, and other arbitragers—clear markets by taking the other side of bets placed by long-term traders.

Over short periods of time, we assume that the bet arrival rate can be approximated by a compound Poisson process. Let γ denote the expected bet arrival rate of independently distributed bets, measured in bets per calendar day. Let \tilde{Q} denote a random variable whose probability distribution represents the signed size of bets; \tilde{Q} is measured in shares (positive for buys, negative for sells) with $E\{\tilde{Q}\} = 0$. The bet arrival rate γ measures market velocity, the rate at which business time passes for a particular stock.¹

Bets can be difficult for researchers to observe. Consider an asset manager who places one bet by purchasing 100,000 shares of IBM stock. The bet might be implemented by placing orders over several days, and each of the orders might be “shredded” into many small trades. Since bets represent independent increments in the intended order flow, the various trades which implement the bet must all be added together to recover the size of the original bet. Thus, individual bets are impossible to reconstruct from publicly disseminated records of time-stamped prices and quantities such as those contained in the Trade and Quote (TAQ) database.

¹Over long periods of time, we assume that the inventories of intermediaries do not grow in an unbounded manner; this requires bets to have small negative autocorrelation. Furthermore, both the bet arrival rate and the distribution of bet size change over longer periods of time as the level of trading activity in a stock increases or decreases.

Similarly, if an analyst issues a buy recommendation to ten different customers and each of the customers quickly places executable orders to buy 10,000 shares, it might be appropriate to think of the ten orders as one bet for 100,000 shares. The bet results from a new idea, which can be shared. The ten individual orders lack statistical independence.

We assume that, on average, each unit of betting volume results in ζ units of total volume, i.e., one unit of betting volume leads to $\zeta - 1$ units of intermediation volume. On a given calendar day, expected trading volume (in shares) is given by $V := \zeta/2 \cdot \gamma \cdot E|\tilde{Q}|$; dividing by two implies that a buy-bet matched to a sell-bet is counted as one unit of volume, not two. We define expected “betting volume” \bar{V} by

$$\bar{V} := \gamma \cdot E\{|\tilde{Q}|\} = \frac{2}{\zeta} \cdot V. \quad (1)$$

We can estimate expected betting volume \bar{V} by combining an estimate of expected market volume V with a value for the “volume multiplier” ζ . If all trades are bets and there are no intermediaries, then $\zeta = 1$, since each unit of trading volume matches a buy-bet with a sell-bet. If a monopolistic specialist intermediates all bets without involvement of other intermediaries, then $\zeta = 2$. If each bet is intermediated by different market makers, each of whom lays off inventory by trading with other market makers, then $\zeta = 3$. If positions are passed around among multiple intermediaries, then $\zeta \geq 4$.

Define “returns volatility” σ as the percentage standard deviation of a stock’s daily returns. Some price fluctuations result from release of information directly without trading, such as overnight news announcements. Let ψ^2 denote the fraction of returns variance σ^2 resulting from order flow imbalances, which are ultimately related to the arrival bets. We define “betting volatility” as the standard deviation of returns resulting from bet-related order flow imbalances:

$$\bar{\sigma} := \psi \cdot \sigma. \quad (2)$$

We can estimate betting volatility $\bar{\sigma}$ by combining an estimate of returns volatility σ with a value for the “volatility multiplier” ψ . Let P denote the price of the stock; then dollar “betting volatility” is $P \cdot \bar{\sigma} = \psi \cdot P \cdot \sigma$.

For simplicity, we assume that ζ and ψ are constant across stocks and time.

Invariance of Bets. Market microstructure invariance hypothesizes that *the distribution of the dollar risk transferred by a bet is the same when the dollar risk it transfers is measured in units of business time.*

A bet of dollar size $P \cdot \tilde{Q}$ generates a standard deviation of dollar mark-to-market gains or losses equal to $P \cdot |\tilde{Q}| \cdot \bar{\sigma} \cdot \gamma^{-1/2}$ in one unit of business time $1/\gamma$. The signed standard deviation $P \cdot \tilde{Q} \cdot \bar{\sigma} \cdot \gamma^{-1/2}$ measures both the direction and the size of the risk transfer resulting from the bet. Instead of measuring the size of a bet as a number of

shares \tilde{Q} , the size of the bet can alternatively be measured as the amount of risk it transfers per unit of business time, which we denote as \tilde{I} . Therefore, \tilde{I} is defined by

$$\tilde{I} := P \cdot \tilde{Q} \cdot \bar{\sigma} \cdot \gamma^{-1/2}. \quad (3)$$

Invariance of bets is the hypothesis that the distribution of the random variable \tilde{I} is the same across stocks and time. In this sense, the distribution of risk transfer \tilde{I} is a market microstructure invariant.

By analogy with bets, we define “trading activity” W as the product of expected dollar trading volume $P \cdot V$ and calendar returns volatility σ , i.e., $W := \sigma \cdot P \cdot V$. Trading activity measures the aggregate dollar risk transferred by all bets during one calendar day. Similarly, define “betting activity” \bar{W} as the product of dollar betting volume $P \cdot \bar{V}$ and betting volatility $\bar{\sigma}$, i.e., $\bar{W} := \bar{\sigma} \cdot P \cdot \bar{V}$.² Given values of the volume multiplier ζ and the volatility multiplier ψ , we can convert more-easily-observed trading activity W into less-easily-observed betting activity \bar{W} using the relationship $\bar{W} = W \cdot 2\psi/\zeta$.

Equations (1) and (3) imply that betting activity \bar{W} can be expressed as a function of the unobservable speed of business time γ as:

$$\bar{W} = \bar{\sigma} \cdot P \cdot \bar{V} = \bar{\sigma} \cdot P \cdot \gamma \cdot E\{|\tilde{Q}|\} = \gamma^{3/2} \cdot E\{|\tilde{I}|\}. \quad (4)$$

Invariance of bets therefore makes it possible to infer the bet arrival rate γ and the average bet size $E\{|\tilde{Q}|\}$ from the level of betting activity \bar{W} , up to some proportionality constant $E\{|\tilde{I}|\}$ which does not vary across stocks. Define $\iota := (E\{|\tilde{I}|\})^{-1/3}$; since \tilde{I} has an invariant probability distribution, ι is a constant. Solving equation (4) for γ in terms of \bar{W} and $E\{|\tilde{Q}|\}$ yields

$$\gamma = \bar{W}^{2/3} \cdot \iota^2, \quad E\{|\tilde{Q}|\} = \bar{W}^{1/3} \cdot \frac{1}{P \cdot \bar{\sigma}} \cdot \iota^{-2}. \quad (5)$$

The shape of the entire distribution of bet size \tilde{Q} can be obtained by plugging γ from equation (5) into equation (3). Expressing bet size \tilde{Q} as a fraction of expected betting volume \bar{V} , we obtain

$$\frac{\tilde{Q}}{\bar{V}} = \bar{W}^{-2/3} \cdot \tilde{I} \cdot \iota. \quad (6)$$

Equations (5) and (6) summarize the implications of invariance for bet size and arrival rate. We test these implications in section 5.

The specific exponents 2/3 and 1/3 in equation (5) have simple intuition. Suppose business time γ speeds up by a factor of 4, but volatility in calendar time $\bar{\sigma}$ does not change. Then volatility per unit of business time $\bar{\sigma} \cdot \gamma^{-1/2}$ decreases by a factor of 2.

²In principle, we could distinguish between \bar{P} and P based on adjustments for transactions fees, fee rebates, taxes, and tick size effects. To keep matters simple, we ignore these issues and effectively assume $\bar{P} = P$.

The invariance principle (3) therefore requires bet size \tilde{Q} to increase by a factor of 2 to keep the distribution of \tilde{I} invariant. The resulting increase in betting volume by a factor of $8 = 4^{3/2}$ can be decomposed into an increase in the number of bets by a factor of $8^{2/3} = 4$ and the size of bets by a factor of $8^{1/3} = 2$.

As betting activity increases, the number of bets increases twice as fast as their size. This specific relationship between the number and size of bets lies at the very heart of invariance. Our hypotheses about transactions costs, price impact, pricing accuracy, and market resiliency are based on the order flow having this specific composition.

Invariance of Transactions Costs. Market microstructure invariance also makes empirical predictions about transactions costs. Market microstructure invariance hypothesizes that *the dollar expected transactions cost of executing a bet is the same function of the size of the bet when its size is measured as the dollar risk it transfers in units of business time.*

Since the risk transferred per unit of business time by a bet of \tilde{Q} shares is measured by $\tilde{I} = P \cdot \tilde{Q} \cdot \bar{\sigma} \cdot \gamma^{-1/2}$, invariance of trading costs implies that there exists an “invariant expected transactions cost function” $C_B(\tilde{I})$ which measures in dollars the expected execution cost of transferring the risk represented by $\tilde{Q} = \tilde{I} \cdot (\bar{\sigma} P \gamma^{-1/2})^{-1}$ shares. The dollar transactions cost function $C_B(\tilde{I})$ is a market microstructure invariant.

Consider the following numerical example. Suppose that a 99th percentile bet in stock A is for \$10 million (e.g., 100,000 shares at \$100 per share) while a 99th percentile bet in stock B is for \$1 million (e.g., 100,000 shares at \$10 per share). Even though the dollar sizes of these bets are different, stock A may be more actively traded than stock B , and therefore its returns volatility per unit of business time is lower. Since both bets occupy the same percentile in the bet size distribution for their respective stocks, the invariance of the distribution of bet size implies that the value of \tilde{I} is the same in both cases. Furthermore, even though the bet in stock A has 10 times the dollar value of the bet in stock B , invariance of transactions costs implies that the expected cost of executing each bet $C_B(\tilde{I})$ is the same in dollars, because both bets transfer the same amount of risk per stock-specific unit of business time. Measured in basis points, however, invariance implies that the transactions cost for Stock B is 10 times greater than for stock A .

Let $C(\tilde{Q})$ denote the stock-specific expected cost of executing a bet of \tilde{Q} shares, expressed as a fraction of the notional value of the bet $|P \cdot \tilde{Q}|$, i.e., in units of 10^{-4} basis points. Define the unconditional expected cost $\bar{C}_B := E\{C_B(\tilde{I})\}$. From equation (3), it follows that the formula for percentage costs is given by

$$C(\tilde{Q}) = \frac{C_B(\tilde{I})}{|P \cdot \tilde{Q}|} = \frac{\bar{C}_B}{E\{|P \cdot \tilde{Q}|\}} \cdot \frac{C_B(\tilde{I})/\bar{C}_B}{|\tilde{I}|/E\{|\tilde{I}|\}}. \quad (7)$$

Let $f(\tilde{I}) := [C_B(\tilde{I})/\bar{C}_B]/[|\tilde{I}|/E\{|\tilde{I}|\}]$ denote the invariant “average cost function” for executing a bet \tilde{I} . This function is the ratio of $C_B(\tilde{I})$ to $|\tilde{I}|$ when both $C_B(\tilde{I})$ and

$|\tilde{I}|$ are expressed as multiples of their means. For example, if I denotes a bet that is 5 times greater than than an average unsigned bet of size $E\{|\tilde{I}|\}$ and such a bet has a dollar transactions cost 10 times greater than the mean cost \bar{C}_B , then $f(I) = 2$.

Let $1/L := \bar{C}_B/E\{|P \cdot \tilde{Q}|\}$ be an asset-specific measure of illiquidity equal to the dollar-volume-weighted expected cost of executing a bet. For an asset manager who places many bets in the same stock, this expresses expected transactions cost as a fraction of the dollar value traded (10^{-4} basis points). Equation (5) yields

$$\frac{1}{L} := \iota^2 \bar{C}_B \cdot \bar{\sigma} \bar{W}^{-1/3} = \iota^2 \bar{C}_B \cdot \left(\frac{\bar{\sigma}^2}{P \cdot \bar{V}} \right)^{1/3}. \quad (8)$$

Recall that $\iota := (E\{|\tilde{I}|\})^{-1/3}$ and \bar{C}_B are constants in this equation.

Plugging the definitions of $f(\tilde{I})$ and $1/L$ into equation (7) makes it possible to write the expected percentage cost function $C(Q)$ as the product of the asset-specific illiquidity measure $1/L$ and an invariant transactions cost function $f(\tilde{I})$:

$$C(\tilde{Q}) = \frac{1}{L} \cdot f(\tilde{I}). \quad (9)$$

This is an important decomposition of transaction cost functions.

When the bet size is measured as a fraction of betting volume \tilde{Q}/\bar{V} , the cost function $C(Q)$ can be expressed conveniently in terms of betting activity \bar{W} and invariant constants ι and \bar{C}_B as

$$C(\tilde{Q}) = \bar{\sigma} \bar{W}^{-1/3} \cdot \iota^2 \bar{C}_B \cdot f\left(\frac{\bar{W}^{2/3}}{\iota} \cdot \frac{\tilde{Q}}{\bar{V}}\right). \quad (10)$$

We test this specification empirically in section 6.

To implement L empirically, it is simpler to define L in terms of expected dollar trading volume $P \cdot V$ and expected returns volatility σ rather than in terms of dollar betting volume $P \cdot \bar{V}$ and betting volatility $\bar{\sigma}$. Using equations (1) and (2), we have

$$\frac{1}{L} = (\zeta \psi^2)^{1/3} \cdot \iota^2 \bar{C}_B \cdot \left(\frac{\sigma^2}{P \cdot V} \right)^{1/3}. \quad (11)$$

The idea that liquidity is related to dollar volume per unit of returns variance $P \cdot V/\sigma^2$ is intuitive. Traders believe that transactions costs are low in markets with high dollar volume and high in markets with high volatility. If the volume multiplier ζ and the volatility multiplier ψ do not vary across stocks, then $L \propto [P \cdot V/\sigma^2]^{1/3}$ becomes a simple index of liquidity.

The liquidity measure $L = [\iota^2 \bar{C}_B]^{-1} \cdot [P \cdot \bar{V}/\bar{\sigma}^2]^{1/3}$ is an intuitive and practical alternative to other measures of liquidity, such as Amihud (2002) and Stambough and Pastor (2003). The value of L can easily be calibrated from price and volume data provided by The Center for Research in Security Prices (CRSP).³

³The liquidity measure L is similar to the definition of “market temperature” $\chi = \bar{\sigma} \cdot \gamma^{1/2}$ in Derman (2002); substituting for γ from equation (5), we obtain $\chi = \iota \cdot [P \cdot \bar{V}]^{1/3} \cdot [\bar{\sigma}]^{4/3} \propto L \cdot \sigma^2$.

Invariance does not imply a specific functional form for function $f(\cdot)$. In our analysis, we focus on two specific functional forms: linear price impact costs and square root price impact costs. For both functional forms, we also allow a constant bid-ask spread cost component. The linear price impact function is consistent with price impact models based on adverse selection, such as Kyle (1985). The square root price impact function is consistent with empirical findings in the econophysics literature, such as Gabaix et al. (2006), although these results are based on single trades rather than bets; some papers in this literature, such as Almgren et al. (2005), find an exponent closer to 0.60 than the square root exponent of 0.50.

For the linear model, we write $f(\tilde{I})$ as the sum of a bid-ask spread component and a linear price impact cost component, $f(\tilde{I}) := (\iota^2 \bar{C}_B)^{-1} \cdot \kappa_0 + (\iota \bar{C}_B)^{-1} \cdot \kappa_I \cdot |\tilde{I}|$, where invariance implies that the bid-ask spread cost parameter κ_0 , the market impact cost parameter κ_I , and the constants ι and \bar{C}_B do not vary across stocks. Since the specific coefficients $\iota^2 \bar{C}_B$ and $\iota \bar{C}_B$ are chosen to cancel out in a nice way, equation (9) implies that the proportional cost function $C(\tilde{Q})$ has the simple form

$$C(\tilde{Q}) = \bar{\sigma} \left(\kappa_0 \cdot \bar{W}^{-1/3} + \kappa_I \cdot \bar{W}^{1/3} \cdot \frac{|\tilde{Q}|}{\bar{V}} \right). \quad (12)$$

When bet sizes are measured as a fraction of expected betting volume and transactions costs are measured in basis points, bid-ask spread costs are decreasing in betting activity \bar{W} and market impact costs are increasing in betting activity \bar{W} . When transactions costs in basis points are further scaled in units of betting volatility $\bar{\sigma}$, equation (12) says that bid-ask spread costs are proportional $\bar{W}^{-1/3}$ and market impact costs are proportional to $\bar{W}^{1/3}$ for a given fraction of betting volume.

For the square root model, we write $f(\tilde{I})$ as the sum of a bid-ask spread component and a square root function of $|\tilde{I}|$, obtaining $f(\tilde{I}) := (\iota^2 \bar{C}_B)^{-1} \kappa_0 + (\iota^{3/2} \bar{C}_B)^{-1} \kappa_I \cdot |\tilde{I}|^{1/2}$, where invariance implies that κ_0 , κ_I , ι , and \bar{C}_B do not vary across stocks. The proportional cost function $C(\tilde{Q})$ from (9) is given by

$$C(\tilde{Q}) = \bar{\sigma} \left(\kappa_0 \cdot \bar{W}^{-1/3} + \kappa_I \cdot \left| \frac{\tilde{Q}}{\bar{V}} \right|^{1/2} \right). \quad (13)$$

When transactions costs are measured in units of betting volatility $\bar{\sigma}$, bid-ask spread costs remain proportional to $\bar{W}^{-1/3}$, but the square root model implies that the betting activity coefficient $\bar{W}^{1/3}$ cancels out of the price impact term. Indeed, the square root is the only function for which invariance leads to the empirical prediction that impact costs (measured in units of returns volatility) depend only on bet size as a fraction of betting activity \tilde{Q}/\bar{V} and not on any other stock characteristics. If there are no bid-ask spread costs so that $\kappa_0 = 0$, then the square root model implies the parsimonious transactions cost function $C(\tilde{Q}) = \bar{\sigma} \cdot \kappa_I \cdot [|\tilde{Q}|/\bar{V}]^{1/2}$.

Sometimes called by practitioners “the Barra model,” a square root price impact model like specification (13) was proposed by Barra (1997) based on empirical regularities observed by Loeb (1983). Grinold and Kahn (1995) use an inventory risk model to derive a square root price impact formula. Gabaix et al. (2006) formalize this approach under the assumptions that (1) orders are executed as a constant fraction of volume and (2) liquidity providers have mean-variance utility functions linear in expected wealth and its standard deviation (not variance).

Invariance of Pricing Accuracy and Market Resiliency. We think of “fundamental value” F as the value to which a stock price would converge if traders continuously invested huge resources acquiring information about its value. Let Σ denote the variance of the log-difference between price and fundamental value, $\Sigma := \text{var}\{\ln(F/P)\}$. Then $\Sigma^{1/2}$ measures the standard deviation of this log-difference as a fraction of the market price, and $\Sigma^{-1/2}$ measures “pricing accuracy.”

Black (1986) defines efficient markets using a similar concept of pricing accuracy; he conjectures that “almost all markets are efficient” in the sense that “price is within a factor 2 of value” at least 90% of the time. Using our notation, Black’s conjecture is satisfied if $\Sigma^{1/2} < \ln(2)/1.64$ in almost all markets, since the probability that a normal distribution is within 1.64 standard deviations of its mean is approximately 90%. If $\Sigma^{1/2}$ becomes smaller, Black would say that the market has become more “efficient.”

Fischer Black’s definition of market efficiency based on pricing accuracy contrasts sharply with the conventional definition of Eugene Fama, who considers a market to be “efficient” if all available information is appropriately reflected in price, even if very little information is available and prices are not very accurate, i.e., $\Sigma^{1/2}$ is large.

The pricing error variance Σ can be converted into time units by scaling it by fundamental volatility σ^2 . The quantity Σ/σ^2 can be further interpreted as the number of years by which the informational content of prices lags behind fundamental value. For example, if a stock’s annual volatility is about 35% and $\Sigma^{1/2} = \ln(2)/1.64$, then prices are about 1.50 years “behind” fundamental value since $(\ln(2)/1.64)^2/0.35^2 \approx 1.50$.

Define “market resiliency” ρ as the mean-reversion parameter (per calendar day) measuring the speed with which a random shock to prices, resulting from execution of an uninformative bet, dies out over time as informative orders drive prices towards fundamental value. The half-life of an uninformative shock to prices is $\rho^{-1} \cdot \ln(2)$.

Market microstructure invariance hypothesizes that *pricing accuracy is the same if its reciprocal is scaled by returns volatility per unit of business time, and market resiliency is the same if it is measured in units of business time*. More specifically, invariance of pricing accuracy implies that the ratio $\Sigma^{1/2}/[\bar{\sigma} \cdot \gamma^{-1/2}]$ is invariant across stocks. Since $1/L \propto \bar{\sigma} \cdot \gamma^{-1/2}$ from equation (11), invariance also implies proportionality between pricing accuracy $\Sigma^{-1/2}$ and liquidity L . Invariance suggests that the ratio ρ/γ is invariant across stocks, making resiliency proportional to the rate at which business time passes.

Using equation (5), invariance implies the following invariance relationships:

$$\Sigma^{1/2} \propto \bar{\sigma} \cdot \bar{W}^{-1/3} \propto \frac{1}{L}, \quad \rho \propto \bar{W}^{2/3} \propto \gamma. \quad (14)$$

For example, when trading activity increases by a factor of 8, resiliency increases by a factor of 4 and pricing accuracy increases by a factor of 2.⁴ Intuitively, the unstated invariant proportionality factors implied by equation (14) should be related to the information content of bets. More informative bets should make markets more efficient and resilient. This intuition is made precise in the structural model examined in the next section.

We do not examine empirically the predictions of equation (14); they are however interesting topics for future research.

Numerical Example. The following example illustrates the invariance hypotheses. Suppose a stock has daily volume of \$40 million and daily returns volatility of 2%. Suppose there are approximately 100 bets per day and the mean size of a bet is \$400,000. A daily volatility of 2% implies a standard deviation of mark-to-market dollar gains and losses equal to \$8,000 per calendar day for the mean bet. Since business time passes at the rate bets arrive into the market, 100 bets per day implies about 1 bet every 4 minutes; the business clock therefore ticks once every 4 minutes. Over a 4-minute period, the standard deviation of returns is 20 basis points ($200/\sqrt{100}$). Thus, the mean bet has a standard deviation of risk transfer of \$800 per unit of business time.

Invariance says that the specific number \$800 is constant across stocks and across time. For example, if the arrival rate of bets increases by a factor of 4, the business clock ticks 4 times faster (once every minute), and the standard deviation of returns per tick on that clock is reduced from 20 basis points to 10 basis points ($200/\sqrt{400}$). For the standard deviation of mark-to-market dollar gains and losses on the mean bet to remain constant at \$800, invariance implies that the dollar size of the mean bet must increase by a factor of 2 from \$400,000 to \$800,000. Thus, holding volatility constant, the bet arrival rate increases by a factor of 4 and the size of bets increases by a factor of 2. This implies that daily volume increases by a factor of 8 from \$40 million to \$320 million. Holding volatility constant, the bet arrival rate increases by a factor proportional the 2/3 power of the factor by which dollar volume changes, and the size of bets increases by the 1/3 power of the factor by which dollar volume changes.

Invariance of transaction costs says that the dollar cost of executing a bet of a given size percentile is the same across different stocks. The dollar costs of executing the mean bet is therefore the same constant across securities and across time and,

⁴If the definition of market efficiency as pricing accuracy in Black (1986) is applied across stocks, invariance suggests a cut-off value of liquidity L^* such that stocks with higher liquidity L have efficient markets and stocks with lower liquidity L have inefficient markets.

for example, equal to \$2,000. Doubling the size of the mean bet from \$400,000 to \$800,000 decreases the cost of \$2,000 measured in basis points from 50 basis points to 25 basis points, i.e., by a factor of 2. The transactions cost of executing a bet, measured in basis points, is therefore inversely proportional to the $1/3$ power of the factor by which dollar volume changes, holding volatility constant.

Similar arguments are valid for bets in different percentiles of the bet size distribution. Suppose, for example, that the standard deviation of mark-to-market gains or losses on a 99th percentile bet is \$10,000 per tick of business time (4 minutes). The size of such a bet is therefore \$5 million in the first stock and \$10 million in the second stock, different by a factor of 2. If the dollar costs of executing 99th percentile bets is equal to \$50,000, then this corresponds to a cost of 100 basis points in the first stock and 50 basis points in the second stock. The percentage cost for the second stock is lower by a factor of 2; this difference is inversely proportional to the $1/3$ power of the factor by which dollar volume changes, holding volatility constant.

Discussion and Other Implications. Invariance is consistent with irrelevance of the units in which time is measured. The values of \tilde{I} , $C_B(\tilde{I})$, $f(\tilde{I})$, and $1/L$ —and therefore the economic content of the predictions of invariance—remain the same regardless of whether researchers measure γ , \bar{V} , $\bar{\sigma}$, and \bar{W} using daily weekly, monthly, or annual units of time. This is unlike some other models, such as ARCH and GARCH.

The values of \tilde{I} and $C_B(\tilde{I})$ are measured in dollars. Invariance relationships can also be applied to an international context in which markets have different currencies or different real exchange rates; they can also be applied across periods of time when the price level is changing significantly. Invariance is consistent with the idea that these nominal values \tilde{I} and $C_B(\tilde{I})$ should be equal to the nominal cost of financial services calculated from the productivity-adjusted wages of finance professionals in the local currency of the given market during the given time period. Since wages are measured in dollars per day and productivity is measured in bets per day, the ratio of wages to productivity is measured in dollars per bet, exactly the same units as \tilde{I} and $C_B(\tilde{I})$. Like fundamental constants in physics, dividing the invariants \tilde{I} and $C_B(\tilde{I})$ by the ratio of wages to productivity makes them dimensionless.

We model market microstructure using invariance in a manner similar to the way modern physicists model turbulence. Kolmogorov (1941a) derived his “two-thirds law” (or “five-thirds law”) for the energy distribution in a turbulent fluid based on dimensional analysis and scaling.⁵

Invariance is also consistent with the Modigliani-Miller irrelevance of leverage and splits. It can be shown that invariance relationships do not change if a company levers up its equity by paying a debt-financed cash dividend or implements a stock split.

⁵We thank an anonymous referee and Sergey N. Smirnov for pointing out the connection to Kolmogorov’s model of turbulence.

Invariance has numerous implications. We outline some of them below to indicate interesting directions for future research without pursuing them further in this paper.

Invariance relationships are based on the implicit assumption that bets are executed at an endogenously determined “natural” speed that trades off the benefits of faster execution against higher transactions costs. It does not rule out the possibility that unusually fast execution of a bet would lead to execution costs higher than the costs implied by the functions $f(\tilde{I})$ and $C_B(\tilde{I})$. For example, it is possible to consider more general invariant cost functions $f(\tilde{I}, T/\gamma)$ and $C_B(\tilde{I}, T/\gamma)$ that depend not only of the size of bets but also on execution horizons T converted from units of calendar time into units of business time T/γ .

Invariance implies that trading liquidity and funding liquidity may be two sides of the same coin. Trading liquidity is measured by L , which is proportional to $\bar{W}^{1/3}/\bar{\sigma}$ and $\gamma^{1/2}/\bar{\sigma}$. A good measure of funding liquidity is the repo haircut that sufficiently protects a creditor from losses if the creditor sells the collateral due to default by the borrower. Such a haircut should be proportional to the volatility of the asset’s return over the horizon during which the collateral would be liquidated. This horizon should be proportional to business time $1/\gamma$, making volatility over the liquidation horizon proportional to $\bar{\sigma} \cdot \gamma^{-1/2}$, which is itself proportional to $1/L$. Thus, invariance suggests that both trading liquidity and funding liquidity can be measured by L .

Asset pricing models often assume that less liquid assets must command a returns premium to compensate investors for illiquidity. Invariance suggests the hypothesis that the returns premium required to compensate an investor for asset illiquidity should be proportional to $1/L$. Indeed, suppose that active long only asset managers are the marginal investors in a universe of stocks with similar turnover rates, and the costs of bets are imposed *pro rata* on the investors. Since market capitalization is proportional to $P \cdot V$ and dollar costs are proportional to the bet arrival rate $\gamma \propto W^{2/3}$, percentage costs in basis points are proportional to $W^{2/3}/(PV) \propto 1/L$. This argument suggests that the proportionality factor may show up both in returns and management fees.

Invariance also leads naturally to implications for the time periods over which institutional investors hold their positions. Active speculative positions are expected to be held for a period of time proportional to the bet arrival rate $\gamma \propto W^{2/3}$.

2 Market Microstructure Invariance as an Implication of a Structural Model

In this section, we derive invariance relationships as endogenous implications of a structural model of informed trading, noise trading, and intermediation (market making).

The model has the following structure. The unobserved “fundamental value” of the stock follows geometric Brownian motion with log-standard-deviation σ_F . There

are two types of traders: informed traders and noise traders. Informed traders and noise traders arrive in the market randomly and trade only once. Each informed trader trades based on a costly informative signal of precision τ ; the price impact of the trade incorporates a fraction θ of this private information into prices. In the spirit of Black (1986), each noise trader trades based on a “fake” signal which has the same unconditional distribution as an informative signal but has no information content. We call the noise traders “Fischer Black noise traders.” When an informed or noise trader arrives in the market, the trader “places a bet” by announcing to competitive market makers the quantity the trader wishes to trade. Conditional on this announced quantity but not knowing whether the trader is informed or not, competitive risk-neutral market makers determine a break-even price at which the trade takes place. Noise traders place bets at a rate which leads to an exogenously given turnover rate η of the exogenous float of N shares. Informed traders place bets at an endogenously determined rate which equates the cost c_I of each private signal to the expected value of trading on that signal, taking into account the price impact of the bet. Our structural model has the flavor of Glosten and Milgrom (1985) and Black (1986).

Invariance relationships come about through the following intuition: Suppose the number of noise traders increases for some exogenous reason. In the structural model, this happens when the share price and therefore market capitalization increases, keeping the share turnover of noise traders constant. To be specific, assume that the number of noise traders increases by a factor of 4. As a result, market depth increases and, consequently, the number of informed traders increases, since their bets now are more profitable. The structural model shows that the number of informed traders eventually increases by a factor of 4 as well, and each of their bets accounts for a 4 times smaller fraction of returns variance. Returns volatility per unit of business time decreases by a factor of 2 (the square root of 4). The structural model shows that pricing accuracy and liquidity both increase by a factor of 2, as a result of which informed traders exactly cover the cost of private signals by submitting bets 2 times as large as before. Overall dollar volume in the market increases by a factor of 8. Thus, the “one-third, two-thirds” intuition comes about: One-third of the increase in dollar volume comes from changes in bet size ($8^{1/3} = 2$) and two-thirds comes from changes in the number of bets ($8^{2/3} = 4$).

To make clear the distinction between endogenous variables and exogenous parameters, we use notation that assumes all exogenous parameters are constants while all endogenous variables are time-varying. In equilibrium, the price $P(t)$ follows an endogenously derived martingale process with returns volatility $\sigma(t)$. The model endogenously determines the rate at which informed bets occur $\gamma_I(t)$, the rate at which bets by noise traders occur $\gamma_U(t)$, the distribution of bet sizes $\tilde{Q}(t)$, pricing accuracy $\Sigma(t)^{-1/2}$, market resiliency $\rho(t)$, market illiquidity $1/L(t)$, and the price impact parameter $\lambda(t)$.

In the remainder of this section, we describe the structural model in more detail,

using notation consistent with the previous section. For simplicity, we assume $\zeta = 2$ and $\psi = 1$ so that $V = \bar{V}$ and $\sigma = \bar{\sigma}$. It is straightforward to adjust the model for $\zeta \neq 2$ and $\psi \neq 1$ by applying equations (1) and (2).

Fundamental Value and Private Information. Let the unobserved “fundamental value” of the asset follow a geometric Brownian motion given by

$$F(t) := \exp[\sigma_F \cdot B(t) - \frac{1}{2} \cdot \sigma_F^2 \cdot t], \quad (15)$$

where $B(t)$ follows standardized Brownian motion with $\text{var}\{B(t + \Delta t) - B(t)\} = \Delta t$, the constant σ_F measures the “Black-Scholes volatility” of the fundamental value, and the term $\frac{1}{2} \cdot \sigma_F^2 \cdot t$ makes $F(t)$ follow a martingale. Assume a zero risk-free rate of interest. Trading takes place over some long finite horizon, ending at a distant date at which all traders receive a payoff equal to the fundamental value of the asset.

The price changes as informed traders and noise traders arrive into the market and anonymously place bets. Risk-neutral market makers set the market price $P(t)$ as the conditional expectation of the fundamental value $F(t)$ given a history of the “bet flow,” which we can assume to be informationally equivalent to the history of prices.

Rather than focusing on the market’s estimate of the fundamental value $F(t)$, we focus instead on the market’s estimate of the Brownian motion $B(t)$, in terms of which $F(t)$ is defined. Let $\bar{B}(t)$ denote the market’s conditional expectation of $B(t)$ based on observing the history of prices. For now, assume that the error $B(t) - \bar{B}(t)$ has approximately a normal distribution with variance denoted $\Sigma(t)/\sigma_F^2$; scaling by σ_F^2 makes the notation $\Sigma(t)$ consistent with the notation for Σ in section 1.

The price is the best estimate of fundamental value when $P(t)$ is given by

$$P(t) = \exp[\sigma_F \cdot \bar{B}(t) + \frac{1}{2} \cdot \Sigma(t) - \frac{1}{2} \cdot \sigma_F^2 \cdot t]. \quad (16)$$

The price has a martingale property and the variance of $\log[F(t)/P(t)]$ is equal to $\Sigma(t)$, i.e. $\Sigma(t)^{-1/2} \Sigma(t)^{-1/2}$ measures “pricing accuracy.”

Informed Traders. Informed traders arrive randomly into the market at endogenously determined rate $\gamma_I(t)$. If an informed trader arrives at time t , he observes one private signal $\tilde{i}(t)$ in addition to the history of prices, then places one and only one bet, which is executed by trading with market makers. Informed signals are assumed to have the form

$$\tilde{i}(t) := \tau^{1/2} \cdot \Sigma(t)^{-1/2} \cdot \sigma_F \cdot [B(t) - \bar{B}(t)] + \tilde{Z}_I(t), \quad (17)$$

where τ is an exogenous constant parameter measuring the precision of the signal and the noise term $\tilde{Z}_I(t) \sim N(0, 1)$ is distributed independently from the history of prices prior to the arrival of the signal. Scaling $\tau^{1/2}$ by $\Sigma(t)^{-1/2} \cdot \sigma_F$, the reciprocal of

the standard deviation of $B(t) - \bar{B}(t)$, makes τ measure the constant signal-to-noise ratio. We assume τ is small enough so that $\text{var}\{\tilde{i}(t)\} = 1 + \tau \approx 1$. Although the signal associated with the n th bet to arrive is different from previous signals, we omit a clarifying subscript n on $\tilde{i}(t)$.

An informed trader calculates his prior estimate $\bar{B}(t)$ of $B(t)$ from market prices. Upon observing a signal $\tilde{i}(t)$, he updates his prior from $\bar{B}(t)$ to $\bar{B}(t) + \Delta\bar{B}_I(t)$. Assuming τ is small, the update is given by

$$\Delta\bar{B}_I(t) \approx \tau^{1/2} \cdot \Sigma(t)^{1/2} \cdot \sigma_F^{-1} \cdot \tilde{i}(t), \quad (18)$$

where $\tau^{1/2} \cdot \Sigma(t)^{1/2} \cdot \sigma_F^{-1}$ is the regression coefficient of $B(t) - \bar{B}(t)$ on $\tilde{i}(t)$. If the signal value were to be fully incorporated into prices, it can be shown that the dollar price change would be equal to

$$E\{F(t) - P(t) \mid \Delta\bar{B}_I(t)\} \approx P(t) \cdot (\exp[\sigma_F \cdot (\Delta\bar{B}_I(t) - \Delta\bar{B}_I(t)^2/2)] - 1) \approx P(t) \cdot \sigma_F \cdot \Delta\bar{B}_I(t). \quad (19)$$

We assume that an informed trader arriving at time t executes a bet of size $\tilde{Q}(t)$. In order to break even in the presence of adverse selection, market makers take the other side of the bet at a price adjusted by $\lambda(t) \cdot \tilde{Q}(t)$, where $\lambda(t)$ is a price impact parameter. The price impact is linear in the size of the bet.

Suppose that a bet of size $\tilde{Q}(t)$ is a linear multiple of $\Delta\bar{B}_I(t)$ in such a way that only a fraction θ of the “fully revealing” impact $P(t) \cdot \sigma_F \cdot \Delta\bar{B}_I(t)$ is incorporated into prices, i.e.,

$$\tilde{Q}(t) = \theta \cdot \lambda(t)^{-1} \cdot P(t) \cdot \sigma_F \cdot \Delta\bar{B}_I(t). \quad (20)$$

If the informed trader were to trade $\tilde{Q}(t)$ shares defined in equation (20) at price $P(t)$ with no price impact costs, then his unconditional expected “paper trading” profits, denoted $\bar{\pi}_I(t)$, would be equal to

$$\bar{\pi}_I(t) := E\{[F(t) - P(t)] \cdot \tilde{Q}(t)\} = \frac{\theta \cdot P(t)^2 \cdot \sigma_F^2 \cdot E\{\Delta\bar{B}_I(t)^2\}}{\lambda(t)}. \quad (21)$$

Conditioning on observing the signal $\tilde{i}(t)$, the expected profits of the informed trader, net of transactions costs $\lambda(t) \cdot \tilde{Q}(t)^2$, are

$$E\{[F(t) - P(t)] \cdot \tilde{Q}(t) - \lambda(t) \cdot \tilde{Q}(t)^2 \mid \Delta\bar{B}_I(t)\} = \frac{\theta \cdot (1 - \theta) \cdot P(t)^2 \cdot \sigma_F^2 \cdot \Delta\bar{B}_I(t)^2}{\lambda(t)}. \quad (22)$$

It is clear that $\theta = 1/2$ maximizes the expected profits of the informed trader and therefore solves the optimization problem of a risk-neutral informed trader. Rather than assuming that informed traders are risk neutral and therefore $\theta = 1/2$ as the equilibrium choice of θ , we will instead allow θ to have an arbitrary exogenous value such that $0 < \theta < 1$. This approach accommodates the possibility that informed traders are risk averse, in which case $\theta < 1/2$ might be optimal; it also accommodates the possibility of information leakage, in which case $\theta > 1/2$ might be optimal. More importantly, it allows us to show that the invariance results derived below depend only on the fact that θ is some constant, not that θ has the specific value $1/2$.

Noise Traders. Noise traders arrive randomly at an endogenously determined rate $\gamma_U(t)$. Each noise trader places one bet which mimics the size distribution of an informed bet, even though it contains no information. In other words, if a noise trader arrives at time t , he receives a signal $\tilde{i}(t)$ which is assumed to have the same unconditional distribution as an informed signal, i.e., $\tilde{i}(t) = \tilde{Z}_U(t) \sim N(0, 1 + \tau) \approx N(0, 1)$, but it is “noise” in the sense that $\tilde{i}(t)$ is distributed independently from the error $\bar{B}(t) - B(t)$ and the history of prices. Noise traders are assumed to place bets at a rate such that a constant fraction η of the market capitalization of the firm turns over on average per day.

Let N denote the number of shares outstanding, let $V(t)$ denote average share volume per day from informed and noise traders combined, and let $\gamma(t) := \gamma_I(t) + \gamma_U(t)$ denote the combined arrival rate of bets by informed traders and noise traders. Let informed trades and noise trades be distributed as the random variable $\tilde{Q}(t)$ in equation (20). Then, expected share volume from noise traders $\eta \cdot N$, total volume $V(t)$, and the fraction of volume from noise traders $\gamma_U(t)/\gamma(t)$ satisfy

$$\gamma_U(t) \cdot E\{|\tilde{Q}(t)|\} = \eta \cdot N, \quad \gamma(t) \cdot E\{|\tilde{Q}(t)|\} = V(t). \quad (23)$$

Market Makers (Intermediaries). Zero-profit, risk-neutral, competitive market makers are assumed to set prices such that the price impact of anonymous trades by informed and noise traders make the price $P(t)$ equal to the conditional expectation of the fundamental value $F(t)$ given the history of all bets. If market makers could observe whether a bet was placed by an informed trader or noise trader, they would multiply the price impact of informed bets $\lambda(t) \cdot \tilde{Q}(t)$ by $1/\theta$ (since informed bets have price impact which reveals only fraction θ of their information content) and they would multiply the price impact of noise bets $\lambda(t) \cdot \tilde{Q}(t)$ by zero (since noise bets have no information content). The probability of an informed bet is $\gamma_I(t)/\gamma(t)$, and the probability of a noise bet is $\gamma_U(t)/\gamma(t)$. In equilibrium, the average impact of a bet must satisfy

$$\lambda(t) \cdot \tilde{Q}(t) = \frac{\gamma_I(t)}{\gamma_I(t) + \gamma_U(t)} \cdot \lambda(t) \cdot \tilde{Q}(t) \cdot \frac{1}{\theta} + \frac{\gamma_U(t)}{\gamma_I(t) + \gamma_U(t)} \cdot \lambda(t) \cdot \tilde{Q}(t) \cdot 0. \quad (24)$$

This implies that the fraction of informed traders and noise traders satisfies

$$\frac{\gamma_I(t)}{\gamma_I(t) + \gamma_U(t)} = \theta. \quad (25)$$

Thus, the fraction of traders who are informed turns out to be to the exogenous constant θ . In the special case when informed traders are risk neutral, $\theta = 1/2$ implies that the number of informed traders is equal to the number of uninformed traders. Note that this ratio is not determined by equations describing the profitability of informed trading but rather by equations describing the adverse selection problem

faced by market makers. Moreover, equations (23) and (25) imply that, in terms of exogenous variables, share volume $V(t)$ is actually a constant given by

$$V(t) = \eta \cdot N / (1 - \theta). \quad (26)$$

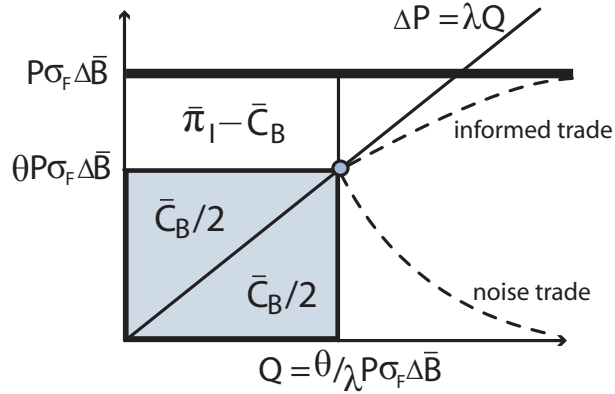
Break-Even Conditions. Equation (20) implies that the unconditional expected price impact cost of an informed bet, denoted $\bar{C}_B(t)$, is given by

$$\bar{C}_B(t) := \lambda(t) \cdot E\{\tilde{Q}(t)^2\} = \frac{\theta^2 \cdot P(t)^2 \cdot \sigma_F^2 \cdot E\{\Delta \bar{B}_I(t)^2\}}{\lambda(t)}. \quad (27)$$

Bets by noise traders have the same expected impact cost since they have the same unconditional distribution as informed bets and are indistinguishable from informed bets from the perspective of market makers. The equilibrium level of costs must allow market makers to break even. Thus, the expected dollar price impact costs that market makers expect to collect from all traders must be equal to the expected dollar paper trading profits of informed traders:

$$(\gamma_I(t) + \gamma_U(t)) \cdot \bar{C}_B(t) = \gamma_I \cdot \bar{\pi}_I(t). \quad (28)$$

Figure 1: Intuition for the Model.



There is price continuation after an informed trade and mean reversion after a noise trade.

Figure 1 illustrates the intuition informally and non-rigorously. Informed traders incorporate only a fraction θ of their information into prices by trading $\tilde{Q}(t)$; they pay transactions costs $\bar{C}_B(t)$ and expect to make $\bar{\pi}_I(t) - \bar{C}_B(t)$ as the price gradually converges to fundamental value $F(t)$ due to the subsequent trading of other informed

traders. These profits are realized at some distant date when the game ends and positions are liquidated at the fundamental value $F(t)$. Noise traders execute orders which incur transactions costs $\bar{C}_B(t)$ but would earn nothing if there were no transactions costs. As in Treynor (1995), the expected losses market makers incur trading with informed traders $\gamma_I(t) \cdot (\bar{\pi}_I(t) - \bar{C}_B(t))$ are equal to their expected gains trading with noise traders $\gamma_U(t) \cdot \bar{C}_B(t)$.

The break-even condition for informed traders yields the rate at which informed traders place bets $\gamma_I(t)$. The expected paper trading profits from trading on a signal $\bar{\pi}_I(t)$ must equal the sum of expected transaction costs $\bar{C}_B(t)$ and the exogenously constant cost of acquiring private information denoted c_I :

$$\bar{\pi}_I(t) = \bar{C}_B(t) + c_I. \quad (29)$$

We let $1/L(t) := \bar{C}_B(t)/E\{|P(t) \cdot \tilde{Q}(t)|\}$ denote the expected percentage cost of executing a bet in basis points, as in section 1.

Diffusion Approximation and Conditional Steady State. The model theoretically implies that the price follows a jump process which changes when a bet arrives. Conditional expectations of $\bar{B}(t)$ calculated based on a random number of bet arrivals during a given time interval are not precisely linear. Moreover, since the price is a nonlinear function of $\bar{B}(t)$, price impact is theoretically nonlinear as well.

To deal with these non-linearities, we assume that the arrival rate of bets is so fast—and the resulting price impact of each bet is so small—that a linear approximation based on a diffusion is appropriate. In the limit as the bet arrival rate $\gamma(t)$ goes to infinity, the conditional expectation $\bar{B}(t)$ becomes an exactly linear function of the history of bets, and the price process becomes a diffusion. This is compatible with assuming that market makers and traders use linear filtering, market makers offer linear supply schedules to traders, traders place bets as linear functions of signals, log-price changes are normally distributed, and price changes are conditionally normally distributed in the sense of an Euler approximation.

As a result of each bet, market makers update their estimate of $\bar{B}(t)$ by $\Delta\bar{B}(t)$ (not to be confused with the update $\Delta\bar{B}_I(t)$ of informed traders). A trade is informed with probability θ and, if informed, incorporates a fraction θ of its information content into prices, leading to an adjustment in $\bar{B}(t)$ of

$$\Delta\bar{B}(t) = \theta \cdot \tau^{1/2} \cdot \Sigma(t)^{1/2} \cdot \sigma_F^{-1} \cdot \left(\tau^{1/2} \cdot \Sigma(t)^{-1/2} \cdot \sigma_F \cdot [B(t) - \bar{B}(t)] + \tilde{Z}_I(t) \right) \quad (30)$$

from equations (17) and (18). A trade is uninformed with probability $1 - \theta$ and, if uninformed, adds noise to $\bar{B}(t)$ of

$$\Delta\bar{B}(t) = \theta \cdot \tau^{1/2} \cdot \Sigma(t)^{1/2} \cdot \sigma_F^{-1} \cdot \tilde{Z}_U(t). \quad (31)$$

When the arrival rate of bets $\gamma(t)$ per day is sufficiently large, the diffusion approximation for the dynamics of the estimate $\bar{B}(t)$ can be written as

$$d\bar{B}(t) = \gamma(t) \cdot \theta^2 \cdot \tau \cdot [B(t) - \bar{B}(t)] \cdot dt + \gamma(t)^{1/2} \cdot \theta \cdot \tau^{1/2} \cdot \Sigma(t)^{1/2} \cdot \sigma_F^{-1} \cdot dZ(t). \quad (32)$$

The first term corresponds to the information contained in informed signals which arrive at rate $\theta \cdot \gamma(t)$. The second term corresponds to the noise contained in all bets arriving at rate $\gamma(t)$. Here, $\gamma(t)^{1/2} \cdot dZ(t)$ is obtained by converting the mixture of $dZ_I(t)$ and $dZ_U(t)$ —with probabilities θ and $1 - \theta$, respectively—into a Brownian motion $dZ(t)$ with the same variance.

Define $\sigma(t)$ by

$$\sigma(t) := \theta \cdot \tau^{1/2} \cdot \Sigma(t)^{1/2} \cdot \gamma(t)^{1/2}. \quad (33)$$

By applying Ito's lemma to the definition of price in equation (16) and using equation (32), we obtain

$$\frac{dP(t)}{P(t)} = \frac{1}{2} \cdot [\Sigma'(t) - \sigma_F^2 + \sigma(t)^2] \cdot dt + \sigma_F \cdot d\bar{B}(t). \quad (34)$$

Market efficiency implies that $P(t)$ must follow a martingale. Its drift therefore must be zero, implying that $\Sigma(t)$ must satisfy the differential equation

$$\frac{d\Sigma(t)}{dt} = \sigma_F^2 - \sigma(t)^2. \quad (35)$$

Now define $d\bar{Z}(t) := \Sigma(t)^{-1} \cdot \sigma(t) \cdot \sigma_F \cdot [B(t) - \bar{B}(t)] \cdot dt + dZ(t)$. The process $\bar{Z}(t)$ is a standardized Brownian motion under the market's filtration, because $\bar{B}(t)$ is the market's best estimate of $B(t)$. Equation (32) can now be written in the simplified form $d\bar{B}(t) = \sigma(t) \cdot \sigma_F^{-1} \cdot d\bar{Z}(t)$. Using this fact and substituting $d\bar{Z}(t)$ for $d\bar{B}(t)$, the price process (34) can be written as the martingale process

$$\frac{dP(t)}{P(t)} = \sigma(t) \cdot d\bar{Z}(t). \quad (36)$$

Thus, $\sigma(t)$ measures returns volatility, consistent with our notation σ throughout this paper.

Returns volatility $\sigma(t)$ measures the rate at which new information is being incorporated into prices. According to equation (33), returns volatility $\sigma(t)$ is stochastic, even though fundamental volatility σ_F is constant. When returns volatility $\sigma^2(t)$ is greater than fundamental volatility σ_F^2 , equation (35) says that the difference represents the rate at which pricing error $\Sigma(t)$ is falling, since new information is being incorporated into prices faster than new fundamental uncertainty is unfolding. When returns volatility $\sigma^2(t)$ is smaller than fundamental volatility σ_F^2 , the difference represents the rate at which pricing error $\Sigma(t)$ is increasing, since new fundamental uncertainty is unfolding faster than information is being incorporated into prices.

When $\Sigma'(t) = 0$, we shall say that $\Sigma(t)$ has reached a “conditional steady state” in which the unfolding of new fundamental volatility and the incorporation of new private information into prices are in balance, i.e., $\sigma(t) = \sigma_F$. We use this terminology because the right-hand side of equation (35) does not converge to a constant due to the fact that the value of $\gamma(t)$, which follows a diffusion, is constantly changing and

this makes $\sigma(t)$ change as well. If $\gamma(t)$ were to remain constant for a long period of time, the value of $\Sigma(t)$ would converge to the conditional steady state given by

$$\Sigma(t) = \frac{\sigma_F^2}{\gamma(t) \cdot \theta^2 \cdot \tau}, \quad (37)$$

obtained from the definition of $\sigma(t)$ in equation (33) and equation (35) when $\Sigma'(t) = 0$. As changes in prices $P(t)$ lead to immediate changes in market capitalization and therefore changes in the arrival rate of bets $\gamma(t)$, the value of $\Sigma(t)$ gradually drifts towards a “conditional steady state” value.⁶ According to equation (37), more accurate signals and more frequent bets make market prices more accurate. The invariance theorem discussed next applies both in the conditional steady state and outside it.

Equations (32) and (33) imply that the difference $B(t) - \bar{B}(t)$ follows the mean-reverting process⁷

$$d[B(t) - \bar{B}(t)] = -\frac{\sigma(t)^2}{\Sigma(t)} \cdot [B(t) - \bar{B}(t)] \cdot dt + dB(t) - \frac{\sigma(t)}{\sigma_F} \cdot dZ(t). \quad (38)$$

As in section 1, let market resiliency $\rho(t)$ denote the mean reversion rate at which pricing errors disappear. We have

$$\rho(t) = \frac{\sigma(t)^2}{\Sigma(t)}. \quad (39)$$

Holding returns volatility constant, resiliency is greater in markets with higher pricing accuracy.

The endogenous quantities in the model are all functions of the two “state variables” $P(t)$ and $\Sigma(t)$, which evolve stochastically according to equations (35) and (36), with stochastic returns volatility given by equation (33). These equations have interesting dynamics due to the fact that business time operates at a faster pace relative to calendar time when market capitalization and therefore trading volume are higher. When trading volume is high, bets arrive quickly and $\Sigma(t)$ moves quickly towards its conditional steady state level; returns volatility remains close to fundamental volatility; $\Sigma(t)$ does not deviate far from its conditional steady state level. When trading volume is low, bets arrive slowly and $\Sigma(t)$ adjusts slowly towards its conditional steady state level; returns volatility may remain below fundamental volatility for extended periods of time.⁸

⁶While it is the *level* of $\gamma(t)$ that follows a diffusion, it is the *derivative* of $\Sigma(t)$ that follows a diffusion. Therefore, as $\gamma(t)$ changes, the value of $\Sigma(t)$ smoothly moves towards a conditional steady state value which it is constantly chasing.

⁷Black (1986) contains the intuition that because transitory noise affects prices, prices have twice the returns variance as fundamental value, and this is associated with mean reversion in returns. Equation (38) shows that, in a steady state, Black’s intuition applies to the log-ratio of prices to fundamental value, not to prices themselves (equation (36)), which have a martingale property. Black (1986) may have been confusing the properties of prices with the properties of the ratio of fundamental value to price.

⁸It can be shown that (with probability one) (1) a stock’s fundamental value will eventually

Solution of the Model with Invariance. The following invariance theorem states that all of the invariance hypotheses in section 1 are implied by this structural model and reveals specific connections among the hypotheses.

Invariance Theorem. *Assume the cost c_I of generating a signal is an invariant constant and let $m := E\{|\tilde{i}(t)|\}$ define an additional invariant constant. Then, the invariance conjectures hold: The dollar risk transferred by a bet per unit of business time $\tilde{I}(t)$ is a random variable with an invariant distribution and the expected cost of executing a bet $\bar{C}_B(t)$ is an invariant constant.⁹*

$$\tilde{I}(t) := \frac{P(t) \cdot \tilde{Q}(t) \cdot \sigma(t)}{\gamma(t)^{1/2}} = \frac{\tilde{Q}(t)}{V(t)} \cdot W(t)^{2/3} \cdot (m \cdot \bar{c}_B)^{1/3} = \bar{c}_B \cdot \tilde{i}(t). \quad (40)$$

$$\bar{C}_B(t) = \bar{c}_B := c_I \cdot \theta / (1 - \theta). \quad (41)$$

The number of bets $\gamma(t)$, their size $\tilde{Q}(t)$, liquidity $L(t)$, pricing accuracy $1/\Sigma(t)^{1/2}$, and market resiliency $\rho(t)$ are related to price $P(t)$, share volume $V(t)$, volatility $\sigma(t)$, and trading activity $W(t) = P(t) \cdot V(t) \cdot \sigma(t)$ by the following invariance relationships, which are consistent with the invariance relationships in equations (5), (6), (8), and (14):

$$\gamma(t) = \left(\frac{\lambda(t) \cdot V(t)}{\sigma(t) \cdot P(t) \cdot m} \right)^2 = \left(\frac{E\{|\tilde{Q}(t)|\}}{V(t)} \right)^{-1} = \frac{(\sigma(t) \cdot L(t))^2}{m^2} = \frac{\sigma(t)^2}{\theta^2 \cdot \tau \cdot \Sigma(t)} = \frac{\rho(t)}{\theta^2 \cdot \tau} = \left(\frac{W(t)}{m \cdot \bar{c}_B} \right)^{2/3}. \quad (42)$$

Here, τ is the precision of a signal, and θ is the fraction of information $\tilde{i}(t)$ incorporated by an informed trade. The price follows a martingale with stochastic returns volatility $\sigma(t) := \theta \cdot \tau^{1/2} \cdot \Sigma(t)^{1/2} \cdot \gamma(t)^{1/2}$.

Proof of Invariance Conjectures and Relationships. Using equations (25), (28), and (29), we obtain equation (41). The value $\bar{c}_B = \theta \cdot c_I / (1 - \theta)$ is constant across stocks, since c_I is constant across stocks by assumption and θ is constant across stocks by assumption (or, alternatively, by proof that $\theta = 1/2$ for risk-neutral informed traders).

Dividing the definition $\tilde{I}(t) := P(t) \cdot \tilde{Q}(t) \cdot \sigma(t) \cdot \gamma(t)^{-1/2}$ by the equation $\bar{C}_B(t) := \lambda(t) \cdot E\{\tilde{Q}(t)^2\}$, plugging in the definitions of $\tilde{Q}(t)$ and $\Delta\tilde{B}_I(t)$ from equations (18) and (20), and using definition (33), we obtain the third equality $\tilde{I}(t) = \bar{c}_B \cdot \tilde{i}(t)$ in equation (40); the first equality in equation (40) is the definition of $\tilde{I}(t)$ and the

become very small (since it follows a geometric Brownian motion which is also a martingale), (2) both the bet arrival rate and returns volatility will eventually converge to zero, (3) and $\Sigma(t)$ will eventually become unboundedly large. This is consistent with the interpretation that almost all stocks are eventually de-listed. As Keynes would say, in the long run, all companies are dead.

⁹In this theorem, “invariance” means that the values of the constants c_I , m , and $\bar{C}_B(t)$ as well as the distribution of $\tilde{I}(t)$ do not vary with time. In a model with different securities, they would not vary across securities either.

second equality (involving $W^{2/3}$) will follow from equation (42) below. Since $\tilde{i}(t)$ is invariant by assumption and \bar{c}_B is invariant by proof, $\tilde{I}(t) = \bar{c}_B \cdot \tilde{i}(t)$ has an invariant distribution. Note that both $\tilde{I}(t)$ and \bar{c}_B are measured in dollars while $\tilde{i}(t)$ represents unit-less information.

We write equation (23) for daily volume, equation (27) for expected costs, and equation (33) for returns volatility using equations (18) and (20) as a system of three equations:

$$\gamma(t) \cdot E\{|\tilde{Q}(t)|\} = V(t), \quad (43)$$

$$\bar{c}_B = \lambda(t) \cdot E\{\tilde{Q}^2(t)\}, \quad (44)$$

$$\gamma(t) \cdot \lambda(t)^2 \cdot E\{\tilde{Q}(t)^2\} = P(t)^2 \cdot \sigma(t)^2. \quad (45)$$

These three equations reflect the economic assumptions important for deriving invariance. Equation (43) says that observable trading volume results from bets. Equation (44) says that the expected price impact cost of a bet is \bar{c}_B (which is constant from equation (41)) assuming the bet has linear price impact costs. Equation (45) says that the price impact of a bet generates returns volatility, also under the assumption that price impact is linear. Note that the concepts of “price impact cost” and “price impact,” which both depend on the linear price impact parameter $\lambda(t)$ in the model, actually represent somewhat different concepts, which happen to be the same in a dealer market model in which markets are semi-strong efficient and dealers make zero profits based on linear signal processing. In the structural model, it is a derived result that price impact and transaction costs are linear.

In the three equations (43), (44), and (45), one can think of $\gamma(t)$, $\lambda(t)$, $E\{\tilde{Q}(t)^2\}$, and $E\{|\tilde{Q}(t)|\}$ as unknown variables to be solved for in terms of known variables $V(t)$, \bar{c}_B , $P(t)$, and $\sigma(t)$. Since there are three equations and four unknowns, we need a fourth equation. The fourth equation is the moment ratio

$$m = \frac{E\{|\tilde{Q}(t)|\}}{[E\{\tilde{Q}(t)^2\}]^{1/2}}. \quad (46)$$

Since $m := E\{|\tilde{i}|\}$, this equation follows from the fact that $\tilde{Q}(t)$ is a linear multiple of $\tilde{i}(t)$ and $\tilde{i}(t)$ has a variance of one. Since $\tilde{i}(t)$ is assumed to have a normal distribution with a variance of one, we have $m = (2/\pi)^{1/2}$. For a different distributional assumptions, m would have a different value. If we think of m as an exogenous parameter, we now have four equations in four unknowns.

Using the definition of m and the definition of trading activity $W(t) = P(t) \cdot V(t) \cdot \sigma(t)$, we can solve equations (43), (44), and (45) for $\gamma(t)$, $E\{|\tilde{Q}(t)|\}$, and $\lambda(t)$, as follows. Multiply the product of (43) and (44) by the square root of (45) and solve for $\gamma(t)$ to obtain

$$\gamma(t) = (m \cdot \bar{c}_B)^{-2/3} \cdot W(t)^{2/3}. \quad (47)$$

Divide the product of (45) and the square of (44) by (43) and solve for $E\{|\tilde{Q}(t)|\}$ using (46) to obtain

$$E\{|\tilde{Q}(t)|\} = (m \cdot \bar{c}_B)^{2/3} \cdot V(t) \cdot W(t)^{-2/3}. \quad (48)$$

Divide the product of (45) and the square root of (44) by (43) and solve for $\lambda(t)$ to obtain

$$\lambda(t) = \left(\frac{m^2}{\bar{c}_B}\right)^{1/3} \cdot \frac{1}{V(t)^2} \cdot W(t)^{4/3}. \quad (49)$$

Equation (48) and the definition of illiquidity $1/L(t) := \bar{c}_B/[P(t) \cdot V(t)]$ imply that $1/L$ satisfies

$$\frac{1}{L(t)} = \left(\frac{m^2}{\bar{c}_B}\right)^{-1/3} \cdot \sigma(t) \cdot W(t)^{-1/3}. \quad (50)$$

Equations (33), (39), and (47) imply that pricing accuracy $1/\Sigma(t)^{1/2}$ and resiliency $\rho(t)$ satisfy

$$\rho(t) = \frac{\sigma(t)^2}{\Sigma(t)} = \left(\frac{1}{m \cdot \bar{c}_B}\right)^{2/3} \cdot \theta^2 \cdot \tau \cdot W(t)^{2/3}. \quad (51)$$

Equations (47) for $\gamma(t)$, (48) for $\tilde{Q}(t)$, (50) for $1/L(t)$, and (51) for $1/\Sigma(t)^{1/2}$ and $\rho(t)$ are summarized in equation (42). They are respectively equivalent to equations (5), (6), (8), and (14) implied by the market microstructure invariance hypotheses in section 1, since $E\{|\tilde{i}|\} = m$ implies $E\{|\tilde{I}|\} = m \cdot \bar{c}_B$ from (40).

Since trading activity $W(t)$ and its components are observable, we can empirically infer values of $\gamma(t)$, $E\{|\tilde{Q}(t)|\}$, $1/L(t)$, $\lambda(t)$, $1/\Sigma(t)^{1/2}$ and $\rho(t)$ from equation (42) if the values of the three constants m , \bar{c}_B , and $\theta^2\tau$ are known.

Discussion. The model adds additional structure that imposes restrictions on the three invariance hypotheses outlined in section 1. These additional assumptions impose a particular structure on the proportionality constants in invariance relationships (5), (6), (9), and (14) and allow us to write these disconnected relationships in a consolidated form of the invariance theorem (42).

The level of trading activity $W(t)$ and its components—prices $P(t)$, share volume $V(t)$, and returns volatility $\sigma(t)$ —are “macroscopic” quantities whose value can be estimated from aggregate market data, e.g., from the CRSP dataset. The bet arrival rate $\gamma(t)$, the bet size $\tilde{Q}(t)$, the average cost of a bet $1/L(t)$, pricing accuracy $\Sigma(t)^{1/2}$, and resiliency $\rho(t)$ are granular “microscopic” quantities whose values are difficult to observe even with data on individual trades by individual traders. Since knowledge of the constants \bar{c}_B , m , and $\tau \cdot \theta^2$ makes it possible to infer microscopic quantities from macroscopic quantities using equation (42), these constants play roles in our structural model somewhat similar to the role played by Boltzmann’s constant or Avogadro’s number in physics.

For example, the structural model implies a particular relationship between the invariance of bet sizes and transaction costs. Equation (40) suggests that the bet size invariant \tilde{I} and transaction costs invariant \bar{c}_B satisfy the restriction $E|\tilde{I}| = m \cdot \bar{c}_B$ or, equivalently, that the standard deviation of \tilde{I} is equal to \bar{c}_B . This restriction follows from the assumption that market makers break even.

The structural model also implies a particular relationship between the invariance of pricing accuracy and the invariance of resiliency. In the model, the result $\rho(t) = \sigma(t)^2/\Sigma(t)$ implies a specific restriction on the proportionality constants in equations (14).¹⁰

In section 1, invariance relationships for bets and transaction costs were derived based on the assumption that the cost of executing a bet \bar{c}_B is constant across stocks. The structural model shows that \bar{c}_B is not the “deepest” structural parameter in the model. The result $\bar{c}_B = \theta \cdot c_I/(1 - \theta)$ implies that \bar{c}_B is constant across stocks if the “deeper” structural parameters c_I and θ are constant across stocks. It is useful to think of the cost of private information c_I as proportional to the average wages of finance professionals, adjusted for their productivity or effort required to generate one bet. The productivity-adjusted wage of a finance professional is therefore a “deeper” parameter than the endogenous cost of executing a bet \bar{c}_B . The invariance relationships in equation (42) result from finance professionals allocating their skills across different markets to maximize the value of trading on the signals they generate. The assumption that distinct bets result from distinct pieces of private information implies a particular level of granularity for both signals and bets.

It is interesting that invariance relationships relating the granularity of bets to their costs depend on the first absolute moment of the distribution of signals m , but not on the precision of signals τ .

The structural model also shows that the invariance of pricing accuracy and resiliency requires stronger assumptions: In addition to \bar{c}_B and m being constant, the informativeness of a bet $\tau \cdot \theta^2$, measured as the product of signal precision τ and the squared fraction of informed traders θ^2 , must be constant across time (or across stocks) as well.

Although the structural model is motivated by the time series properties of a single stock as its market capitalization changes, the model applies cross-sectionally across different securities under the assumption that the exogenously assumed cost of a private signal c_I , the shape of the distribution of signals m , and the informativeness of bets $\tau \cdot \theta^2$ are constant across all markets.

¹⁰This restriction suggests an empirical strategy for calibrating Black’s measure of market efficiency, which is difficult to observe directly. A value for Σ can be inferred indirectly from an estimate of ρ , which can be obtained by examining how fast price effects from noise trades die out over time (or from examining how long traders hold actively managed positions based on a more general model in which positions are liquidated after some fraction of their information content is revealed in prices). For example, if a stock’s annual volatility is about 35% and $\Sigma^{1/2} = \ln(2)/1.64$, then prices are about 1.50 years “behind” fundamental value, i.e., $\Sigma/\sigma^2 = (\ln(2)/1.64)^2/0.35^2 \approx 1.50$, as discussed in section 1. The error $B(t) - \bar{B}(t)$ in equation (38) mean-reverts to zero at rate $\rho = \sigma^2 \cdot \Sigma^{-1}$, or $\rho = 0.35^2/(\ln(2)/1.64)^2 = 0.69$ per year. This implies that the half-life of the price impact of a noise trade is $\ln(2)/\rho \approx 1$. Black (1986) could, therefore, have equivalently defined an efficient market as one in which the half-life of the price impact of noise trades is less than one year.

Robustness of Assumptions. Our structural model makes restrictive assumptions which lead to specific properties of equilibrium. Private signals are normally distributed. Bets are linear functions of private signals and are therefore also normally distributed. Price impact is linear in bet size. Informed traders place one bet in a dealer market setting rather than shredding bets into many pieces and executing them at an equilibrium speed over time. Dealers do not make profits; there are no bid-ask spread costs; there are no “effective bid-ask spread” costs related to immediate reversal of temporary price impact in executing bets.

The empirical results we are about to describe are not consistent with the “linear-normal” features of the structural model, but are consistent with the more general invariance hypotheses concerning bet size and transactions costs in section 1. In section 5, we find that the size of unsigned bets closely fits a log-normal distribution, not a normal distribution. In section 6, we find that a linear price impact model predicts transaction costs reasonably well, but a square root model predicts transaction costs better than a linear model.

We conjecture that it should be possible to modify our structural model to accommodate non-normally distributed bet size, non-linear price impact, and dynamic execution of bets at an equilibrium speed proportional to the rate at which business time unfolds.

Although the structural model assumes that fundamental volatility σ_F , shares outstanding N , and noise-trader turnover rate η are constants, it is also straightforward to modify the model so that these quantities vary over time or across stocks.

We provide a discussion of the invariance hypotheses in section 1 separately from the discussion of the structure model in this section because the invariance hypotheses are likely to hold more generally than under the somewhat restrictive assumptions of our structural model. The structural model is to be interpreted as a “proof of concept,” while the invariance hypotheses should apply more generally.

3 Microstructure Invariance in the Context of the Market Microstructure Literature

Market microstructure invariance builds a bridge from theoretical models of market microstructure to empirical tests of those models. Theoretical microstructure models usually suggest measures of liquidity based on the idea that order flow imbalances move prices. By scaling business time to be proportional to the rate at which bets arrive, market microstructure invariance imposes cross-sectional (or time-series) restrictions which make it easier to implement liquidity measures based on order flow imbalances.

Many theoretical models use game theory to model trading. These models typically make specific assumptions about the risk aversion of traders, 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 sides of trades. Some models emphasize adverse selection, such as Treynor (1995), Kyle (1985), Glosten and Milgrom (1985), and Back and Baruch (2004); some models emphasize inventory dynamics, such as Grossman and Miller (1988) and Campbell and Kyle (1993); some models emphasize both, such as Grossman and Stiglitz (1980) and Wang (1993).

While these theoretical models are all based on the idea that order flow imbalances move prices (with particular parameters depending on specifics of each model), it has proven difficult to infer from these models precise empirical implications. Theoretical models usually provide neither a unified framework for mapping the theoretical concept of an order flow imbalance into its empirical measurements nor precise predictions concerning how price impact varies across different stocks.

Instead, researchers have taken an approach based on *ad hoc* empirical intuition. For example, price changes can be regressed on imperfect empirical proxies for order flow imbalances—e.g., the difference between uptick and downtick volume, popularized by Lee and Ready (1991)—to obtain market impact coefficients, which can then be related to stock characteristics such as market capitalization, trading volume, and volatility. Breen, Hodrick and Korajczyk (2002) is an example of this approach. A voluminous empirical literature describes how the rate at which orders arrive in calendar time, the dollar size of orders, the market impact costs, and bid-ask spread costs vary across different assets. For example, Brennan and Subrahmanyam (1998) estimate order size as a function of various stock characteristics. Hasbrouck (2007) and Holden, Jacobsen and Subrahmanyam (2015) provide surveys of this empirical literature.

In contrast to this literature, microstructure invariance generates precise empirically testable predictions about how the size of bets, arrival rate of bets, market impact costs, and bid-ask spread costs vary across assets with different levels of trading activity. These predictions are consistent with intuition shared by many models. The unidentified parameters in theoretical models show up as invariant constants (e.g., $E\{|\tilde{I}|\}$ and \tilde{C}_B), which can be calibrated from data.

In this sense, microstructure invariance is a modeling principle applicable to different models, not a model itself. It compliments theoretical models by making it easier to test them empirically.

Applying Invariance to the Model of Kyle (1985). Consider, for example, the continuous-time theoretical model of Kyle (1985). The market depth formula $\lambda = \sigma_V / \sigma_U$ in that model measures market depth (in units of dollars per share per share) as the ratio of the standard deviation of stock price changes σ_V (measured in dollars per share per square root of units of time) to the standard deviation in order flow imbalances σ_U (measured in shares per square root of units of time). This formula asserts that price fluctuations result from the linear impact of order flow imbalances. The market depth formula itself does not depend on specific assumptions

about interactions among market makers, informed traders, and noise traders. An empirical implementation of the market impact formula $\lambda = \sigma_V/\sigma_U$ should *not* be considered a test of the specific assumptions of the model of Kyle (1985), such as the existence of a monopolistic informed trader who trades smoothly and patiently in a context where less patient liquidity traders trade more aggressively and market makers set stock prices efficiently. Instead, empirical implementation of the formula $\lambda = \sigma_V/\sigma_U$ attempts the more general task of measuring a market impact coefficient λ based on the assumption that price fluctuations result from the linear impact of order flow innovations, a property shared by many models, including the structural model in section 2.

Measuring the numerator σ_V is much more straightforward than measuring the denominator σ_U . The value of σ_V is easily inferred from a stock price and returns volatility. In the context of our structural model (where $\sigma = \bar{\sigma}$), we have $\sigma_V = \bar{\sigma} \cdot P$.

Measuring the denominator σ_U is difficult because the connection between observed trading volume and order flow imbalances is not straightforward. Intuitively, σ_U should be related to trading volume in some way. The continuous-time model provides no help concerning what this relationship is; in the Brownian motion model of Kyle (1985), trading volume is infinite. Without some approach for measuring σ_U , the model is untestable. In the context of our structural model, order flow imbalances result from random discrete decisions by traders to change stock holdings so that the standard deviation of order imbalances is given by $\sigma_U = \gamma^{1/2} \cdot [E\{\tilde{Q}(t)^2\}]^{1/2}$. This calculation is also consistent with the spirit of other models, such as Glosten and Milgrom (1985) and Back and Baruch (2004).

The formulas for the numerator and denominator imply that the price impact of a bet of X shares, expressed as a fraction of the value of a share P , is given by

$$\frac{\lambda}{P^2} \cdot (X \cdot P) = \frac{\sigma_V}{\sigma_U} \cdot \frac{X}{P} = \bar{\sigma} \gamma^{-1/2} \cdot \frac{X}{[E\{\tilde{Q}(t)^2\}]^{1/2}}. \quad (52)$$

Thus, a one-standard deviation bet $X = [E\{\tilde{Q}(t)^2\}]^{1/2}$ has a price impact $\bar{\sigma} \gamma^{-1/2}$ equal to one standard deviation of returns volatility $\bar{\sigma}$ measured over a time interval of length $1/\gamma$, i.e., the expected time between bet arrivals. Empirical tests of this formula require assumptions about \tilde{Q} and γ , which decompose the order flow so that the standard deviation of order flow imbalances can be calculated.

This formula can be tested empirically using restrictions imposed by microstructure invariance. Using equations (5) and (6) to determine how γ and moments of \tilde{Q} vary with observable volume and volatility, we find that the price impact cost of an order of dollar size $X \cdot P$, as a fraction of the value traded, is

$$\frac{\lambda}{P^2} \cdot (X \cdot P) := \frac{\bar{\sigma}}{P \cdot \gamma^{1/2} \cdot (E\tilde{Q}^2)^{1/2}} \cdot (X \cdot P) = \frac{E\{|\tilde{I}|\}^{2/3}}{[E\{\tilde{I}^2\}]^{1/2}} \cdot \frac{\bar{\sigma}}{P \cdot \bar{V}} \cdot \bar{W}^{1/3} \cdot (X \cdot P). \quad (53)$$

The percentage price impact is proportional to $\bar{W}^{1/3} \cdot \bar{\sigma}/(P \cdot \bar{V})$, which itself is proportional to the illiquidity measure $1/L^2$. As a consequence of the hypothesis, which

implies that the size and number of bets vary with the $1/3$ and $2/3$ exponents of trading activity, the distribution of \tilde{I} invariant. Since this makes the proportionality constant $[E\{\tilde{I}^2\}]^{-1/2} \cdot E\{|\tilde{I}|\}^{2/3}$ invariant as well, implementation of the market impact formula (53) requires calibration of only one proportionality constant for all stocks and all time periods. Note that this constant does not depend on the units of time in which variables are measured, because \tilde{I} itself does not depend on time units.

As an alternative to invariance, the formula $\lambda = \sigma_V/\sigma_U$ can be implemented empirically by imposing different assumptions concerning the connection between σ_U and trading volume. For example, we can assume that the expected arrival rate of bets is some unknown constant, the same for all stocks; this will imply that σ_U is proportional to volume V and illiquidity measure in formula (52) is proportional to $\sigma/(P \cdot V)$. We obtain

$$\frac{\lambda}{P^2} \cdot (X \cdot P) := \frac{\bar{\sigma}}{P \cdot \gamma^{1/2} \cdot (E\tilde{Q}^2)^{1/2}} \cdot (X \cdot P) \propto \frac{\bar{\sigma}}{P \cdot \bar{V}} \cdot (X \cdot P). \quad (54)$$

This empirical implementation of the formula can be thought of as the illiquidity ratio in Amihud (2002). Indeed, Amihud’s illiquidity ratio is the time-series average of the daily ratios of the absolute value of percentage returns to dollar volume. To the extent that dollar volume is relatively stable across time and returns are drawn from the same distribution, illiquidity ratio is effectively proportional to $\sigma/(P \cdot V)$. Although this is a logically consistent way to connect theory with empirical implementation, it is unrealistic to assume that the most actively traded and least actively traded stocks have the same number of bets per day; empirical intuition suggests that stocks with high levels of trading activity have more bets per day than stocks with low levels of trading activity. We are aware of no empirical studies which claim that the number of orders or bets in different stocks is the same.

The same issue can be addressed by thinking about time units. Unlike our illiquidity measure $1/L = \iota^2 \bar{C}_B \cdot [P \cdot \bar{V}/\bar{\sigma}^2]^{-1/3}$, the Amihud ratio $\sigma/(P \cdot V)$ has time units: σ^2 and $P \cdot V$ have the same time units, but σ and $P \cdot V$ do not; Amihud’s ratio thus depends on the time horizon at which volume and volatility are measured. To keep the left-hand side consistent with the right-hand side of equation (54), the proportionality constant in that equation must change when time units are changed. Furthermore, if the invariance-implied market impact formula (53) is correct, then Amihud’s market impact formula (54) theoretically implies a different proportionality constant for every stock. This problem can be “fixed”—i.e., the same proportionality coefficient can be obtained for every stock using Amihud’s approach—if data for each stock is sampled at a different stock-specific frequency appropriate to the stock’s level of trading activity. Invariance implies that the appropriate sampling frequency should be proportional to $1/\gamma$, which is proportional to $W^{-2/3}$.

Illiquidity ratios calculated using data sampled at the same calendar time frequencies, as proposed originally and implemented in many empirical studies, implicitly rely on the unrealistic assumption that the standard deviation of order imbalances is pro-

portional to trading volume. In contrast, our illiquidity measure $1/L$ does not depend on time units, and therefore it does not matter over what time horizons its components are measured; even if different horizons are used for different stocks, its value will be the same.

4 Data

Portfolio Transitions Data. We test the empirical implications of market microstructure invariance using a proprietary dataset 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 2,552 portfolio transitions executed over the period 2001-2005 for U.S. clients. A portfolio transition may involve orders for hundreds of individual stocks. Each order is a stock-transition pair potentially executed over multiple days using a combination of internal crosses, external crosses, and open-market transactions.

The portfolio transitions dataset contains fields identifying the portfolio transition; its starting and ending dates; the stock traded; the trade date; the number of shares traded; a buy or sell indicator; the average execution price; the pre-transition benchmark price (closing price the day before the transition trades began); commissions; SEC fees; and a trading venue indicator distinguishing among internal crossing networks, external crossing networks, open market transactions, and in-kind transfers.

When old “legacy” and new “target” portfolios overlap, positions are transferred from the legacy to the new portfolio as “in-kind” transfers. For example, if the legacy portfolio holds 10,000 shares of IBM stock and the new portfolio holds 4,000 shares of IBM, then 4,000 shares are transferred in-kind and the balance of 6,000 shares is sold. The in-kind transfers do not incur transactions costs and have no effect on our empirical analysis. The 6,000 shares sold constitute one “portfolio transition order,” even if the 6,000 shares are sold over multiple days.

We augment the portfolio transitions data with stock price, returns, and volume data from CRSP. 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 from January 2001 through December 2005 are included in the sample. ADRs, REITs, and closed-end funds are excluded. Also excluded are stocks with missing CRSP information necessary to construct variables used for empirical tests, transition orders in high-priced Berkshire Hathaway class A shares, and transition observations which appeared to contain typographical errors and obvious inaccuracies. Since it is unclear from the data whether adjustments for dividends and stock splits are 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, there are 439,765 observations (“orders”), including 201,401 buy orders and 238,364 sell orders.

CRSP Data: Prices, Volume, and Volatility. For each of the transition-stock observations ($i = 1, \dots, 439765$), we collect data on pre-transition benchmark price, expected volume, and expected volatility. The benchmark price, denoted $P_{0,i}$, is the closing price for the stock the evening before the first trade is made in any of the stocks in the portfolio transition. A proxy for expected daily trading volume, denoted V_i (in shares), is the average daily trading volume for the stock in the previous full pre-transition calendar month.

The expected volatility of daily returns, denoted σ_i for order i , is calculated using past daily returns in two different ways.

First, for each security j and each calendar month m , we estimate the monthly standard deviation of returns $\sigma_{j,m}$ as the square root of the sum of squared daily returns for the full calendar month m (without de-meaning or adjusting for autocorrelation). We define $\sigma_i = \sigma_{j,m}/N_m^{1/2}$, where j corresponds to the stock traded in order i , m is the previous full calendar month preceding order i , and N_m is the number of CRSP trading days in month m .

Second, to reduce effects from the positive skewness of the standard deviation estimates, we estimate for each stock j a third-order moving average process for the changes in $\ln(\sigma_{j,m})$ for all months m over the entire period 2001-2005: $(1 - L)\ln(\sigma_{j,m}) = \Theta_{j,0} + (1 - \Theta_{j,1}L - \Theta_{j,2}L^2 - \Theta_{j,3}L^3)u_{j,m}$. Letting $y_{j,m}$ denote the estimate of $\ln(\sigma_{j,m})$ and \hat{V}_j the variance of the prediction error, we alternatively define the conditional forecast for the volatility of daily returns by $\sigma_i = \exp(y_{j,m} + \hat{V}_j/2)/N_m^{1/2}$, where m is the *current* full calendar month for order i .

These volatility estimates can be thought of as instrumental variables for true expected volatility. While below we report results using the second definition of σ_i based on the log-ARIMA model, these results remain quantitatively similar when we use the first definition of σ_i based on simple historical volatility during the preceding full calendar month.

Except to the extent that the ARIMA model uses in-sample data to estimate model parameters, we use the pre-transition variables known to the market before portfolio transition trades are executed in order to avoid any spurious effects from using contemporaneous variables.

Descriptive Statistics. Table 1 reports descriptive statistics for traded securities in panel A and for individual transition orders in panel B. The first column reports statistics for all securities in aggregate; the remaining ten columns report statistics for stocks in ten dollar volume groups. Instead of dividing the securities into ten deciles with the same number of securities in each decile, 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 of dollar trading volume. Group 10 approximately corresponds to the universe of S&P 100 stocks. The top five groups approximately cover the universe of S&P 500 stocks. Narrower percentile bands 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 descriptive statistics for traded securities. For the entire sample, the median daily volume is \$18.72 million, ranging from \$1.13 million for the lowest volume group to \$212.85 million for the highest volume group. The median volatility is 1.93 percent per day, ranging from 1.76 percent in the highest volume decile to 2.16 in the lowest decile. Since there is so much more cross-sectional variation in dollar volume than in volatility across stocks, the variation in trading activity across stocks is related mostly to variation in dollar volume. Trading activity differs by a factor of 150 between stocks in the lowest group and stocks in the highest group, and this variation creates statistical power helpful in determining how transactions costs and order sizes vary with trading activity.

The median quoted bid-ask spread, obtained from the transition dataset, is 12.04 basis points; its mean is 25.42 basis points. From lowest volume group to highest volume group, the median spread declines monotonically from 40.96 to 4.83 basis points, by a factor of 8.48. A back-of-the-envelope calculation based on invariance suggests that spreads should decrease approximately by a factor of $150^{1/3} \approx 5.31$ from lowest to highest volume group. The difference between 5.31 and 8.48 is partially explained by differences in returns volatility across the volume groups and warrants further investigation. The monotonic decline of almost one order of magnitude is potentially large enough to generate significant statistical power in estimates of a bid-ask spread component of transactions costs based on implementation shortfall.

Panel B of table 1 reports properties of portfolio transition order sizes. The average order size is 4.20% of average daily volume, declining monotonically across the ten volume groups from 16.23% in the smallest group to 0.49% in the largest group, by a factor of 33.12. The median order is 0.57% of average daily volume, also declining monotonically from 3.33% in the smallest group to 0.14% in the largest group, by a factor of 23.79. The invariance hypothesis implies that order sizes should decline by a factor of approximately $150^{2/3} \approx 28.23$, a value which matches the data closely. The medians are much smaller than the means, indicating that distributions of order sizes are skewed to the right. We show below that the distribution of order sizes closely fits a log-normal.

The average trading cost (estimated based on implementation shortfall, as explained below) is 16.79 basis points per order, ranging from 44.95 basis points in the lowest volume group to 6.16 basis points in the highest volume group. Invariance suggests that these costs should fall by a factor of $150^{1/3} \approx 5.31$, somewhat smaller

than the actual decline. The cost estimates exclude commissions and SEC fees.¹¹

One portfolio transition typically contains orders for dozens or hundreds of stocks; it typically takes several days to execute all of the orders. About 60% of orders are executed during the first day of a portfolio transition. Since transition managers often operate under a “cash-in-advance” constraint—using proceeds from selling stocks in a legacy portfolio to acquire stocks in a target portfolio—sell orders tend to be executed slightly faster than buy orders (1.72 days versus 1.85 days). In terms of dollar volume, about 41%, 23%, 15%, 7%, and 5% of dollar volume is executed on the first day through the fifth days respectively. The two longest transitions in the sample were executed over 18 and 19 business days. The time frame for a portfolio transition is usually set before its actual implementation begins.

5 Empirical Tests Based on Order Sizes

Market microstructure invariance predicts that the distribution of $\bar{W}^{2/3} \cdot \tilde{Q}/\bar{V}$ does not vary across stocks or time (see equation (6)). We test these predictions using data on portfolio transition orders, making the identifying assumption that portfolio transition orders are proportional to bets.

Portfolio Transitions and Bets. Since bets are statistically independent intended orders, bets can be conceptually difficult for researchers to observe. Consider, for example, a trader who makes a decision on Monday to make one bet to buy 100,000 shares of stock, then implements the bet by purchasing 20,000 shares on Monday and 80,000 shares on Thursday. To an econometrician, this one bet for 100,000 shares may be difficult to distinguish from two bets for 20,000 shares and 80,000 shares respectively. In the context of a portfolio transition, identifying a bet is easier because the size of the order for 100,000 shares is known and recorded on Monday, even if the order is executed over several subsequent days.

Portfolio transition orders may not have a size distribution matching precisely the size distribution of typical bets. Transition orders may be smaller than bets if transitions tend to liquidate a portion of an asset manager’s positions or larger than bets if transitions liquidate the sum of bets made by the asset manager in the past. When both target and legacy portfolios hold long positions in the same stock, the portfolio transition order may represent the difference between two bets.

Let X_i denote the unsigned number of shares transacted in portfolio transition order i , $i = 1, \dots, 439765$. The quantity X_i sums shares traded over multiple days,

¹¹The SEC fee represents a cost of about 0.29 basis points and does not vary much across volume groups. From lowest to highest volume group, these costs fall by a factor of 7.30. The average commission is 7.43 basis points, declining monotonically from 14.90 basis points for the lowest group to 2.68 basis points for the highest group. Since commissions may be negotiated for the entire transition, the allocation of commission costs to individual stocks is an accounting exercise with little economic meaning.

excluding in-kind transfers.

We make the identifying assumption that, for some constant δ which does not vary across stocks with different characteristics such as trading activity, the distribution of scaled portfolio transition orders $\delta \cdot X_i$ is the same as the distribution of unsigned bets in the same stock at the same time, denoted $|\tilde{Q}|$. If $\delta = 1$, the distribution of portfolio transition orders matches the distribution of bets.

The Empirical Hypotheses of Invariance and Log-Normality for the Size Distribution of Portfolio Transition Orders. Let $W_i := V_i \cdot P_i \cdot \sigma_i$ and $\bar{W}_i := \bar{V}_i \cdot P_i \cdot \bar{\sigma}_i$ denote trading activity and betting activity, respectively, for the stock in transition order i . Under the identifying assumption that portfolio transition orders are proportional to bets, invariance of bets implies invariance of portfolio transition orders. Specifically, replacing \tilde{Q} with X_i in equation (6) implies that the distribution of $\bar{W}_i^{2/3} \cdot X_i / \bar{V}_i$ does not vary with stock characteristics such as volume, volatility, stock price, or market capitalization.

To facilitate intuitive interpretation of parameter estimates, we scale observations by a hypothetical “benchmark stock” with share price P^* of \$40, daily volume V^* of one million shares, and volatility σ^* of 2% per day, implying $W^* = 40 \cdot 10^6 \cdot 0.02$. This benchmark stock would belong to the bottom tercile of S&P 500 (volume group 7 in table 1).

Combining invariance of portfolio transition orders with equations (1) and (2) to convert the betting activity variables \bar{W}_i and \bar{V}_i into trading activity variables W_i and V_i and taking logs, invariance implies the empirically testable relationship

$$\ln \left(\left[\frac{W_i}{W^*} \right]^{2/3} \cdot \frac{X_i}{V_i} \right) = \ln(\bar{q}) + \tilde{\epsilon}_i. \quad (55)$$

Under the identifying assumptions that the volume multiplier ζ , the volatility multiplier ψ , and the deflator δ do not vary across observations, $\ln(\bar{q})$ is an invariant constant $\ln(\bar{q}) = E\{\ln(|\tilde{Q}^*|/V^*)\} - \ln(\delta)$ and $\tilde{\epsilon}_i$ is a zero-mean error with the same invariant distribution as $\ln(|\tilde{I}|) - E\{\ln(|\tilde{I}|)\}$.¹² Adjustment by W^* in equation (55) scales each observation on the left side so that it has the same invariant distribution as the log of a hypothetical portfolio transition order in the benchmark stock, expressed as a fraction of its expected daily volume.

We will also examine the stronger log-normality hypothesis—not implied by microstructure invariance—that the distribution of unsigned order sizes adjusted for trading activity $[W_i/W^*]^{2/3} \cdot X_i/V_i$ has a log-normal distribution, i.e., $\tilde{\epsilon}_i$ in equation (55) has a normal distribution. The log-normality hypothesis implies that the right-side of equation (55) is characterized by two invariant constants, the mean $\ln(\bar{q})$ and the variance of $\tilde{\epsilon}_i$.

Next, we implement several tests to examine this hypothesis.

¹²More generally, $\ln(\bar{q}) := E\{\ln(|\tilde{Q}^*|/V^*)\} - 1/3 \ln(\zeta_i/\zeta^*) - 2/3 \ln(\psi_i/\psi^*) - \ln(\delta_i)$.

The Graphical Relationship Between Order Sizes and Trading Activity.

One way to examine the invariance hypothesis is to plot the log of order size as a fraction of average daily volume $\ln[X_i/V_i]$ against the log of scaled trading activity $\ln[W_i/W^*]$. Figure 4 presents a cloud of points for all 400,000+ portfolio transition orders. The line $\ln[X_i/V_i] = -5.71 - 2/3 \cdot \ln[W_i/W^*]$ is also shown for comparison. The slope of this line is fixed at $-2/3$, as implied by invariance; the intercept, estimated from an OLS regression, is the sample mean. On the horizontal axis, zero represents the log of trading activity in the benchmark stock; on the vertical axis, zero represents orders for 100% of expected daily volume. The shape of the “super-cloud” conforms well with the invariance hypothesis in that the slope of $-2/3$ is close to the shape of the plotted points and there is only little evidence of heteroscedasticity.

Log-Normal Order Size Distribution for Volume and Volatility Groups.

When the portfolio transition orders are sorted into different groups based on characteristics such as dollar volume, volatility, stock price, and turnover, the joint hypotheses of invariance and log-normality imply that the means and variances of $\ln[[W_i/W^*]^{2/3} \cdot X_i/V_i]$ for each group should match the mean and variance of the pooled sample. The pooled sample mean of $\ln[[W_i/W^*]^{2/3} \cdot X_i/V_i]$ is -5.71 ; the pooled sample variance is 2.53 . The pooled sample mean is an estimate of $\ln(\bar{q})$; the pooled sample variance is an estimate of the variance of the error $\tilde{\epsilon}_i$.¹³

To examine this hypothesis visually, we plot the empirical distributions of the left-hand side of equation (55), $\ln[[W_i/W^*]^{2/3} \cdot X_i/V_i]$, for selected volume and volatility groups. As before, we define ten dollar volume groups with thresholds corresponding to the 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of NYSE dollar volume. We define five volatility groups with thresholds corresponding to the 20th, 40th, 60th, and 80th percentiles of returns standard deviation for NYSE stocks. On each plot, we superimpose the bell-shaped density function $N(-5.71, 2.53)$ matching the mean and variance of the pooled sample.

Figure 2 shows plots of the empirical distributions of $\ln[[W_i/W^*]^{2/3} \cdot X_i/V_i]$ for volume groups 1, 4, 7, 9, and 10 and for volatility groups 1, 3, and 5. Consistent with the invariance hypothesis, these fifteen distributions of W -adjusted order sizes are all visually strikingly similar to the superimposed normal distribution. Results for the remaining 35 subgroups also look very similar and therefore are not presented in this paper. The visual similarity of the distributions is reflected in the similarity of their first four moments. For the 15 volume-volatility groups, the means range from -6.03 to -5.41 , close to the mean of -5.71 for the pooled sample. The variances range from 2.23 to 2.90 , also close to the variance of 2.53 for the pooled sample. The skewness ranges from -0.21 to 0.10 , close to skewness of zero for the normal distribution. The kurtosis ranges from 2.73 to 3.38 , also close to the kurtosis of 3 for a normal random variable. These results suggest that it is reasonable to assume that unsigned order

¹³There is not much difference in the distributions of buy and sell orders. For buy orders, the mean is -5.70 and the variance is 2.51 ; for sell orders, the mean is -5.71 and the variance is 2.55 .

sizes have a log-normal distribution. Scaling order sizes by $[W/W^*]^{2/3}$, as implied by the invariance hypothesis, adjusts the means of the distributions so that they visually appear to be similar.

Despite the visual similarity, a Kolmogorov-Smirnov test rejects the hypothesis that all fifty empirical distributions are generated from the same normal distribution. The standard deviation of the means across bins is larger than implied by a common normal distribution. Microstructure invariance does not describe the data perfectly, but it makes a good benchmark from which the modest deviations seen in these plots can be investigated in future research.

Figure 3 further examines log-normality by focusing on the tails of the distributions of portfolio transition orders. For each of the five volume groups 1, 4, 7, 9, and 10, panel A shows quantile-quantile plots of the empirical distribution of $\ln[[W_i/W^*]^{2/3} \cdot X_i/V_i]$ versus a normal distribution with the same mean and variance. The more similar these empirical distributions are to a normal distribution, the closer the plots should be to the 45-degree line. Panel B shows logs of ranks based on scaled order sizes. Under the hypothesis of log-normality, the right tail should be quadratic. A straight line in the right tail implies a power law. Both panels show that the empirical distributions are similar to a normal distribution, except in the far right and left tails.

In panel A, the smallest orders in the left tails tend to be smaller than implied by a normal distribution. These observations are economically insignificant. Most of them represent one-share transactions in low-price stocks (perhaps the result of coding errors in the data). There are too few such orders to have a meaningful effect on our statistical results.

In panel A, the largest orders in the right tails are much more important economically. On each subplot, a handful of positive outliers (out of 400,000+ observations) do not appear to fit a normal distribution. The largest orders in low-volume stocks appear to be smaller than implied by a normal distribution, and the largest orders in high-volume stocks appear to be larger than implied by a normal distribution.

The finding that the largest orders in low-volume stocks are smaller than implied by a log-normal may be explained by reporting requirements. When an owner acquires more than 5% of the shares of a publicly traded company, the SEC requires information to be reported on Schedule 13D. To avoid reporting requirements, large institutional investors may intentionally acquire fewer shares when intended holdings would otherwise exceed the 5% reporting threshold. Indeed, all 400,000+ portfolio transition orders are for amounts smaller than 4.5% of shares outstanding. A closer examination reveals that the five largest orders for low-volume stocks accounts for about 2%, 3%, 4%, 4%, and 4% of shares outstanding, respectively, just below the 5% threshold. The largest order in high-volume stocks is for about 1% of shares outstanding.

To summarize, we conclude that the distribution of portfolio transition order sizes appears to conform closely to—but not exactly to—the invariance hypothesis.

Furthermore, the distribution of order sizes appears to be similar to—but not exactly equal to—a log-normal.

OLS Estimates of Order Size. The order size predictions from equation (55) can also be tested using a simple log-linear OLS regression

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + \tilde{\epsilon}_i. \quad (56)$$

Invariance of bets implies $\alpha_0 = -2/3$.

To adjust standard errors of OLS estimates of α_0 for positive contemporaneous correlation in transition order sizes across different stocks, the 439,765 observations are pooled by week over the 2001-2005 period into 4,389 clusters across 17 industry categories. The double clustering by weeks and industries conservatively adjusts standard errors for large portfolio transitions that may involve hundreds of relatively large orders, executed during the course of a week and potentially concentrated in particular industries.¹⁴

Table 2 presents estimates for the OLS coefficients in equation (56). 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 OLS regressions in which the parameters are estimated separately for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells.

For the entire sample, the estimate for α_0 is $\hat{\alpha}_0 = -0.62$ with standard error of 0.009. Economically, the point estimate for α_0 is close to the value $-2/3$ predicted by the invariance hypothesis, but the hypothesis $\alpha_0 = -2/3$ is strongly rejected ($F = 25.31, p < 0.0001$) because the standard error is very small.

When the sample is broken down into NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells, it is interesting to note that the estimated coefficients for buy orders, -0.63 for NYSE and -0.71 for NASDAQ, are closer to $-2/3$ than the coefficients for sell orders, -0.59 for both NYSE and NASDAQ. Since portfolio transitions tend to be applied to long-only portfolios, sell orders tend to represent liquidations of past bets. If the size distribution of sell orders depends on past values of volume and volatility—not current values—there is an errors-in-variables problem related to past trading activity being used as a noisy version of current trading activity. This will bias the absolute values of coefficient estimates downwards, consistent with the absolute values of the coefficient estimates for NYSE and NASDAQ sell orders being less than $2/3$.

¹⁴A potential econometric issue with the log-linear specification in equation (56) is that taking the log of order size as a fraction of average daily volume may create large negative outliers from tiny, economically meaningless orders, with an inordinately large influence on coefficient estimates. Since we have shown above that the shape of the distribution of scaled order sizes closely matches a log-normal, these tiny orders are expected to have only a negligible distorting effect on estimates.

Quantile Estimates of Order Sizes. Table 7 in the Appendix presents quantile regression results for equation (56) based on the 1st (smallest orders), 5th, 25th, 50th, 75th, 95th, and 99th percentiles (largest orders). The corresponding quantile estimates for α_0 are -0.65 , -0.64 , -0.61 , -0.62 , -0.61 , -0.64 , and -0.63 , respectively. Although the hypothesis $\alpha_0 = -2/3$ is rejected due to small standard errors, all quantile estimates are economically close to the value of $-2/3$ predicted by the invariance hypothesis.

Model Calibration and Its Economic Interpretation. Under the invariance and log-normality hypotheses, we can calibrate the distribution of bet sizes by imposing the restriction $\alpha_0 = -2/3$ on equation (56). Thus, only the constant term in the regression needs to be estimated.

The results of this calibration exercise are presented in table 3. The estimated constant term, -5.71 , is the previously reported sample mean of $\ln(\bar{q})$ in equation (55). The mean-square error, 2.53 , is the previously reported sample variance of $\tilde{\epsilon}_i$ in equation (55).

The R^2 (with zero degrees of freedom) is 0.3149 ; the log of trading activity $\ln(W/W^*)$, with the coefficient $\alpha_0 = -2/3$ imposed by invariance, explains a significance percentage of the variation of order size as a fraction of volume X_i/V_i .

When the parameter α_0 is estimated rather than held fixed, changing α_0 from the predicted value of $\alpha_0 = -2/3$ to the estimated value of $\hat{\alpha}_0 = -0.62$ increases the R^2 from 0.3149 (table 3) to 0.3167 (table 2), a modest increase of 0.0018 . Although statistically significant, the addition of one degree of freedom does not add much explanatory power.

We relax the specification further by allowing the coefficients on the three components of trading activity—volatility σ_i , price $P_{0,i}$, and volume V_i —as well as monthly turnover rate ν_i to vary freely:

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + b_1 \cdot \ln \left[\frac{\sigma_i}{0.02} \right] + b_2 \cdot \ln \left[\frac{P_{0,i}}{40} \right] + b_3 \cdot \ln \left[\frac{V_i}{10^6} \right] + b_4 \cdot \ln \left[\frac{\nu_i}{1/12} \right] + \tilde{\epsilon}_i. \quad (57)$$

This regression imposes on $\ln[W_i/W^*]$ the coefficient $\alpha_0 = -2/3$ predicted by invariance and then allows the coefficients b_1 , b_2 , b_3 , b_4 on the three components of W_i and turnover rate to vary freely. The invariance hypothesis implies $b_1 = b_2 = b_3 = b_4 = 0$. Table 3 reports that increasing the degrees of freedom from one to four increases the R^2 of the regression from 0.3167 to 0.3229 , an increase of 0.0062 . Although again statistically significant, the improvement in R^2 is again modest. Invariance explains much—but not quite all—of the variation in portfolio transition order size across stocks that can be explained by all four variables.

The point estimates for the coefficient on volatility of $\hat{b}_1 = 0.42$, the coefficient on price of $\hat{b}_2 = 0.24$, the coefficient on share volume of $\hat{b}_3 = 0.06$, the coefficient on turnover rate of $\hat{b}_4 = -0.18$ are all statistically significant, with standard errors of

0.040, 0.019, 0.010, and 0.015, respectively (see table 8 in the Appendix). The coefficients on volatility and price are significantly positive, indicating that order size—as a fraction of average daily volume—does not decrease with increasing volatility and price as fast as predicted by the invariance hypothesis. The statistically significant positive coefficient on volume may be partially offset by a statistically significant negative coefficient on turnover rate.

Discussion. The documented log-normality of bet size is strikingly different from the typical assumptions of microstructure models, where innovations in order flow from noise traders are distributed as a normal, not a log-normal or power law. Although normal random variables are a convenient modeling device—they allow conditional expectations to be linear functions of underlying jointly normally distributed variables—their implications are qualitatively very different.

The estimated log-mean of -5.71 implies that a median portfolio transition order size is equal to 0.33% of expected daily volume for the benchmark stock, since $\exp(-5.71) \approx 0.0033$. The estimated log-variance of 2.53 implies that a one standard deviation increase in order size is a factor of 4.90 for all stocks, since $\exp(2.53^{1/2}) \approx 4.90$.

We next explain why the log-variance of 2.53 also implies that a large fraction of trading volume and an even larger fraction of returns variance come from large bets.

Let $\eta(z)$ and $N(z)$ denote the PDF and CDF, respectively, of a standardized normal distribution. Define $F(\bar{z}, p)$ by $F(\bar{z}, p) := \int_{z=\bar{z}}^{\infty} \exp(p \cdot \sqrt{2.53} \cdot z) \cdot \eta(z) \cdot dz$. It is easy to show that $F(\bar{z}, p) = \exp(p^2 \cdot 2.53/2) \cdot (1 - N(\bar{z} - p \cdot \sqrt{2.53}))$. This implies that the fraction of the p th moment of order size arising from bets greater than \bar{z} standard deviations above the log-mean is given by $F(\bar{z}, p)/F(-\infty, p) = 1 - N(\bar{z} - p \cdot \sqrt{2.53})$.

Plugging $p = 1$, we find that bets larger than \bar{z} standard deviations above the log-mean (median) generate a fraction of total trading volume given by $1 - N(\bar{z} - \sqrt{2.53})$. Bets larger than the 50th percentile generate 94.41% of trading volume ($\bar{z} = 0$). Bets larger than $\sqrt{2.53}$ standard deviations above the log-mean (median) bet size—i.e., the largest 5.39% of bets—generate 50% of trading volume ($\bar{z} = \sqrt{2.53}$).

Plugging $p = 2$, we find that bets larger than \bar{z} standard deviations above the log-mean bet size contribute a fraction of total returns variance given by $1 - N(\bar{z} - 2 \cdot \sqrt{2.53})$ under the assumption that the contribution of bets to price variance is proportional to their squared size (implied by linear price impact). Bets greater than the 50th percentile generate 99.93% of returns variance ($\bar{z} = 0$). Bets larger than $2 \cdot \sqrt{2.53} = 3.18$ standard deviations above the log-mean—i.e., the largest 0.07% of bets—generate 50% of returns variance ($\bar{z} = 2 \cdot \sqrt{2.53}$).

Under the assumptions $\zeta/2 = \psi = \delta = 1$ (stronger than our identifying assumptions), the estimates of mean and variance imply that the benchmark stock has about 85 bets per day for each of the 252 trading days in a calendar year. These estimates then imply that the 1,155 largest bets out of 21,420 bets generate approximately half of the trading volume during one year, and the 15 largest bets generate approximately

half of returns variance during one year.

Rare large bets may not only account for a significant percentage of returns variance but may also account for some of the stochastic time series variation in volatility. We conjecture that the pattern of short term volatility associated with execution of rare large bets may depend on the speed with which such bets are executed. Large market disturbances such as the stock market crash of 1929 and 1987, the liquidation of Jerome Kerviel’s rogue trades by Société Générale, and the flash crash of May 6, 2010, could have been induced by execution of gigantic bets.

Another implication of log-normality may be greater kurtosis in the empirical distribution of price changes than a normal distribution would suggest. Given the estimated log-variance of 2.53, the excess kurtosis of one bet has the enormous value of $\exp(10)$, or about 22,000. Thus, excess kurtosis in daily price changes may be influenced more by the kurtosis of individual bets than by the random number of bets arriving each day.

Invariance implies a different way of thinking about trading data from that in the “time change” literature, which goes back to Mandelbrot and Taylor (1967) and Clark (1973). Mandelbrot and Taylor (1967) begin with the intuition that the distribution of price changes is a stable distribution, i.e., a distribution such that a linear combination of two independent random variables has the same shape, up to location and scale parameters. Since it has fatter tails than a normal distribution, it is confined to be a stable Pareto distribution. Following this line of research, the econophysics literature—such as Gopikrishnan et al. (1998), Plerou et al. (2000), and Gabaix et al. (2006)—estimates different power-laws for the probability distributions of different variables and searches for price-formation models consistent with those distributions. Whether order size follows a power law or a log-normal distribution is an interesting question for future research.

Clark (1973) suggests as an alternative hypothesis that the distribution of daily price changes is subordinated to a normal distribution with a time clock linked to a log-normally distributed trading volume. The log-normal distribution is neither stable nor infinitely divisible; the sum of random variables with independent log-normal distributions is not log-normal. Thus, if daily price changes can be described by Clark’s hypothesis, neither half-day price changes nor weekly price changes will be described by the same hypothesis.

In some sense, our approach seems to be closer to Mandelbrot and Taylor (1967), who imagine orders of different sizes arriving in the market, with business time linked to their arrival rates rather than to trading volume.

Empirical regularities similar to those implied by invariance can be inferred from the previous literature. Bouchaud, Farmer and Lillo (2009) report, for example, that the number of TAQ prints per day is proportional to market capitalization raised to powers between 0.44 to 0.86. Under the assumption that volatility and turnover rates are stable across stocks as shown in table 1, the midpoint 0.65 of that interval is close to the value of $2/3$ implied by invariance for the number of bets per day.

The log-normality of bet size may be related to the log-normality of assets under management for financial firms. Schwarzkopf and Farmer (2010) study the size of U.S. mutual funds and find that its distribution closely conforms to a log-normal with log-variance of about 2.50, similar to our estimates of log-variance for portfolio transition orders. Their annual estimates of log-variance are stable during the twelve years from 1994 to 2005, ranging from 2.43 to 2.59. For years 1991, 1992, and 1993, the log-variance estimates of 1.51, 1.98, and 2.09 are slightly lower, probably because many observations are missing from the CRSP U.S. mutual funds dataset for those years.

As discussed by Aitchison and Brown (1957), log-normal distributions can be found in many areas of natural science. For example, Kolmogorov (1941*b*) proves mathematically that the probability distribution of the sizes of particles under fragmentation converges over time to a log-normal.

6 Empirical Tests Based on Transactions Costs

To examine statistically whether transactions costs conform to the predictions of market microstructure invariance in equation (9), we use the concept of implementation shortfall developed by Perold (1988). Specifically, we estimate costs by comparing the average execution prices of portfolio transition orders with closing prices the evening before any portfolio transition orders begin to be executed. Our tests measure implicit transactions costs resulting from bid-ask spreads and market impact; they exclude explicit transactions costs such as commissions and fees.

Portfolio Transitions and Implementation Shortfall. In portfolio transitions, quantities to be traded are known precisely before trading begins, these quantities are recorded accurately, and all intended quantities are executed. In other trading situations, quantities intended to be traded may not be recorded accurately, and orders may be canceled or quantities may be revised in response to price movements after trading begins. When orders are canceled after prices move in an unfavorable direction or when order size is increased after prices move in a favorable direction, implementation shortfall may dramatically underestimate actual transactions costs. Portfolio transitions data are not subject to these concerns.

Portfolio transition trades are unlikely to be based on short-lived private information about specific stocks because decisions to undertake portfolio transitions and their timing likely result from regularly scheduled meetings of investment committees and boards of plan sponsors, not from fast-breaking private information in the hands of fund managers. Transactions cost estimates are therefore unlikely to be biased upward as a result of short-lived private information being incorporated into prices while orders are being executed.

These properties of portfolio transitions are not often shared by other data. Con-

sider a dataset built up from trades by a mutual fund, a hedge fund, or a proprietary trading desk at an investment bank. In such samples, the intentions of traders may not be recorded in the dataset. For example, a dataset 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 actual intention to buy another 200,000 shares after the first 100,000 shares are bought. Furthermore, trading intentions may not coincide with realized trades because the trader changes his mind as market conditions change. Indeed, traders often condition their trading strategies on prices by using limit orders or canceling orders, thus hard-wiring into their strategies a selection bias problem for using such data to estimate transactions costs. The dependence of actually traded quantities on prices usually makes it impossible to use implementation shortfall in a meaningful way to estimate market depth and bid-ask spreads from data on trades only. Portfolio transitions data are particularly well suited for using implementation shortfall to measure transactions costs because portfolio transitions data avoid these sources of statistical bias.

The Empirical Hypotheses of Invariance and a Power Function Specification for Transaction Costs. For each unsigned transition order X_i , let $\mathbb{I}_{BS,i}$ denote a buy-sell indicator variable which is +1 for buy orders and -1 for sell orders. For transition order i , let C_i denote the expected transactions cost as a fraction of the value transacted. Let S_i denote the implementation shortfall, defined by $S_i = \mathbb{I}_{BS,i} \cdot (P_{ex,i} - P_{0,i})/P_{0,i}$, where $P_{ex,i}$ is the average execution price of order i and $P_{0,i}$ is its benchmark price. Implementation shortfall is positive when orders are unusually costly and negative when orders are unusually cheap.

Invariance imposes the restriction that the unobserved transactions cost C_i has the form given in equation (10):

$$C_i = \bar{\sigma}_i \cdot \bar{W}_i^{-1/3} \cdot \left(\iota^2 \cdot \bar{C}_B \cdot f\left(\frac{\bar{W}_i^{2/3}}{\iota} \cdot \frac{\mathbb{I}_{BS,i} \cdot X_i}{\bar{V}_i}\right) \right). \quad (58)$$

Under the stronger hypothesis that the cost function has a power specification for market impact costs, invariance implies the generalization of equations (12) and (13)

$$C_i = \bar{\sigma}_i \cdot \bar{W}_i^{-1/3} \cdot \left(\kappa_0 + \kappa_I \cdot \left(\bar{W}_i^{2/3} \cdot \frac{X_i}{\bar{V}_i} \right)^z \right), \quad (59)$$

where $z = 1$ for the linear specification and $z = 1/2$ for the square root specification.

Next, we test whether the cost functions can be in fact represented as the product of $\bar{\sigma}_i \cdot \bar{W}_i^{-1/3}$ and an invariant function of $\bar{W}_i^{2/3} \cdot [X_i/\bar{V}_i]$.

The predictions invariance makes about transactions costs can be expressed in terms of a non-linear regression. To justify nonlinear regression estimation, we can think of implementation shortfall as representing the sum of two components: (1) the transactions costs incurred as a result of order execution and (2) the effect of other

random price changes between the time the benchmark price is set and the time the trades are executed. If we make the identifying assumption that the implementation shortfall from the portfolio transition dataset is an unbiased estimate of the transactions cost, we can think of modeling the other random price changes as an error in a regression of implementation shortfall on transactions costs.

For example, suppose that while one portfolio transition order is being executed, there are 99 other bets being executed at the same time. The temporary and permanent price impact of executing the portfolio transition order shows up as a transactions cost, while the temporary and permanent price impact of the other 99 unobserved bets being executed shows up as other random price changes. Since the portfolio transition order is one of 100 bets being simultaneously executed, the R^2 of the regression is likely to be about 0.01.

To further develop a non-linear regression framework for testing invariance, we need to make several adjustments.

First, using equation (1) and equation (2), we replace the “bar” variables $\bar{\sigma}$, \bar{V} , and \bar{W} with observable variables σ , V , and W and with potentially unobservable constants. We also incorporate the assumption that portfolio transition orders are proportional to bets.

Second, since we want the error in our regression of implementation shortfall on transaction costs be positive when the stock price is moving up and negative when the stock price is moving down, we multiply both the implementation shortfall S_i and the transactions cost function $C(X_i)$ by the buy-sell indicator. The regression specification can be then written $\mathbb{I}_{BS,i} \cdot S_i = \mathbb{I}_{BS,i} \cdot C_i + \tilde{\epsilon}_i$. Note that $\mathbb{I}_{BS,i} \cdot S_i = (P_{ex,i} - P_{0,i})/P_{0,i}$ since $\mathbb{I}_{BS,i}^2 = 1$. This gives us a non-linear regression of the form

$$\mathbb{I}_{BS,i} \cdot S_i = \mathbb{I}_{BS,i} \cdot \left[\frac{\psi}{\psi^*} \right]^{2/3} \left[\frac{\zeta}{\zeta^*} \right]^{1/3} \left[\frac{\sigma_i}{\sigma^*} \right] \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot f(I_i \cdot \delta^{-1})/L^* + \tilde{\epsilon}_i, \quad (60)$$

where $I_i := \phi^{-1} \cdot (W_i/W^*)^{2/3} \cdot X_i/V_i$, with invariant constant ϕ obtained from equation (6) and illiquidity measure for the benchmark stock $1/L^* := \iota^2 \cdot \bar{C}_B \cdot \bar{\sigma}^* \cdot [\bar{W}^*]^{-1/3}$ obtained from equation (8).¹⁵ Note that $f(I_i \cdot \delta^{-1})/L^*$ denotes the invariant cost function for the benchmark stock, expressed as a fraction of notional value, similar to equation (9).

Third, since W_i , X_i , and V_i are observable, the quantity $\phi \cdot I_i$ is observable. The quantity I_i itself in equation (60), however, is not observable because the constant ϕ is defined in terms of potentially unobservable constants ι , δ , ψ , and ζ . To estimate the nonlinear regression equation (60), we substitute for $f(\cdot)$ a different function $f^*(\cdot)$ defined by $f^*(x) = (\psi/\psi^*)^{2/3} \cdot (\zeta/\zeta^*)^{1/3} \cdot f(\phi^{-1}\delta^{-1}x)$. Using $x = \phi \cdot I_i$, the right side of equation (60) becomes a simpler expression in terms of observable data, with various potentially unobserved constants incorporated into the definition of f^* , whose functional form is to be estimated from the data. Under the identifying assumptions

¹⁵More specifically, $\phi := \delta^{-1}\iota\psi^{-2/3}(\zeta/2)^{-1/3}(W^*)^{-2/3}$.

$\psi = \psi^*$ and $\zeta = \zeta^*$, we have $f^*(\phi I_i) = f(I_i \cdot \delta^{-1})$, where $\phi \cdot I_i := (W_i/W^*)^{2/3} \cdot X_i/V_i$ is observable. The unobserved constants hidden in ϕ affect the economic interpretation of the scaling of the estimated functional form for $f^*(.)$, but they do not otherwise affect the estimation itself.

Fourth, the variance of errors in the regression is likely to be proportional in size to the variance of returns and the execution horizon. On average, portfolio transition orders tend to be executed in about one day. To correct for heteroscedasticity resulting from differences in return volatility, we divide both the right and left sides by return volatility σ_i/σ^* , where $\sigma^* = 0.02$. Indeed, this adjustment makes the root mean squared error of the resulting regression approximately equal to 0.02.

Fifth, to control for the economically and statistically significant influence that general market movements have on implementation shortfall, we add the CRSP value-weighted market return $R_{mkt,i}$ on the first day of the transition to the right side of the regression equation. To the extent that portfolio transition orders are sufficiently large to move the entire U.S. stock market, this adjustment will result in understated transactions costs by measuring only the idiosyncratic component of transactions costs. It is an interesting subject for future research to investigate how large trades in multiple stocks affect general market movements.

Upon making these two changes and using the definition of $f^*(.)$, regression equation (60) becomes

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt,i} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot \left[\frac{W_i}{W^*} \right]^\alpha \cdot f^*(\phi I_i)/L^* + \tilde{\epsilon}_i. \quad (61)$$

where invariance implies $\alpha = -1/3$. One of our tests is designed to examine this prediction.

Scaling by W^* makes the function $f^*(\phi I_i)/L^*$ in equation (61) measure the transactions cost for the benchmark stock in terms of the observable value ϕI_i . Although invariance itself does not specify a function form for $f^*(.) / L^*$, the regression places strong cross-sectional restrictions on the shape of the transactions cost function. In addition to the restriction $\alpha = -1/3$, it requires that the same function $f^*(\phi I_i)/L^*$ with $\phi I_i = (W_i/W^*)^{2/3} \cdot X_i/V_i$ for order i be used for all stocks. We test this prediction as well.

We do not undertake separate estimates of transactions cost parameters for internal crosses, external crosses, and open market transactions. Such estimates would be difficult to interpret due to selection bias resulting from transition managers optimally choosing trading venues to minimize costs.

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

Dummy Variable Regression. In our first test, we fix $\alpha = -1/3$ in transaction costs regression equation (61), estimate function $f^*(.) / L^*$ using dummy variables,

and examine calibrated functions across ten volume groups. Invariance predicts those functions to be similar. The test does not put restrictions on the specific functional forms of $f^*(.)$.

We sort all 439,765 orders into 100 order size bins of equal size based on the value of the “invariant” order size $\phi \cdot I_i = [W_i/W^*]^{2/3} \cdot [X_i/V_i]$. As before, we also place each order into one of ten volume groups based on average dollar trading volume in the underlying stock $P_i V_i$, with thresholds corresponding to the 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of NYSE dollar volume. As shown in section 5, the distribution of ϕI_i is approximately invariant across volume groups; specifically, across all volume groups $k = 1, \dots, 10$, each bin h has a similar number of observations and similar magnitudes for ϕI_i .

In the regression equation (61), we replace the function $f^*(\phi I_i)/L^*$ with 1,000 dummy variables $D_i^*(k, h)$, $k = 1, \dots, 10$ and $h = 1, \dots, 100$, where $D_i^*(k, h) = 1$ if bet i belongs to volume group k based on dollar volume $P_i V_i$ and to order size bin h based on ϕI_i ; otherwise $D_i^*(k, h) = 0$. We then estimate 1,000 coefficients $f^*(k, h)/L^*$, $k = 1, \dots, 10$, $h = 1, \dots, 100$ for the dummy variables using a separate OLS regression for each of the volume groups, $k = 1, \dots, 10$,

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot \sum_{h=1}^{100} D_i^*(k, h) \cdot f^*(k, h)/L^* + \tilde{\epsilon}_i. \quad (62)$$

For each volume group k , the 100 dummy variable coefficients $f^*(k, h)/L^*$ (where $h = 1, \dots, 100$) track the shape of function $f^*(.) / L^*$, without imposing any particular restrictions on its functional form. Invariance predicts that the ten values of the coefficients $f^*(k, h)/L^*$, $k = 1, \dots, 10$ should be the same for each order size bin h , $h = 1, \dots, 100$. In other words, $f^*(k, h)/L^*$ is predicted not to depend on volume-group index k .

Figure 5 shows ten plots, one for each of the ten volume groups, with the 100 estimated coefficients for the dummy variables plotted as solid dots in each plot. On each plot, we also superimpose the 95% confidence intervals for 100 dummy variable coefficients estimated based on the pooled sample (dotted lines). The superimposed confidence bands help to assess the degree of similarity between cost functions estimated separately based on observations in each volume bin.

On each of the ten plots, the horizontal and vertical axes are scaled in the same way to facilitate comparison. On the horizontal axis, we plot the value for order-size bin h equal to the log of the average ϕI_i for observations in that size bin and corresponding volume group k .

On the right vertical axis, we plot the values of the dummy variable coefficients $f^*(k, h)/L^*$ quantifying for the benchmark stock the cost function as a fraction of notional value, scaled in basis points. To make deviations of cost patterns from invariants visually obvious, we have effectively scaled cost functions as suggested by invariance using regression (62): We multiply orders sizes X_i/V_i by $(W_i/W^*)^{2/3}$

and divide implementation shortfalls S_i by $L^*/L = (\sigma_i/\sigma^*) \cdot (W_i/W^*)^{-1/3}$. Here $1/L^* := \iota^2 \cdot \bar{C}_B \cdot \bar{\sigma}^* \cdot [\bar{W}^*]^{-1/3}$ is the illiquidity measure for the benchmark stock from equation (60). The invariance hypothesis implies that the 100 points plotted for each of the 10 volume groups will describe the same underlying cost function when the vertical axis is scaled according to invariance.

On the left vertical axis, we plot actual average transactions cost $f^*(k, h)/L^k$ as a fraction of notional value, scaled in basis points. For each volume group k , this scaling reverses invariance-based scaling by multiplying estimated coefficients $f^*(k, h)/L^*$ by L^*/L^k , where $1/L^k$ is the illiquidity measure for orders in volume group k given by $1/L^k := \iota^2 \cdot \bar{C}_B \cdot \bar{\sigma}_{med}^k \cdot (\bar{W}_{med}^k)^{-1/3}$, with $\bar{\sigma}_{med}^k$ denoting median betting volatility and \bar{W}_{med}^k denoting median betting activity for volume group k .

Without appropriate scaling, the data do not reveal their invariant properties. The actual costs on the left vertical axes vary significantly across volume groups. In the low volume group, costs range from -220 basis points to 366 basis points; in the high volume group, costs range from -33 basis point to 55 basis points, 7 times less than in the low volume group.

After applying “invariance” scaling, however, our plots appear to be visually consistent with the invariance hypothesis. For all ten subplots in figure 5, the estimated dummy variable coefficients on the right vertical axes are very similar across volume groups. They also line up along the superimposed confidence band.

Moving from low-volume groups to high-volume groups, these estimates also become visually more noisy. For low-volume group 1, dummy variable estimates lie within the confidence band, very tightly pinning down the estimated shape for the function $f^*(.) / L^*$. For high-volume group 10, many dummy variable estimates lie outside of the confidence band, with 11 observations above the band and about 40 observations below the band. These patterns suggest that the statistical power of our tests concerning transactions costs comes mostly from low-volume groups.

Invariance suggests that orders might be executed over horizons inversely proportional to the speed of business time, implying very slow executions for large orders in stocks with low trading activity. Portfolio transitions are, however, usually implemented within a clearly defined tight calendar time frame, which has the effect of speeding up the “natural” execution horizon for stocks with low trading activity. When transition orders are executed over a fixed number of calendar days, the execution in business time is effectively faster for low-volume stocks and slower for high-volume stocks. When a transition order in a low-volume stock is being executed, there are therefore probably fewer other bets being executed at the same time; this makes the R^2 of the regression higher. Over the same period of calendar time, more bets are being executed for the high-volume stocks, making the R^2 lower than for low-volume stocks. The more patient business-time pace of execution for high-volume stocks may explain why the dummy variable estimates are noisier for high-volume stocks than for low-volume stocks. This may also explain why the execution costs of high volume stocks appear to be slightly less expensive than low-volume stocks.

Transactions Cost Estimates in Non-Linear Regression. Next, we test the hypothesis $\alpha = -1/3$ in transaction cost regression equation (61) while simultaneously calibrating a specific functional form for the cost function $f^*(\cdot)/L^*$. We assume that this function has a particular parametric functional form equal to the sum of a constant bid-ask spread term and a market impact term which is a power of $\phi \cdot I$, similar to equation (59). For this particular specification, the non-linear regression (61) can be written as

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot \kappa_0^* \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_1} + \mathbb{I}_{BS,i} \cdot \kappa_I^* \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} \cdot \left[\frac{\phi I_i}{0.01} \right]^z + \tilde{\epsilon}_i, \quad (63)$$

where $\phi I_i/0.01 = [W_i/W^*]^{2/3} \cdot X_i/[0.01V_i]$. The explanatory variables are scaled so that, for the benchmark stock, execution of one percent of daily volume has price impact cost of κ_I^* and fixed bid-ask spread of κ_0^* , both measured as a fraction of the value traded (with units of 10^{-4} equivalent to basis points). Equation (63) nests for empirical testing both the linear model of equation (12) ($z = 1$) and the square root model of equation (13) ($z = 1/2$).

First, we report estimates of the six parameters (β_{mkt} , z , α_1 , κ_0^* , α_2 , and κ_I^*) in equation (63) using non-linear regression. Second, we calibrate the three-parameter linear impact model of equation (12) with parameters (β_{mkt} , κ_0^* , κ_I^*) by imposing the additional invariance restrictions $\alpha_1 = \alpha_2 = -1/3$ and the linear cost restriction $z = 1$. Third, we also calibrate the three-parameter square root model of equation (13) with parameters (β_{mkt} , κ_0^* , κ_I^*) by imposing the alternate restriction $z = 1/2$. Finally, we examine a twelve-parameter generalization of equation (63) which replaces powers α_1 and α_2 of trading activity W with powers of volatility σ , price P , volume V , and turnover η . Although statistical tests reject invariance, the results indicate that the predictions of invariance are economically significant, with the square root version of invariance explaining transactions costs better than the linear version.

The parameter estimates for the six parameters β_{mkt} , κ_0^* , z , α_1 , κ_I^* , α_2 in the non-linear regression (63) are reported in table 4.

For the coefficient β_{mkt} , which multiplies the market return R_{mkt} , the estimate is $\hat{\beta}_{mkt} = 0.65$ with standard error 0.013. The fact that $\hat{\beta}_{mkt} < 1$ suggests that many transition orders are executed early on the first day.¹⁶

The point estimate of the estimated bid-ask spread exponent is $\hat{\alpha}_1 = -0.49$, with standard error 0.050, three standard errors lower than the predicted value $\alpha_1 = -1/3$. In comparison with invariance, this result implies higher spread costs for less actively traded stocks and lower spread costs for more actively traded stocks. Note, however, that the factor involving α_1 is multiplied by $2 \cdot \kappa_0^*$, and κ_0^* is only of marginal statistical significance since it differs from zero by about two standard errors; excluding the bid-ask spread component of prices reduces the R^2 from 0.1010 to 0.1006 (not reported in table). This result may have something to do with the minimum tick size of one cent being a binding constraint for some securities.

¹⁶The fact that $\hat{\beta}_{mkt} = 0.65$ is close to $2/3$ is a coincidence; it is not implied by invariance.

The point estimate for α_2 is $\hat{\alpha}_2 = -0.32$ with standard error 0.015. Since invariance implies $\alpha_2 = -1/3$, this result strongly supports invariance. When the four parameters are estimated separately for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells, the estimated coefficients for α_2 are -0.40 , -0.33 , -0.41 , and -0.29 , respectively.

The estimate for the market impact curvature parameter z is $\hat{z} = 0.57$ with standard error 0.039. This suggests that a square root specification ($z = 1/2$) may describe observed transactions costs better than a linear specification ($z = 1$). Note that invariance does not place any restrictions on the parameter z itself.

The point estimate of the market impact coefficient κ_I^* is $\hat{\kappa}_I^* = 10.69 \cdot 10^{-4}$ with standard error $1.376 \cdot 10^{-4}$. The estimates of κ_I^* are higher for buy orders than for sell orders ($12.08 \cdot 10^{-4}$ versus $9.56 \cdot 10^{-4}$ for NYSE; $12.33 \cdot 10^{-4}$ versus $9.34 \cdot 10^{-4}$ for NASDAQ).

Note that the point estimate of the market impact coefficient κ_I^* is different from zero with much greater statistical significance than the bid-ask spread coefficient κ_0^* ; this is consistent with the interpretation that market impact costs are of considerably greater economic importance than bid-ask spread costs.

Results for the five-parameter linear specification, regression equation (63) with the parameter restriction $z = 1$ (linear impact), are reported in table 9 in the Appendix. The estimate of the bid-ask spread cost κ_0^* is $6.28 \cdot 10^{-4}$ with standard error $0.890 \cdot 10^{-4}$, and the estimate of the exponent α_1 is $\hat{\alpha}_1 = -0.39$ with standard error 0.020. The estimate of the market impact cost κ_I^* is $2.73 \cdot 10^{-4}$ with standard error 0.252. The estimate of the exponent α_2 is $\hat{\alpha}_2 = -0.31$ with standard error of 0.028; thus, under the restriction of linear price impact ($z = 1$), the additional restriction imposed by invariance ($\alpha_2 = 1/3$) is not statistically rejected.

Model Calibration. Next, we calibrate transaction cost models, under the assumption of invariance and the assumption of either linear or square root specification for the cost function.

Table 5 presents estimates for the three parameters β_{mkt} , κ_0^* , and κ_I^* in equation (63), imposing the invariance restrictions $\alpha_1 = \alpha_2 = -1/3$ and also imposing either a linear transactions cost model $z = 1$ or a square root model $z = 1/2$.

In the linear specification with $z = 1$, the point estimate for market impact cost $\hat{\kappa}_I^*$ is equal to $2.50 \cdot 10^{-4}$, and the point estimate for bid-ask spread cost $\hat{\kappa}_0^*$ is equal $8.21 \cdot 10^{-4}$.

In the square root specification with $z = 1/2$, the point estimate for market impact cost $\hat{\kappa}_I^*$ is equal to $12.08 \cdot 10^{-4}$, and the point estimate for half bid-ask spread $\hat{\kappa}_0^*$ is equal to $2.08 \cdot 10^{-4}$.

For the benchmark stock, these estimates imply that the total cost of a hypothetical trade of one percent of daily volume incurs a cost of about 10.71 basis points in the linear model and 14.16 basis points in the square root model.

The benchmark stock would belong to volume group 7, and the corresponding

average quoted spread in table 1 for that group is 12.04 basis points. The implied spread estimate of about 16.42 basis points for the linear model is close to the quoted spread; the implied spread estimate of 4.16 basis points for the square root model may be biased downwards due to collinearity between the constant term and the square root term in the regression in the region close to zero.

Economic Interpretation. We examine the economic significance of our results by comparing the R^2 of different specifications for transaction costs models.

The R^2 is equal to 0.0847 in the transaction cost regressions with market return only (not reported). This implies that a substantial part of realized transaction costs is explained by overall market dynamics. The transaction costs models improve the R^2 s by only one or two percent.

A comparison of the R^2 s in table 4 and table 5 provides strong support for the invariance hypothesis. When the coefficient on W/W_* is fixed at the invariance-implied value of $-1/3$ and only two transactions cost parameters κ_I^* and κ_0^* are estimated (table 5), the R^2 is 0.0991 for a linear specification and 0.1007 for a square root specification. The square root specification performs better than the linear specification. Compared with the square root specification, adding the three additional parameters α_1, α_2 and z modestly increases the R^2 from 0.1007 to 0.1010 (table 5). The modest increase strongly supports the economic importance of invariance.

We also consider a more general specification with eleven estimated coefficients. The exponents on the three components of trading activity W_i (volatility σ_i , price $P_{0,i}$, volume V_i) as well as the exponent on the monthly turnover ν_i are allowed to vary freely. The estimated regression equation is

$$\begin{aligned} \mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = & \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} \mathbb{I}_{BS,i} \cdot \kappa_0^* \cdot \left[\frac{W_i}{W_*} \right]^{-1/3} \cdot \frac{\sigma_i^{\beta_1} \cdot P_{0,i}^{\beta_2} \cdot V_i^{\beta_3} \cdot \nu_i^{\beta_4}}{(0.02)(40)(10^6)(1/12)} + \\ & + \mathbb{I}_{BS,i} \cdot \kappa_I^* \cdot \left[\frac{\phi I_i}{0.01} \right]^z \cdot \left[\frac{W_i}{W_*} \right]^{-1/3} \cdot \frac{\sigma_i^{\beta_5} \cdot P_{0,i}^{\beta_6} \cdot V_i^{\beta_7} \cdot \nu_i^{\beta_8}}{(0.02)(40)(10^6)(1/12)} + \tilde{\epsilon}_i. \end{aligned} \quad (64)$$

Because the exponents on the W -terms are set to be $-1/3$, the invariance hypothesis predicts $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$.

Table 5 shows that despite increasing the number of estimated parameters from four to eleven, the R^2 in the aggregate regression increases from 0.1010 to only 0.1016. The estimates of $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$, and β_8 are shown in table 11 in the Appendix. The estimates of $\beta_1, \beta_2, \beta_3, \beta_4$ are often statistically significant, but these explanatory variables are multiplied by statistically insignificant coefficient κ_0^* . Almost all estimates of $\beta_5, \beta_6, \beta_7$, and β_8 are statistically insignificant, both for the pooled sample as well as the four sub-samples.

In all three specifications, separate regressions for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells suggest that price impact costs are higher for buy orders than for sell orders. This is consistent with the hypothesis, discussed Obizhaeva

(2009), that the market believes that buy orders—in particular, buy orders in portfolio transitions—contain more information than sell orders.

OLS Estimates for Quoted Spread. Finally, we present results of statistical tests based on the data on quoted bid-ask spread for portfolio transition orders.

Since invariance implies that bid-ask spread costs are proportional to $\bar{\sigma} \cdot \bar{W}^{-1/3}$, intuition suggests that quoted spreads may also have this invariant property. As a supplement to our empirical results on transactions costs, we test this prediction using data on quoted spreads, supplied in the portfolio transition data as pre-trade information for each transition order.

Let s_i denote the dollar quoted spread for order i . Applying equation (12) or (13) to quoted spreads, we obtain

$$s_i/P_i \propto \bar{\sigma}_i \cdot \bar{W}_i^{-1/3}. \quad (65)$$

Using equation (1) and equation (2), we can write the log-linear OLS regression

$$\ln \left[\frac{s_i}{P_i \cdot \sigma_i} \right] = \ln \bar{s} + \alpha_3 \cdot \ln \left[\frac{W_i}{W^*} \right] + \tilde{\epsilon}_i, \quad (66)$$

where invariance implies $\alpha_3 = -1/3$. The constant term $\ln \bar{s} := \ln[s^*/(40 \cdot 0.02)] + 2/3 \ln(\psi/\psi^*) - 1/3 \ln(\zeta^*/\zeta)$ quantifies the dollar spread s^* for the benchmark stock as a fraction of dollar volatility $P^* \cdot \sigma^*$, under the identifying assumptions $\zeta = \zeta^*$ and $\psi = \psi^*$.

Table 6 presents the regression results. The point estimate $\hat{\alpha}_3 = -0.35$ has standard error 0.003. For sub-samples of NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells, the estimates are -0.31 , -0.32 , -0.40 , and -0.39 , respectively. Although the hypothesis $\alpha_3 = -1/3$ is usually rejected statistically, the estimates are economically close to the value of $-1/3$ predicted by invariance. The point estimate of $\ln \bar{s}$ is equal to -3.07 , implying a quoted spread of $\exp(-3.07) \cdot 0.02 \approx 9 \cdot 10^{-4}$ for the benchmark stock. This number is similar to the median spread of 8.12 basis points for volume group 7 in table 1.

It can be shown that an implicit spread proportional to $\bar{\sigma} \cdot \bar{W}^{-1/3}$, as implied by invariance, provides a better proxy for the actually incurred spread costs than the quoted spread itself. When regression equation (63) is estimated with linear impact $z = 1$, using only the 436,649 observations for which quoted bid-ask spread data is supplied, we find the R^2 is equal to 0.0992. Now replace the invariance-implied spread cost proportional to $\bar{\sigma} \cdot \bar{W}^{-1/3}$ with the quoted half spread $1/2 \cdot s_i/P_i$ in equation (12). The estimated equation is

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot h \cdot \frac{1}{2} \cdot \frac{s_i}{P_{0,i}} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot \kappa_I \cdot \left[\frac{\phi I_i}{0.01} \right] \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} + \tilde{\epsilon}_i. \quad (67)$$

We find that the R^2 drops from 0.0992 to 0.0976 (table 10 in the Appendix). The point estimate of the coefficient on the quoted half-spread coefficient is $\hat{h} = 0.71$. The estimates are equal to 0.61, 0.74, 0.61, and 0.75, when estimated for NYSE Buys, NYSE Sells, NASDAQ Buys, and NASDAQ Sells, respectively.

One interpretation of the estimate of 0.71 is that transition managers incur as a transactions cost only 71% of the quoted half-spread. The values are consistent with the intuition in Goettler, Parlour and Rajan (2005) that endogenously optimizing traders capture a fraction of the bid-ask spread by mixing between market orders and limit orders. Another interpretation is that noise in the quoted spread relative to the “true” spread biases the coefficient towards zero and reduces the explanatory power of the regression.

Bouchaud, Farmer and Lillo (2009) and Dufour and Engle (2000) report that the quoted bid-ask spread is proportional to the standard deviation of percentage returns between trades; this result is implied by microstructure invariance under the assumption that the rate at which trades occur is proportional to the rate at which bets arrive. Stoll (1978a) proposes a theory that the percentage dealer bid-ask spread in NASDAQ stocks is proportional to variables including the dealer holding period and the returns variance of the stock; this captures the spirit of invariance if the dealer holding period is proportional to the rate at which bets arrive. Stoll (1978b) also tests this theory using data on dealer spreads for NASDAQ stocks, and his estimates are consistent with our findings as well.¹⁷

Discussion. Figure 6 plots the estimated coefficients for the 100 dummy variables, along with their 95% confidence intervals, estimated from the dummy variable regression equation (62) by pooling the data across all 10 volume groups. The linear and square root cost functions with parameters calibrated in table 5 are superimposed. The linear specification is $2.50 \cdot 10^{-4} \cdot \phi I / 0.01 + 8.21 \cdot 10^{-4}$ (solid black line), and the square root specification is $12.07 \cdot 10^{-4} \cdot \sqrt{\phi I / 0.01} + 2.08 \cdot 10^{-4}$ (solid grey line). Both specifications result in estimates economically close to each other.

Consistent with the higher reported R^2 for the square root model than the linear model in table 5, the square root specification fits the data slightly better than the linear specification, particularly for large orders in the order size bins from 90th to 99th percentiles. Consistent with our results, most studies find that total price impact is best described by a concave function.¹⁸ For example, Almgren et al. (2005) obtain

¹⁷Stoll (1978b) reports an R^2 of approximately 0.82 in an OLS regression of percentage bid-ask spread on the logs of various variables including dollar volume, stock price, returns variance, turnover, and number of dealers. Using the standard deviations and correlation matrix for the variables (p. 1165), it can be shown that imposing coefficients of $-1/3$ on dollar volume and $+1/3$ on returns variance (to mimic the definition of $1/L$), while imposing coefficients of zero on all other explanatory variables except a constant term, results in an R^2 equal to 0.66. This result is similar to our results in table 6.

¹⁸Since we plot the log of order size on the horizontal axis but do not take the log of the transaction cost on the vertical axis (to make standard errors have similar magnitudes for different observations),

an estimate $\hat{z} = 0.60$ for their sample of almost 30,000 U.S. stock orders executed by Citigroup between 2001 and 2003; this is comparable to our estimate of $\hat{z} = 0.56$ when the constraint $\alpha_1 = \alpha_2 = -1/3$ is imposed in regression equation (63). To differentiate temporary impact from permanent impact of earlier executed trades, Almgren et al. (2005) assume a particular execution algorithm with a constant rate of trading. We do not quantify these cost components separately but rather focus on total costs.

Intuition might suggest that for gigantic orders, the square root model would predict dramatically lower transactions costs than the linear model, making it easy to distinguish the predictions of one model from the other. As the superimposed estimated linear and square root cost functions for the ten plots in figure 5 make clear, both specifications estimate similar transactions costs for the bin representing the largest 1% of orders (because the graphs of the linear and square root functions in figure 6 cross near the bin representing the largest 1% of orders). Furthermore, the transaction cost dummy variable for the largest 1% of orders fits both the linear and square root models well. For the largest 1% of orders in the highest volume group in figure 5, the estimated dummy variable fits the higher cost estimates of the linear model better than the square root model.

7 Implications

The invariance relationships (5), (6), (9), and (14) are like a structural model which describes the implications of market microstructure for bet size, transactions costs, efficiency and resiliency. The model is fully specified by constants describing the moments of \tilde{I} and the shape of the un-modeled function $C_B(\cdot)$, which determines the constant \bar{C}_B . These constants can be inferred from the estimates in section 5 and section 6, but their economic interpretation depends on assumptions about the volume multiplier ζ , the volatility multiplier ψ , and the deflator δ .

Our empirical tests provide not only evidence in favor of the invariance hypotheses but also inputs for the calibration process. Our empirical results can be summarized as follows. The distribution of portfolio transition orders $|\tilde{X}|$ —expressed as a fraction of volume—is approximately a log-normal. It is therefore fully described by two parameters, the log-mean for the benchmark stock estimated to be -5.71 and the log-variance estimated to be 2.53 (table 3). The following formula shows how these estimates can be extrapolated to stocks with other levels of trading activity $W = \sigma \cdot P \cdot V$ and volume V :

$$\ln \left[\frac{|\tilde{X}|}{V} \right] \approx -5.71 - \frac{2}{3} \cdot \ln \left[\frac{W}{(0.02)(40)(10^6)} \right] + \sqrt{2.53} \cdot \tilde{Z}, \quad \tilde{Z} \sim N(0, 1) \quad (68)$$

both the linear model and the concave square root model show up as exponential functions; the graph of the linear model is more convex than the graph of the square root model.

$$\gamma = 85 \cdot \left[\frac{W}{(0.02)(40)(10^6)} \right]^{2/3}. \quad (69)$$

The last equation for the number of bets γ follows directly from equation (68) under the assumption that the volume multiplier $\zeta = 2$ and the portfolio transition size multiplier $\delta = 1$. These two equations fully describe the order flow process.

Our empirical results also suggest that transactions cost functions can be described by either a linear model or a square root model. Since both models also have a constant bid-ask spread term, each model is described by two parameters. For an order of 1% of average daily volume in the benchmark stock, the estimates imply market impact costs of $\kappa_I = 2.50 \cdot 10^{-4}$ and spread costs of $\kappa_0 = 8.21 \cdot 10^{-4}$ for the linear model as well as market impact costs of $\kappa_I = 12.08 \cdot 10^{-4}$ and spread costs of $\kappa_0 = 2.08 \cdot 10^{-4}$ for the square root model (table 5). The following formulas show how these estimates can be extrapolated to execution costs of an order of X shares for stocks with other levels of trading activity W , volume V , and volatility σ :

$$C(X) = \frac{\sigma}{0.02} \left(\frac{8.21}{10^4} \cdot \left[\frac{W}{(0.02)(40)(10^6)} \right]^{-1/3} + \frac{2.50}{10^4} \cdot \left[\frac{W}{(0.02)(40)(10^6)} \right]^{1/3} \frac{X}{(0.01)V} \right). \quad (70)$$

$$C(X) = \frac{\sigma}{0.02} \left(\frac{2.08}{10^4} \cdot \left[\frac{W}{(0.02)(40)(10^6)} \right]^{-1/3} + \frac{12.08}{10^4} \cdot \left[\frac{X}{(0.01)V} \right]^{1/2} \right). \quad (71)$$

These two equations fully describe the transactions cost models.

To summarize, formulas (68), (69), (70), and (71) provide a simple way to calculate the number of bets, different percentiles of bet sizes, and transactions costs. The only stock-specific inputs required are expected dollar volume and volatility.¹⁹

Our results can also be used to calibrate the distribution of \tilde{I} and the invariant cost function $C_B(\cdot)$, which further implies specific quantitative relationships concerning various market microstructure variables such as the number of bets per day, bet sizes, spread, liquidity, pricing accuracy, and resiliency as functions of easily observable trading activity and its components. These implications ultimately depend on values assigned to the volatility multiplier ψ , the volume multiplier ζ , and the deflator δ .

In a more complicated exercise left for future research, this handful of parameters would make it possible to triangulate the value of parameters measuring the fraction of trading volume due to long-term investors rather than intermediaries $1/\zeta$, the fraction of returns volatility generated by bets ψ , and the ratio of the size of bets to the size of portfolio transition orders δ . In the future, careful thinking about calibration of

¹⁹If there is a concern that the volume multiplier ζ and the volatility multiplier ψ are different from the ones relevant for our sample of portfolio transitions, then a simple adjustment to the formulas must be implemented. First, one needs to deflate trading volume and trading volatility in those formulas by appropriate multipliers in order to write those formulas in terms of betting volume and betting volatility for observations in portfolio transitions data. Second, one needs to plug into the modified formulas betting volume and betting volatility appropriate for the market, in the context of which calculations are made.

the invariants and estimation of multipliers will be necessary to sharpen predictions based on invariance principles.

8 Conclusion

We have shown that the predictions based on market microstructure invariance are economically consistent with results estimate from portfolio transitions data for U.S. equities. We conjecture that predictions based on invariance may be found to hold in other data as well, such as spread, quote, and transactions data in the TAQ dataset; levels and changes in holdings recorded in 13-F filings of institutional investment managers; institutional trades reported in the Ancerno dataset; and other datasets. For example, we conjecture that data on news articles can help to show that information flows take place in the same business time as trading.

We conjecture that predictions of market microstructure invariance may generalize to other markets such as bond markets, currency markets, and futures markets, as well as to other countries. Whether market microstructure invariance applies to other markets poses an interesting set of issues for future research.

We do not expect invariance to hold perfectly across different markets and different times periods. Differences in trading institutions across markets might make the volume multiplier and the volatility multiplier vary across markets. We expect transactions costs, particularly bid-ask spread costs (but perhaps not market impact costs), to be influenced by numerous institutional features, such as government regulation (e.g., short sale restrictions, customer order handling rules), transactions taxes, competitiveness of market making institutions, efficiency of trading platforms, market fragmentation, technological change, and tick size. For example, if minimum tick size rules affect bid-ask spread costs, we believe that market microstructure invariance can be used as a benchmark against which the effect of tick size on bid-ask spread costs can be evaluated.

To conclude, market microstructure invariance implies simple scaling laws which lead to sharp statistical hypothesis about bet size and transaction costs; it is consistent with a simple structural model from which specific versions of these scaling laws can be derived in closed form; and its implications explain an economically significant portion of the variation in portfolio transition order size and transaction costs when the scaling laws are imposed on the data. The scaling laws enable us to derive simple operational formulas describing order size distributions and transaction costs; thus, they provide simple benchmarks from which past research can be evaluated and open up new lines of research into market microstructure.

References

Aitchison, J., and J. A. C. Brown. 1957. *The Lognormal Distribution with Special*

Reference to Its Uses in Economics. Cambridge University Press.

- Allais, Maurice.** 1956. "Explication des Cycles Economiques Par Un Modele Non Linéaire a Régulation Retardée." *Metroeconomica*, 8(1): 4–83.
- Allais, Maurice.** 1966. "A Restatement of the Quantity Theory of Money." *American Economic Review*, 56(5): 1123–1157.
- Almgren, Robert, Chee Thum, Emmanuel Hauptmann, and Hong Li.** 2005. "Direct Estimation of Equity Market Impact." *Risk*, 18(7): 58–62.
- Amihud, Yakov.** 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *The Journal of Financial Markets*, 5: 31–56.
- Ané, Thierry, and Helyette Geman.** 2000. "Order Flow, Transaction Clock, and Normality of Asset Returns." *The Journal of Finance*, 55(5): 2259–2284.
- Back, Kerry, and Shmuel Baruch.** 2004. "Information in Securities Markets: Kyle meets Glosten and Milgrom." *Econometrica*, 72: 433–465.
- Barra.** 1997. In *Market Impact Model Handbook*. Berkley, CA: Barra.
- Black, Fischer.** 1986. "Noise." *The Journal of Finance*, 41: 529–543.
- Bouchaud, Jean-Philippe, Dooyne J. Farmer, and Fabrizio Lillo.** 2009. "How Markets Slowly Digest Changes in Supply and Demand." In *Handbook of Financial Markets: Dynamics and Evolution*. 47–160. North Holland.
- Breen, William, Laurie Simon Hodrick, and Robert A. Korajczyk.** 2002. "Predicting Equity Liquidity." *Management Science*, 48: 470–83.
- Brennan, Michael, and Avanidhar Subrahmanyam.** 1998. "The Determinants of Average Trade Size." *The Journal of Business*, 71(1): 1–25.
- Campbell, John Y., and Albert S. Kyle.** 1993. "Smart Money, Noise Trading, and Stock Price Behavior." *The 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.
- Derman, Emanuel.** 2002. "The Perception of Time, Risk and Return during Periods of Speculation." *Quantitative Finance*, 2: 282–296.
- Dufour, Alfonso, and Robert F. Engle.** 2000. "Time and the Price Impact of a Trade." *The Journal of Finance*, 55(6): 2467–2598.

- Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and Eugene H. Stanley.** 2006. "Institutional Investors and Stock Market Volatility." *The Quarterly Journal of Economics*, 121: 461–504.
- Glosten, Lawrence R., and Paul R. Milgrom.** 1985. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *The Journal of Financial Economics*, 14: 71–100.
- Goettler, Ronald L., Christine A. Parlour, and Uday Rajan.** 2005. "Equilibrium in a Dynamic Limit Order Market." *The Journal of Finance*, 60(5): 2149–2192.
- Gopikrishnan, Parameswaran, Martin Meyer, Luis A. Amaral, and Eugene H. Stanley.** 1998. "Inverse Cubic Law for the Distribution of Stock Price Variations." *The European Physical Journal*, 3: 139–140.
- Grinold, Richard C., and Ronald N. Kahn.** 1995. In *Active Portfolio Management*. Probus Publishing.
- Grossman, Sanford J., and Joseph E. Stiglitz.** 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review*, 70(3): 393–408.
- Grossman, Sanford J., and Merton H. Miller.** 1988. "Liquidity and Market Structure." *The Journal of Finance*, 43(3): 617–633.
- Hasbrouck, Joel.** 1999. "Trading Fast and Slow: Security Market Events in Real Time." Working Paper, New York University.
- Hasbrouck, Joel.** 2007. *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading*. Oxford University Press, USA.
- Holden, Craig W, Stacey E Jacobsen, and Avanidhar Subrahmanyam.** 2015. "The empirical analysis of liquidity." *Foundations and Trends in Finance*, forthcoming.
- Kolmogorov, Andrei N.** 1941a. "The Local Structure of Turbulence in Incompressible Viscous Fluids at Very Large Reynolds Numbers." *Doklady Akademii Nauk SSSR*, 30: 301–305. Reprinted in *Proceedings of the Royal Society London A* 434, 9-13, 1991.
- Kolmogorov, Andrei N.** 1941b. "On The Log-Normal Distribution of Particle Sizes during Break-up Process." *Doklady Akademii Nauk SSSR*, 31(2): 99–101.
- 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." *The Journal of Finance*, 46: 733–746.

- Loeb, Thomas.** 1983. "Trading Cost: The Critical Link between Investment Information and Results." *Financial Analysts Journal*, 39: 39–44.
- Mandelbrot, Benoît, and Howard M. Taylor.** 1967. "On the Distribution of Stock Price Differences." *Operations Research*, 15(6): 1057–1062.
- Modigliani, Franco, and Merton H. Miller.** 1958. "The Cost of Capital, Corporation Finance and the Theory of Investment." *American Economic Review*, 48: 261–297.
- Obizhaeva, Anna.** 2009. "Portfolio Transitions and Price Dynamics." Working Paper, University of Maryland.
- Perold, André.** 1988. "The Implementation Shortfall: Paper vs. Reality." *Journal of Portfolio Management*, 14(3): 4–9.
- Plerou, Vasiliki, Parameswaran Gopikrishnan, Luis A. Amaral, Xavier Gabaix, and Eugene H. Stanley.** 2000. "Economic Fluctuations and Anomalous Diffusion." *Physical Review E*, 62: 3023–3026.
- Schwarzkopf, Yonathan, and Doynne J. Farmer.** 2010. "Empirical Study of the Tails of Mutual Fund Size." *Physical Review E*, 81: 066113.
- Stambough, Robert F., and Lubos Pastor.** 2003. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy*, 111: 642–685.
- Stoll, Hans R.** 1978a. "The supply of dealer services in securities markets." *The Journal of Finance*, 33(4): 1133–1151.
- Stoll, Hans R.** 1978b. "The pricing of security dealer services: An empirical study of NASDAQ stocks." *The Journal of Finance*, 33(4): 1153–1172.
- 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).
- 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.

panel A: Properties of Securities.

	All	1	2	3	4	5	6	7	8	9	10
$\text{Med}(V \cdot P) \times 10^6$	18.72	1.13	5.10	9.92	15.93	23.87	31.41	42.12	60.25	101.60	212.85
$\text{Med}(\sigma) \times 10^2$	1.93	2.16	2.04	1.94	1.98	1.90	1.86	1.80	1.78	1.77	1.76
$\text{Med}(S_{prd}) \times 10^4$	12.04	40.96	18.72	13.70	12.02	10.32	9.42	8.12	7.21	5.92	4.83
$\text{Med}(Turn) \times 10^2$	12.72	6.58	12.36	14.80	15.95	16.80	16.36	15.37	14.41	14.92	11.12

panel B: Properties of Portfolio Transitions Orders.

	All	1	2	3	4	5	6	7	8	9	10
$\text{Avg}(X/V) \times 10^2$	4.20	16.23	4.54	2.62	1.83	1.37	1.18	1.08	0.88	0.69	0.49
$\text{Med}(X/V) \times 10^2$	0.57	3.33	1.36	0.79	0.53	0.40	0.34	0.30	0.25	0.20	0.14
$\text{Avg}(X/Cap) \times 10^4$	1.72	3.55	2.68	2.04	1.59	1.26	1.06	0.91	0.72	0.56	0.37
$\text{Med}(X/Cap) \times 10^4$	0.35	0.98	0.80	0.58	0.42	0.32	0.27	0.23	0.19	0.15	0.09
$\text{Avg } C(X) \times 10^4$	16.79	44.95	21.46	14.53	12.62	11.70	5.58	9.27	3.99	7.37	6.16
$\text{Avg Comm} \times 10^4$	7.43	14.90	9.30	7.86	7.00	6.15	5.49	4.93	4.34	3.62	2.68
$\text{Avg SEC fee} \times 10^5$	2.90	3.26	3.02	3.00	2.85	2.84	2.76	2.76	2.73	2.68	2.56
# Obs	439,765	71,000	68,689	41,238	49,000	28,073	29,330	29,778	34,409	40,640	47,608
# Stks	2,583	1,108	486	224	182	106	126	90	102	81	78

Table reports the characteristics of securities and transition orders. Panel A shows the median average daily dollar volume (in millions of \$), the median daily volatility (in percents), the median percentage spread (in basis points), the median monthly turnover rate (in percents). panel B shows the average and median order size (in percents of daily volume and in basis points of market capitalization) as well as average implementation shortfall (in basis points), the average commission (in basis points), and the average SEC fee for sell orders (in percents per 10 basis points). The thresholds of ten volume groups correspond to 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of dollar volume for common stocks listed on the NYSE. Group 1 (Group 10) contains orders in stocks with lowest (highest) dollar trading volume. The sample ranges from January 2001 to December 2005.

Table 2: OLS Estimates of Order Size.

	All	NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
$\ln [\bar{q}]$	-5.67 (0.017)	-5.68 (0.023)	-5.63 (0.018)	-5.75 (0.035)	-5.65 (0.032)
α_0	-0.62 (0.009)	-0.63 (0.011)	-0.59 (0.008)	-0.71 (0.019)	-0.59 (0.015)
R^2	0.3167	0.2587	0.2646	0.4298	0.3542
$Q^*/V^* \cdot \delta^{-1} \times 10^{-4}$	34.62	34.14	35.98	31.80	34.78
#Obs	439,765	131,530	150,377	69,871	87,987

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

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

Each observation corresponds to transition order i with order size X_i , benchmark price $P_{0,i}$, expected daily volume V_i , expected daily volatility σ_i , trading activity W_i . \bar{q} is the measure of order size such that Q^*/V^* measures the median order size for a benchmark stock. The benchmark stock has daily volatility of 2%, share price of \$40, and daily volume of one million shares. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.

Table 3: OLS Estimates for Order Size: Model Calibration.

		NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
<i>Restricted Specification: $\alpha_0 = -2/3, b_1 = b_2 = b_3 = b_4 = 0$</i>					
$\ln [\bar{q}]$	-5.71 (0.019)	-5.70 (0.023)	-5.68 (0.019)	-5.70 (0.042)	-5.77 (0.039)
$Q^*/V^* \cdot \delta^{-1} \times 10^{-4}$	33.13	33.46	34.14	33.46	31.20
MSE	2.53	2.61	2.54	2.32	2.56
R^2	0.3149	0.2578	0.2599	0.4278	0.3479
<i>Unrestricted Specification With 5 Degrees of Freedom: $\alpha_0 = -2/3$.</i>					
R^2	0.3229	0.2668	0.2739	0.4318	0.3616
#Obs	439,765	131,530	150,377	69,871	87,987

Table presents the estimates $\ln \bar{q}$ and the mean squared error (MSE) for the regression:

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + b_1 \cdot \ln \left[\frac{\sigma_i}{0.02} \right] + b_2 \cdot \ln \left[\frac{P_{0,i}}{40} \right] + b_3 \cdot \ln \left[\frac{V_i}{10^6} \right] + b_4 \cdot \ln \left[\frac{\nu_i}{1/12} \right] + \tilde{\epsilon}_i.$$

with α_0 restricted to be $-2/3$ as predicted by invariance hypothesis and $b_1 = b_2 = b_3 = 0$. Each observation corresponds to transition order i with order size X_i , benchmark price $P_{0,i}$, expected daily volume V_i , expected daily volatility σ_i , trading activity W_i , and monthly turnover rate ν_i . \bar{q} is the measure of order size such that Q^*/V^* measures the corresponding percentile of order size for a benchmark stock. The benchmark stock has daily volatility of 2%, share price of \$40, and daily volume of one million shares. The R^2 s are reported for restricted specification with $\alpha_0 = -2/3, b_1 = b_2 = b_3 = b_4 = 0$ as well as for unrestricted specification with coefficients $\ln \bar{q}$ and b_1, b_2, b_3, b_4 allowed to vary freely. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.

Table 4: Transactions Cost Estimates in Non-Linear Regression.

	All	NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
β_{mkt}	0.66 (0.013)	0.63 (0.016)	0.62 (0.016)	0.76 (0.037)	0.78 (0.036)
$\kappa_0^* \times 10^4$	1.77 (0.837)	-0.27 (2.422)	1.14 (1.245)	0.77 (4.442)	3.55 (1.415)
α_1	-0.49 (0.050)	-0.37 (1.471)	-0.50 (0.114)	0.53 (1.926)	-0.44 (0.045)
$\kappa_I^* \times 10^4$	10.69 (1.376)	12.08 (2.693)	9.56 (2.254)	12.33 (2.356)	9.34 (2.686)
z	0.57 (0.039)	0.54 (0.056)	0.56 (0.062)	0.44 (0.051)	0.63 (0.086)
α_2	-0.32 (0.015)	-0.40 (0.037)	-0.33 (0.029)	-0.41 (0.035)	-0.29 (0.037)
R^2	0.1010	0.1118	0.1029	0.0945	0.0919
#Obs	439,765	131,530	150,377	69,871	87,987

Table presents the estimates for β_{mkt} , z , α_1 , κ_0^* , α_2 , and κ_I^* in the regression:

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot \kappa_0^* \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_1} + \mathbb{I}_{BS,i} \cdot \kappa_I^* \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} \cdot \left[\frac{\phi I_i}{0.01} \right]^z + \tilde{\epsilon}_i. \quad (72)$$

where $\phi I_i / 0.01 = X_i / (0.01 V_i) \cdot (W_i / W^*)^{2/3}$. S_i is implementation shortfall. R_{mkt} is the value-weight market return for the first day of transition. 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 2% per day, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order i . κ_I^* is the market impact costs of executing a trade of one percent of daily volume in a benchmark stock, and κ_0^* is the effective spread cost. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.

Table 5: Transactions Costs: Model Calibration.

	NYSE			NASDAQ	
	All	Buy	Sell	Buy	Sell
<i>Linear Model: $z = 1, \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0.$</i>					
$\kappa_0^* \times 10^4$	8.21 (0.578)	7.19 (1.122)	6.77 (0.794)	9.18 (1.563)	9.27 (0.781)
$\kappa_I^* \times 10^4$	2.50 (0.190)	3.37 (0.370)	1.92 (0.265)	3.46 (0.395)	2.46 (0.327)
R^2	0.0991	0.1102	0.1012	0.0926	0.0897
<i>Square Root Model: $z = 1/2, \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0.$</i>					
$\kappa_0^* \times 10^4$	2.08 (0.704)	-1.31 (1.278)	0.92 (0.926)	2.28 (2.055)	4.65 (0.824)
$\kappa_I^* \times 10^4$	12.08 (0.742)	15.65 (1.218)	11.10 (1.298)	13.50 (1.456)	10.41 (1.207)
R^2	0.1007	0.1116	0.1027	0.0941	0.0911
<i>Unrestricted Specification With 12 Degrees of Freedom.</i>					
R^2	0.1016	0.1121	0.1032	0.0957	0.0944
#Obs	439,765	131,530	150,377	69,871	87,987

Table presents the estimates κ_0^* and κ_I^* for the regression:

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} \mathbb{I}_{BS,i} \cdot \kappa_0^* \cdot \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot \frac{\sigma_i^{\beta_1} \cdot P_{0,i}^{\beta_2} \cdot V_i^{\beta_3} \cdot \nu_i^{\beta_4}}{(0.02)(40)(10^6)(1/12)} +$$

$$+ \mathbb{I}_{BS,i} \cdot \kappa_I^* \cdot \left[\frac{\phi I_i}{0.01} \right]^z \cdot \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot \frac{\sigma_i^{\beta_5} \cdot P_{0,i}^{\beta_6} \cdot V_i^{\beta_7} \cdot \nu_i^{\beta_8}}{(0.02)(40)(10^6)(1/12)} + \tilde{\epsilon}_i.$$

where invariant $\phi I_i/0.01 = X_i/(0.01 V_i) \cdot (W_i/W^*)^{2/3}$. S_i is implementation shortfall. R_{mkt} is the value-weight market return for the first day of transition. 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 2% per day, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order i . κ_I^* is the market impact costs of executing a trade of one percent of daily volume in a benchmark stock, and κ_0^* is the effective spread cost. The R^2 s are reported for restricted specification as well as for unrestricted specification with twelve coefficients $\beta_{mkt}, z, \kappa_I^*, \kappa_0^*, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8$ allowed to vary freely. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.

Table 6: OLS Estimates of Log of Quoted Spread.

	All	NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
$\ln \bar{s}$	-3.07 (0.008)	-3.09 (0.008)	-3.08 (0.008)	-3.04 (0.013)	-3.04 (0.012)
α_3	-0.35 (0.003)	-0.31 (0.003)	-0.32 (0.003)	-0.40 (0.004)	-0.39 (0.004)
R^2	0.4744	0.3545	0.3964	0.5516	0.5721
$e^{\ln \bar{s}} \cdot 0.02 \times 10^4$	9.28	9.10	9.19	9.57	9.57
#Obs	434,920	130,700	149,197	68,833	86,190

Table presents the estimates $\ln \bar{s}$ and α_3 for the regression:

$$\ln \left[\frac{s_i}{P_i \cdot \sigma_i} \right] = \ln \bar{s} + \alpha_3 \cdot \ln \left[\frac{W_i}{W^*} \right] + \tilde{\epsilon}_i,$$

Each observation corresponds to order i . The left-hand side variable is the logarithm of the quoted bid-ask spread $s_i/P_{0,i}$ as a fraction of expected return volatility σ_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 2% per day, price \$40 per share, and trading volume of one million shares per day. The median percentage spread for a benchmark stock is $\exp(\ln \bar{s}) \cdot 0.02$. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.

Figure 2: Invariant Order Size Distribution.

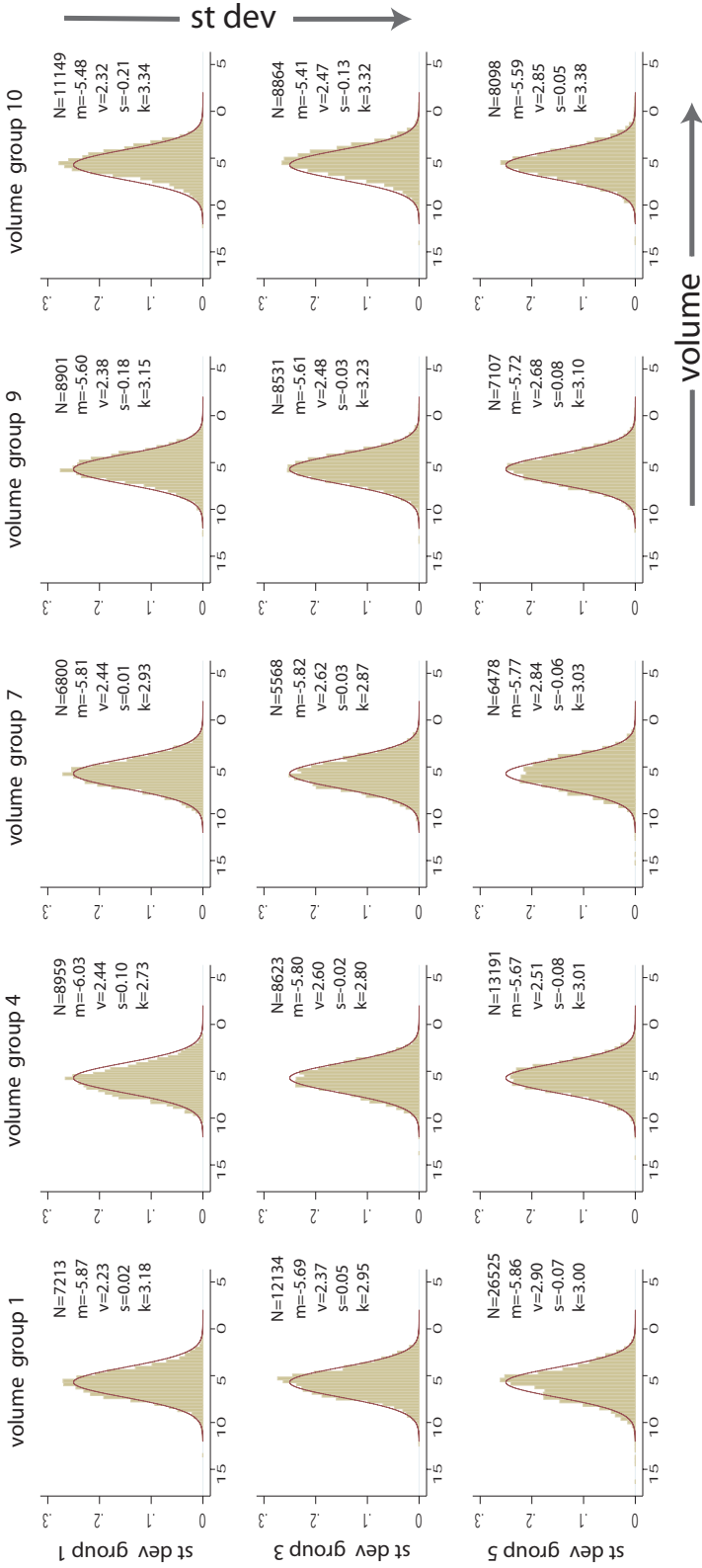
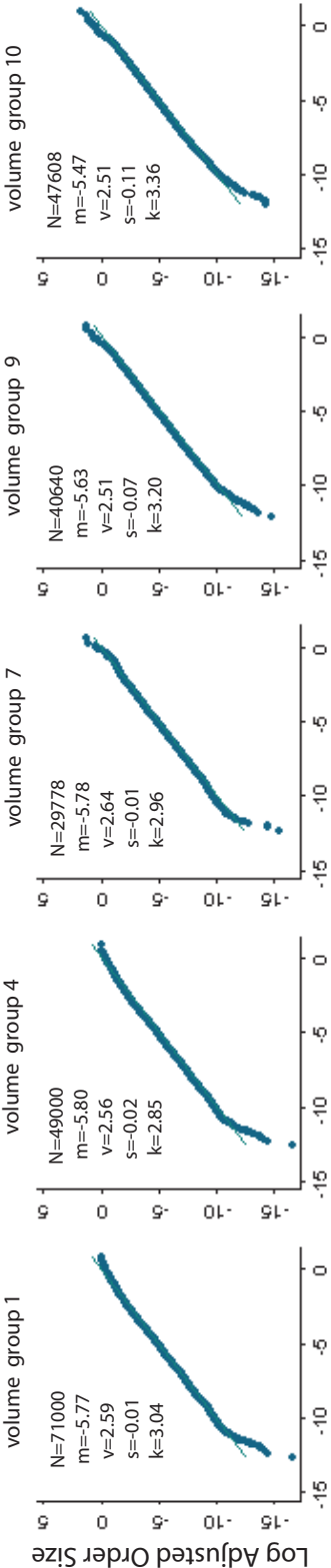


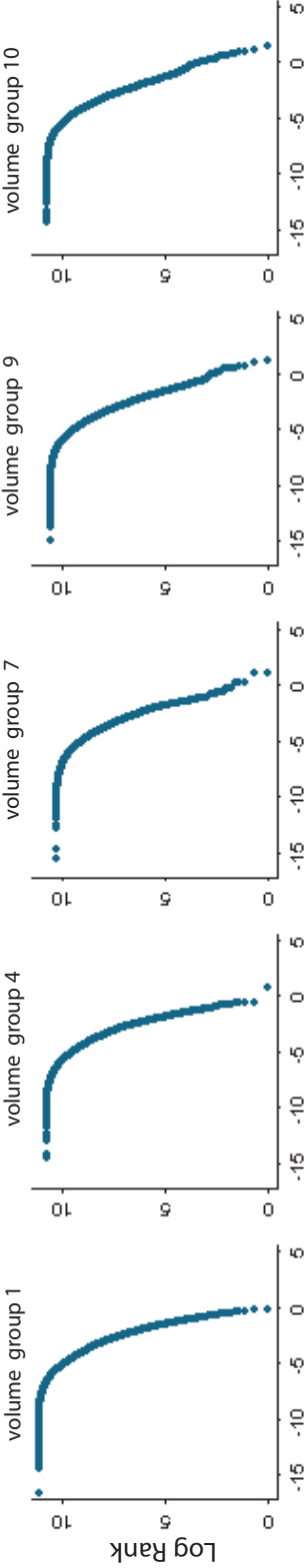
Figure shows distributions of $\ln(\tilde{X}/V) + 2/3 \ln(W_i/W^*)$ for stocks sorted into 10 volume groups and 5 volatility groups (only volume groups 1, 4, 7, 9, 10 and volatility groups 1, 3, 5 are reported). X_i is an order size in shares, V_i is the average daily volume in shares, and W_i is the trading activity. The thresholds of ten volume groups correspond to 30th, 50th, 60th, 70th, 80th, 85th, 90th, and 95th percentiles of dollar volume for common stocks listed on the NYSE. Volume group 1 (group 10) contains orders in stocks with lowest (highest) dollar volume. The thresholds of five volatility groups correspond 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 -5.71 and variance of 2.54 is imposed on each subplot. The common mean and variance are calculated as the mean and variance of distribution over the entire sample. The sample ranges from January 2001 to December 2005.

Figure 3: Invariant Order Size Distribution.

Panel A: Quantile-to-Quantile Plot for Empirical and Lognormal Distribution.

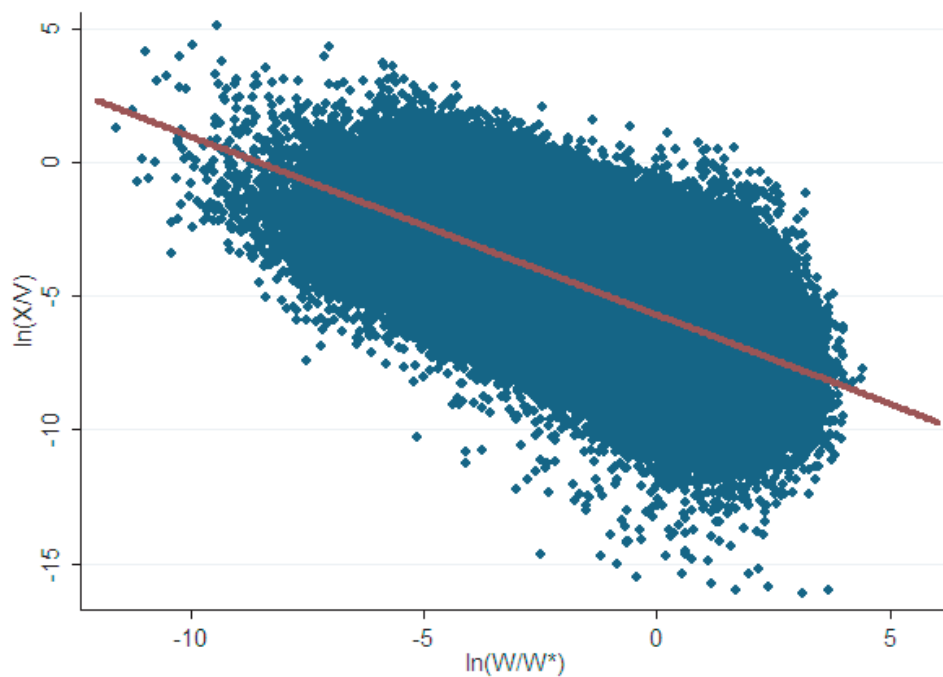


Panel B: Logarithm of Ranks against Quantiles of Empirical Distribution.



panel A shows quantile-quantile plots of empirical distributions of $\ln(\tilde{X}/V) + 2/3 \ln(W_i/W^*)$ and a normal distribution for stocks sorted into 10 volume groups (only volume groups 1, 4, 7, 9, 10 are reported). panel B depicts the logarithm of ranks based on that distribution. 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. Each subplot shows the number of observations (N), the mean (m), the variance (v), the skewness (s), and the kurtosis (k) of a depicted distribution. There are 400,000+ data points. The sample ranges from January 2001 to December 2005.

Figure 4: Order Size and Trading Activity.



The figure plots $\ln(X_i/V_i)$ on the vertical axis against $\ln(W_i/W^*)$ on the horizontal axis, where X_i is portfolio transition order size in shares, V_i is average daily volume in shares, and $W_i = P_i \cdot V_i \cdot \sigma_i$ is trading activity. The fitted line is $\ln(X_i/V_i) = -5.705 - 2/3 \cdot \ln(W_i/W^*)$, where the intercept is estimated from an OLS regression with the slope fixed at $-2/3$. There are 400,000+ data points from January 2001 to December 2005.

Figure 5: Invariant Transactions Cost Functions.

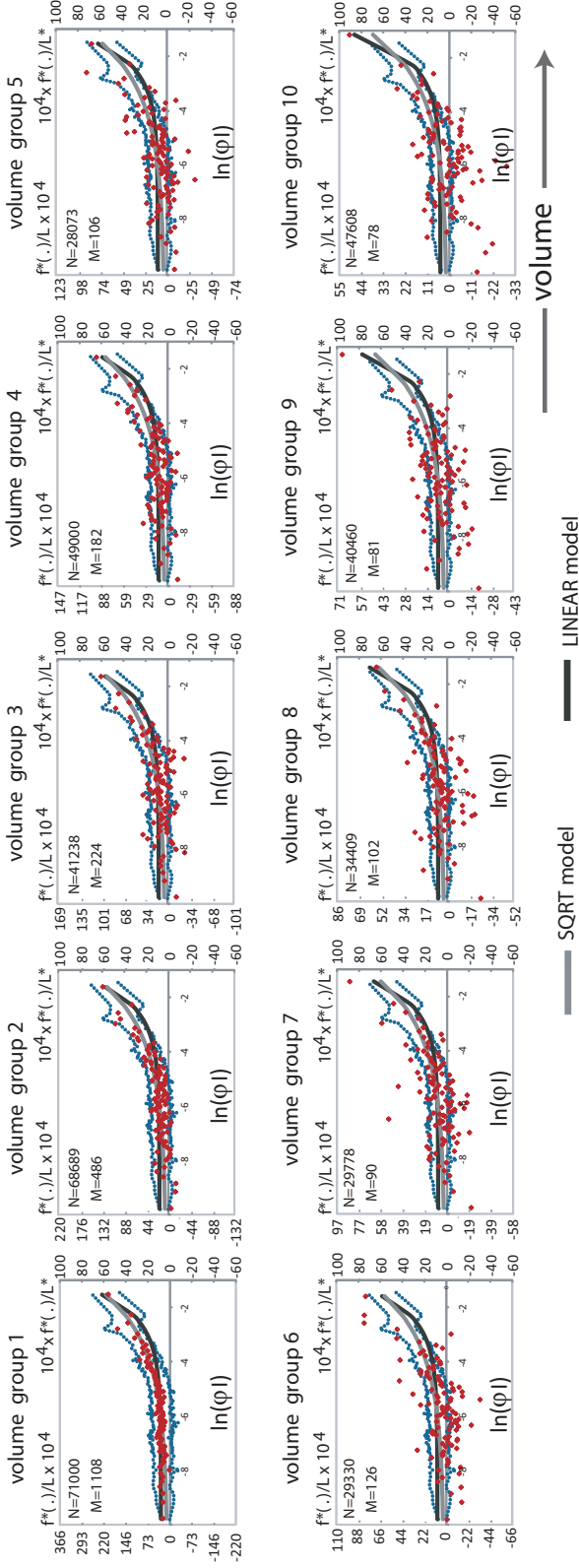


Figure shows estimates of transactions cost functions for stock sorted into 10 volume groups. On the horizontal axis, there are 100 equally spaced bins based on re-scaled order sizes, $\phi I = \tilde{X} / V \cdot (W_i / W^*)^{2/3}$. For each volume group $k = 1, \dots, 10$, the subplot contains 100 estimates of dummy variables $f^*(k, h) / L^*$, $h = 1, \dots, 100$ from regression (62). On the right-hand side vertical axis, there are units of scaled transactions cost $f^*(.) / L^*$ for a benchmark stock. On the left-hand side vertical axis, there are units of actual transactions cost $f^*(.) / L^k$ for a benchmark stock, where $1 / L^k$ is the illiquidity measure for orders in volume group k . The 95th percent confidence interval estimates based on the entire sample are imposed on each subplot (blue dotted lines). The common linear and square root functions are imposed on each subplot with the parameter estimated on the entire sample. A linear function is $2.50 \cdot 10^{-4} \cdot \phi I / 0.01 + 8.21 \cdot 10^{-4}$ (black solid line). A square root function is $12.07 \cdot 10^{-4} \cdot \sqrt{\phi I / 0.01} + 2.08 \cdot 10^{-4}$ (grey solid line). The thresholds of ten volume groups correspond to 30th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles of dollar volume for common stocks listed on the NYSE. Group 1 (group 10) contains orders in stocks with lowest (highest) dollar volume. Each subplot also shows the number of observations N and the number of stocks M (for the last month). The sample ranges from January 2001 to December 2005.

Figure 6: Transactions Cost Functions.

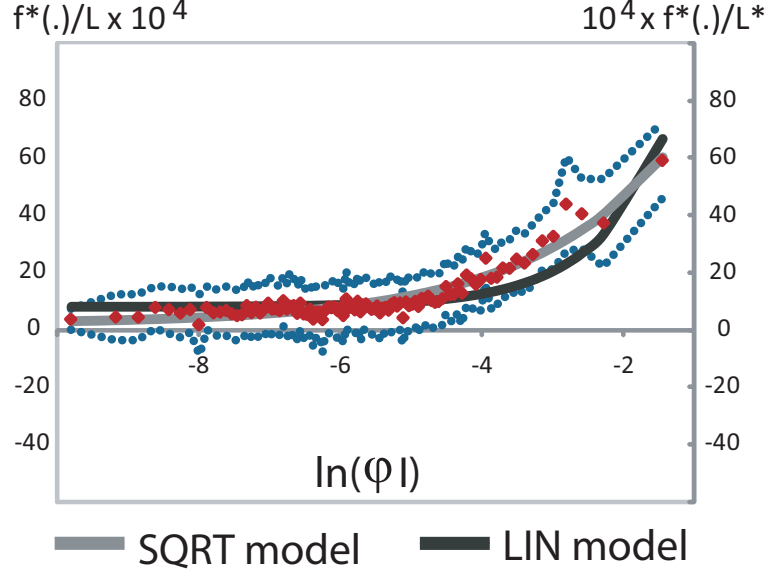


Figure shows estimates of transactions cost functions based on entire sample. On the horizontal axis, there are 100 equally spaced bins based on re-scaled order sizes, $\phi I = \tilde{X}/V \cdot (W_i/W^*)^{2/3}$. The plot contains 100 estimates $f^*(k, h)/L^*, h = 1, \dots, 100$ from the regression

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot \sum_{h=1}^{100} D_i^*(k, h) \cdot f^*(k, h)/L^* + \tilde{\epsilon}_i.$$

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 vertical axis presents estimated transactions cost invariant $f^*(.) / L$ in basis points. The 95th percent confidence interval are superimposed (dotted lines). A linear function is $2.50 \cdot 10^{-4} \cdot \phi I / 0.01 + 8.21 \cdot 10^{-4}$ (black solid line). A square root function is $12.07 \cdot 10^{-4} \cdot \sqrt{\phi I / 0.01} + 2.08 \cdot 10^{-4}$ (grey solid line). The sample ranges from January 2001 to December 2005.

Part I

Appendix

Table 7: Quantile Estimates of Order Size.

	p1	p5	p25	p50	p75	p95	p99
$\ln [\bar{q}]$	-9.37 (0.008)	-8.31 (0.006)	-6.73 (0.004)	-5.66 (0.003)	-4.59 (0.004)	-3.05 (0.006)	-2.05 (0.009)
α_0	-0.65 (0.005)	-0.64 (0.003)	-0.61 (0.002)	-0.62 (0.002)	-0.61 (0.002)	-0.64 (0.003)	-0.63 (0.005)
Pseudo R^2	0.1621	0.1534	0.1650	0.1727	0.1795	0.1949	0.2232
$Q^*/V^* \cdot \delta \times 10^{-4}$	0.85	2.46	11.95	34.83	101.53	473.59	1287.35
#Obs	439,765	439,765	439,765	439,765	439,765	439,765	439,765

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

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

Each observation corresponds to transition order i with order size X_i , benchmark price $P_{0,i}$, expected daily volume V_i , expected daily volatility σ_i , trading activity W_i . \bar{q} is the measure of order size such that Q^*/V^* measures the corresponding percentile of order size for a benchmark stock. The benchmark stock has daily volatility of 2%, share price of \$40, and daily volume of one million shares. The standard errors are shown in parentheses. The sample ranges from January 2001 to December 2005.

Table 8: OLS Estimates for Order Size: Model Calibration.

		NYSE		NASDAQ	
	All	Buy	Sell	Buy	Sell
<i>Restricted Specification: $\alpha_0 = -2/3, b_1 = b_2 = b_3 = b_4 = 0$</i>					
$\ln [\bar{q}]$	-5.71 (0.019)	-5.70 (0.023)	-5.68 (0.019)	-5.70 (0.042)	-5.77 (0.039)
$Q^*/V^* \cdot \delta \times 10^4$	33.13	33.46	34.14	33.46	31.20
MSE	2.53	2.61	2.54	2.32	2.56
R^2	0.3149	0.2578	0.2599	0.4278	0.3479
<i>Unrestricted Specification With 5 Degrees of Freedom: $\alpha_0 = -2/3$.</i>					
$\ln [\bar{q}]$	-5.53 (0.019)	-5.55 (0.026)	-5.48 (0.019)	-5.77 (0.051)	-5.48 (0.047)
b_1	0.42 (0.040)	0.47 (0.050)	0.53 (0.043)	0.19 (0.094)	0.33 (0.087)
b_2	0.24 (0.019)	0.17 (0.021)	0.29 (0.017)	0.04 (0.049)	0.33 (0.040)
b_3	0.06 (0.010)	0.06 (0.012)	0.07 (0.009)	-0.06 (0.026)	0.07 (0.021)
b_4	-0.18 (0.015)	-0.24 (0.020)	-0.22 (0.017)	-0.02 (0.040)	-0.11 (0.032)
R^2	0.3229	0.2668	0.2739	0.4318	0.3616
#Obs	439,765	131,530	150,377	69,871	87,987

Table presents the estimates and the mean squared error (MSE) for the regression:

$$\ln \left[\frac{X_i}{V_i} \right] = \ln [\bar{q}] + \alpha_0 \cdot \ln \left[\frac{W_i}{W^*} \right] + b_1 \cdot \ln \left[\frac{\sigma_i}{0.02} \right] + b_2 \cdot \ln \left[\frac{P_{0,i}}{40} \right] + b_3 \cdot \ln \left[\frac{V_i}{10^6} \right] + b_4 \cdot \ln \left[\frac{\nu_i}{1/12} \right] + \tilde{\epsilon}_i.$$

with α_0 restricted to be $-2/3$ as predicted by invariance hypothesis and $b_1 = b_2 = b_3 = 0$. Each observation corresponds to transition order i with order size X_i , benchmark price $P_{0,i}$, expected daily volume V_i , expected daily volatility σ_i , trading activity W_i , and monthly turnover rate ν_i . \bar{q} is the measure of order size such that Q^*/V^* measures the corresponding percentile of order size for a benchmark stock. The benchmark stock has daily volatility of 2%, share price of \$40, and daily volume of one million shares. The R^2 s are reported for restricted specification with $\alpha_0 = -2/3, b_1 = b_2 = b_3 = b_4 = 0$ as well as for unrestricted specification with coefficients $\ln \bar{q}$ and b_1, b_2, b_3, b_4 allowed to vary freely. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.

Table 9: Transactions Cost Estimates in Non-Linear Regression with Linear Impact.

	All	NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
β_{mkt}	0.66 (0.013)	0.63 (0.016)	0.62 (0.016)	0.77 (0.037)	0.77 (0.036)
$\kappa_0^* \times 10^4$	6.28 (0.890)	6.51 (1.600)	5.43 (1.154)	5.94 (2.147)	6.54 (1.501)
α_1	-0.40 (0.020)	-0.36 (0.048)	-0.39 (0.029)	-0.44 (0.051)	-0.40 (0.031)
$\kappa_I^* \times 10^4$	2.73 (0.252)	2.63 (0.460)	2.10 (0.346)	3.69 (0.663)	3.13 (0.765)
α_2	-0.31 (0.028)	-0.45 (0.038)	-0.31 (0.041)	-0.32 (0.056)	-0.28 (0.058)
R^2	0.0993	0.1105	0.1014	0.0931	0.0901
#Obs	439,765	131,530	150,377	69,871	87,987

Table presents the estimates for β_{mkt} , α_1 , κ_0^* , α_2 , and κ_I^* in the regression:

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot \kappa_0^* \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_1} + \mathbb{I}_{BS,i} \cdot \kappa_I^* \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} \cdot \left[\frac{\phi I_i}{0.01} \right]^z + \tilde{\epsilon}_i. \quad (73)$$

where $z = 1$ and $\phi I_i / 0.01 = X_i / (0.01 V_i) \cdot (W_i / W^*)^{2/3}$. S_i is implementation shortfall. R_{mkt} is the value-weight market return for the first day of transition. 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 2% per day, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order i . κ_I^* is the market impact costs of executing a trade of one percent of daily volume in a benchmark stock, and κ_0^* is the effective spread cost. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.

Table 10: Transactions Cost Estimates in Non-Linear Regression with Quoted Spread.

	All	NYSE		NASDAQ	
		Buy	Sell	Buy	Sell
β_{mkt}	0.65 (0.013)	0.63 (0.016)	0.62 (0.015)	0.76 (0.036)	0.77 (0.037)
$\kappa_I \times 10^4$	2.95 (0.261)	2.97 (0.504)	2.24 (0.366)	3.76 (0.700)	2.95 (0.749)
α_2	-0.32 (0.029)	-0.44 (0.036)	-0.32 (0.039)	-0.37 (0.053)	-0.33 (0.060)
h	0.71 (0.053)	0.61 (0.110)	0.74 (0.094)	0.61 (0.127)	0.75 (0.073)
R^2	0.0976	0.1094	0.1010	0.0891	0.0872
#Obs	436,649	131,100	149,600	69,218	86,731

Table presents the estimates for β_{mkt} , κ_I , α_2 , and h in the regression:

$$\mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot h \cdot \frac{1}{2} \cdot \frac{s_i}{P_{0,i}} \cdot \frac{(0.02)}{\sigma_i} + \mathbb{I}_{BS,i} \cdot \kappa_I \cdot \left[\frac{\phi I_i}{0.01} \right] \cdot \left[\frac{W_i}{W^*} \right]^{\alpha_2} + \tilde{\epsilon}_i.$$

where invariant $I_i = \frac{X_i}{(0.01)V_i} \cdot \left[\frac{W_i}{W^*} \right]^{2/3}$. Each observation corresponds to order i . $\mathbb{I}_{BS,i}$ is a buy/sell indicator, S_i is implementation shortfall, R_{mkt} is the value-weight market return for the first day of transition. 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 2% per day, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order i . κ_I is the market impact costs of executing a trade of one percent of daily volume in a benchmark stock. $s_i/P_{0,i}$ is the quoted spread. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.

Table 11: Transactions Costs: Model Calibration.

		NYSE		NASDAQ	
	All	Buy	Sell	Buy	Sell
<i>Linear Model: $z = 1, \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0.$</i>					
β_{mkt}	0.6571 (0.0135)	0.6308 (0.0159)	0.6195 (0.0158)	0.7693 (0.0371)	0.7771 (0.0365)
$\kappa_0 \times 10^4$	8.2134 (0.5776)	7.1934 (1.1215)	6.7698 (0.7943)	9.1832 (1.5627)	9.2658 (0.7811)
$\kappa_I \times 10^4$	2.5003 (0.1903)	3.3663 (0.3700)	1.9220 (0.2650)	3.4614 (0.3953)	2.4629 (0.3267)
R^2	0.0991	0.1102	0.1012	0.0926	0.0897
<i>Square Root Model: $z = 1/2, \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0.$</i>					
β_{mkt}	0.6552 (0.0134)	0.6285 (0.0158)	0.6192 (0.0159)	0.7598 (0.0365)	0.7782 (0.0364)
$\kappa_0 \times 10^4$	2.0763 (0.7035)	-1.3091 (1.2779)	0.9167 (0.9264)	2.2844 (2.0554)	4.6530 (0.8244)
$\kappa_I \times 10^4$	12.0787 (0.7416)	15.6544 (1.2177)	11.0986 (1.2979)	13.5025 (1.4564)	10.4063 (1.2069)
R^2	0.1007	0.1116	0.1027	0.0941	0.0911
<i>Unrestricted Specification With 12 Degrees of Freedom.</i>					
β_{mkt}	0.66 (0.013)	0.63 (0.016)	0.62 (0.015)	0.76 (0.036)	0.78 (0.036)
$\kappa_0^* \times 10^4$	0.94 (0.675)	-0.05 (0.124)	0.47 (0.556)	1.55 (1.698)	1.61 (1.148)
β_1	-0.43 (0.147)	-2.47 (0.890)	-1.08 (0.392)	-0.44 (0.489)	-0.46 (0.131)
β_2	0.17 (0.072)	2.87 (1.230)	0.23 (0.231)	0.20 (0.127)	0.11 (0.109)
β_3	-0.56 (0.159)	1.85 (0.754)	-0.47 (0.296)	-0.47 (0.238)	-0.49 (0.155)
β_4	0.62 (0.173)	0.13 (0.620)	0.49 (0.490)	0.49 (0.313)	0.58 (0.175)
$\kappa_I^* \times 10^4$	9.36 (1.307)	11.61 (2.471)	10.93 (1.804)	8.88 (3.340)	5.00 (2.033)
z	0.58 (0.041)	0.54 (0.039)	0.52 (0.042)	0.58 (0.094)	0.63 (0.083)
β_5	0.02 (0.135)	-0.11 (0.192)	0.36 (0.229)	-0.17 (0.252)	-0.23 (0.242)
β_6	-0.14 (0.061)	-0.11 (0.113)	0.03 (0.100)	-0.27* (0.120)	-0.22 (0.113)
β_7	0.01 (0.037)	-0.07 (0.050)	0.04 (0.052)	0.00 (0.099)	-0.16 (0.100)
β_8	0.08 (0.067)	0.07 (0.086)	-0.11 (0.101)	0.08 (0.143)	0.39 (0.153)
R^2	0.1016	0.1121	0.1032	0.0957	0.0944
#Obs	439,765	131,530	150,377	69,871	87,987

Table presents the estimates for the regression:

$$\begin{aligned} \mathbb{I}_{BS,i} \cdot S_i \cdot \frac{(0.02)}{\sigma_i} = & \beta_{mkt} \cdot R_{mkt} \cdot \frac{(0.02)}{\sigma_i} \mathbb{I}_{BS,i} \cdot \kappa_0^* \cdot \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot \frac{\sigma_i^{\beta_1} \cdot P_{0,i}^{\beta_2} \cdot V_i^{\beta_3} \cdot \nu_i^{\beta_4}}{(0.02)(40)(10^6)(1/12)} + \\ & + \mathbb{I}_{BS,i} \cdot \kappa_I^* \cdot \left[\frac{\phi I_i}{0.01} \right]^z \cdot \left[\frac{W_i}{W^*} \right]^{-1/3} \cdot \frac{\sigma_i^{\beta_5} \cdot P_{0,i}^{\beta_6} \cdot V_i^{\beta_7} \cdot \nu_i^{\beta_8}}{(0.02)(40)(10^6)(1/12)} + \tilde{\epsilon}_i. \end{aligned}$$

where $\phi I_i/0.01 = X_i/(0.01V_i) \cdot (W_i/W^*)^{2/3}$. S_i is implementation shortfall. R_{mkt} is the value-weight market return for the first day of transition. 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 2% per day, price \$40 per share, and trading volume of one million shares per day. X_i is the number of shares in the order i . κ_I^* is the market impact costs of executing a trade of one percent of daily volume in a benchmark stock, and κ_0^* is the effective spread cost. The R^2 s are reported for restricted specification as well as for unrestricted specification with twelve coefficients $\beta_{mkt}, z, \kappa_I^*, \kappa_0^*, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8$ allowed to vary freely. The standard errors are clustered at weekly levels for 17 industries and shown in parentheses. The sample ranges from January 2001 to December 2005.