Market Microstructure Invariance: A Dynamic Equilibrium Model

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

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

First Draft: January 17, 2016
This Draft: January 31, 2019†

Abstract

Invariance relationships are derived in a dynamic, infinite-horizon, equilibrium model of adverse selection with risk-neutral informed traders, noise traders, risk-neutral market makers, and endogenous information production. Scaling laws for bet size and transaction costs require the assumption that the effort required to generate one bet does not vary across securities and time. Scaling laws for pricing accuracy and market resiliency require the additional assumption that private information has the same signal-to-noise ratio across markets. Prices follow a martingale with endogenously derived stochastic volatility. Returns volatility, pricing accuracy, market depth, and market resiliency are closely related to one another. The model solution depends on two state variables: stock price and hard-to-observe pricing accuracy. Invariance makes predictions operational by expressing them in terms of log-linear functions of easily observable variables such as price, volume, and volatility.

Keywords: Market microstructure, invariance, liquidity, bid-ask spread, market impact, transaction costs, market efficiency, efficient markets hypothesis, pricing accuracy, resiliency, order size.

---

*Albert S. Kyle has worked as a consultant for various companies, exchanges, and government agencies. He is a non-executive director of a U.S.-based asset management company.

†This paper was previously part of the manuscript "Market Microstructure Invariants: Theory and Empirical Tests" (October 17, 2014). The authors thank Agostino Capponi, Jeongmin Lee, Mark Loewenstein, and Shri Santosh for helpful comments.

Microstructure invariance defines a dimensionless liquidity measure L which is proportional to the cube root of the ratio of dollar trading volume to percentage returns variance. Invariance predicts that average bet size scales with L, the arrival rate of bets scales with the product of returns variance and L^2, and percentage transaction costs scale with $1/L$. Kyle and Obizhaeva (2016) derived these scaling laws from ad hoc empirical conjectures. Kyle and Obizhaeva (2017) obtained the same hypotheses from dimensional analysis and leverage neutrality. In both derivations, information and adverse selection play no role.

This paper derives the same scaling laws as endogenous implications of a microfounded economic model of trading with adverse selection. Explicit modeling of private information also makes it possible to obtain new invariance relationships linking the informativeness of prices and market resiliency to the liquidity measure L. Obviously, predictions about the informativeness of prices intrinsically require a dynamic model in which prices play an informational role and therefore cannot be derived using the previous two approaches.

Invariance relationships are implied by a model based on the following assumptions. The unobserved fundamental value of the stock follows geometric Brownian motion. Informed and noise traders arrive randomly, obtain a private signal about the fundamental value by paying a fixed dollar cost, and then trade once. Each informed trader obtains an informative signal. Each noise trader obtains a fake signal with the same unconditional distribution as an informative signal but with not information content; noise traders believe themselves to be informed. The number of noise traders is assumed to adjust so that they turn over the outstanding share float at an exogenously given rate. The number of informed traders adjusts endogenously, with new informed traders entering the market as long as they expect to make nonnegative profits, net of transaction costs and net of costs of acquiring signals. Without knowing whether traders are informed or not, competitive risk-neutral market makers update their expectations and take the other side of each bet at break-even prices.

We solve for an approximate linear equilibrium in which the number of traders, trading strategies, market liquidity, volatility, market resiliency, and the informativeness of prices are log-linear functions of two state variables. These state variables are the current stock price—which is perfectly correlated with market capitalization and share trading volume—and the current *pricing error variance*, defined as the log-variance of the ratio of prices to fundamental value. From the perspective of market makers, the price follows a martingale even though each trader believes his own trading moves the price toward fundamental value. Although fundamental volatility is assumed to be constant, returns volatility turns out to be stochastic and increasing in the size of the pricing error variance. A conditional steady state occurs when returns volatility and fundamental volatility coincide, so that new fundamental uncertainty unfolds at the same rate prices incorporate private information.

Our model's theoretical predictions are difficult to test empirically because the state variable measuring pricing accuracy requires quantifying how far prices are from fundamentals. Market microstructure invariance makes the predictions empirically testable by making it possible to infer pricing accuracy from observable market characteristics such as liquidity L.

Market microstructure invariance is based on the intuition that financial markets are in some fundamental sense similar to each other, except that they operate at different speeds. In active liquid markets, business time runs quickly; in inactive illiquid markets, business time runs slowly. Business time, which is hard to observe, is related to the speed with which bets, or new investment ideas, arrive to the marketplace. In our model, the numbers of traders and bet arrival rate are endogenous variables which change as state variables change.

We show that two ad hoc empirical invariance conjectures of Kyle and Obizhaeva (2016) hold exactly in an approximate linear equilibrium. First, *bet size invariance* says that the distribution of the dollar risks transferred by bets does not vary when measured in business time. This hypothesis implies that average dollar bet size is proportional to L, and bets arrive at a rate proportional to the product of returns variance and L^2. Second, *transaction cost invariance* says that the expected dollar market impact cost of a bet is an unvarying function of the dollar risks that bets transfer in business time. This hypothesis implies that percentage transactions costs are proportional to $1/L$.

The model leads to two new invariance hypotheses. Define *pricing accuracy* as the reciprocal of the standard deviation of the log-distance between the market price and fundamental value. *Pricing accuracy invariance* says that pricing accuracy is invariant if the pricing error standard deviation is scaled by returns volatility per unit of business time. This hypothesis implies that pricing accuracy is proportional to L. Define *market resiliency* as the rate at which uninformative shocks to prices decay in calendar time or, equivalently, as the rate at which prices converge towards fundamental value in calendar time. *Market resiliency invariance* says that market resiliency is invariant if it is measured per unit of business time. This hypothesis implies that market resiliency is proportional to the product of returns variance and L^2.

Invariance relationships imply scaling laws which link difficult-to-observe microstructure characteristics—such as bet size, number of bets, market depth, market liquidity, pricing accuracy, and market resiliency—to more easily observable dollar volume and returns volatility. Business time passes at a rate proportional to the two-thirds power of *trading activity*, defined as the product of dollar volume and returns volatility. The difference between prices and fundamental value is difficult to quantify empirically. The model predicts that the percentage standard deviation of the difference between prices and fundamental value is proportional to illiquidity $1/L$, which is itself an observable function of dollar volume and returns variance. These scaling laws are summarized in Theorem 2, equations (53) and Corollary 2, equations (69).

These invariance relationships are simple implications of general properties likely shared by many models (see equations (54)–(57)). First, trading volume is defined as the sum of all bets. Second, order flow moves prices and induces returns volatility; unconditional long-term price impact is linear in the information content of bets. Third, the dollar cost of acquiring an informative signal—which in equilibrium equals the dollar price impact cost of the bet—is the same across assets and time. Fourth, the distributions of bet size and signals have the same shape across markets. The invariance of bet size and invariance of market depth require that the dollar effort cost to generate a signal is invariant (across assets and time). Invariance of pricing accuracy and resiliency requires the additional assumption that the signal-to-noise ratio of an informative signal is invariant.

The model clarifies the intuition behind invariance relationships and the seemingly obscure exponents of one-third and two-thirds. Suppose returns volatility remains constant and equal to fundamental volatility, but market capitalization increases due to increases in prices. When market capitalization is higher, more traders execute bets, dollar volume is higher, the market becomes more efficient in the sense that market depth increases and the distance between prices and fundamentals shrinks. Traders must execute bets of larger sizes in order to make enough profits to cover the *same* dollar costs of producing a private signal. Pricing accuracy does adjust to changes in dollar volume, but only slowly, since returns volatility per bet and the average log-distance between prices and fundamentals decreases only half as much as the rate of increase in the number of bets. Traders thus must scale up the size of their bets by the same rate, by half as much as the increase in the number of bets. This implies a one-to-two ratio between an increase in the size of bets and their arrival rate. Alternatively, since dollar volume is the product of the number of bets and their average dollar size, the number of bets increases log-linearly with the two-thirds power of trading volume and their average dollar size increases linearly with one-third power of trading volume.

The model highlights the difference between two definitions of market efficiency. On the one hand, the model assumes that our markets are efficient in the sense that prices follow a martingale, consistent with Fama (1970) and LeRoy (1989). On the other hand, we derive endogenously how pricing errors vary as functions of paths of trading volume and volatility, consistent with the idea of Black (1986) that market efficiency should be based on the accuracy of prices. In our model, pricing error variance is shown to be inversely proportional to market resiliency and the rate at which bets arrive.

Our model blends together several traditional strands of the market microstructure literature. The models resembles the model of Kyle (1985) by assuming linear trading intensity, linear price impact, normally distributed random variables, and zero-profit market makers; yet it is different because the assumed linear trading strategies and pricing updates are only approximately, not exactly, optimal. The model resembles the models of Glosten and Milgrom (1985) and Back and Baruch (2004) in that orders arrive sequentially and are processed by market makers one at a time; it differs by assuming that traders may choose to buy or sell any quantity, not just one round lot. Like Treynor (1971) and Black (1986), the model assumes that noise traders trade on uninformative, fake signals; noise traders believe they are informed traders even though they are trading on noise. Unlike Kyle, Obizhaeva and Wang (2018), traders do not smooth out their trades over time but instead trade only once. The issues discussed in our paper are relevant for all theoretical models regardless of their specific modeling assumptions.

Several recent empirical studies confirm empirical predictions based on invariance. Kyle and Obizhaeva (2016) find scaling laws for sizes and trading costs of portfolio transition orders executed in the U.S. stock market. Kyle and Obizhaeva (2017) document scaling laws for quoted bid ask spreads and the number of trades for Russian and U.S. stocks. Kyle, Obizhaeva and Tuzun (2017) discuss scaling laws for the number of trades and their sizes in tick-by-tick transaction data in the U.S. stock market. Bae et al. (2014) document scaling laws in the number of switching points in transactions in the South Korean stock market. Kyle et al. (2010) find scaling laws in the number of news articles about the U.S. firms. Andersen et al. (2015) find scaling laws in intraday data for U.S. stock index futures. All of these scaling laws have empirical exponents close to the values of one-third and two-thirds predicted by the model. Consistent with invariance of pricing accuracy, Farboodi, Matray and Veldkamp (2017) show that actively traded stocks are more accurate predictors of future earnings than the prices of smaller, less actively traded stocks. A common invariant structure, hidden in the data, seems to be revealed by these studies examining data from different angles.

This paper is structured as follows. Section 1 presents a setup with a dynamic model of trading. Section 2 introduces a number of market characteristics and reviews market microstructure invariance. Section 3 derives the approximate linear solution. Section 4 discusses how to derive invariance relationships in the context of this equilibrium model. Section 5 shows that many scaling laws can be derived based on a simple four-equation meta-model. Section 6 discusses invariance implications for market efficiency, pricing accuracy, and resiliency. Section 7 discusses characteristics of steady state, approximations, an exactly linear model, and other issues. Section 8 concludes. Appendix A discusses approximations. Appendix B contains proofs.

1 Setup of a Dynamic Model of Trading

This section describes the assumptions and defines the equilibrium for a dynamic model of sequential speculative trading.

Setup. There are three types of traders: informed traders, noise traders, and market makers. They exchange a risky asset for a risk-free numeraire asset whose returns are normalized to zero. The unobserved fundamental value of a risky asset $F(t)$ evolves over time due to continuous unmodeled changes in production processes, consumer tastes, costs of materials, prices of outputs, competitor strategies, and other market conditions. Let $F(t)$ follow a geometric Brownian motion with fundamental volatility σ_F,

$$F(t) := F_0 \cdot \exp\left(\sigma_F \cdot B(t) - \tfrac{1}{2} \cdot \sigma_F^2 \cdot t\right), \tag{1}$$

where $B(t)$ denotes a standardized Brownian motion with $B(t+h)-B(t) \sim \mathcal{N}(0,h)$ for $t \geq 0$ and $h \geq 0$, $B(0)$ is normally distributed, and the initial value F_0 is a known constant. The term $\frac{1}{2} \cdot \sigma_F^2 \cdot t$ adjusts for convexity so that the fundamental value follows an exponential martingale. Trading takes place until some distant date at which traders receive a payoff equal to the fundamental value. Let N denote outstanding shares.

Traders arrive into the market sequentially at endogenous times t_n, for $n = 1, 2, \ldots$. Let $\mathcal{T}(t) = \{t_1, \ldots, t_n : t_n < t\}$ denote the set of all transaction times before time t. Each trader anonymously places a bet by announcing a quantity to trade $Q(t_n)$, trades once at price $P(t_n^+)$, and then does not trade again. The other side of the bet is taken by market makers. The *trading history* $\mathcal{H}(t)$ is the public information set consisting of past trade times, quantities, and prices: $\mathcal{H}(t) := \{\langle t_n, Q(t_n), P(t_n^+)\rangle : t_n \in \mathcal{T}(t)\}$. Market makers, informed traders, and noise traders observe this history. Let $\mathrm{E}_t[\ldots]$ and $\mathrm{Var}_t[\ldots]$ denote the expectation and variance operators conditional on $\mathcal{H}(t)$, which excludes information about a bet possibly arriving at date t itself.

A trader may be either informed or uninformed. Informed traders arrive into the market at rate $\gamma_I(t)$, and noise traders arrive at rate $\gamma_U(t)$. The combined arrival rate $\gamma(t) := \gamma_I(t) + \gamma_U(t)$ is an instantaneous arrival rate which depends on the history of trading up to time t and varies over time, even between bet arrival times. If $\gamma(t)$ changes very little over some time interval, then the waiting time between bets has approximately an exponential distribution over this interval, implying $\mathrm{E}_{t_n}[t_{n+1} - t_n] \approx 1/\gamma(t_n)$. Market microstructure invariance, which we discuss at the end of Section 2, implements the intuition that the expected arrival rate of bets $\gamma(t)$ sets the pace of business time in the market, and markets differ from one other due to differences in the speed with which trading evolves.

Instead of focusing on the market's estimate of the fundamental value itself, it is more convenient to focus on the estimates of the Brownian motion $B(t)$, in terms of which the fundamental value is defined in equation (1). Let $\bar{B}(t)$ denote the market's conditional expectation of $B(t)$:

$$\bar{B}(t) := \mathrm{E}_t[B(t)]. \tag{2}$$

The difference $B(t)-\bar{B}(t)$ measures the estimation error. Let $\Sigma(t)$ denote the conditional variance of $\sigma_F \cdot \big(B(t)-\bar{B}(t)\big)$:

$$\Sigma(t) := \mathrm{Var}_t\left[\sigma_F \cdot \big(B(t)-\bar{B}(t)\big)\right]. \tag{3}$$

All traders can infer $\bar{B}(t)$ and the scaled error variance $\Sigma(t)$ from the trading history.

Both informed traders and noise traders believe they are informed. Each trader pays an exogenously fixed cost $\bar{c}_I$ to observe a private signal $i(t)$. An informed trader's signal has exogenously fixed precision $\bar{\tau}$, with $0<\bar{\tau}<<1$. A noise trader's signal is pure noise; it contains no information about the difference between the current price and fundamental value. The private signal has the form

$$i(t) = \begin{cases} \bar{\tau}^{1/2} \cdot \dfrac{\sigma_F \cdot \big(B(t)-\bar{B}(t)\big)}{\Sigma^{1/2}(t)} + (1-\bar{\tau})^{1/2} \cdot Z_I(t) & \text{if an informed trader,} \\ Z_U(t) & \text{if a noise trader,} \end{cases} \tag{4}$$

The random variables $Z_I(t)$ and $Z_U(t)$ are pure noise assumed to be distributed $\mathcal{N}(0,1)$ independently from the trading history $\mathcal{H}(t)$ and fundamental value $F(t)$. The definition of $\Sigma(t)$ in equation (3) implies that signals $i(t)$ of both informed and noise traders have the same unconditional distribution $\mathcal{N}(0,1)$.

This particular specification for signals is important for obtaining invariance relationships. In equation (4), scaling $B(t)-\bar{B}(t)$ by the time-varying factor $\Sigma^{1/2}(t)\cdot\sigma_F^{-1}$ insures that an informed trader's signal has a constant signal-to-noise ratio $\bar{\tau}/(1-\bar{\tau})$. Each bet therefore incorporates the same amount of information into prices in the sense that it reduces the error variance by a constant fraction proportional to $\bar{\tau}$. Without such scaling, the percentage reduction in error variance would vary with pricing error. Invariance relationships require the reduction in error variance to be constant.[1]

Equilibrium. The trading strategy $\hat{Q}(t,i)$ determines the size of a bet at time t as a function of the trader's information $\mathcal{H}(t)$ and signal i. Informed traders are rational profit maximizers. Noise traders are irrational in a sense that they trade on noise as if they were informed. When a trader generates a signal $i(t_n)$ at date t_n, the trader places a bet of size $Q(t_n) := \hat{Q}\big(t_n\, ,\, i(t_n)\big)$.

The pricing rule $\hat{P}(t,Q)$ determines the price set by market makers at time t as a function the market makers' information $\mathcal{H}(t)$ and the size of an arriving bet Q. At any time t, the pricing

[1]For example, if $i(t) := \tau^{1/2}\cdot\big(B(t)-\bar{B}(t)\big)+(1-\tau)^{1/2}\cdot Z_I(t)$, then the signal-to-noise ratio of an informed trade would be equal to $\tau\cdot\Sigma(t)\cdot\sigma_F^{-2}/(1-\tau)$ and would depend on $\Sigma(t)/\sigma_F^2$, with a smaller reduction in error variance when $\Sigma(t)$ is larger or σ_F is smaller. Kyle, Obizhaeva and Wang (2018) show that this particular way to model information—in which each informative signal reduces error variance by a constant fraction $\bar{\tau}$—can be naturally extended to continuous information flow.

rule $\hat{P}(t,.)$ defines the limit order book. When a bet of size $Q(t_n)$ arrives at time t_n, it is executed at trade price $P(t_n^+) = \hat{P}\big(t_n\,,\,Q(t_n)\big)$.[2]

Define a conditional expected paper-trading profit function

$$\hat{\pi}(t,i,Q) := \mathrm{E}_t\Big[\big(F(t)-P(t)\big)\cdot Q\,\Big|\,\text{informed signal } i(t)=i\Big], \tag{5}$$

which expresses a trader's expected profits from trading quantity Q at time t given information $\mathcal{H}(t)$ with pre-trade benchmark mid-price $P(t)$ and signal i believed to be informative.

Define the dollar price impact cost function

$$\hat{C}_B(t,Q) := \big(\hat{P}(t,Q)-P(t)\big)\cdot Q, \tag{6}$$

which expresses the dollar cost of executing a bet of arbitrary quantity Q placed at time t conditional on $\mathcal{H}(t)$. Perold (1988) calls this measure of transaction costs expected *implementation shortfall*; it compares the actual execution price $\hat{P}(t,Q)$ with the pre-trade benchmark $P(t)$ under the assumption that entire bet Q is executed. Since adverse selection makes bets move prices, the execution price $\hat{P}(t,Q)$ for trading Q is different from the pre-trade mid-price $P(t)$.

Definition 1. An *equilibrium* is a dynamic trading strategy $\hat{Q}(t,.)$, a pricing rule $\hat{P}(t,.)$, an arrival rate for informed traders $\gamma_I(t)$, and an arrival rate for noise traders $\gamma_U(t)$, all in the information set $\mathcal{H}(t)$, such that the following four conditions hold for all dates $t>0$:

1. *Profit Maximization*: The trading strategy $\hat{Q}(t,.)$ maximizes a trader's expected trading profits at time t, net of market impact costs:

$$\hat{Q}(t,i) = \arg\max_{Q}\ \Big[\hat{\pi}(t,i,Q)-\hat{C}_B(t,Q)\Big]. \tag{7}$$

2. *Market Efficiency*: The pricing rule $\hat{P}(t,.)$ defines a price equal to the conditional expectation of the fundamental value, given public information $\mathcal{H}(t)$ available before time t and information contained in a bet of size Q:

$$\hat{P}(t,Q) = \mathrm{E}_t\Big[F(t)\,\Big|\,\hat{Q}\big(t\,,\,i(t)\big)=Q\Big]. \tag{8}$$

3. *Free Entry*: At any time t, net of information cost $\bar{c}_I$ and market impact costs, informed

[2] Equation (4) not only describes signals generated at trade dates $t=t_n$ but also describes signals that could have been generated at non-trade dates $t \neq t_n$, in which case the trade size would have been $\hat{Q}\big(t\,,\,i(t)\big)$ and the price would have been $\hat{P}\big(t\,,\,\hat{Q}\big(t\,,\,i(t)\big)\big)$ if a trade had occurred.

and noise traders expect to break even if they buy a signal and then trade on it optimally:

$$\bar{c}_I = \mathrm{E}_t\Big[\hat{\pi}\big(t\,,\,i(t)\,,\,\hat{Q}\big(t\,,\,i(t)\big)\big) - \hat{C}_B\big(t\,,\,\hat{Q}\big(t\,,\,i(t)\big)\big)\Big]. \tag{9}$$

4. *Noise Trader Turnover*: Noise traders are expected to trade a rate which turns over the float N at exogenous rate η at all dates t:

$$\gamma_U(t)\cdot \mathrm{E}_t\Big[\big|\hat{Q}\big(t\,,\,i(t)\big)\big|\Big] = \eta\cdot N. \tag{10}$$

The profit maximization condition incorporates the assumption that both informed traders and noise traders are strategic and risk neutral; they believe themselves to be informed with probability one. The market efficiency condition incorporates the assumption that market makers are competitive and risk neutral, trading at prices which earn zero profits conditional on public information, including the size of the arriving bet. It is also based on the assumption that market makers do not know whether they trade with an informed trader or noise trader. The free entry condition says not only that traders break even when they trade but also that would have broken even if they had traded at times when they did not trade. The noise trader turnover condition implies that noise traders randomly choose when to trade and might trade at any time.

Linear Approximations. The assumptions that bets are drawn from a mixture of informed and noise traders and the function mapping the Brownian motion $B(t)$ into liquidation value $F(t)$ is nonlinear would make equilibrium $\hat{Q}(t,i)$ and $\hat{P}(t,Q)$ nonlinear in i and Q, respectively. Since intuition suggests that equilibrium may be almost—but not quite—linear, we will work with approximate linear equilibria in which the trading strategy and pricing rule are linear.

Therefore, let $\beta(t)$ and $\lambda(t)$ be stochastic processes, depending on information $\mathcal{H}(t)$, which define a linear trading strategy of the form

$$\hat{Q}(t,i) = \beta(t)\cdot i \tag{11}$$

and a linear pricing rule of the form

$$\hat{P}(t,Q) = P(t) + \lambda(t)\cdot Q. \tag{12}$$

The justification for linear trading strategies and pricing rule is based on the following two approximations. First, the prior conditional error $\sigma_F\cdot\big(B(t)-\bar{B}(t)\big)$ is approximately normally distributed. Second, each valuation update to $P(t)$ is approximately linear in the signal. We discuss these approximations in more detail in Section 7.3.

Definition 2. An *approximate linear equilibrium* is described by four randomly time-varying quantities $\beta(t)$, $\lambda(t)$, $\gamma_I(t)$, and $\gamma_U(t)$, all depending on $\mathcal{H}(t)$, such that a linear trading strategy $\hat{Q}(t,i)$ of the form (11) and a linear pricing rule $\hat{P}(t,Q)$ of the form (12) satisfy the four conditions for an equilibrium as approximations.[3]

In an approximate linear equilibrium, market makers take the other side of a bet of size $Q(t) = \beta(t) \cdot i(t)$ at adjusted price $P(t^+) = P(t) + \lambda(t) \cdot Q(t)$. The price impact is linear in $Q(t)$, and $\lambda(t)$ is an endogenous parameter measuring linear price impact. In an approximate linear equilibrium, the market efficiency condition and the law of iterated expectations imply that $P(t)$ is approximately a martingale.

We will show that there exists a unique approximate linear equilibrium, which can be easily characterized in closed form. In this equilibrium, market microstructure invariance conjectures hold exactly.

2 Market Characteristics and Microstructure Invariance

The main purpose of this paper is to show that market microstructure invariance can arise endogenously in an equilibrium model of adverse selection. In the paradigm of invariance, all markets work in a similar way in the sense that traders play the same trading game; yet, in different markets this game evolves at a different pace reflected in market liquidity.

This section defines several endogenous market characteristics like dollar volume, volatility, trading activity, price impact, liquidity, pricing accuracy, and resiliency, which vary greatly across assets and across time. It then briefly reviews the main assumptions of market microstructure invariance and discusses how a particular scaling of characteristics to adjust for existing differences in the pace of business time can make them similar across markets.

Volume $V(t)$ and volatility $\sigma(t)$. For formulating invariance principles, the important concepts are instantaneous expected volume and volatility. Define expected instantaneous trading *volume* $V(t)$ by

$$V(t) := \lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \mathrm{E}_t \left[\sum_{n: t \le t_n \le t+\Delta t} \left| Q(t_n) \right| \right] \approx \gamma(t) \cdot \mathrm{E}_t \left[\left| Q(t) \right| \right]. \tag{13}$$

Instantaneous expected *dollar volume* is $P(t) \cdot V(t)$.

[3] The specific way in which traders make approximations shows up as equation (37) in Appendix 3.1.

Define instantaneous expected *returns variance* $\sigma^2(t)$ as the product of the rate at which bets are expected to arrive and the contribution the price impact of each bet is expected to make to returns variance:

$$\sigma^2(t) := \lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \mathrm{Var}_t\left[\frac{P(t+\Delta t) - P(t)}{P(t)}\right] \approx \gamma(t) \cdot \mathrm{E}_t\left[\left(\frac{\lambda(t) \cdot Q(t)}{P(t)}\right)^2\right]. \tag{14}$$

Volatility $\sigma(t)$ is the square root of returns variance $\sigma^2(t)$. In an approximate linear equilibrium, percentage returns variance $\mathrm{Var}[\Delta P(t)/P(t)]$ and log returns variance $\mathrm{Var}\big[\ln\big(P(t+\Delta t)/P(t)\big)\big]$ are approximately the same.[4]

Trading Activity $W(t)$. Another important concept is a measure of calendar-time *trading activity* $W(t)$, defined as the product of dollar volume $P(t) \cdot V(t)$ and volatility $\sigma(t)$:

$$W(t) := P(t) \cdot V(t) \cdot \sigma(t). \tag{15}$$

It measures the standard deviation of the dollar change in the mark-to-market value of an entire day's trading volume; this is an empirical measure of the rate at which the market transfers risks.[5]

Trading activity is a good observable measure of risk transfer. It takes into account that assets differ in how risky they are. For low-volatility assets, even a large dollar volume may ultimately correspond only to an insignificant amount of risk transferred. Unlike share volume $V(t)$, trading activity is neutral with respect to splits. Unlike dollar trading volume $P(t) \cdot V(t)$ and share volume $V(t)$, trading activity $W(t)$ is leverage neutral: If a firm increases its leverage by paying out a debt-financed cash dividend per share that is equal to half of the stock price, then the value of the stock halves and its return volatility $\sigma(t)$ doubles. Since each share represents the same risk and dollar volume $P(t) \cdot V(t)$ halves, trading activity $W(t)$ remains unchanged if risk transfer does not change. This is why $W(t)$ is a better measure of trading activity than notional value or share volume. The concept of leverage neutrality, an essential feature of the invariance framework, is further discussed by Kyle and Obizhaeva (2017).

Expected costs $C_B(t)$ and profits $\pi(t)$. Let $C_B(t)$ denote the expected dollar price impact cost of executing a bet of optimal size at time t, conditional on past information $\mathcal{H}(t)$ but uncondi-

[4]We ignore the remote possibility that a gigantic negative bet Q might lead to negative prices if $P(t)+\lambda(t)\cdot Q < 0$.

[5]Trading activity has units of dollars $\times$ days$^{-3/2}$. Since business time has units of days^{-1}, mapping trading activity $W(t)$ into business time $\gamma(t)$ will ultimately require scaling trading activity by a dollar denominated quantity—related to trading costs or the cost of private information in the economic model—and then taking a 2/3 power. This is where the scaling parameters of 1/3 and 2/3 come from in the scaling laws.

tional on the new signal $i(t)$:

$$C_B(t) := \mathrm{E}_t\left[\hat{C}_B\big(t\,,\,\hat{Q}\big(t\,,\,i(t)\big)\big)\right]. \tag{16}$$

Let $\pi(t)$ denote the expected paper-trading profits of a trader at time t, conditional on past information $\mathcal{H}(t)$ but unconditional on the new signal $i(t)$:

$$\pi(t) := \mathrm{E}_t\Big[\hat{\pi}\big(t\,,\,i(t)\,,\,\hat{Q}\big(t\,,\,i(t)\big)\big)\Big]. \tag{17}$$

Liquidity Measure $L(t)$. Market practitioners measure liquidity in basis points. Define *illiquidity* $1/L(t)$ as the average dollar cost of executing bets $C_B(t)$ in equation (16) as a fraction of expected dollar pretrade bet size $\mathrm{E}_t\Big[\big|P(t)\cdot Q(t)\big|\Big]$:

$$\frac{1}{L(t)} := \frac{C_B(t)}{\mathrm{E}_t\Big[\big|P(t)\cdot Q(t)\big|\Big]}. \tag{18}$$

This dimensionless quantity measures the dollar-volume-weighted expected price impact cost of executing a bet as a fraction of the expected dollar value traded. It quantifies the average transaction costs for the market. For example, if the average dollar cost of executing bets is $C_B(t) = \$2000$ and the average bet is one million dollars, then $1/L(t) = 0.0020$ is a dimensionless fraction with the interpretation that dollar-weighted average market impact costs are 20 basis points. Market *liquidity* $L(t)$ can vary greatly across markets even though dollar market impact costs $C_B(t)$ are the same.

Pricing Accuracy $\Sigma^{-1/2}(t)$. Prices fluctuate around fundamentals. If prices are above fundamentals, then prices will tend to decrease over time towards fundamentals, as trading gradually incorporates private information into prices. If prices are below fundamentals, then informed trading will tend to make prices increase over time in the direction of fundamentals.

Recall that the standard deviation of the pricing error $\Sigma^{1/2}(t)$ is defined as $\mathrm{Var}^{1/2}\left[\sigma_F\cdot\big(B(t)-\bar{B}(t)\big)\right]$. Given our assumption about the conditional error being approximately a normal random variable in equation (3), we have

$$\begin{aligned}
P(t) &= \mathrm{E}_t\big[F(t)\big] \\
&= F_0\cdot\exp\big(\sigma_F\cdot\bar{B}(t)\big)\cdot\mathrm{E}_t\Big[\exp\Big(\sigma_F\cdot\big(B(t)-\bar{B}(t)\big)-\tfrac{1}{2}\cdot\sigma_F^2\cdot t\Big)\Big] \\
&\approx F_0\cdot\exp\Big(\sigma_F\cdot\bar{B}(t)+\tfrac{1}{2}\cdot\Sigma(t)-\tfrac{1}{2}\cdot\sigma_F^2\cdot t\Big).
\end{aligned} \tag{19}$$

The law of iterated expectations implies that $P(t)$ is approximately a martingale.

From equations (1) and (19), we find that $\Sigma^{1/2}(t)$ measures the standard deviation of the log-difference between fundamental value $F(t)$ and prices $P(t)$:

$$\Sigma^{1/2}(t) = \mathrm{Var}_t^{1/2}\Big[\ln\Big(\frac{F(t)}{P(t)}\Big)\Big]. \tag{20}$$

It has the interpretation as a log-percentage *pricing error*, or the average percentage difference between observable prices and unobservable fundamentals. Its reciprocal $\Sigma^{-1/2}(t)$ measures *pricing accuracy*. If the price deviates further from fundamental value, it is less accurate: the pricing error $\Sigma^{1/2}(t)$ is larger, and pricing accuracy $\Sigma^{-1/2}(t)$ is smaller.

Market Resiliency $\boldsymbol{\rho(t)}$. Define *market resiliency* $\rho(t)$ as the rate at which the estimation error $B(t)-\bar{B}(t)$ decays over time. Market resiliency measures the speed with which a random shock to prices—or estimation error resulting from execution of an uninformative bet—dies out over time as informative bets drive prices back towards fundamental value. Let $B_{\text{err}}(t) := B(t)-\bar{B}(t)$ denote the unobserved error at time t. Of course, we have $\mathrm{E}_t[B_{\text{err}}(t)] = 0$ by definition. In an approximate linear equilibrium, conditional expectations are approximately linear; therefore, resiliency $\rho(t)$ can be defined as the linear regression of innovations in the unobserved estimation error $B_{\text{err}}(t+\Delta t) - B_{\text{err}}(t)$ on its most recent unobserved level $B_{\text{err}}(t)$:[6]

$$\mathrm{E}_t\left[B_{\text{err}}(t+\Delta t) - B_{\text{err}}(t) \,|\, B_{\text{err}}(t)\right] \approx -\rho(t)\cdot B_{\text{err}}(t)\cdot \Delta t \qquad \text{for small } \Delta t. \tag{21}$$

The instantaneous half-life of the price impact of a noise trade is approximately $\ln(2)/\rho(t)$.

We will show in Section 6 that the concept of pricing error is closely related to the concept of market resiliency. They are two sides of the same coin; resiliency $\rho(t)$ is greater in markets with higher pricing accuracy $\Sigma^{-1/2}(t)$.

Moment ratios of Bet Sizes $\boldsymbol{|Q(t)|}$. Define the moment ratio $m(t)$, relating expected unsigned bet size $\mathrm{E}_t\big[|Q(t)|\big]$ and the standard deviation of signed bet size $\left(\mathrm{E}_t[Q^2(t)]\right)^{1/2}$, as

$$m(t) = \frac{\mathrm{E}_t\big[|Q(t)|\big]}{\left(\mathrm{E}_t[Q^2(t)]\right)^{1/2}}. \tag{22}$$

In the context of invariance, this particular moment ratio is important because the moment of $Q(t)$ in the numerator is related to trading volume $V(t)$ while the moment of $Q(t)$ in the denominator is related to returns volatility $\sigma(t)$.

[6] In equation (21), Δt denotes a small time interval, not the arrival time between bets. Between bet arrivals, $\rho(t)$ is approximately constant but actually increases slightly as pricing error $\Sigma^{1/2}(t)$ increases slightly.

Define the constant $\bar{m}$ as

$$\bar{m} := \mathrm{E}_t\left[\left|i(t)\right|\right]. \tag{23}$$

Since bets are linear in signals $i(t)$ in an approximate linear equilibrium and $\mathrm{E}_t[i^2(t)] = 1$, the model implies $m(t) = \bar{m}$.

Invariance Conjectures as Empirical Hypotheses. Market microstructure invariance is a collection of empirical hypotheses describing how expected bet size, bet arrival rate, trading costs, pricing accuracy, and resiliency depend on dollar volume and returns variance.

Broadly speaking, market microstructure invariance is the hypothesis that markets look the same when examined in business time. The rate of bet arrivals sets the business-time clock, specific for each market. In active, liquid markets bets arrive at a fast rate; in an inactive, illiquid markets bets arrive at a slow rate. Bets, also called meta-orders, may be executed as block trades in a dealer market or shredded into many small trades executed over time on securities exchanges. Our model describes a dealer market in which trades $Q(t_n)$ correspond to bets, and business time passes at rate $\gamma(t)$.

The starting point for market microstructure invariance is a set of two empirical conjectures about distributions of bet sizes and transaction costs functions in trading games.

Invariance conjectures begin with measuring the risk transferred by bets in business time. The dollar size of a bet is $P(t)\cdot Q(t)$, and the return standard deviation per unit of business time is $\sigma(t)/\gamma^{1/2}(t)$. The risk transferred by a bet per unit of business time, denoted $I(t)$, can be defined by

$$I(t) := P(t)\cdot Q(t)\cdot \frac{\sigma(t)}{\gamma^{1/2}(t)}. \tag{24}$$

The quantity $I(t)$ has units of dollars. Conditional on $\mathcal{H}(t)$, the quantities $P(t)$, $\sigma(t)$, and $\gamma^{1/2}(t)$ are known, but $Q(t)$ is random because $i(t)$ is random; thus, $I(t)$ is random as well.

Bet size invariance hypothesizes that the probability distribution of the risk transferred by a bet, $I(t)$, is invariant across markets and across time, when measured in dollars and in business time. This means there exists some invariant random variable I^* such that $I(t) \overset{d}{=} I^*$ for all t. Since $Q(t) = \beta(t)\cdot i(t)$ and $i(t)$ has a mean of zero and variance of one, bet size invariance implies that trading intensity $\beta(t)$ changes endogenously over time so that bets on average transfer the same dollar risks in business time.

Transaction cost invariance hypothesizes that the expected dollar price impact cost of executing a bet is an invariant function of the dollar risk transferred by the bet per unit of business time. This means there exists an invariant price impact cost function $C^*(.)$ such that $\hat{C}_B\big(t\ ,\ Q(t)\big) = C^*\big(I(t)\big)$ where $I(t) \equiv P(t)\cdot Q(t)\cdot \frac{\sigma(t)}{\gamma^{1/2}(t)}$ with probability one for all t. Price impact cost functions are invariant across markets and across time if costs are measured in dollars rather than basis points and if order sizes are measured in terms of dollar risks they transfer in business time rather than nominal dollar value or shares.

We also introduce two new invariance principles related to the accuracy and resiliency of prices. These invariance principles require an economic model such as ours, in which bets and prices convey information.

First, *invariance of pricing accuracy* hypothesizes that the pricing error standard deviation is invariant when scaled by volatility in business time. This means that $\Sigma^{1/2}(t)$ is proportional to $\sigma(t)/\gamma^{1/2}(t)$ with an invariant constant of proportionality. Pricing accuracy $\Sigma^{-1/2}(t)$ is inversely proportional to the standard deviation of the price impact of one bet $\sigma(t)/\gamma^{1/2}(t)$ in business time. It takes the same number of bets for prices to catch up with fundamentals, under the assumption that fundamentals will not be changing over time.

Second, *invariance of market resiliency* hypothesizes that market resiliency $\rho(t)$ is invariant in business time. This means that $\rho(t)$ is proportional to $\gamma(t)$ with an invariant proportionality constant.

Implied Scaling Laws and Invariant Parameters. From the invariance conjectures about bet sizes and transaction costs, Kyle and Obizhaeva (2016) derive a number of scaling laws for how bet size, number of bets, market depth, bid-ask spread, and other variables of interest relate to the products of dollar volume $V(t) \cdot P(t)$ and returns volatilities $\sigma(t)$ with different powers of one-third and two-thirds. It is also possible to derive similar scaling laws for pricing accuracy and resiliency. We prove in Section 4 that the invariance conjectures as well as the implied scaling laws are endogenous implications of an approximate linear equilibrium.

The scaling laws are exact implications of the assumption that a small subset of the exogenous parameters are *invariant* in the sense that they do not vary across time. The two most important invariant parameters are the cost of a signal $\bar{c}_I$ and the precision of an informative signal $\bar{\tau}$. These particular parameters are of obvious importance in an economic model of costly private information. Invariance of these two parameters implies that the cost of private information per unit of precision, $\bar{c}_I/\bar{\tau}$, is invariant. Invariance of these parameters also implies that private information arrives in chunks which cost $\bar{c}_I$.

The model also has two important but less visible implicit invariant parameters. The first parameter is $\bar{m}$, defined as $\bar{m} = \mathrm{E}_t\left[\left|i(t)\right|\right]$ in equation (23). Since $i(t) \sim \mathcal{N}(0,1)$, we have $\bar{m} = \sqrt{2/\pi} \approx 0.7979$. For various other distributions, $\bar{m}$ can have any value such that $0 < \bar{m} \leq 1$.[7] The

[7]Under different distributions for $i(t)$, Jensen's inequality implies $0 < \bar{m} \leq 1$. The maximum value $\bar{m} = 1$ is attained if and only if $i(t)$ is a binomial random variable with equally likely values of $+1$ and -1. We do not replace $\bar{m}$ with its value implied by a normal distribution, because the solution to the model does not depend on the normality assumption except for its effect on $\bar{m}$. Our results do not change if the distribution of signals $i(t)$ is changed to a different distribution with mean of zero and variance of one.

model essentially hardwires invariance of the moment ratio for bet sizes $m(t)$, since $m(t) = \bar{m}$. To make it clear that our results do not change if the distribution of signals $i(t)$ is changed to a different distribution with mean of zero and variance of one, we keep $\bar{m}$ as an invariant parameter.

The second invariant parameter also relates to private information. The assumption that informed and noise traders are risk neutral will lead to the implication that they trade to incorporate half of their private information into prices. To make it clear that our derivation of invariance properties depends only on traders incorporating an invariant fraction of their private information into prices and not on the particular fraction $1/2$ that is implied by risk neutrality, we implicitly generalized the equilibrium concept slightly by making the assumption that traders multiply their optimal risk neutral quantity by the fraction $2 \cdot \bar{\theta}$ and therefore incorporate a fraction $\bar{\theta}$ of their private information into prices. The parameter $\bar{\theta}$ is an invariant exogenous parameter, with baseline value $\bar{\theta} = 1/2$ corresponding to explicit model assumptions.[8]

Some exogenous parameters are not important for generating invariance hypotheses. For example, the model assumes that shares outstanding N, noise trader turnover rate η, and fundamental volatility σ_F are constant across time. The invariance conjectures would still hold if these parameters varied over time. To distinguish exogenous parameters which are important for obtaining invariance from exogenous parameters which are not important, we place a bar over the important parameters $\bar{c}_I$, $\bar{\tau}$, $\bar{m}$, and $\bar{\theta}$ and omit a bar from the other parameters N, η, and σ_F.[9]

As stock prices change, the market itself changes. When the price—and therefore market capitalization—increases significantly, dollar trading volume increases, traders arrive more frequently and place larger bets, market resiliency is higher, returns volatility is higher than fundamental volatility, trading incorporates information into prices faster than fundamental uncertainty is unfolding, the pricing error variance is shrinking, the market is becoming more liquid, and price dynamics quickly converges to a conditional steady state. When prices and market capitalization are falling, trading volume is falling, traders are arriving less frequently, traders are placing smaller bets, market resiliency and returns volatility are falling, trading is not incorporating information into prices as fast as fundamental uncertainty is unfolding, the pricing error variance is widening, the market is becoming less liquid, and the price dynamics may remain far from the conditional steady state for an extended periods of time.

[8] See equation (40) below. The reader can think of $\bar{\theta}$ as an abbreviation for the fraction $1/2$. Using the notation $\bar{\theta}$ is a device for keeping track of how the invariant parameter $\bar{\theta} = 1/2$ affects the equilibrium.

[9] Nothing in the model changes if N, η, and σ_F are possibly stochastic functions of time. Using the notation N, η, and σ_F instead of $N(t)$, $\eta(t)$, and $\sigma_F(t)$ is a simple device for distinguishing exogenous from endogenous parameters. The absence of a time parameter t indicates an exogenous variable.

3 Characterization of Approximate Linear Equilibrium

It is straightforward to characterize the unique approximate linear equilibrium in closed form:

Theorem 1 (Characterization of Approximate Linear Equilibrium)**.** *There exists a unique approximate linear equilibrium characterized by the four endogenous parameters $\lambda(t)$, $\beta(t)$, $\gamma_I(t)$, $\gamma_U(t)$, which are the following functions of the state variables $P(t)$, $\Sigma(t)$ and the exogenous parameters $\bar{\tau}$, $\bar{c}_I$, $\bar{m}$, $\bar{\theta}$, σ_F, η, and N:*

$$\lambda(t) = \frac{\bar{\theta}\cdot(1-\bar{\theta})\cdot\bar{\tau}}{\bar{c}_I}\cdot P^2(t)\cdot\Sigma(t), \tag{25}$$

$$\beta(t) = \frac{\bar{c}_I}{(1-\bar{\theta})\cdot\bar{\tau}^{1/2}}\cdot\frac{1}{P(t)\cdot\Sigma^{1/2}(t)}, \tag{26}$$

$$\gamma_I(t) = \bar{\theta}\cdot\gamma(t), \qquad \gamma_U(t) = (1-\bar{\theta})\cdot\gamma(t), \qquad \textit{where} \qquad \gamma(t) = \frac{\bar{\tau}^{1/2}\cdot\eta\cdot N}{\bar{c}_I\cdot\bar{m}}\cdot P(t)\cdot\Sigma^{1/2}(t). \tag{27}$$

At times $t \neq t_n$ when bets do not arrive, the price $P(t)$ is constant, and error variance increases at the rate fundamental volatility unfolds: $\mathrm{d}\Sigma(t)/\mathrm{d}t = \sigma_F^2$. At times t_n when bets arrive, the price $P(t_n)$ and error variance $\Sigma(t_n)$ jump, following the difference equation system

$$P(t_n^+) = P(t_n) + \bar{\theta}\cdot\bar{\tau}^{1/2}\cdot P(t_n)\cdot\Sigma^{1/2}(t_n)\cdot i(t_n), \tag{28}$$

$$\Sigma(t_n^+) \approx \Sigma(t_n)\cdot\left(1-\bar{\theta}^2\cdot\bar{\tau}\right), \qquad \textit{with} \qquad \Sigma(t_n) = \Sigma(t_{n-1}^+) + \sigma_F^2\cdot(t_n - t_{n-1}). \tag{29}$$

Corollary 1. *In a approximate linear equilibrium, the endogenous variables $V(t)$, $\pi(t)$, $C_B(t)$, and $m(t)$ are the following functions of the exogenous parameters η, N, $\bar{\theta}$, $\bar{m}$, and $\bar{c}_I$:*

$$V(t) = \frac{\eta\cdot N}{1-\bar{\theta}}, \qquad \pi(t) = \frac{\bar{c}_I}{1-\bar{\theta}}, \qquad C_B(t) = \bar{c}_B := \frac{\bar{\theta}}{1-\bar{\theta}}\cdot\bar{c}_I, \qquad m(t) = \bar{m}. \tag{30}$$

The endogenous variables $\mathrm{E}_t\left[|Q(t)|\right]$, $\mathrm{E}_t\left[Q^2(t)\right]$, $\gamma(t)$, $\gamma_I(t)$, $\gamma_U(t)$, $1/L(t)$, and $\rho(t)$ vary randomly through time as the following functions:

$$\mathrm{E}_t\big[\left|Q(t)\right|\big] = \frac{\bar{c}_I\cdot\bar{m}}{(1-\bar{\theta})\cdot\bar{\tau}^{1/2}}\cdot\frac{1}{P(t)\cdot\Sigma^{1/2}(t)}, \tag{31}$$

$$\mathrm{E}_t\big[Q^2(t)\big] = \frac{\bar{c}_I^2}{(1-\bar{\theta})^2\cdot\bar{\tau}}\cdot\frac{1}{P^2(t)\cdot\Sigma(t)}, \tag{32}$$

$$\frac{1}{L(t)} = \frac{\bar{\theta} \cdot \bar{\tau}^{1/2}}{\bar{m}} \cdot \Sigma^{1/2}(t), \tag{33}$$

$$\sigma^2(t) = \frac{\bar{\theta}^2 \cdot \bar{\tau}^{3/2}}{\bar{c}_I \cdot \bar{m}} \cdot \eta \cdot N \cdot P(t) \cdot \Sigma^{3/2}(t), \tag{34}$$

$$\rho(t) = \frac{\bar{\theta}^2 \cdot \bar{\tau}^{3/2}}{\bar{c}_I \cdot \bar{m}} \cdot \eta \cdot N \cdot P(t) \cdot \Sigma^{1/2}(t). \tag{35}$$

The endogenous quantities in the model are all functions of the two state variables $P(t)$ and $\Sigma(t)$, which change randomly due to arrival of bets and realization of fundamental uncertainty. Before discussing the intuition of this equilibrium, we outline the main steps of the proof and point out some of its interesting properties. The rest of this section proves Theorem 1. Details of the proof of Theorem 1 and proofs of Corollary 1 are presented in Appendices B.1 and B.2.

3.1 Proof of Theorem 1

We first derive a key equation for each of the four equilibrium conditions. The solution of these four equations implies the values for $\beta(t)$, $\lambda(t)$, $\gamma_I(t)$, and $\gamma_U(t)$ in equations (25), (26), and (27). We then discuss the two-variable difference-equation system (28) and (29).

1. Profit Maximization. A strategic informed trader chooses a bet size $Q(t)$ to maximize his profits net of market impact costs by solving the problem

$$Q(t) = \arg\max_Q \mathrm{E}_t\Big[\big(F(t) - P(t)\big) \cdot Q - \hat{C}_B(t, Q) \;\Big|\; i(t)\Big]. \tag{36}$$

The conditional estimate of the fundamental value using private signal $i(t)$ and the history of prices, including the most recently observed price, is

$$\mathrm{E}_t\Big[F(t) - P(t) \;\Big|\; \text{informed } i(t)\Big] \approx P(t) \cdot \bar{\tau}^{1/2} \cdot \Sigma^{1/2}(t) \cdot i(t). \tag{37}$$

In an approximate linear equilibrium, price impact satisfies $\hat{C}_B(t, Q) = \lambda(t) \cdot Q(t)$. The linear pricing rule makes the problem quadratic:

$$Q(t) = \arg\max_Q \Big[\big(\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t) \cdot i(t) - \lambda(t) \cdot Q\big) \cdot Q\Big]. \tag{38}$$

The solution to the linear first-order condition is

$$Q(t) = \beta(t) \cdot i(t), \qquad \text{where} \qquad \beta(t) = \frac{\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t)}{2 \cdot \lambda(t)}. \tag{39}$$

Equation (39) says that the informed trader trades to incorporate exactly one half of his information into prices. Generalizing the definition of equilibrium and assuming that the trader incorporates a fraction $\bar{\theta}$ of his information into price, with $0 < \bar{\theta} < 1$, changes equation (39) to the first key equation

$$\beta(t) = \frac{\bar{\theta} \cdot \bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t)}{\lambda(t)}. \tag{40}$$

This approach accommodates the possibility that 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. The generalization makes it possible to show that the invariance results derived below do not depend on the specific value $\bar{\theta} = 1/2$ implied by risk neutral profit maximization.

2. Pricing Rule. Conditional on observing a bet of size $Q(t)$, market makers infer that the bet has a probability $\gamma_I(t)/\gamma(t)$ of being informed and a probability $\gamma_U(t)/\gamma(t)$ of being noise. This inference follows from the fact that informed bets and noise bets arrive anonymously and are drawn from the same unconditional distribution $\mathcal{N}(0, \beta^2(t))$. Market makers can infer the signal $i(t)$ from size of the bet $Q(t) = \beta(t) \cdot i(t)$. Since the price update implied by an informed bet is $\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t) \cdot i(t)$ and the price update implied by a noise bet is zero, the market makers update prices as

$$\begin{aligned} &\mathrm{E}_t\big[F(t) - P(t) \bigm| Q(t)\big] \\ &\quad = \frac{\gamma_I(t)}{\gamma(t)} \cdot \mathrm{E}_t\big[F(t) - P(t) \bigm| \text{informed } Q(t)\big] + \frac{\gamma_U(t)}{\gamma(t)} \cdot \mathrm{E}_t\big[F(t) - P(t) \bigm| \text{noise } Q(t)\big]. \\ &\quad = \frac{\gamma_I(t)}{\gamma(t)} \cdot \bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t) \cdot i(t). \end{aligned} \tag{41}$$

This implies the pricing rule $\hat{P}(t, Q) = P(t) + \lambda(t) \cdot Q$, where $\lambda(t)$ satisfies the second key equation

$$\lambda(t) = \frac{\gamma_I(t)}{\gamma_I(t) + \gamma_U(t)} \cdot \frac{\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t)}{\beta(t)}. \tag{42}$$

3. Free Entry. The free entry condition says that the expected profits of both informed traders and noise traders, net of market impact costs $C_B(t) = \mathrm{E}_t[\lambda(t) \cdot Q^2(t)]$ and costs of information $\bar{c}_I$, are equal to zero. Plugging the optimal demand (40) into the maximized profits of traders, net of price impact costs, then using $\mathrm{E}_t[i^2(t)] = 1$, yields the third key equation

$$\frac{\bar{\theta} \cdot (1 - \bar{\theta}) \cdot \big(\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t)\big)^2}{\lambda(t)} = \bar{c}_I. \tag{43}$$

The expected trading profits, calculated before the signal $i(t)$ is realized and net of transaction costs, must equal the cost of information $\bar{c}_I$.

Since a noise trade may occur at any time, traders must be indifferent between trading and not trading at any time as well. Thus, equation (43) must hold at all times, both when trades occur and when trades do not occur. Intuitively, it implies that market liquidity adjusts continuously to make informed and noise traders indifferent between trading and not trading.

4. Noise Traders. Noise traders generate share volume at rate $\gamma_U(t) \cdot \mathrm{E}_t\left[\left|Q(t)\right|\right] = \eta \cdot N$. Since $Q(t) = \beta(t) \cdot i(t)$ and $\mathrm{E}_t\left[\left|i(t)\right|\right] = \bar{m}$, the expected size of a bet is

$$\mathrm{E}_t\left[\left|Q(t)\right|\right] = \beta(t) \cdot \bar{m}. \tag{44}$$

This implies the fourth key equation

$$\gamma_U(t) = \frac{\eta \cdot N}{\beta(t) \cdot \bar{m}}. \tag{45}$$

Solution of Four-Equation System. The four key log-linear equations (40), (42), (43), and (45) involve only the four endogenous parameters $\beta(t)$, $\lambda(t)$, $\gamma_I(t)$, and $\gamma_U(t)$ given in Theorem 1; exogenous parameters; and the state variables $P(t)$ and $\Sigma(t)$. These four equations can be easily solved for the endogenous parameters to obtain equations (25), (26), and (27) as follows: (1) Solve equation (43) for $\lambda(t)$. (2) Use the result to solve equation (40) for $\beta(t)$. (3) Use the solutions for $\beta(t)$ to solve equation (45) for $\gamma_U(t)$. (4) Use the solutions for $\lambda(t)$, $\beta(t)$, and $\gamma_U(t)$ to solve equation (42) for $\gamma_I(t)$. The solution is an approximate linear equilibrium because the trader's second order condition ($\lambda(t) > 0$) holds.

State Variables $\boldsymbol{P(t)}$ and $\boldsymbol{\Sigma(t)}$. In an approximate linear equilibrium, the state variables $P(t)$ and $\Sigma(t)$ are sufficient statistics for describing the market's information at date t. The price $P(t)$ simultaneously describes the market's estimate of fundamental value $F(t)$ and noise trader dollar volume $\eta \cdot N \cdot P(t)$.[10] Pricing error variance $\Sigma(t)$ is a natural way to describe the importance of adverse selection in the market.

When bets do not arrive, the market obtains no new information about fundamental value and therefore the price $P(t)$ is constant, but fundamental uncertainty continues to unfold so that $\mathrm{d}\Sigma(t) = \sigma_F^2 \cdot \mathrm{d}t$. When a bet arrives, the price changes by $\lambda(t_n) \cdot \beta(t_n) \cdot i(t_n)$, and the error variance $\Sigma(t_n)$ is reduced by fraction $\bar{\theta}^2 \cdot \bar{\tau}$. If market makers could tell whether each bet was informed or uninformed, then the error variance would decrease by fraction τ with probability

[10] If the model were changed to make noise trader share volume $\eta \cdot N$ randomly time varying, the equilibrium would change only cosmetically, but price $P(t)$ and $\eta \cdot N \cdot P(t)$ would become two separate state variables.

θ when an informed bet arrived and would remain unchanged with probability $1-\theta$ when an uninformed bet arrived; the percentage reduction would on average be $\theta \cdot \tau$, not $\theta^2 \cdot \tau$. Since market makers cannot distinguish informed bets from uninformed bets, the price impact of a bet is smaller (multiplied by additional factor θ) and the proportional variance reduction is only $\theta^2 \cdot \tau$, as reflected in equation (29).

The solutions for $\lambda(t)$ and $\beta(t)$ in equations (25) and (26) imply equations (28) and (29). The price follows a martingale with stochastic returns volatility $\sigma(t)$, which depends on $P(t)$ and $\Sigma(t)$, both of which are stochastic. Thus, $\sigma^2(t)$ is stochastic even though the innovation variance of fundamentals σ_F^2 is constant. This completes the proof of Theorem 1. The proof of Corollary 1 is in Appendix B.2.

3.2 Intuition and Properties of Equilibrium

Figure 1 illustrates the intuition. Informed traders strategically incorporate a fraction $\bar{\theta}$ of their information into prices by trading $Q(t)$, and the price jumps by $\bar{\theta} \cdot \mathrm{E}_t\big[F(t) - P(t) \mid Q(t)\big]$, which is equal to $\lambda(t) \cdot Q(t)$. Informed traders incur transaction costs $C_B(t)$ and expect to make trading profits of $\pi(t) - C_B(t)$ as the price gradually converges to expected fundamental value $\mathrm{E}_t\big[F(t) \mid Q(t)\big]$ due to the subsequent trading of other informed traders. These profits are realized at some distant date when the game ends and all positions are liquidated at the expected fundamental value. In contrast, noise traders execute orders which also incur expected dollar transaction costs $C_B(t)$, but they lose money since, on average, after their trades the price converges back to pre-trade levels.

The solution is characterized by two break-even conditions and a third property related to the market efficiency condition.

First, market makers break even on average. As Treynor (1971) describes, the expected losses market makers incur trading with informed traders $\pi(t) - C_B(t)$ must on average be equal to their expected gains from trading with noise traders $C_B(t)$. Since informed traders and noise traders arrive at a rate $\gamma_I(t)$ and $\gamma_U(t)$, respectively, this leads to the equilibrium condition

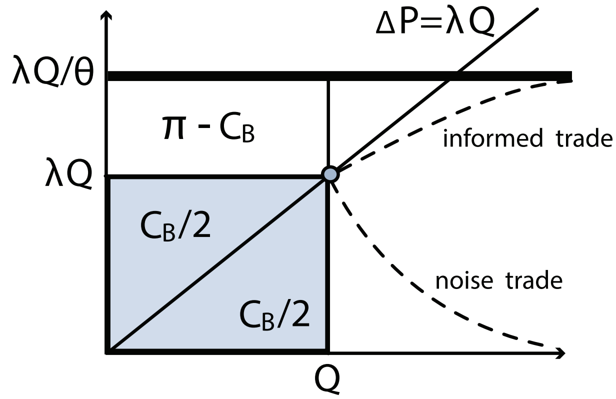
$$\gamma_I(t) \cdot \big(\pi(t) - C_B(t)\big) = \gamma_U(t) \cdot C_B(t). \tag{46}$$

Second, the free entry condition implies the break-even condition for traders

$$\bar{c}_I + C_B(t) = \pi(t). \tag{47}$$

On average, expected profits have to cover costs of obtaining a signal and executing a bet.

Figure 1: Intuition for the Model.



The figure illustrates price dynamics in response to arrival of a bet. The horizontal axis measures the size of a bet $Q(t)$, and the vertical axis represents its price impact $\Delta P(t) = \lambda(t) \cdot Q(t)$. In the equilibrium, the trader executes $Q(t) = \beta(t) \cdot i(t)$ After informed bets, the price continues to increase to fully incorporate the information content of the bet into prices to $\lambda(t) \cdot Q(t)/\bar{\tau}$. After noise bets, the price reverses to its initial level. Both noise and informed traders pay $\hat{C}_B\big(t\,,\,Q(t)\big)$ as market impact costs to market makers, but informed traders also earn trading profits of $\hat{\pi}\big(t\,,\,Q(t)\big)$.

Third, interestingly, profit maximization (40) and the pricing rule (42) do not imply a solution for $\beta(t)$ and $\lambda(t)$; the system is overdetermined. Instead, these two equations imply that the fraction of informed traders $\gamma_I(t)/\gamma(t)$ is equal to the fraction of the informed trader's information which is incorporated into prices $\bar{\theta} = 1/2$:[11]

$$\frac{\gamma_I(t)}{\gamma_I(t) + \gamma_U(t)} = \bar{\theta}. \tag{48}$$

Since $\bar{\theta}$ and $\bar{c}_I$ are constant, the above three equations imply the invariance of expected dollar costs:

$$C_B(t) = \bar{c}_B := \frac{\bar{\theta}}{1-\bar{\theta}} \cdot \bar{c}_I, \tag{49}$$

For the baseline case $\bar{\theta} = 1/2$, an approximate linear equilibrium implies that the price impact cost of a bet $C_B(t) = \bar{c}_B$ is exactly equal to the cost of a signal $\bar{c}_I$. The invariance of the moment

[11]In the continuous model of Kyle (1985), the price impact parameter $\lambda(t)$ is not identified from the market efficiency conditions either. Instead, market depth is pinned down by a condition stating that all volatility results from trading, or equivalently, that the error variance disappears by the end of the game, since the informed trader has pushed prices all the way to fundamental value.

ratio $m(t) = \bar{m}$ is almost hardwired into the model due to linearity of demand. The invariance of both $C_B(t)$ and $m(t)$ are essential for the non-obvious scaling laws described in Section 4.

4 Invariance Theorem

We have shown that the solution can be presented in terms of the two state variables $P(t)$ and $\Sigma(t)$. Even though it is easy to observe the price, the error variance is hard to estimate. This makes it difficult to use equations (25)–(35) as a basis for operational quantitative predictions about financial variables. In this section, we show how to solve this problem by expressing the model's predictions in terms of easily observable variables $P(t)$, $V(t)$, and $\sigma(t)$; this ultimately leads to market microstructure invariance.

The core result of this paper is that both the invariance conjectures and scaling laws hold in an approximate linear equilibrium. While the model hardwires linearity and normal distributions, the invariance hypotheses and scaling laws they imply are non-obvious implications of the model's assumptions. To state this in a self-contained way, we summarize notation and state the result as a theorem.

The invariant parameters are the cost of a signal $\bar{c}_I$, the precision of a signal $\bar{\tau}$, the fraction $\bar{\theta}$ of information in signal $i(t)$ incorporated into prices by an informed trade, and the moment ratio $\bar{m} = E_t\left[\left|i(t)\right|\right]$. The baseline model assumes $\bar{\theta} = 1/2$ for risk neutral traders and $\bar{m} = \sqrt{2/\pi} \approx 0.7979$ for normally distributed signals $i(t)$.

Theorem 2 (Invariance in an Approximate Linear Equilibrium)**.** *Bet size invariance holds in the sense that the dollar risk $I(t)$ transferred by a bet per unit of business time has an invariant distribution $\bar{c}_B \cdot i(t)$, where $i(t) \sim \mathcal{N}(0,1)$ and $\bar{c}_B$ is the invariant expected cost $C_B(t)$ of executing a bet:*

$$I(t) = \frac{P(t)\cdot Q(t)\cdot\sigma(t)}{\gamma^{1/2}(t)} = \frac{Q(t)}{V(t)}\cdot W^{2/3}(t)\cdot(\bar{m}\cdot\bar{c}_B)^{1/3} = \bar{c}_B\cdot i(t), \tag{50}$$

$$C_B(t) = \lambda(t)\cdot \mathrm{E}_t\left[Q^2(t)\right] = \bar{c}_B := \bar{c}_I\cdot\frac{\bar{\theta}}{1-\bar{\theta}}. \tag{51}$$

Transaction cost invariance holds in the sense that the expected dollar cost $\hat{C}_B(Q,t)$ of executing a bet of size Q is an invariant quadratic function of the dollar risk I this bet transfers in units of business time:

$$\hat{C}_B(t,Q) := \frac{1}{\bar{c}_B}\cdot I^2, \qquad \text{where} \qquad I \equiv \frac{P(t)\cdot Q\cdot\sigma(t)}{\gamma^{1/2}(t)}. \tag{52}$$

The number of bets $\gamma(t)$, the size of bets $Q(t)$, market impact $\lambda(t)$, liquidity $L(t)$, pricing accuracy $\Sigma^{-1/2}(t)$, and market resiliency $\rho(t)$ are related to easily observable 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 scaling laws:

$$\left(\frac{W(t)}{\bar{m}\cdot\bar{c}_B}\right)^{2/3} = \gamma(t) = \left(\frac{\mathrm{E}_t\left[|Q(t)|\right]}{V(t)}\right)^{-1} = \left(\frac{\lambda(t)\cdot V(t)}{\sigma(t)\cdot P(t)\cdot\bar{m}}\right)^2 = \frac{\left(\sigma(t)\cdot L(t)\right)^2}{\bar{m}^2} = \frac{\sigma^2(t)}{\bar{\theta}^2\cdot\bar{\tau}\cdot\Sigma(t)} = \frac{\rho(t)}{\bar{\theta}^2\cdot\bar{\tau}}. \quad (53)$$

Proof. See Appendix B.3. □

Since trading activity $W(t) := P(t)\cdot V(t)\cdot\sigma(t)$ is observable, the scaling laws (53) provide a way to measure the number of bets $\gamma(t)$, the expected size of bets $Q(t)$, market impact $\lambda(t)$, liquidity $L(t)$, pricing accuracy $\Sigma^{-1/2}(t)$, and market resiliency $\rho(t)$ in terms of the $2/3$ power of $W(t)$ with $\bar{c}_B$, $\bar{m}$, and $\bar{\theta}^2\cdot\bar{\tau}$ as invariant proportionality coefficients. These equations directly correspond to the empirical hypotheses and scaling laws proposed in Kyle and Obizhaeva (2016).[12] The empirical predictions about pricing accuracy and resiliency are new. The pricing error $\Sigma^{1/2}(t)$ is proportional to $\sigma(t)/W^{1/3}(t)$, and resiliency $\rho(t)$ is proportional to $W^{2/3}(t)$.

The first four scaling laws for number of bets $\gamma(t)$, the size of bets $Q(t)$, market impact $\lambda(t)$, and liquidity $L(t)$ require invariance of the market impact cost of a bet $\bar{c}_B = \bar{\theta}\cdot\bar{c}_I/(1-\bar{\theta})$ and the moment ratio of bet sizes $\bar{m}$. Scaling laws for pricing accuracy $\Sigma^{-1/2}(t)$ and market resiliency $\rho(t)$ additionally require invariance of the informativeness of a bet $\bar{\theta}^2\cdot\bar{\tau}$.

The parameters $\bar{c}_B$, $\bar{m}$, and $\bar{\theta}^2\cdot\bar{\tau}$ can be estimated empirically as the intercepts in regressions of logs of the corresponding variables on logs of trading activity. For example, equation (53) implies that the number of bets $\gamma(t)$ is proportional to easily observable $W^{2/3}(t)$ with the proportionality coefficient $(\bar{m}\cdot\bar{c}_B)^{-2/3}$. Thus, one can generate quantitative predictions about $\gamma(t)$ if one either knows values of the parameters $\bar{m}$ and $\bar{c}_B$ or, alternatively, estimates the value of $(\bar{m}\cdot\bar{c}_B)^{-2/3}$ by regressing $\ln(\gamma(t))$ on $\ln\left(W^{2/3}(t)\right)$ and using the estimate of the intercept.

Although the model describes 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 $\bar{c}_I$, the shape of the distribution of signals $\bar{m}$, and the informativeness of bets $\bar{\theta}^2\cdot\bar{\tau}$ are constant across all markets. The possible economic mechanism is intuitive. Suppose that the cost of private signals $\bar{c}_I$ is proportional to the average wages of finance professionals, adjusted for their productivity or effort required to generate one bet. They optimally allocate skills across different markets to maximize the value of trading on the private signals that they generate. In equilibrium, the average

[12] Equation (50) directly corresponds to bet size invariance. Equation (52) directly corresponds to transaction costs invariance; equation (51) is the unconditional version of the same statement. Equation (53) summarizes empirical implications about bet arrival rate, bet size, and price impact. The bet arrival rate $\gamma(t)$ is proportional to $W^{2/3}(t)$; the size of bets as a fraction of volume $E_t\left[|Q(t)/V(t)|\right]$ is proportional to $W^{-2/3}(t)$; and market liquidity $L(t)$ is proportional to $W^{1/3}(t)/\sigma(t)$. Since $E_t\left[|i(t)|\right] = \bar{m}$ implies $E_t\left[|I|\right] = \bar{m}\cdot\bar{c}_B$ from (50), one can easily check that equations (B-6) for bet arrival rate $\gamma(t)$, (B-7) for bet size $Q(t)$, and (B-10) for illiquidity $1/L(t)$ are *exactly equivalent* to invariance equations (7), (8), and (15) in Kyle and Obizhaeva (2016).

cost of generating a private signal $\bar{c}_I$ is likely to be similar across markets.

Price, volatility, and volume are public, *macroscopic* quantities in the sense that, for a specific asset at a specific time, these quantities are aggregate statistics describing the interaction of all of the traders in the market, and their values can be estimated from aggregate market data. The distribution of bet size $Q(t)$, bet arrival rate $\gamma(t)$, the average cost of a bet $1/L(t)$, pricing accuracy $\Sigma^{1/2}(t)$, and resiliency $\rho(t)$, and the price impact or information content of individual bets are, by contrast, *microscopic* quantities in the sense that they are statistics describing individual bets, and their values are difficult to observe. Invariance helps to link together macroscopic and microscopic quantities.

The constants $\bar{c}_B$, $\bar{m}$, and $\bar{\theta}^2 \cdot \bar{\tau}$ in our structural model play the role somewhat similar to the role played by Boltzmann's constant or Avogadro's number in physics. Theoretical models help to fill in detail and connect these constants to deep parameters of the model.

5 A Four-Equation Meta-Model

Is all of the machinery of the dynamic model necessary to derive invariance relationships? We will show that an unconditional version of the invariance hypotheses describing mean bet size and average transaction costs relies on only four simple equations; therefore, only a subset of the dynamic model's structure is required.

We call these four structural properties a *meta-model* because these properties are likely to be shared by many other equilibrium models:

1. *Volume Equation*: Trading volume results from bets. Since bets of average size $\mathrm{E}_t\left[\left|Q(t)\right|\right]$ arrive at rate $\gamma(t)$, share trading volume $V(t)$ satisfies

$$\gamma(t) \cdot \mathrm{E}_t\left[\left|Q(t)\right|\right] = V(t). \tag{54}$$

2. *Volatility Equation*: The dynamic model implies that returns volatility results from the linear price impact of bets. Since one bet moves prices by $\lambda(t) \cdot Q(t)$ dollars and bets arrive at rate $\gamma(t)$, the calendar-time variance of dollar price change $\sigma^2(t) \cdot P^2(t)$ satisfies

$$\gamma(t) \cdot \lambda^2(t) \cdot \mathrm{E}_t\left[Q^2(t)\right] = \sigma^2(t) \cdot P^2(t). \tag{55}$$

3. *Price Impact Cost Equation*: Since each bet moves prices by $\lambda(t) \cdot Q(t)$ and thus incurs a price impact cost $\lambda(t) \cdot Q^2(t)$, the expected dollar price impact cost of a bet $C_B(t)$ satisfies

$$\lambda(t) \cdot \mathrm{E}_t\left[Q^2(t)\right] = C_B(t). \tag{56}$$

4. *Moment Equation*: Expected unsigned bet size $\mathrm{E}_t\left[\left|Q(t)\right|\right]$ and the standard deviation of signed bet size $\left(\mathrm{E}_t\left[Q^2(t)\right]\right)^{1/2}$ are related by a moment ratio $m(t)$ satisfying

$$\frac{\mathrm{E}_t\left[\left|Q(t)\right|\right]}{\left(\mathrm{E}_t\left[Q^2(t)\right]\right)^{1/2}} = m(t). \tag{57}$$

The four structural equations (54)–(57) define a meta-model in the sense that they define structural properties that may be shared by many models of market microstructure without filling in details which may differ across models. The second equation says that order flow moves prices; the three other equations are simply definitions.

All four meta-model equations hold in an approximate linear equilibrium. The volume equation assumes that market makers take the other side of each bet, so that $V(t)$ simultaneously measures buy volume, sell volume, and market maker volume; it is the same as equation (13). The volatility equation is consistent with linear price impact of bets, but equation (55) does not itself imply linear price impact because it is an unconditional assertion about price impact, not a conditional assertion; it is implied by equation (14). The price impact cost equation is consistent with price impact costs being quadratic, but equation (56) itself does not imply quadratic costs because it is an unconditional assertion about the variance of $Q(t)$, not a conditional assertion about its shape as a function of $Q(t)$; it is the same as equation (16). The moment equation defines $m(t)$ as a moment ratio depending on the shape, but not the scaling of the distribution of bet size; it is implied by equation (22).

As we discussed in Section 3.2, the dynamic model implies invariance of two variables $C_B(t) = \bar{c}_B :== \frac{\bar{\theta}}{1-\bar{\theta}} \cdot \bar{c}_I$ and $m(t) = \bar{m}$ in meta-model equations (56) and (57). The meta-model, combined with the invariance of $C_B(t)$ and $m(t)$, directly implies scaling laws.

Theorem 3 (Invariance and Meta-Model)**.** *If $C_B(t) = \bar{c}_B$ and $m(t) = \bar{m}$, then the four meta-model equations* (54), (55), (56)*, and* (57) *are a log-linear system which can be solved for the four parameters $\gamma(t)$, $\lambda(t)$, $\mathrm{E}_t\left[\left|Q(t)\right|\right]$, and $\mathrm{E}_t\left[Q^2(t)\right]$ in terms of $P(t)$, $V(t)$, $\sigma(t)$, $\bar{c}_B := \frac{\bar{\theta}}{1-\bar{\theta}} \cdot \bar{c}_I$, and $\bar{m}$, as in Theorem 2.*

Proof. See Appendix B.4. □

Except for the two equalities for pricing error $\Sigma(t)$ and resiliency $\rho(t)$, all other invariance results in Theorem 2 are derived based on the four meta-model equations (54)–(57) combined with the results $C_B(t) = \bar{c}_B$ and $m(t) = \bar{m}$. Invariance relationships therefore represent general properties inherent to many microstructure models of speculative trading. Yet, the two invariance relationships related to pricing accuracy and market resiliency require the full machinery of the dynamic model of adverse selection because these variables are missing from the meta-model equations.

The structural meta-model implies a particular relationship between the invariance of bet sizes and transaction costs hypotheses. Since the four meta-model equations reference the first and second moments of unsigned bet size but not other moments, the meta-model has implications related only to the first two moments of bet sizes.

First, if $C_B(t) = \bar{c}_B$ and $m(t) = \bar{m}$, then meta-model equations (55), (56), and (57) imply a specific connection between mean unsigned risk transfer $I(t)$ and invariant transaction cost $\bar{c}_B$:

$$\mathrm{E}_t\Big[\big|I(t)\big|\Big] = \bar{m} \cdot \bar{c}_B. \tag{58}$$

Intuitively, this economic restriction follows from the assumption that market makers break even.

Second, if $C_B(t) = \bar{c}_B$ and $m(t) = \bar{m}$, then meta-model equations (55) and (56) lead to another restriction that connects the second moment of the bet size invariant $\mathrm{E}_t\big[I^2(t)\big]$ and the cost invariant $\bar{c}_B$,

$$\mathrm{E}_t\big[I^2(t)\big] = \bar{c}_B^2. \tag{59}$$

Since signals, and therefore bets, are normally distributed, the dynamic model implies that I^* is distributed $\mathcal{N}(0\,,\,\bar{c}_B^2)$ and $\mathrm{E}_t\big[|I^*|\big] = \bar{m} \cdot \bar{c}_B$. This distribution is invariant because $\bar{c}_B$ and $\bar{m}$ are shown to be invariant constants.

The restrictions (58) and (59) impose a particular structure on the proportionality constants in invariance relationships. It is this structure that ultimately allows us to link to one another disconnected scaling relationships, derived in Kyle and Obizhaeva (2016), and write them in a consolidated one-line form of equation (53) in the invariance Theorem 2.

6 Market Efficiency, Pricing Accuracy, and Resiliency

There are two different definitions of market efficiency. Our model helps to clarify the sharp distinction between them.

Eugene Fama conceptualizes a market to be efficient if all available information is appropriately reflected in price; this implies that prices—adjusted for the risk-free rate, dividend yield, and risk premium—follow a martingale, regardless of how much information is available overall in the market. In our model, prices in equation (28) are always efficient in the sense of Fama's definition because prices are martingales with respect to public information.

Fischer Black (1986) conceptualizes market efficiency as the accuracy with which prices estimate fundamental value. In our model, pricing accuracy $\Sigma^{-1/2}(t)$ is directly related to this concept of market efficiency because its reciprocal measures the standard deviation of the log-distance between the fundamental value and the price. As pricing accuracy varies endogenously over time, the log-distance between prices and fundamentals may be either large or small; higher capitalization (higher $P(t)$) is associated with more bets and greater efficiency in the sense of Black's definition.

Black 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. The market becomes more efficient if the standard deviation of the log-distance $\Sigma^{1/2}(t)$ between observable prices and unobservable fundamentals becomes smaller. Since the probability that a normal distribution is within 1.64 standard deviations of its mean is approximately 90%, Black's conjecture holds formally when a 1.64 standard deviation event does not deviate from the mean by more than a factor of 2. In the context of our model, Black would say that markets are efficient if $\Sigma^{1/2}(t) < \ln(2)/1.64 \approx 0.42$.

It is convenient to scale the pricing error variance $\Sigma(t)$ by annual returns variance $\sigma^2(t)$ so that $\Sigma(t)/\sigma^2(t)$ quantifies the number of years by which the informational content of prices lags behind fundamental value given current level of returns volatility. If prices are less accurate and returns volatility is lower (larger $\Sigma(t)$ and smaller $\sigma^2(t)$), it takes more years for prices to catch up with fundamentals. For example, suppose a stock's annual volatility is $\sigma(t) = 0.35$ and $\Sigma^{1/2}(t) = \ln(2)/1.64$. Then, since $\Sigma(t)/\sigma^2(t) = (\ln(2)/1.64)^2/0.35^2 \approx 1.50$, this implies that prices are about 1.50 years behind fundamental value. On average, it would take about 1.50 years of 35% annual returns volatility for prices to converge to fundamental value under the counterfactual assumption that the fundamental value would remain frozen in time.

In practice, it is difficult to observe directly Black's measure of market efficiency $\Sigma^{-1/2}(t)$ because fundamental value is unobservable. Yet, it is possible to infer $\Sigma^{-1/2}(t)$ indirectly from a closely related, easier-to-observe measure of *market resiliency*, also discussed by Black.

Market resiliency $\rho(t)$ is the mean-reversion parameter (per calendar year) measuring the speed with which a random shock to prices—resulting from execution of an uninformative bet—dies out over time, as informative bets drive prices back towards fundamental values. If resiliency is approximately constant over a given time period, then the half-life of an uninformative shock to prices must be equal to $\ln(2)/\rho(t)$.

Black (1986) intuited that since transitory noise affects prices, returns variance is twice as large as innovations in fundamental value, and this implies mean reversion in returns. Black's intuition is incorrect because prices have a martingale property due to efficient pricing. In fact, his intuition applies to the log-ratio of prices to fundamental value. It is the difference between prices and fundamentals that exhibits mean reversion, not prices themselves. If prices become disconnected from fundamentals permanently, the presence of a bubble creates arbitrage opportunities for traders without private information. In our model, trading based on private information gradually drives prices towards fundamental value, preventing bubbles. Since prices in our model follow a martingale (by assumption), our model reflects intuition different from Shiller (1992), who describes an inefficient market with excess volatility and predictable mean reversion over long time periods.

More formally, equations (3), (27), (34), and (35) yield the following results for market resiliency $\rho(t)$, volatility $\sigma(t)$, and pricing error variance $\Sigma(t)$:

$$\Sigma(t) = \mathrm{Var}_t\left[\sigma_F \cdot \left(B(t) - \bar{B}(t)\right)\right], \tag{60}$$

$$\sigma^2(t) = \bar{\theta}^2 \cdot \bar{\tau} \cdot \Sigma(t) \cdot \gamma(t), \tag{61}$$

$$\rho(t) = \bar{\theta}^2 \cdot \bar{\tau} \cdot \gamma(t) = \frac{\sigma^2(t)}{\Sigma(t)}. \tag{62}$$

Equation (62) is important. It illustrates the relationship among these three variables. Market resiliency $\rho(t)$ is greater in markets with higher pricing accuracy $\Sigma^{-1/2}(t)$ and higher returns volatility $\sigma(t)$.

Equation (62) suggests an empirical strategy for calibration of pricing accuracy $\Sigma(t)$ from an estimate of resiliency $\rho(t)$. The latter can be obtained by examining how fast the temporary price impact of noise trades dies out over time. In the previous example, if volatility is 35% per year and $\Sigma^{1/2}(t) = \ln(2)/1.64$, then $\Sigma(t)/\sigma^2(t) \approx 1.50$ and prices are about 1.50 years behind the fundamental value. The error $B(t) - \bar{B}(t)$ in equation (21) (as well as prices $P(t)$) mean-reverts at rate $\rho(t) = \sigma^2(t)/\Sigma(t) = 0.35^2/(\ln(2)/1.64)^2 = 0.69$ per year. This implies that the half-life of the price impact of noise trades is equal to $\ln(2)/\rho(t) \approx 1$ year. Thus, Black (1986) could have equivalently defined an efficient market where "price is within a factor 2 of value" as a market where "the half-life of the price impact of noise trades is less than one year."

The empirical strategy of using $\rho(t)$ to infer pricing accuracy $\Sigma^{-1/2}(t)$ also makes it possible to infer the information content of one bet $\bar{\theta}^2 \cdot \bar{\tau}$. Equation (62) imply $\bar{\theta}^2 \cdot \bar{\tau} = \rho(t)/\gamma(t)$. To illustrate the concept, suppose it is known that Black's marginally efficient stock with $\rho(t) \approx 0.69$ per year has about 100 bets per day, or 25,000 bets per year based on 250 trading days per year. It immediately follows that $\bar{\theta}^2 \cdot \bar{\tau} \approx 0.69/25000 = 0.28 \times 10^{-4}$. From equation (29), the invariance of resiliency and pricing accuracy then implies that, in any market and at any time, one bet reduces the error variance of prices $\Sigma(t)$ by about 0.0028 percent.

Also, equation (33) suggests an alternative empirical strategy for calibration of pricing accuracy $\Sigma^{-1/2}(t)$ from an estimate of liquidity $L(t)$, where the latter can be proxied by percentage bid-ask spread:

$$\frac{1}{L(t)} = \left(\frac{\bar{\theta}^2 \cdot \bar{\tau}}{\bar{m}^2} \cdot \Sigma(t)\right)^{1/2}. \tag{63}$$

Pricing accuracy $\Sigma^{-1/2}(t)$ is higher when liquidity $L(t)$ is higher. If $\Sigma^{1/2}(t) = \ln(2)/1.64$, $\bar{m} = 0.7979$, and $\bar{\theta}^2 \cdot \bar{\tau} \approx 0.69/5000 = 0.28 \times 10^{-4}$, then the average percentage transaction costs are $1/L(t) = 0.0028$ or 28 basis points.

7 Model Discussion and Robustness

The model has a simple structure carefully designed to formulate empirical hypotheses relating the dynamics of market liquidity to the informativeness of prices. It has infinite horizon and small number of parameters with natural empirical interpretation. The model suggests an empirically implementable measure of liquidity which satisfies invariance hypothesis and changes gradually with slowly moving returns variance to dollar volume ratio.

7.1 Conditional Steady State.

The price $P(t)$ follows a martingale with stochastic returns variance $\sigma^2(t)$. The error variance of prices $\Sigma(t)$ has a drift and changes much more slowly than the price. It is possible to show that the percentage change of error variance has the following two moments

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \mathrm{E}_t\left[\frac{\Sigma(t+\Delta t) - \Sigma(t)}{\Sigma(t)}\right] = \frac{1}{\Sigma(t)} \cdot \left(\sigma_F^2 - \sigma^2(t)\right), \tag{64}$$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \mathrm{Var}_t\left[\frac{\Sigma(t+\Delta t) - \Sigma(t)}{\Sigma(t)}\right] = \bar{\theta}^2 \cdot \bar{\tau} \cdot \sigma^2(t). \tag{65}$$

Unfolding fundamental uncertainty increases error variance at rate σ_F^2, while information being incorporated into prices reduces it at rate $\sigma^2(t)$. When returns volatility $\sigma(t)$ is greater (smaller) than fundamental volatility σ_F, information is being incorporated into prices faster (slower) than fundamental uncertainty is unfolding, and pricing error variance $\Sigma(t)$ in equation (64) is shrinking (increasing) at a rate $\sigma_F^2 - \sigma^2(t)$. The variance is equal to $\bar{\theta}^2 \cdot \bar{\tau} \cdot \sigma^2(t)$, which is much smaller than the returns variance of prices $\sigma^2(t)$, because $0 < \bar{\theta}^2 \cdot \bar{\tau} << 1$.

If the two forces are in balance, then the pricing error variance remains constant at a level $\Sigma^*(t)$, which we call a *conditional steady state.*

$$\Sigma^*(t) := \frac{\sigma_F^2}{\gamma(t) \cdot \bar{\theta}^2 \cdot \bar{\tau}} = \frac{\sigma_F^2}{\bar{\theta}^2 \cdot \bar{\tau}} \cdot \left(\frac{\bar{c}_B \cdot \bar{m}}{\sigma_F \cdot P(t) \cdot V(t)}\right)^{2/3}. \tag{66}$$

The first equation can be proved by substitution of fundamental variance σ_F^2 for market return variance $\sigma^2(t)$ in equations (61). The second equation can be proved by substitution of fundamental variance σ_F^2 for market return variance $\sigma^2(t)$ in equations (34) and then plugging $\bar{c}_B$ and $V(t)$ from equation (30).

The conditional steady state $\Sigma^*(t)$ does not represent a steady state in the usual sense; it represents the level to which $\Sigma(t)$ would converge over time if market capitalization were not changing, as proxied by the price $P(t)$ since shares outstanding N do not change. Keeping σ_F fixed, more accurate signals $\bar{\theta}^2 \cdot \bar{\tau}$ and more frequent bets $\gamma(t)$ make steady-state error variance $\Sigma^*(t)$ smaller and market prices more accurate.

Looking at financial markets from a bird eye's view, our model presents the following picture of what happens when prices change. Changes in prices $P(t)$ immediately lead to changes in market capitalization $P(t) \cdot N$, changes in returns volatility $\sigma(t)$ in equation (34), and changes in the arrival rate of bets $\gamma(t)$ in equation (27). The value of $\Sigma(t)$ gradually drifts in equation (29) towards a conditional steady-state level of $\Sigma^*(t)$, which it is constantly chasing, but never fully converges to, since the steady-state level is itself constantly changing with changes in $P(t)$.

When price is high, corresponding to both dollar capitalization and dollar trading volume being high, then returns volatility is high, bets arrive quickly, and $\Sigma(t)$ moves quickly towards its conditional steady state level fast; returns volatility remains close to fundamental volatility; and $\Sigma(t)$ does not deviate far from its conditional steady-state level. When prices are low and dollar trading volume is low, bets arrive slowly and $\Sigma(t)$ adjusts only slowly towards its conditional steady-state level; returns volatility may remain below fundamental volatility for extended periods of time.

The properties of exponential martingales imply with probability one that (1) the values of $F(t)$ and $P(t)$ will eventually converge to zero, (2) both the bet arrival rate and returns volatility will eventually converge to zero, and (3) pricing error variance $\Sigma(t)$ will eventually become unboundedly large. The model makes realistic predictions that trading volume in any given stock eventually dies out, and at any point in time, much of the volume in the market consists of trading in a small number of active stocks. 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.

Invariance relationships are consistent with the following mechanism of market adjustment to shocks to market capitalization. Suppose that, beginning in a conditional steady state with $\sigma(t) = \sigma_F$, the price suddenly rises by a factor of 8. Theorem 1 shows what happens in the short run. The price change increases the arrival rate of bets $\gamma(t)$, return variance $\sigma^2(t)$, and market resiliency $\rho(t)$ by a factor of 8, but leaves market accuracy $\Sigma^{-1/2}(t)$, market liquidity $L(t)$, percentage market impact $\lambda(t)/P^2(t)$ and average dollar bet size $\mathrm{E}_t\left[\left|P(t) \cdot Q(t)\right|\right]$ initially unchanged, while reducing share bet size $Q(t)$ by a factor of 8 and increasing $\lambda(t)$ by a factor of 8^2. The effect of the increase in $\gamma(t)$, $\rho(t)$, and $\sigma^2(t)$ is balanced out by the drop in $Q(t)$ and increase in $\lambda(t)$, so that invariance relationships continue to hold.

In the long run, the high arrival rate of bets makes market prices more accurate through equation (29), eventually reducing pricing errors $\Sigma^{1/2}(t)$ by a factor of 2 ($8^{1/3} = 2$) to the new conditional steady-state level described in equation (66). In the new conditional steady state, pricing accuracy $\Sigma^{-1/2}(t)$, liquidity $L(t)$, and average bet size $\mathrm{E}_t\left[\left|P(t) \cdot Q(t)\right|\right]$ are 2 times higher than before ($8^{1/3} = 2$); the arrival rate of bets $\gamma(t)$ and resiliency slow down but remain 4 times higher than before ($8^{2/3} = 4$); returns variance is equal to its conditional steady state level $\sigma^2(t) = \sigma_F^2$. The invariance exponents of 1/3 and 2/3 for $\mathrm{E}_t\left[\left|P(t) \cdot Q(t)\right|\right]$ and $\gamma(t)$ are reflected in this new steady state with adjusted pricing accuracy.

Thus, the invariance holds both in the steady-state and outside of the steady state.

7.2 Liquidity.

Liquidity $L(t)$ can be expressed in two different ways. First, equation (33) implies that liquidity $L(t)$ is proportional to pricing accuracy $\Sigma^{-1/2}(t)$:

$$L(t) = \frac{\bar{m}}{\bar{\theta} \cdot \bar{\tau}^{1/2}} \cdot \Sigma^{-1/2}(t). \tag{67}$$

Liquidity is proportional to how much information has been incorporated into prices from past trading and has nothing to do with how fast information is being incorporated into prices at the current moment t. It suggests that liquidity should not vary a great deal over short periods of time because pricing error $\Sigma^{1/2}(t)$ changes gradually due to steadily unfolding fundamental volatility σ_F and only a small reduction in error variance $\Sigma(t)$ upon arrival of each bet.

Second, equation (53) implies that liquidity $L(t)$ can be expressed as a function of dollar volume $P(t) \cdot V(t)$ and variance $\sigma^2(t)$, expected at a particular moment in time:[13]

$$L(t) = \left(\frac{\bar{m}^2 \cdot P(t) \cdot V(t)}{\bar{c}_B \cdot \sigma^2(t)} \right)^{1/3}. \tag{68}$$

Expected returns variance $\sigma^2(t)$ measures how fast information is expected to be incorporated into prices at a particular point in time, and it is known to change over the trading day. Liquidity is not a function of fundamental volatility σ_F.

Equations (67) and (68), taken together, generate an important empirical prediction. For liquidity $L(t)$ to be relatively constant over time, even when volatility $\sigma(t)$ and prices $P(t)$ are time varying, the ratio of instantaneous expected dollar volume $P(t) \cdot V(t)$ to instantaneous

[13] Kyle and Obizhaeva (2016) and Kyle and Obizhaeva (2017) use the same expression for $L(t)$ as equation (68).

expected returns variance $\sigma^2(t)$ must vary slowly, preserving proportionality to slowly varying pricing accuracy $\Sigma^{-1/2}(t)$. By imposing this strong volume-volatility relationship on the equilibrium price discovery process, the dynamic model allows both equations (67) and (68) to be valid simultaneously.

Corollary 2. *Scaling laws* (53) *can be expressed in terms of* $L(t)$ *instead of trading activity* $W(t)$*:*

$$L^2(t) = \left(\frac{\bar{m}^2 \cdot P(t) \cdot V(t)}{\bar{c}_B \cdot \sigma^2(t)}\right)^{2/3} = \left(\frac{\bar{m}^2 \cdot W(t)}{\bar{c}_B \cdot \sigma^3(t)}\right)^{2/3} = \frac{\bar{m}^2 \cdot \gamma(t)}{\sigma^2(t)} = \left(\frac{\mathrm{E}\left[|P(t) \cdot Q(t)|\right]}{\bar{c}_B}\right)^2 = \frac{P^2(t) \cdot \bar{m}}{\lambda(t) \cdot \bar{c}_B} = \frac{\bar{m}^2}{\bar{\theta}^2 \cdot \bar{\tau}} \cdot \frac{1}{\Sigma(t)} = \frac{\bar{m}^2}{\bar{\theta}^2 \cdot \bar{\tau}} \cdot \frac{\rho(t)}{\sigma^2(t)}. \tag{69}$$

Proof. Equation (69) is a direct implication of equations (68) and (53). □

Corollary 2 says that $P(t) \cdot Q(t)$ is proportional to $L(t)$, $\gamma(t)$ is proportional to $\sigma^2(t) \cdot L^2(t)$, $\lambda(t)$ is proportional to $P^2(t)/L(t)$, $\Sigma^{1/2}(t)$ is proportional to $1/L(t)$, and $\rho(t)$ is proportional to $\sigma^2(t) \cdot L^2(t)$.

7.3 Approximations

To obtain a close-form solution, we make several assumptions involving approximations. First, we assume that the estimation error $B(t) - \bar{B}(t)$ is approximately normally distributed. This assumption makes the filtering problem of an informed trader linear when the signal of an informed trader is jointly normally distributed with the valuation error. This assumption is an approximation because each price increment is a mixture—not a sum or an average—of trades by either informed traders or noise traders.

Second, we assume that an informed trader chooses a quantity to trade which is linear in the estimate of the information content of the private signal. This assumption makes the quantity $Q(t)$ observed by market makers jointly normally distributed with the valuation error and thus justifies linear filtering by market makers. This assumption is an approximation because a linear approximation to the exponential function associated with geometric Brownian motion is used.

Third, we assume that the market makers choose a price impact parameter $\lambda(t)$ so that price impact is a linear function of the quantity traded $Q(t)$. This assumption makes price changes approximately normally distributed and justifies linear filtering. It is an approximation because the geometric Brownian motion assumption implies that price impact should be nonlinear.

For empirically reasonable parameter values describing publicly traded stocks with reasonably active trading volume, we believe that all of these approximations involve economically inconsequential errors. Proving this formally is a topic for future research, but simulations in Appendix A provide some supportive evidence.[14]

Our paper illustrates that invariance can be derived in the context of an equilibrium model. Empirical evidence is certainly more consistent with more general empirical hypotheses about bet sizes and transaction costs, rather than the properties of our structural linear-normal model. For example, Kyle and Obizhaeva (2016) find that the sizes of unsigned bets closely fit a log-normal distribution with log-variance of 2.50, not a normal distribution. A square root price impact model often predicts transaction costs better than a linear model, although both models predict transaction costs reasonably well if the linear model is supplemented with a constant bid-ask spread cost. We conjecture that it may 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, but this will make the model much less tractable.

7.4 An Exactly Linear Model.

Nonlinearity of the model raises the question whether a modified version of the model, which is exactly linear, would yield similar results. The following alternative assumptions describe a model with an exactly linear equilibrium:

- Fundamental value follows Brownian motion, not a geometric Brownian motion. This makes the valuation formula linear, with $F(t) = F_0 + F_1 \cdot \sigma_F \cdot B(t)$ and $P(t) = F_0 + F_1 \cdot \sigma_F \cdot \bar{B}(t)$ for some fixed constants F_0 and F_1. [15]

- Informed bets and noise bets arrive anonymously, in matched pairs, at non-stochastic time intervals $\Delta t = 1/\gamma_U(t) = 1/\gamma_I(t) = 2/\gamma(t)$. This changes the mixture of bets to their linear combination.

- Traders are risk-neutral monopolists, and they optimally incorporate exactly half of their information into prices, $\bar{\theta} = 1/2$.

[14]We choose to model fundamentals as a geometric Brownian motion rather than arithmetic Brownian motion because geometric Brownian motion makes it possible to describe the time series properties of the life cycle of a stock in an empirically realistic manner. For example, when a stock's market capitalization increases, returns volatility increases somewhat in the short run while pricing accuracy and liquidity increase in the long run. Since the meta-model equations are largely the same, both approaches generate invariance relationships. While a potential disadvantage of our approach is that we rely on linear approximations, we believe the approximate linear equilibrium is very close to an exact non-linear equilibrium.

[15]The new exogenous constant F_1, with units dollars/share, is needed so that the exogenous parameter σ_F continues to have the same units day$^{-1/2}$ as an our main model. Without loss of generality, one may assume $F_1 = 1$ dollar/share.

In the modified model, market makers observe two bets at the same time; they know one bet is informed and one is noise, but they do not know which is which. Under this batching assumption, the dollar pricing error $F_1 \cdot \sigma_F \cdot \left(B(t) - \bar{B}(t)\right)$ and the two bets are jointly normally distributed. This makes linear projections exactly the same as conditional expectations, not an approximation as has been assumed so far. The error $B(t) - \bar{B}(t)$ is exactly normally distributed, and conditional expectations are exactly linear.

In the modified model, the explicit solutions in Theorem 1 and Corollary 1 as well as the invariance results in Theorem 2 and the relation between returns variance, resiliency, and business time continue to hold almost exactly as before, except for some minor differences. These results and proofs are presented in Appendix B.5.

In the exactly linear model, the conditional steady state becomes an actual steady state. Over time, the value of $\Sigma(t)$ converges to the steady state value $\Sigma^* = \Sigma^*(t)$ given by equation (B-37) with volume $V(t) = 2 \cdot \eta \cdot N$ being a constant: In this steady state, trading intensity $\beta(t)$, market impact $\lambda(t)$, the bet arrival rate $\gamma(t)$, and dollar pricing error $\Sigma(t) \cdot P^2(t)$ are constants; bet size has an unchanging distribution $Q(t)$. Liquidity $L(t)$ and volatility $\sigma^2(t)$ change so that $P^2(t)/L(t)$ and $\sigma^2(t) \cdot P^2(t)$ remain constant.[16]

Invariance relationships no longer show up in the time series because endogenous parameters like $\gamma(t)$, $\beta(t)$, $\lambda(t)$ are constant, but they show up in the cross-section when different assets have different value for σ_F, η, and N. Since the exactly linear model is not significantly easier to describe than the approximate linear equilibrium, we have chosen to emphasize more realistic approximate linear equilibrium in the main part of our paper. This makes it easier to show why invariance relationships hold both outside of a conditional steady state and in a steady state, when endogenous parameters are not constant.

8 Conclusion

The dynamic structural model described in this paper is to be interpreted as a proof of concept that invariance hypotheses and scaling laws may be derived in the context of a reasonable, well-specified theoretical model of speculative trading based on adverse selection.

The derivation of invariance relationships relies mostly on the four meta-model equations (54)–(57). These equations capture generic properties of models of speculative trading: (1) or-

[16] Models with constant dollar volatility, common in a CARA-normal framework, formally imply that prices $P(t)$ may become negative or prices may become so high that volatility goes to zero. To deal with this possibility, it is straightforward to modify the model so that the firm accumulates or disposes of cash by having very deep-out-of-the-money rights offerings or paying dividends as needed. Adding these leverage changes to the model requires some accounting notation but does not change the underlying economics at all. In particular, changes in leverage do not change $W(t)$ and therefore invariance relationships continue to hold due to higher market capitalization offsetting lower volatility, and vice versa, when cash distributions or equity issuance change leverage.

der flow creates volume and induces volatility, (2) the expected dollar transaction costs of a bet are invariant across assets and time, and (3) the ratio of moments of bet size distributions is stable across assets and time. We therefore conjecture that more general invariance relationships can be obtained in the context of other market microstructure models as well.

References

Andersen, Torben G., Oleg Bondarenko, Albert S. Kyle, and Anna A. Obizhaeva. 2015. "Intraday Trading Invariance in the E-Mini S&P 500 Futures Market." Working Paper. Available at `http://dx.doi.org/10.2139/ssrn.2693810`.

Back, Kerry, and Shmuel Baruch. 2004. "Information in Securities Markets: Kyle Meets Glosten and Milgrom." *Econometrica,* 72: 433–465.

Bae, Kyounghun, Albert S. Kyle, Eun Jung Lee, and Anna A. Obizhaeva. 2014. "Invariance of Buy-Sell Switching Points." Working Paper. Available at `http://dx.doi.org/10.2139/ssrn.2730770`.

Black, Fischer. 1986. "Noise." *Journal of Finance,* 41(3): 529–543.

Fama, Eugene. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance,* 25(2): 383–417.

Farboodi, Maryam, Adrien Matray, and Laura Veldkamp. 2017. "Where Has All the Big Data Gone?" Working Paper. Available at `http://people.stern.nyu.edu/lveldkam/pdfs/FinEfficiencyFMV.pdf`.

Glosten, Lawrence R., and Paul R. Milgrom. 1985. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics,* 14(1): 71–100.

Kyle, Albert S. 1985. "Continuous Auctions and Insider Trading." *Econometrica,* 53(6): 1315–1335.

Kyle, Albert S., and Anna A. Obizhaeva. 2016. "Market Microstructure Invariance: Empirical Hypotheses." *Econometrica,* 84(4): 1345–1404.

Kyle, Albert S., and Anna A. Obizhaeva. 2017. "Dimensional Analysis, Leverage Neutrality, and Market Microstructure Invariants." Working Paper. Available at `http://dx.doi.org/10.2139/ssrn.2785559`.

Kyle, Albert S., Anna A. Obizhaeva, and Tugkan Tuzun. 2017. "Microstructure Invariance in U.S. Stock Market Trades." Working Paper. Available at SSRN: `https://ssrn.com/abstract=1107875` or `http://dx.doi.org/10.2139/ssrn.1107875`.

Kyle, Albert S., Anna A. Obizhaeva, and Yajun Wang. 2018. "Smooth Trading with Overconfidence and Market Power." *Review of Economic Studies*, 85: 611–662.

Kyle, Albert S., Anna A. Obizhaeva, Nitish Sinha, and Tugkan Tuzun. 2010. "News Articles and Equity Trading." Working Paper. Available at `http://dx.doi.org/10.2139/ssrn.1786124`.

LeRoy, Stephen F. 1989. "Efficient Capital Markets and Martingales." *Journal of Economic Literature*, 27(4): 1583–1621.

Perold, André F. 1988. "The Implementation Shortfall: Paper versus Reality." *Journal of Portfolio Management*, 14(3): 4–9.

Shiller, Robert. 1992. *Market Volatility.* MIT Press.

Treynor, Jack. 1995. "The Only Game in Town." *Financial Analysts Journal*, 51(1): 81–83. Reprint of Walter Bagehot (pseud.), 1971, "The Only Game in Town." *Financial Analysts Journal*, 22(2): 12–14, 22.

Appendix A Approximation for the Distribution of Errors

The approximate linear solution relies on the assumption that errors $\sigma_F \cdot \big(B(t) - \bar{B}(t)\big)$ are distributed approximately as a normal distribution $\mathcal{N}\big(0, \Sigma(t)\big)$. This appendix analyzes the robustness of this assumption using numerical simulations.

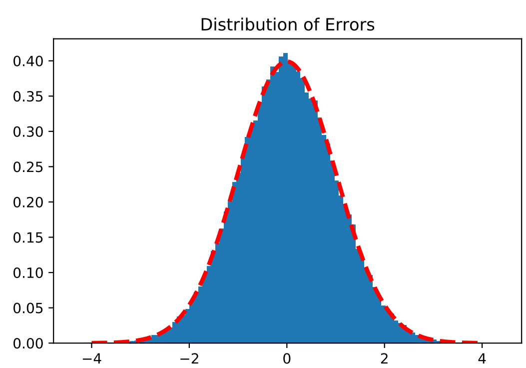
Assume the following about the model parameters: The annual volatility of fundamentals is $\sigma_F = 0.35$, the fraction of informed traders is $\bar{\theta} = 0.5$, the unconditional dollar costs of producing a signal is $\bar{c}_B = \$2000$, the moment ratio is $\bar{m} = 0.80$ corresponding to normal random variables, the number of shares outstanding is $N = 250$ million, the annualized turnover of noise traders is $\eta = 0.5$, and the precision of private information is $\bar{\tau} = 0.000552$, which is obtained as $\rho(t)/(\gamma(t) \cdot \bar{\theta}^2)$ under the assumption that annual resiliency satisfies $\rho(t) = 0.69$ and there are $\gamma(t) = 5000$ bets executed per year (i.e., 20 bets per day over 250 days).

We simulate 100,000 scenarios with the arrival of 50,000 bets, which approximately corresponds to a ten-year time period. As the initial starting point for each scenario at time t_0, we assume that fundamental $F(t_0)$ is \$40, price $P(t_0)$ is \$40, error variance $\Sigma(t_0)$ is 0.42^2, and the expected number of bets $\gamma(t_0)$ is 5000 per year.

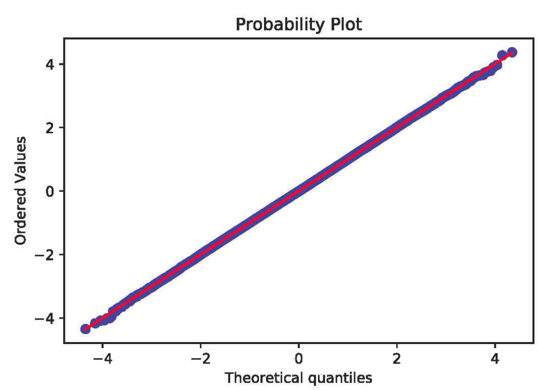
On the kth step, assume that a trader arrives after time $\Delta t = 1/\gamma(t_k)$. This trader is an informed or noise trader with equal probability $1/2$. If a trader is informed, then he observes a signal $i(t_k) = \frac{\bar{\tau}^{1/2}}{\Sigma^{1/2}(t_k)} \cdot \sigma_F \cdot \big(B(t_k) - \bar{B}(t_k)\big) + (1-\bar{\tau})^{1/2} \cdot Z_I(t)$, as defined in equation (4); using equations (1) and (19), the current error $\sigma_F \cdot \big(B(t_k) - \bar{B}(t_k)\big)$ can be inferred from current fundamentals and prices as $\ln\big(F(t_k)/P(t_k)\big) + 0.5 \cdot \Sigma(t_k)$. If a trader is a noise trader, then he observes a signal $i(t_k) = Z_U(t_k)$, as defined in equation (4). We next update the arrival rate of bets $\gamma(t_{k+1})$ using equation (27), the fundamental value $F(t_{k+1})$ using equation (1), the share price $P(t_{k+1})$ and the error variance $\Sigma(t_{k+1})$ using recursive equations (28) and (29).

After 100,000 of such updates, we calculate the final distribution of errors $\sigma_F \cdot \big(B(t_{K+1}) - \bar{B}(t_{K+1})\big)$ at time t_{K+1} as $\ln\big(F(t_{K+1})/P(t_{K+1})\big) + 0.5 \cdot \Sigma(t_{K+1})$ and then scale it by $\Sigma^{1/2}(t_{K+1})$. This simulated distribution of standardized errors is a proxy for the distribution of pricing errors, which we assume to be close to a standardized normal in our approximate linear solution.

The figure shows that the simulated distribution of steady-state scaled errors between prices and fundamentals $\sigma_F \cdot \big(B(t_{K+1}) - \bar{B}(t_{K+1})\big)/\Sigma^{1/2}(t_{K+1})$ indeed does not differ much from the standardized normal distribution. Panel A shows the histogram of these simulated scaled errors, and panel B shows the quantile-to-quantile plot of the simulated distribution against the standardized normal distribution with the zero mean and unit variance. Both figures suggest that the normal approximation is reasonable. Even the formal Kolmogorov–Smirnov test produces the p-value of $p = 0.51$ and does not reject the normality assumption. In our simulations,



Distribution of Scaled Errors



Quantile-Quantile Plot

the median error variance $\Sigma(t_{K+1})$ is equal to 0.0564, the median price is \$35.97, the median fundamental value is \$35.22, and the median number of bets is 15,682 per year. This is consistent with the median conditional steady-state error variance $\Sigma^*(t_{K+1}) := \frac{\sigma_F^2}{\gamma(t_{K+1}) \cdot \bar{\theta}^2 \cdot \bar{\tau}}$ in equation (66), which is equal to 0.0567. Since the median $\Sigma^{1/2}(t_{K+1})$ is equal to 0.23, and it is less than $\ln(2)/1.64$ or 0.42, the simulated market is efficient in the sense of Fischer Black.

Appendix B Proofs

B.1 Details of the Proof of Theorem 1

The proof of Theorem 1 in Section 3.1 relies on equation (37), which we prove below. When a trader observes a signal $i(t)$, he thinks that the signal is informative and linearly updates the estimate of $B(t)$ by

$$\Delta\bar{B}(t) := \mathrm{E}_t\Big[B(t) - \bar{B}(t) \;\Big|\; \text{informed } i(t)\Big] \approx \frac{\bar{\tau}^{1/2} \cdot \Sigma^{1/2}(t)}{\sigma_F} \cdot i(t). \tag{B-1}$$

Since linear filtering approximates a conditional expectation, the coefficient of $i(t)$ is a linear regression coefficient defined as the ratio of $\mathrm{Cov}_t\big[i(t)\,,\, B(t) - \bar{B}(t)\big]$ to $\mathrm{Var}_t[i(t)] = 1$.

Conditional on observations of $i(t)$, the difference $\sigma_F \cdot \big(B(t) - \bar{B}(t)\big)$ is distributed normally with the mean of $\sigma_F \cdot \Delta\bar{B}(t)$ and variance of $\Sigma(t) - \sigma_F^2 \cdot \mathrm{Var}_t\,[\Delta\bar{B}(t)]$. Then, the following chain of approximate equalities holds,

$$\begin{aligned}
\mathrm{E}_t\,[F(t) - P(t) \,|\, \text{informed } i(t)] &= P(t) \cdot \mathrm{E}_t\Big[\exp\Big(\sigma_F \cdot \big(B(t) - \bar{B}(t)\big) - \tfrac{1}{2} \cdot \Sigma(t)\Big) - 1 \,\Big|\, \text{informed } i(t)\Big] \\
&\approx P(t) \cdot \Big(\exp\Big(\sigma_F \cdot \Delta\bar{B}(t) - \tfrac{1}{2} \cdot \sigma_F^2 \cdot \mathrm{Var}[\Delta\bar{B}(t)]\Big) - 1\Big) \\
&\approx P(t) \cdot \sigma_F \cdot \Delta\bar{B}(t) + \text{error}.
\end{aligned} \tag{B-2}$$

The first line of this equation uses equations (1) for $F(t)$ and (19) for $P(t)$. The second line of the next equation then follows from $\mathrm{E}[\exp(x)] = \exp\big(\mathrm{E}[x] + \frac{1}{2}\mathrm{Var}[x]\big)$ when x is normally distributed. The second line is exact if $B(t) - \bar{B}(t)$ is exactly jointly normally distributed with the zero-mean informed signal $i(t)$. The third line is a Taylor series approximation to the exponential function which keeps second-order terms. The fourth line sets the second-order terms to zero using the approximation $\mathrm{Var}[\Delta\bar{B}(t)] \approx \Delta\bar{B}^2(t)$, which is exact in expectation. It implies that a revision $\Delta\bar{B}(t)$ to the estimate of $\bar{B}(t)$ changes prices $P(t)$ by approximately $P(t) \cdot \sigma_F \cdot \Delta\bar{B}(t)$.

Equations (B-2) and (B-1) imply that the trader's update to fundamental value is approximately linear in the signal $i(t)$:

$$\mathrm{E}_t\,[F(t) - P(t) \,|\, \text{informed } i(t)] \approx P(t) \cdot \bar{\tau}^{1/2} \cdot \Sigma^{1/2}(t) \cdot i(t). \tag{B-3}$$

The difference between an (exact nonlinear) equilibrium and an approximate linear equilibrium is that traders and market makers use the linear approximation in equation (B-3) instead of the exact, potentially nonlinear conditional expectation. This completes the proof.

B.2 Proof of Corollary 1

It is easy to prove equation (30): (1) Use equations (13), (11), (26), and (27) to solve for $V(t)$. (2) Since $\hat{C}_B(t,Q) = \lambda(t)\cdot Q(t) = \lambda(t)\cdot\beta(t)^2$, use equations (25) and (26) to solve for $C_B(t)$. (3) Since traders incorporate $\bar{\theta}$ of information into price, $\pi(t) = C_B(t)/\bar{\theta}$. (4) Use equations (22) and (23) to solve for $m(t)$.

The endogenous variables in equations (31)–(34) can be obtained as follows: (1) Solve equations (27) and (30) for $\mathrm{E}_t\big[|Q(t)|\big]$. (2) Solve equations (11) and (26) for $\mathrm{E}_t\big[Q^2(t)\big]$. (3) Solve equations (18), (30), and (31) for $L(t)$. (4) Solve equations (14), (27) and (32) for $\sigma^2(t)$.

To derive equation (35) for $\rho(t)$, write the changes in the unobserved estimation error $B_{\mathrm{err}}(t) := B(t)-\bar{B}(t)$ as

$$\begin{aligned}
&\mathrm{E}_t\left[B_{\mathrm{err}}(t+\Delta t)-B_{\mathrm{err}}(t)\,|\,B_{\mathrm{err}}(t)\,,\,i(t)\right] \approx -\theta\cdot\frac{\bar{\tau}^{1/2}\cdot\Sigma^{1/2}(t)}{\sigma_F}\cdot i(t)\\
&= -\theta\cdot\frac{\bar{\tau}^{1/2}\cdot\Sigma^{1/2}(t)}{\sigma_F}\cdot\left(\theta\cdot\bar{\tau}^{1/2}\cdot\frac{\sigma_F\cdot\big(B(t)-\bar{B}(t)\big)}{\Sigma^{1/2}(t)}+\theta\cdot(1-\bar{\tau})^{1/2}\cdot Z_I(t)+(1-\theta)\cdot Z_U(t)\right).
\end{aligned} \tag{B-4}$$

Normality is preserved because a mixture of normal variables with the same mean and variance has a normal distribution. Joint normality is not preserved because higher order co-moments are affected by the mixture of distributions; this makes results approximations. The first equation is similar to equation (B-1), except a factor θ reflects the probability of signal being informative. The second equation is obtained by using equation (4). Compare equations (21) and (B-4) and use $\Delta t = 1/\gamma(t)$ to get

$$\rho(t) = \theta^2\cdot\tau\cdot\gamma(t). \tag{B-5}$$

Then, plug in equation (27) to solve for $\rho(t)$. This completes the proof of the corollary.

B.3 Proof of Theorem 2

Equation (51) follows from equation (30). Equation (50) follows from equations (26), (27), and (34) using equation (11) and definition (15). Equation (52) follows from plugging equations (25), (14), and (50) into $\hat{C}_B(t,Q) = \lambda(t)\cdot Q^2(t)$. It is easy to prove equation (53) by using equations (25) through (35) and equation (15). This completes the proof.

B.4 Proof of Theorem 3

Invariance relationships in equation (53), formulated in terms of $C_B(t) = \bar{c}_B$ and $m(t) = \bar{m}$, can be derived based on the four structural economic equations (54)–(57), which define a meta-model. In this system of four equations, one can think of $\gamma(t)$, $\lambda(t)$, $E_t\left[Q^2(t)\right]$, and $E_t\left[\left|Q(t)\right|\right]$ as unknown variables to be solved for in terms of known variables $V(t)$, $P(t)$, $\sigma(t)$, $C_B(t) = \bar{c}_B$, and $m(t) = \bar{m}$.

Using the definition of trading activity $W(t) = P(t) \cdot V(t) \cdot \sigma(t)$, solve the system of four equations (54)–(57) for four unknowns $\gamma(t)$, $E_t\left[\left|Q(t)\right|\right]$, $\lambda(t)$, and $E_t\left[\left|Q(t)\right|\right]$ as follows. Divide (55) by the squared product of (56) and (57) and use (54) to solve for $\gamma(t)$, obtaining

$$\gamma(t) = (\bar{m} \cdot \bar{c}_B)^{-2/3} \cdot W^{2/3}(t). \tag{B-6}$$

Plug (B-6) into (54) to solve for $E_t\left[\left|Q(t)\right|\right]$:

$$E_t\left[\left|Q(t)\right|\right] = (\bar{m} \cdot \bar{c}_B)^{2/3} \cdot V(t) \cdot W^{-2/3}(t). \tag{B-7}$$

Multiply (56) by the square of (57) and use (B-7) to solve for $\lambda(t)$:

$$\lambda(t) = \left(\frac{\bar{m}^2}{\bar{c}_B}\right)^{1/3} \cdot \frac{1}{V^2(t)} \cdot W^{4/3}(t). \tag{B-8}$$

Plug (B-7) into (57) to solve for $E_t\left[Q^2(t)\right]$:

$$\mathrm{E}_t\left[Q^2(t)\right] = \left(\frac{\bar{c}_B^2}{\bar{m}}\right)^{2/3} \cdot V^2(t) \cdot W^{-4/3}(t). \tag{B-9}$$

Equation (B-7) and the definition of illiquidity $1/L(t) := \bar{c}_B / \left(E_t\left[\left|P(t) \cdot Q(t)\right|\right]\right)$ imply

$$\frac{1}{L(t)} = \left(\frac{\bar{m}^2}{\bar{c}_B}\right)^{-1/3} \cdot \sigma(t) \cdot W^{-1/3}(t). \tag{B-10}$$

Then, equations (B-6) for $\gamma(t)$, (B-7) for $\left|Q(t)\right|$, (B-8) for $\lambda(t)$, and (B-10) for $1/L(t)$ can be combined as in equation (53). This completes the proof of Theorem 3.

B.5 Exactly Linear Model with Brownian motion and Batched Bets

In the exactly linear model, risk-neutral informed and noise bets arrive anonymously, in batched pairs, at stochastic time intervals $\Delta t = 1/\gamma_U(t) = 1/\gamma_I(t) = 2/\gamma(t)$. For bets arriving at time t, an informed trader's signal is denoted $i_I(t)$, a noise trader's signal is denoted $i_U(t)$, and a signal which might be either informed or uninformed is denoted $i(t)$. The fundamental value follows arithmetic, not geometric, Brownian motion,[17]

$$F(t) := F_0 + F_1 \cdot \sigma_F \cdot B(t), \tag{B-11}$$

where $B(t)$ denotes a standardized Brownian motion with $B(t+h) - B(t) \sim \mathcal{N}(0,h)$ for $t \geq 0$ and $h \geq 0$, $B(0)$ is normally distributed, the initial value F_0 and the sensitivity parameter F_1 are known constant with units dollars/share. The error variance is defined as

$$\Sigma(t) = \mathrm{Var}_t\left[\frac{F(t) - P(t)}{P(t)}\right] = \mathrm{Var}_t\left[\frac{F_1 \cdot \sigma_F \cdot \left(B(t) - \bar{B}(t)\right)}{P(t)}\right]. \tag{B-12}$$

Here $\Sigma(t)$ is percentage error variance, slightly different from log-error-variance in equation (20).

Slightly modified versions of Theorem 1, Corollary 1, and Theorem 2 continue to hold. Minor changes in notation are needed to deal with batching of orders. The main substantive differences are that bets are exactly linear functions of signals, signals and fundamental value are exactly jointly normally distributed, and the error $B(t) - \bar{B}(t)$ is exactly normally distributed.

Theorem 4 (Characterization of Exactly Linear Equilibrium)**.** *There exists a unique linear equilibrium characterized by the four endogenous parameters $\lambda(t)$, $\beta(t)$, $\gamma_I(t)$, $\gamma_U(t)$, which are the following functions of the state variables $P(t)$, $\Sigma(t)$ and the exogenous parameters $\bar{\tau}$, $\bar{c}_I$, $\bar{m}$, $F_1 \cdot \sigma_F$, η, and N:*

$$\lambda(t) = \frac{\bar{\tau}}{4 \cdot \bar{c}_I} \cdot P^2(t) \cdot \Sigma(t), \tag{B-13}$$

$$\beta(t) = \frac{2 \cdot \bar{c}_I}{\bar{\tau}^{1/2}} \cdot \frac{1}{P(t) \cdot \Sigma^{1/2}(t)}, \tag{B-14}$$

$$\gamma_I(t) = \gamma_U(t) = \frac{\gamma(t)}{2}, \qquad \text{where} \qquad \gamma(t) = \frac{\eta \cdot N}{\bar{c}_I \cdot \bar{m}} \cdot \bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t). \tag{B-15}$$

At times $t \neq t_n$ when no bet arrives, the price $P(t)$ is constant, and error variance increases at the rate fundamental volatility unfolds: $\mathrm{d}\Sigma(t)/\mathrm{d}t = \left(F_1 \cdot \sigma_F / P(t)\right)^2$*. At times t_n when a pair of bets arrive, the price $P(t_n)$ and error variance $\Sigma(t_n)$ jump, following the difference equation system*

$$P(t_n^+) = P(t_n) + \frac{\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t_n)}{2} \cdot \left(i_I(t_n) + i_U(t_n)\right), \tag{B-16}$$

[17] Arithmetic Brownian motion implies that fundamental value may eventually be negative. To deal with this issue in a realistic manner, it would be possible to assume that the firm makes capital calls to add cash to its capital structure, keeping the price positive. Since the price would change, market liquidity $L(t)$ would change, but leverage neutrality implies that these changes have no effect on the dollar risk transferred by bets and therefore no effect on dollar market impact costs. Since capital calls would necessitate more complicated mathematical notation, we do not deal with this issue in this paper.

$$\Sigma(t_n^+) = \Sigma(t_n)\left(1 - \frac{\bar{\tau}}{2}\right), \qquad \textit{with} \qquad \Sigma(t_n) = \Sigma(t_{n-1}^+) + \frac{F_1^2 \cdot \sigma_F^2}{P^2(t)} \cdot (t_n - t_{n-1}). \tag{B-17}$$

Corollary 3. *In an exactly linear equilibrium, the endogenous variables $V(t)$, $\pi(t)$, $C_B(t)$, and $m(t)$ are the same functions of the exogenous parameters η, N, $\bar{\theta}$, $\bar{m}$, and $\bar{c}_I$ as in Corollary 1 with $\bar{\theta} = 1/2$. The endogenous variables $\mathrm{E}_t\left[\left|Q(t)\right|\right]$, $\mathrm{E}_t\left[Q^2(t)\right]$, $\gamma(t)$, $\gamma_I(t)$, $\gamma_U(t)$, $1/L(t)$, and $\rho(t)$ vary randomly through time as the slightly modified functions*

$$\mathrm{E}_t\left[\left|Q(t)\right|\right] = \frac{2 \cdot \bar{c}_I \cdot \bar{m}}{\bar{\tau}^{1/2}} \cdot \frac{1}{P(t) \cdot \Sigma^{1/2}(t)}, \tag{B-18}$$

$$\mathrm{E}_t\left[Q^2(t)\right] = \frac{4 \cdot \bar{c}_I^2}{\bar{\tau}} \cdot \frac{1}{P^2(t) \cdot \Sigma(t)}, \tag{B-19}$$

$$\frac{1}{L(t)} = \frac{\bar{\tau}^{1/2}}{2 \cdot \bar{m}} \cdot \Sigma^{1/2}(t), \tag{B-20}$$

$$\sigma^2(t) = \frac{\bar{\tau}^{3/2}}{4 \cdot \bar{c}_I \cdot \bar{m}} \cdot \eta \cdot N \cdot P(t) \cdot \Sigma^{3/2}(t), \tag{B-21}$$

$$\rho(t) = \frac{\bar{\tau}^{3/2}}{4 \cdot \bar{c}_I \cdot \bar{m}} \cdot \eta \cdot N \cdot P(t) \cdot \Sigma^{1/2}(t). \tag{B-22}$$

Proof of Theorem 4. The proof is similar to the proof of Theorem 1. It starts with deriving the system of four equations and then solving it for $\beta(t)$, $\lambda(t)$, $\gamma_I(t)$, and $\gamma_U(t)$.

First, derive the profit maximization condition. The risk-neutral informed or noise trader maximizes profits net of market impact costs by solving the problem

$$Q(t) = \arg\max_Q \mathrm{E}_t\left[\left(F(t) - \hat{P}(t,Q)\right) \cdot Q \;\middle|\; i(t)\right]. \tag{B-23}$$

Since each trader thinks that his signal contains information and

$$\mathrm{E}_t\left[F(t) - P(t) \,|\, \text{informed } i(t)\right] = \bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t) \cdot i(t), \tag{B-24}$$

the optimization problem (B-23) is exactly quadratic, not approximately quadratic as in equation (B-3). Its first-order condition yields

$$Q(t) = \beta(t) \cdot i(t), \qquad \text{where} \qquad \beta(t) = \frac{\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t)}{2 \cdot \lambda(t)}. \tag{B-25}$$

Second, derive the pricing rule. Informed bets and noise bets arrive in pairs. Market makers observe $Q_I(t)+Q_U(t)$, the sum of an informed bet $Q_I(t) := \beta(t)\cdot i_I(t)$ and a noise bet $Q_U(t) := \beta(t)\cdot i_U(t)$. Since they do not know which bet contains information, they update prices as

$$\begin{aligned}
&\mathrm{E}_t\big[F(t)-P(t) \mid Q_I(t)+Q_U(t)\big] \\
&\quad = \mathrm{E}_t\Big[F_1\cdot\sigma_F\cdot\big(B(t)-\bar{B}(t)\big) \;\Big|\; \bar{\tau}^{1/2}\cdot\frac{F_1\cdot\sigma_F\cdot\big(B(t)-\bar{B}(t)\big)}{P(t)\cdot\Sigma^{1/2}(t)}+(1-\bar{\tau})^{1/2}\cdot Z_I(t)+Z_U(t)\Big] \\
&\quad = \frac{1}{2\cdot\beta(t)}\cdot\bar{\tau}^{1/2}\cdot P(t)\cdot\Sigma^{1/2}(t)\cdot\big(Q_I(t)+Q_U(t)\big).
\end{aligned} \tag{B-26}$$

This implies the pricing rule $\hat{P}(\dots)$ and market depth $\lambda(t)$ given by

$$\hat{P}\big(t\,,\,Q_I(t)+Q_U(t)\big)=P(t)+\lambda(t)\cdot\big(Q_I(t)+Q_U(t)\big), \qquad \text{where} \qquad \lambda(t)=\frac{1}{2}\cdot\frac{\bar{\tau}^{1/2}\cdot P(t)\cdot\Sigma^{1/2}(t)}{\beta(t)}. \tag{B-27}$$

The conditions (B-25) and (B-27) are equivalent to each other; both define the product of $\beta(t)\cdot\lambda(t)$, but not the coefficients $\beta(t)$ and $\lambda(t)$ separately.

Third, the free entry condition says that the expected profits of an informed trader, net of market impact costs $C_B(t)=\mathrm{E}_t[\lambda(t)\cdot Q_I^2(t)]$ and costs of information $\bar{c}_I$, are equal to zero.

$$\bar{c}_I=\mathrm{E}_t\Big[\big(F(t)-P(t)\big)\cdot Q_I(t)-\lambda(t)\cdot Q_I^2(t)\Big]. \tag{B-28}$$

Intuitively, equation (B-29) implies that market liquidity adjusts continuously to make informed traders and noise traders indifferent between trading and not trading at any time. Since $Q(t)=\beta(t)\cdot i(t)$ from equation (B-25) and $\mathrm{E}_t[i^2(t)]=1$, free entry implies the third key equation

$$\frac{\big(\bar{\tau}^{1/2}\cdot P(t)\cdot\Sigma^{1/2}(t)\big)^2}{4\cdot\lambda(t)}=\bar{c}_I. \tag{B-29}$$

This equation defines $\lambda(t)$. Then, equation (B-25) or (B-27) yields the solution for $\beta(t)$.

Fourth, noise traders generate share volume at rate $\gamma_U(t)\cdot\mathrm{E}_t\big[|Q(t)|\big]=\eta\cdot N$. Bets arriving in pairs implies $\gamma_U(t)=\gamma_I(t)$. Since $\mathrm{E}_t\big[|i(t)|\big]=\bar{m}$, the expected size of a bet is

$$\mathrm{E}_t\big[|Q(t)|\big]=\beta(t)\cdot\bar{m}. \tag{B-30}$$

This implies the equation for the arrival rate of traders

$$\gamma_U(t)=\gamma_I(t)=\frac{\gamma(t)}{2}=\frac{\eta\cdot N}{\beta(t)\cdot\bar{m}}. \tag{B-31}$$

The four key log-linear equations (B-25), (B-27), (B-29), and (B-31) yield solutions for $\beta(t)$, $\lambda(t)$, $\gamma_U(t) = \gamma_I(t) = \gamma(t)/2$ in Theorem 4.

In an exactly linear equilibrium, the state variables $P(t)$ and $\Sigma(t)$ are sufficient statistics for describing the market's information at date t. When bets do not arrive, the market obtains no new information about fundamental value and therefore the price $P(t)$ is constant, but fundamental uncertainty continues to unfold so that $\mathrm{d}\Sigma(t) = \frac{F_1^2 \cdot \sigma_F^2}{P^2(t)} \cdot \mathrm{d}t$. When a bet arrives, the price changes by $\lambda(t_n) \cdot \big(Q_1(t_n) + Q_2(t_n)\big) = \lambda(t_n) \cdot \beta(t_n) \cdot \big(i_I(t_n) + i_U(t_n)\big)$, and the error variance $\Sigma(t_n)$ is reduced by fraction $\bar{\tau}/2$. The sum of two bets, one of which has precision of $\bar{\tau}$ and the other contains no information, effectively has a precision of $\bar{\tau}/2$. The solutions for $\lambda(t)$ and $\beta(t)$ in equations (B-13) and (B-14) imply equations (B-16) and (B-17).

□

Proof of Corollary 3. Equations (B-15) and $\mathrm{E}_t\big[|i(t)|\big] = \bar{m}$ imply that, in terms of exogenous variables, share volume $V(t)$ constitutes a constant fraction of shares outstanding N:

$$V(t) = 2 \cdot \eta \cdot N. \tag{B-32}$$

The endogenous variables in equations (B-18)–(B-21) can be obtained as follows: (1) Solve $Q(t) = \beta(t) \cdot i(t)$ and equation (B-14) for $\mathrm{E}_t\big[|Q(t)|\big]$ and $\mathrm{E}_t\big[Q^2(t)\big]$. (2) Solve definition (18) together with equations (B-13), (B-18), and (B-19) for $L(t)$. (3) Solve equations (14), (B-13), (B-15) and (B-19) for $P^2(t) \cdot \sigma^2(t)$.

To derive equation (B-22) for $\rho(t)$, write the changes in the unobserved estimation error $B_{\mathrm{err}}(t) := B(t) - \bar{B}(t)$ conditional on the sum of informed and noise signals $i_I(t_n) + i_U(t_n)$ as

$$\begin{aligned} \mathrm{E}_t\left[B_{\mathrm{err}}(t+\Delta t) - B_{\mathrm{err}}(t) \,|\, B_{\mathrm{err}}(t)\,,\, i_I(t_n) + i_U(t_n)\right] &= -\frac{\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t)}{2 \cdot F_1 \cdot \sigma_F} \cdot \big(i_I(t_n) + i_U(t_n)\big) \\ &= -\frac{\bar{\tau}^{1/2} \cdot P(t) \cdot \Sigma^{1/2}(t)}{2 \cdot F_1 \cdot \sigma_F} \cdot \left(\bar{\tau}^{1/2} \cdot \frac{F_1 \cdot \sigma_F \cdot \big(B(t) - \bar{B}(t)\big)}{P(t) \cdot \Sigma^{1/2}(t)} + (1-\bar{\tau})^{1/2} \cdot Z_U(t) + Z_I(t)\right). \end{aligned} \tag{B-33}$$

The first equation, which is similar to the approximate equation (B-1), is exact, not an approximation. The second equation is obtained by using equation (4). Compare equations (21) and (B-33) and use $\Delta t = 2/\gamma(t)$ to obtain

$$\rho(t) = \frac{\bar{\tau}}{2} \cdot \frac{\gamma(t)}{2}. \tag{B-34}$$

Then, plug in equation (B-15) to solve for $\rho(t)$. In the exactly linear model in which risks are modeled using arithmetic Brownian motion, the relationship (62) of variance $\sigma^2(t)$, resiliency

$\rho(t)$, and business time $\Delta t = 2/\gamma(t)$ becomes

$$\rho(t) = \frac{\bar{\tau}}{2} \cdot \frac{\gamma(t)}{2} = \frac{\sigma^2(t)}{\Sigma(t)}. \tag{B-35}$$

□

In the exactly linear model, the conditional steady state is an actual steady state! The dynamics of $\Sigma(t)$ is affected by realization of fundamental uncertainty at a rate $\frac{F_1^2 \cdot \sigma_F^2}{P^2(t)} \cdot \Delta t$ and incorporation of private information into prices at a rate $\sigma^2(t) \cdot \Delta t$:

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \mathrm{E}_t \left[\Sigma(t + \Delta t) - \Sigma(t) \right] = \frac{F_1^2 \cdot \sigma_F^2}{P^2(t)} - \sigma^2(t). \tag{B-36}$$

In a steady state, returns volatility $\sigma(t)$ and fundamental volatility $F_1 \cdot \sigma_F$ are related by $\sigma(t) \cdot P(t) = F_1 \cdot \sigma_F$. The price $P(t)$ follows a Brownian motion and volatility $\sigma(t)$ is stochastic, but their product is equal to non-stochastic $F_1 \cdot \sigma_F$.

Over time, the value of $\Sigma(t)$ converges to the steady state value $\Sigma^* = \Sigma^*(t)$ given by

$$\Sigma^*(t) := \frac{4 \cdot F_1^2 \cdot \sigma_F^2}{\gamma(t) \cdot \bar{\tau} \cdot P^2(t)} = \frac{4}{\bar{\tau}} \cdot \left(\frac{F_1^2 \cdot \sigma_F^2 \cdot \bar{c}_B \cdot \bar{m}}{V(t)} \right)^{2/3} \cdot \frac{1}{P^2(t)}. \tag{B-37}$$

The first equation is obtained from equation (B-35) by substituting market volatility $\sigma(t) \cdot P(t)$ for fundamental volatility $F_1 \cdot \sigma_F$. The second equation is obtained from equation (B-21) by substitution market volatility $\sigma(t) \cdot P(t)$ for fundamental volatility $F_1 \cdot \sigma_F$, $\bar{c}_B = \bar{c}_I$, and $V(t)$ from equation (B-32).

In the steady state, dollar error variance $\Sigma^*(t) \cdot P^2(t)$, volume $V(t) = 2 \cdot \eta \cdot N$, trading intensity $\beta(t)$, market impact $\lambda(t)$, bet arrival rate $\gamma(t)$, and resiliency $\rho(t)$ are constant; bet size has an unchanging distribution; returns volatility $\sigma(t)$ and fundamental volatility σ_F are related by $\sigma(t) \cdot P(t) = F_1 \cdot \sigma_F$, and liquidity $L(t)$ changes so that $P(t)/L(t)$ remains constant.

Theorem 2, Theorem 3, and the four meta-model equations (54)–(57) hold as before, but with $\bar{\theta} = 1/2$. The volatility equation (B-21) implies two ways to describe liquidity $L(t)$:

$$L(t) = \frac{2 \cdot \bar{m}}{\bar{\tau}^{1/2} \cdot \Sigma^{1/2}(t)} \qquad \text{and} \qquad L(t) = \left(\frac{\bar{m}^2 \cdot P(t) \cdot V(t)}{\bar{c}_B \cdot \sigma^2(t)} \right)^{1/3}. \tag{B-38}$$