PRICE MANIPULATION AND QUASI-ARBITRAGE[1]

By Gur Huberman and Werner Stanzl

In an environment where trading volume affects security prices and where prices are uncertain when trades are submitted, quasi-arbitrage is the availability of a series of trades which generate infinite expected profits with an infinite Sharpe ratio. We show that when the price impact of trades is permanent and time-independent, only linear price-impact functions rule out quasi-arbitrage and thus support viable market prices. When trades have also a temporary price impact, only the permanent price impact must be linear while the temporary one can be of a more general form. We also extend the analysis to a time-dependent framework.

Keywords: Market microstructure, price impact, arbitrage, market viability, gain-loss ratio.

1. INTRODUCTION

In any markets, trades can affect prices. In financial markets, the same individual can buy and subsequently sell the same security. In principle, then, a trader in a financial market can manipulate prices by buying and then selling the same security, with the expectation of earning a positive profit from such a manipulation.

This paper takes the perspective of a market watcher who has no opinion on the direction of security price movements but is an excellent student of the relation between trades and price changes. In fact, he has estimated that relation and is tempted to exploit this knowledge to his advantage. In a market in which prices are uncertain when orders are submitted, he attempts to implement a quasi-arbitrage which is a trading strategy that produces infinite expected profits with an infinite Sharpe ratio. What are the possible relations between price changes and trades that rule out quasi-arbitrage for this market watcher?

The absence of quasi-arbitrage is tantamount to market viability. A market is viable if the optimal demand of an agent with a mean/standard-deviation utility function exists. This result resembles that of Dybvig and Ross (1987) for the classical pricing model. They show that the absence of pure arbitrage is equivalent to the existence of an optimum for an agent who prefers more to less.

The dependence of price on trade size has an immediate as well as a permanent component. The price-impact function is the immediate price reaction to trading volume, including both temporary and permanent effects. The price-update function is the permanent effect of trade size on future prices. This paper's main result is the characterization of the price-update function under time independence. Specifically, price-update functions that admit no quasi-arbitrage possibilities and thus ensure viable markets are linear in trade size.

Recent empirical papers assume in addition to time independence that the price-impact and price-update functions are the same and suggest nonlinear price-update functions. Examples include Hasbrouck (1991), Hausman et al. (1992), and Kempf and Korn (1999). Interpreted in light of this work, these empirical results imply the feasibility of price manipulation.

Alternatively, our work calls into question either the time independence underlying much of the empirical work or the identification of the price-impact with the price-update function. (Holthausen et al. (1987), Gemmill (1996), or Keim and Madhavan (1996) make exactly the distinction between temporary and permanent effects of trades on prices.)

Anticipating some of this paper's results, Black (1995) imagines equilibrium exchanges where only limit orders labeled by levels of urgency are traded, and informally argues that price moves at each urgency level should be roughly proportional to order sizes to avoid "arbitrage." Presumably, he had in mind a time-independent framework where trades have a permanent price impact only. This paper does not address limit orders, but provides a formal proof of Black's conjecture when only market orders are allowed.

The standard justification of a price-impact function in an environment with asymmetrically informed agents is that information is impounded into prices through trades. Kyle (1985) is the leading example. Such models assume linear price-impact functions for tractability. This paper argues that linearity is justified in an environment that rules out quasi-arbitrage. It thereby selects which price-impact and price-update functions qualify for equilibrium.

The framework of this paper can also be used to evaluate Black's (1995) conjecture that the Kyle (1985) model allows price manipulation. We demonstrate that this is wrong. Underlying these results is a characterizing condition of the absence of price manipulation which governs how liquidity is allowed to vary over time. Furthermore, we prove that under certain conditions the equilibrium price-impact functions in Kyle have to be linear when time between trades is small.

The recent incomplete-markets, asset-pricing literature suggests to impose bounds on either the Sharpe ratio (see Cochrane and Saa-Requejo (2000)) or the gain-loss ratio (see Bernardo and Ledoit (2000)) to rule out "good deals." These bounds are then used to calculate price bounds for the assets in the economy. The framework of this paper can also embed such restrictions on prices. All results derived here for the absence of quasi-arbitrage are qualitatively the same as those for imposing no good deals with infinite gain-loss ratio.

Also, a simple corollary of our analysis is that ruling out price manipulation in a pure risk-neutral world has the same implications for the shape of the price-impact and price-update functions as the absence of quasi-arbitrage. A risk-neutral price manipulation embodies a trading strategy that generates positive expected profits.

Allen and Gale (1992) examine risk-neutral price manipulation in the Glosten and Milgrom (1985) framework. They construct equilibria in which uninformed agents profitably manipulate the security price. This price manipulation does not erode the equilibrium because traders are volume-constrained. Jarrow (1992) investigates whether a large trader, whose trades move the price in an otherwise complete and continuous-time market, can make profits from price manipulation. He gives several examples of pure arbitrage and states a sufficient condition that precludes arbitrage. The setups and results of both papers differ substantially from ours.

The rest of the paper is organized as follows. Section 2 introduces the model and describes the properties of prices. It also defines price manipulation and quasi-arbitrage. Section 3 finds necessary and sufficient conditions for both the absence of price manipulation and the absence of quasi-arbitrage, when the price-impact and price-update functions are time-independent. Section 4 investigates time-dependent price-impact and price-update functions and discusses the Kyle (1985) model. Section 5 treats multi-asset price dynamics. Section 6 studies the relationship between price impact and the gain-loss ratio, and Section 7 concludes. All proofs are relegated to the Appendix.

2. PRICE MODEL AND MARKET CONDITIONS

The first subsection describes the trading environment and how trades affect price changes.

The second takes the viewpoint of a trader and defines price manipulation and quasi-arbitrage. The third explains why ruling out quasi-arbitrage is necessary for market viability. The fourth is more technical and lays out possible assumptions regarding the probability distribution of security prices.

2.1. *Price Impact and Dynamics*

Consider a trader of a single asset over N periods in the time interval $(0,1]$. (The multi-asset case will be discussed in Section 5.) The asset can be bought or sold via market orders at times $n\triangle_N$, $1 \leq n \leq N$, where $\triangle_N \equiv 1/N$. In each period n, the initial price of the asset is given by the price quote $\tilde{p}_{n,N}$. In the absence of uncertainty, the trader, who trades the quantity $q_{n,N}$, has to pay a total of $p_{n,N}q_{n,N}$ with $p_{n,N} = \tilde{p}_{n,N} + P_{n,N}(q_{n,N})$. The initial price for the next period will be the quote $\tilde{p}_{n+1,N} = \tilde{p}_{n,N} + U_{n,N}(q_{n,N})$. A positive (negative) $q_{n,N}$ indicates a purchase (sale). The *price-impact function* $P_{n,N}$ measures the immediate price reaction to the trade $q_{n,N}$, including both the permanent and the temporary price impact, while the *price-update function* $U_{n,N}$ describes only the trade's permanent price impact. Hence, the temporary price impact is the difference $P_{n,N} - U_{n,N}$.

In every period competitive liquidity providers stand ready to fill the order of the trader. They set quotes and transaction prices. The price-impact and price-update functions represent their price reaction to trade size. Since the price-impact functions include the permanent price impact, any new information contained in an order is incorporated into the transaction price instantaneously.

Price-impact and price-update functions may be the consequence of either asymmetric information (see Glosten and Milgrom (1985), Kyle (1985), or Easley and O'Hara (1987, 1992)) or inventory costs (e.g., Amihud and Mendelson (1980) and Ho and Stoll (1981)). We neither model here how beliefs about the fundamental value of the asset are formed, nor do we postulate a specific dynamic for the inventory of the liquidity providers. Rather, we assume that we are given a certain set of price-impact and price-update functions that are already the outcome of these processes. The market can be in an equilibrium or a disequilibrium; it only has to meet the assumptions that we make here regarding the trading environment, but, other than that, can be very general. In particular, the price-impact and price-update functions may not only incorporate the fundamental value of the asset and inventory effects, but also other exchange-specific characteristics that affect prices and are sensitive to trading volume.

The trader does not possess superior information about the value of the asset, but can trade any amount he wishes. Others, called the *trading crowd*, may also submit orders. Their trade sizes are random, from the trader's perspective. All orders are submitted simultaneously at the end of each period. In addition, news that reveal value-relevant information arrive randomly. To incorporate both types of uncertainty, the price process is augmented with stochastic terms as follows.

After the most recent trades at time $(n-1)\triangle_N$, the price change due to public news, $\varepsilon_{n,N}$, is revealed at the beginning of period n and the price is updated to $\tilde{p}_{n,N}$, taking into account both the last trades and the latest news. Since the next trades take place only at time $n\triangle_N$ and the trading crowd does not reveal its intended orders, the trader knows $\tilde{p}_{n,N}$ and $\varepsilon_{n,N}$ before his trade in period n, but not the net order size of the trading crowd denoted by $\eta_{n,N}$. Our setup should best capture real trading activity: while it is unlikely that new information occurs at the moment of submitting a trade, other traders' order submissions not known to a trader are likely to happen.

For simplicity, the price-impact and price-update functions, $P_{n,N}$ and $U_{n,N}$, are both assumed to be finite and deterministic functions of the total trading volume $q_{n,N}+\eta_{n,N}$ only. Hence, the structure outlined above gives rise to the following price dynamics:

$$(1)\quad \tilde{p}_{n,N} = \tilde{p}_{n-1,N} + U_{n-1,N}(q_{n-1,N} + \eta_{n-1,N}) + \varepsilon_{n,N},$$

$$p_{n,N} = \tilde{p}_{n,N} + P_{n,N}(q_{n,N} + \eta_{n,N}),$$

for $1 \leq n \leq N$, $\tilde{p}_{0,N} > 0$, where $p_{n,N}$ denotes the transaction price at time $n\triangle_N$. Recall that the prices given by (1) are set by the liquidity providers.

All random variables in (1) are defined on the same probability space, $(\Omega, \mathcal{F}, \mathbb{P})$. Both, the $\varepsilon_{n,N}$'s and the $\eta_{n,N}$'s are *i.i.d.* processes with zero expected values, and are independent of each other. The information of the liquidity providers before orders have been submitted at time $n\triangle_N$ is represented by the sigma-algebra $\mathcal{F}^{(N)}_{n\triangle_N}$ and includes the knowledge of $\{\tilde{p}_{j,N}\}_{j=1}^{n}$, $\{p_{j,N}\}_{j=1}^{n-1}$, $\{q_{j,N}+\eta_{j,N}\}_{j=1}^{n-1}$, and $\{\varepsilon_{j,N}\}_{j=1}^{n}$. The transaction price $p_{n,N}$ and the total volume $q_{n,N}+\eta_{n,N}$ are known only after the trades and hence belong to $\mathcal{F}^{(N)}_{(n+1)\triangle_N}$. Thus, the liquidity providers observe the total trading volume, but never know its decomposition. Such providers resemble the market makers in Kyle (1985).

The trader, in contrast, knows in addition to $\mathcal{F}^{(N)}_{n\triangle_N}$ his own orders $\{q_{j,N}\}_{j=1}^{n}$ and extracts the past orders of the trading crowd $\{\eta_{j,N}\}_{j=1}^{n-1}$ from $\{q_{j,N}+\eta_{j,N}\}_{j=1}^{n-1}$, at time $n\triangle_N$. However, he can compute $\eta_{n,N}$ only after the trades have taken place in that period. Conditional expectations, denoted by $\mathbb{E}_{n,N}$, are with respect to the private information of the trader, $\mathcal{G}^{(N)}_{n\triangle_N}$.

The roles of the trader and the trading crowd can be interpreted in two ways. Either one thinks of the trader as being a large monopolistic trader and the trading crowd as being noise traders, or one sees the trader as a representative agent who believes that others' orders are noise. The model is a stylized model of microstructure. In most financial markets, participants can trade continuously, whereas here (as in other microstructure models, such as Kyle (1985)) they trade via a series of equally spaced auctions. Actual order flows or even signed order flows seem to violate the *i.i.d.* assumption assumed about the $\eta_{n,N}$'s. (See, e.g., Hasbrouck and Ho (1987), and Madhavan, Richardson and Rommans (1997).) To accommodate these empirical regularities one may want to interpret the $\eta_{n,N}$'s as the unexpected demand of the trading crowd. This interpretation leaves the analysis intact and is consistent with the observed empirical regularities.

The functions $P_{n,N}$ and $U_{n,N}$ can integrate bid-ask spreads and the market depths associated with these spreads. For example, suppose that at time $n\triangle_N$ the bid and the ask prices are $\tilde{p}_{n,N} \mp s_n/2$ for up to S_n shares. In addition, assume that after the trade at time $n\triangle_N$ the new mid-point quote equals the average of the previous mid-point quote and the last transaction price, plus $\varepsilon_{n+1,N}$. This situation can be modeled by setting $P_{n,N}(q) = s_n/2$ if $0 < q \leq S_n$, $= -s_n/2$ if $-S_n \leq q < 0$, $= 0$ if $q = 0$, and $U_{n,N} = P_{n,N}/2$. What $P_{n,N}$ and $U_{n,N}$ look like for trades larger than S_n will depend on other features of the trading environment, such as the liquidity of the limit-order book and the upstairs market (if they exist), the trading protocols, and so on.

The trading interval $(0, 1]$ represents a short-term time horizon such as a day or a week. Here, the time between trades, $\triangle_N$, is exogenously fixed by the exchange before trading starts; later sections also allow the traders to choose $\triangle_N$. Fixed transactions costs are charged by intermediaries: $c(k) > 0$ specifies the total fixed costs of k trades.

All variables and functions depend on the number of trades available, N, or equivalently, on the length of the inter-trading interval, $\triangle_N$. The sensitivity of the price-impact and price-update functions to trading volume may change when $\triangle_N$ decreases. Section 2.4 states the pertinent assumptions.

The ranges of the trader's order size, $q_{n,N}$, and the crowd's trading volume, $\eta_{n,N}$, are $\mathcal{D}_M$ and $\mathcal{D}_\eta$, respectively. The domain of the price-impact and price-update functions, $\mathcal{D}$, is thus $\mathcal{D}_M + \mathcal{D}_\eta$.

A relatively tractable special case of (1) is

$$(2) \quad \tilde{p}_{n,N} = \alpha\tilde{p}_{n-1,N} + (1-\alpha)p_{n-1,N} + \varepsilon_{n,N},$$

$$p_{n,N} = \tilde{p}_{n,N} + P_{n,N}(q_{n,N} + \eta_{n,N}),$$

where $\alpha \in [0,1]$. The price dynamics (2) can be obtained by setting $U_{n,N} = (1-\alpha)P_{n,N}$ in (1). The trader faces a price quote that is a convex combination of the previous quote and the transaction price of the last trade. In this case, temporary and permanent price changes are closely linked. This will allow the derivation of stronger conditions that are implied by the absence of quasi-arbitrage or the absence of price manipulation. When $\alpha = 0$, i.e., $U_{n,N} = P_{n,N}$, then (2) simplifies to

$$(3) \quad p_{n,N} = p_{n-1,N} + U_{n,N}(q_{n,N} + \eta_{n,N}) + \varepsilon_{n,N},$$

implying that the price change is a function of the current trades and randomness only, i.e., it does not depend on history. Further, the transaction price at time $n\triangle_N$ equals the price quote at time $(n+1)\triangle_N$ minus $\varepsilon_{n+1,N}$, and each trade has only a permanent impact on the security price.

Observe that the Kyle (1985) model can be recovered from (3), by setting $\varepsilon_{n,N} = 0$ and making the $U_{n,N}$'s linear functions. Thus, the price model in (1) generalizes Kyle's model.

2.2. *Definition of Price Manipulation and Quasi-Arbitrage*

What are reasonable prices for the market model introduced in the previous section? Equivalently, what do reasonable price-impact and price-update functions look like? Usually, the absence of arbitrage is invoked to find the set of viable asset prices. If arbitrage were feasible and scalable, agents would want to trade an infinite amount of shares over a finite time horizon and thus would cause the market to collapse. Two crucial assumptions underlying the classical arbitrage theory are that traders know the prices at which they can trade at any time, and that investments can be scaled without affecting prices.

Since we relax both of these assumptions, imposing the absence of pure arbitrage is less suitable here. First, if prices are unknown before the trades, pure arbitrages are hard to implement. Prices may fall after a sell order has been submitted or rise after a buy order has been announced; such adverse price movements make it difficult to avoid states in which losses occur with positive probability. Second, the dependence of prices on trade sizes limits the profitability of pure arbitrage. In this case, the existence of arbitrage does not necessarily imply the breakdown of a market. Hence, to identify the prices which support a viable market, we do not employ a no pure-arbitrage condition but use the concept of price manipulation and quasi-arbitrage as defined and discussed next.

A sequence of trades $q^{0,N} \equiv \{q_{n,N}\}_{n=1}^{N}$ is a *round-trip trade* if the sum of all these trades is zero, i.e., $\sum_{n=1}^{N} q_{n,N} = 0$. This sequence may contain zero trades and thus the actual number of trades, $T(q^{0,N})$, may be strictly less than N. The profit of a round-trip trade is given by

$$(4) \quad \pi(q^{0,N}) \equiv -\sum_{n=1}^{N} p_{n,N} q_{n,N} - c(T(q^{0,N})).$$

We now define price manipulation and quasi-arbitrage as follows.

Definition 1: A *(risk-neutral) price manipulation* is a round-trip trade $q^{0,N}$ with the expected value $\mathbb{E}[\pi(q^{0,N})] > 0$.

Definition 2: An *unbounded (risk-neutral) price manipulation* is a sequence $\{q_m^{0,N}\}_{m=1}^{\infty}$ of round-trip trades with $\lim_{m\to\infty}\mathbb{E}[\pi(q_m^{0,N})] = \infty$.

Definition 3: A *quasi-arbitrage* is an unbounded price manipulation $\{q_m^{0,N}\}_{m=1}^{\infty}$ which satisfies $\lim_{m\to\infty}\mathbb{E}[\pi(q_m^{0,N})]/Std[\pi(q_m^{0,N})] = \infty$, where $Std[\pi(q_m^{0,N})]$ denotes the standard deviation of $\pi(q_m^{0,N})$.

A quasi-arbitrage is a sequence of round-trip trades that exhibits not only infinite expected profits but also infinite expected profits per unit of risk, since the ratio $\mathrm{SR}(q^{0,N}) \equiv \mathbb{E}[\pi(q^{0,N})]/Std[\pi(q^{0,N})]$ ($\mathrm{SR}(0) \equiv 0$) can be interpreted as the "Sharpe ratio" of the trading profits. The standard deviation is allowed to converge to infinity as long as the expected value grows at a faster rate. However, if a quasi-arbitrage is divisible, meaning that any fraction of the quasi-arbitrage can be held, then its risk can be eliminated asymptotically. This is because one can always find a sequence of portfolio weights $\{\theta_m\}_{m=1}^{\infty}$ such that $\lim_{m\to\infty}\mathbb{E}[\theta_m\pi(q_m^{0,N})] = \infty$ and $\lim_{m\to\infty} Std[\theta_m\pi(q_m^{0,N})] = 0$, for example, $\theta_m = 1/(\sqrt{\mathrm{SR}(q_m^{0,N})}Std[\pi(q_m^{0,N})])$.

Price manipulation and quasi-arbitrage have been defined here for a fixed $\triangle_N$. Sections 3 and 5 allow the trader to select $\triangle_N$. For this case, Definitions 2 to 3 need to be slightly generalized: replace $\{q_m^{0,N}\}_{m=1}^{\infty}$ by $\{q_m^{0,N_m}\}_{m=1}^{\infty}$ to permit the trader to choose round-trip trades based on different $\triangle_{N_m}$'s.

A divisible quasi-arbitrage resembles asymptotic arbitrage as introduced in Huberman (1982) for the APT. An asymptotic arbitrage is a sequence of zero-cost investments that produces an infinite average return in the limit, while the variance of the returns falls to zero. In Huberman asymptotic arbitrage could occur when the number of assets becomes infinite; asymptotic arbitrage here could occur when the trading volume goes to infinity.

Quasi-arbitrage is a weaker constraint than δ-arbitrage as introduced in Ledoit (1995). Translated into this framework, a δ-arbitrage is a round-trip trade $q^{0,N}$ for which the Sharpe ratio $\mathrm{SR}(q^{0,N})$ is larger than δ. Obviously, the absence of δ-arbitrage excludes quasi-arbitrage.

2.3. *Market Viability*

We consider a market not viable if the optimal demand for the trader does not exist. To determine when this occurs, the trader's utility first needs to specified. We assume here that his utility, $u(\mathrm{SR}(q^{0,N}))$, depends only on the Sharpe ratio of his trading profits. His optimal demand is therefore given by $\arg\max_{q^{0,N}} u(\mathrm{SR}(q^{0,N}))$, where $q^{0,N}$ is such that prices never become negative in expectation. The function u is non-decreasing and never becomes flat, that is, for every y there exists a $y' > y$ with $u(y') > u(y)$. In case the trader is risk-neutral, his utility is simply given by $u(\mathbb{E}[\pi(q^{0,N})])$.

We can now make more precise what we mean under market viability.

Definition 4: A market is *weakly viable* if the optimal demand for the trader exists, and *strongly viable* if the trader's optimal demand exists uniquely and is zero.

While price-manipulation is impossible in strongly viable markets, limited price manipulation is feasible in weakly viable ones. Some market models may allow price manipulation in equilibrium and hence are only weakly viable. Of course, any equilibrium implies the existence of the optimal demand for the trader. Thus, weak viability seems a minimal requirement for any theoretical or econometric model aiming to study or estimate the relation between price changes and trading volume.

Weak viability rules out quasi-arbitrage. This follows immediately from the assumptions on u and the definition of quasi-arbitrage. Whether the converse holds, is not obvious. We will identify conditions under which the absence of quasi-arbitrage implies either strong or weak viability.

2.4. *Market Classification*

We aim at using a market model as general as possible. To this end, we will distinguish three markets depending on whether the first and second moments of the security prices exist, on how the second moments evolve when the time between trades becomes small, and on whether the trader can choose $\triangle_N$. The first two markets study price manipulation and only require the existence of the asset prices' first moments. One of them allows $\triangle_N$ to be endogenous in order to embed unbounded price manipulation. The third market focuses on quasi-arbitrage and needs in addition the existence of the second moments and a law that governs how these change when the inter-trading interval varies.

The *expected price-impact* and *expected price-update functions*, defined by $\hat{P}_{n,N}(q) \equiv \mathbb{E}[P_{n,N}(q + \eta_{n,N})]$ and $\hat{U}_{n,N}(q) \equiv \mathbb{E}[U_{n,N}(q + \eta_{n,N})]$, $q \in \mathcal{D}_M$, respectively, satisfy, if they exist, $\hat{P}_{n,N}(q) \geq 0$ and $\hat{U}_{n,N}(q) \geq 0$ if $q \geq 0$, and $\hat{P}_{n,N}(q) \leq 0$ and $\hat{U}_{n,N}(q) \leq 0$ if $q < 0$, in all three markets. Therefore, purchases are expected to have a positive impact on the price while sales have a negative one (the expected temporary impact of a trade, $\hat{P}_{n,N}(q) - \hat{U}_{n,N}(q)$, can have any sign, though). Fixed costs, $c(k)$, are assumed to be proportional to k^e in either market, where $e < 2$.

If the functions $\hat{P}_{n,N}$ and $\hat{U}_{n,N}$ exist for all $q \in \mathcal{D}_M$ and $1 \leq n \leq N$, and if $\triangle_N$ is fixed by the exchange, then we call the market described by the price process (1) $\mathcal{M}_1(\triangle_N)$. Hence, the prices in this market all have first moments. The market $\mathcal{M}_1(\triangle_N)$ provides a minimal environment to study price manipulation.

In the second market, $\mathcal{M}_1$, not only all $\hat{P}_{n,N}$'s and $\hat{U}_{n,N}$'s exist for $N \geq 1, 1 \leq n \leq N$, but also the trader can choose $\triangle_N$ (at time zero). Furthermore, we require that $\hat{U}_{n,N}(q) \geq -\hat{U}_{n,N}(-q)$ for all non-negative $q \in \mathcal{D}_M$. This inequality says that purchases have an expected price update no smaller than sales. It can be interpreted in various ways. One argument is that sales often occur because of liquidity shocks and thus have less informational content. Or, purchases have to have at least as big a price impact as sales, on average, because otherwise the price would be expected to drop to zero in case buys and sells are equally likely. In market $\mathcal{M}_1$ we will examine unbounded price manipulation.

Finally, the market $\mathcal{M}_1$ shall be labeled $\mathcal{M}_2$ if also the variances $V_P(q, n, N) \equiv Var[P_{n,N}(q+\eta_{n,N})]$, $V_U(q, n, N) \equiv Var[U_{n,N}(q + \eta_{n,N})]$, and $\sigma^2_\varepsilon(N) \equiv Var[\varepsilon_{1,N}]$ (recall that $Var[\varepsilon_{n,N}]$ is constant for $1 \leq n \leq N$) exist for all $q \in \mathcal{D}_M$, $N \geq 1$, and $1 \leq n \leq N$, and if the variances, as a function of N, do not grow faster than linearly as $N \to \infty$, for each $q \in \mathcal{D}_M$ and n (see Appendix A for details). We will determine under which conditions quasi-arbitrage is absent in market $\mathcal{M}_2$.

Standard models typically assume that σ^2_ε asymptotically evolves as $1/N$ (e.g., see Merton (1990)), which say that the total variance of the asset price during a fixed time horizon is evenly divided between the N per-period variances of that asset. In contrast, this work permits more general behavior of the variances. For instance, $\mathcal{M}_2$ admits σ^2_ε to be constant, which allows the total variance, $\sigma^2_\varepsilon(N)N$, to be linearly increasing in the number of trades. Such an assumption is appropriate if market volatility rises due to a higher trading intensity. At most, the per-period variances in $\mathcal{M}_2$ can grow linearly in N, or put differently, the total variances during the time interval $(0, 1]$ can grow no more than quadratically in N.

The conditions introduced above are not restrictive and examples of the three markets can be quite easily constructed. In any case, $\mathcal{M}_2$ is contained in $\mathcal{M}_1$, which, in turn, is a subset of $\mathcal{M}_1(\triangle_N)$, for all $\triangle_N$.

3. SINGLE-ASSET TIME-INDEPENDENT PRICE IMPACT

The price-impact and price-update functions are time-independent if $P_{n,N} = P$ and $U_{n,N} = U$ for all $N \geq 1$ and $1 \leq n \leq N$.

3.1. *Necessary Conditions for the Absence of Quasi-Arbitrage*

We first show that the absence of price manipulation in market $\mathcal{M}_1(\triangle_N)$ requires the price-update function to be linear. To this end, allow for any trade size for the trader and assume the crowd's trades are normally distributed. The arguments below offer an outline of the proof using four steps. The formal proof is in Appendix B.

Claim 1: The expected price-update function must be symmetric, i.e., $\hat{U}(q) = -\hat{U}(-q)$. To show this, note that either $\hat{U}(q) > -\hat{U}(-q)$ or $\hat{U}(q) < -\hat{U}(-q)$ for a $q > 0$ would invite price manipulation. In the former case (where purchasing q units has a stronger impact on the price update than selling q units), a trader could buy q shares in each of the first m periods and then sell q shares in each of the subsequent m periods ($2m \leq N$). The expected profit of this round-trip strategy satisfies

$$(5) \quad \mathbb{E}[\pi(q^{0,N})] = \frac{m^2}{2} q[\hat{U}(q) + \hat{U}(-q)] - \frac{m}{2} q[\hat{U}(-q) - \hat{U}(q) + 2(\hat{P}(q) - \hat{P}(-q))] - c(2m).$$

(Two additional technical assumptions (see conditions (C1) and (C3) in Appendix B) ensure that this and all round-trip trades below induce only non-negative expected prices.) Hence, if $\triangle_N$ is sufficiently small, then $\mathbb{E}[\pi(q^{0,N})] > 0$ and there is price manipulation.

In the second case, the reverse strategy (first selling q shares in each of the first m periods and then buying back q shares in each of periods $m+1$ to $2m$) would yield $\mathbb{E}[\pi(q^{0,N})] > 0$ for sufficiently small $\triangle_N$.

It is straightforward to check that $\hat{U}(0) = 0$, which concludes the proof of Claim 1.

Claim 2: $\hat{U}$ is continuous everywhere except possibly at the origin, i.e., $\lim_{j\to\infty}\hat{U}(q_j) = \hat{U}(q)$, $q \neq 0$, when $\lim_{j\to\infty} q_j = q$. To sketch the idea of this part of the proof, consider the following example. Suppose that the price-update function has an upward jump at $q > 0$, that is, $\lim_{q_j \downarrow q} \hat{U}(q_j) > \hat{U}(q)$. The strategy of buying $q_j > q$ shares in each of the first m periods and selling q shares in each of the following m periods, and selling the remaining shares at time $2m+1$, where q_j is chosen arbitrarily close to q, yields $\mathbb{E}[\pi(q^{0,N})] > 0$ for sufficiently small $\triangle_N$. Due to the jump, the updating reacts less to sales than to buys, causing the average selling price to exceed the average purchasing price. Appendix B demonstrates that for any possible type of jump price manipulations can be found.

Claim 3: $\hat{U}$ is linear, since either $\hat{U}(q) > \hat{U}(1)q$ or $\hat{U}(q) < \hat{U}(1)q$ would induce price manipulation, for an arbitrary q. To see this, consider the first case and note that $q > 0$ can be assumed to be a rational number. Now, buying q shares in each of the first m periods and then selling one share in each of the following mq periods (mq can be chosen to be an integer) constitutes a price manipulation for sufficiently small $\triangle_N$, because the selling moves the price down by less than the degree to which the buying shifts the price upwards. The second inequality can be rejected analogously.

Claim 4: U is linear. For this purpose, define $R_U(q) \equiv U(q) - \hat{U}(q)$. Then, Claim 3 implies

$\mathbb{E}[R_U(q+\eta_{1,N})] = 0$ for all q, which in turn has $R_U = 0$, as a consequence, thanks to the normality of $\eta_{1,N}$. Actually, the equality $U(q) = \hat{U}(q)$ holds only $L(\mathbf{R})$-almost everywhere, where $L(\mathbf{R})$ is the Lebesgue measure on $\mathbf{R}$.

Exactly the same strategy can be applied to prove that also the absence of unbounded price manipulation in $\mathcal{M}_1$ and the absence of quasi-arbitrage in $\mathcal{M}_2$ each imply the linearity of the price-update function. For example, let us verify that the first price manipulation used to prove Claim 1 can be extended to a quasi-arbitrage in market $\mathcal{M}_2$ if $\hat{U}(q) > -\hat{U}(-q)$. The expected profit, $\mathbb{E}[\pi(q^{0,N})]$, is of order m^2 by (5), while its standard deviation is only of order m^a, $a < 2$, by inequality (10) in Appendix B. Thus, when the trader chooses $\triangle_N \to 0$, $q^{0,N}$ becomes a quasi-arbitrage. Similarly, all other price manipulations above can be augmented to quasi-arbitrages if U deviates from linearity. For the details of these arguments and the proof of the following proposition see Appendix B.

Proposition 1: *Suppose that any trade size is allowed for the trader and that either*

(i) $\mathbb{P}[\eta_{1,N} = 0] = 1$ *(zero net trades of the crowd) or*

(ii) the crowd's trades are normally distributed.

Then, the absence of price manipulation in $\mathcal{M}_1(\triangle_N)$ *requires* U *to be linear with non-negative slope,* $L(\mathbf{R})$ *– a.e., for all sufficiently small* $\triangle_N$*. The linearity of* U *with non-negative slope is also implied by each the absence of unbounded price manipulation in* $\mathcal{M}_1$ *and the absence of quasi-arbitrage in* $\mathcal{M}_2$*.*

Let us now generalize the previous results in two ways. First, the domains $\mathcal{D}$, $\mathcal{D}_M$, $\mathcal{D}_\eta$, and $\mathcal{D}_\varepsilon$ need no longer be $\mathbf{R}$. Instead, each of them can be an arbitrary *symmetric* set (a set $\mathcal{D}$ is symmetric if $0 \in \mathcal{D}$ and $q \in \mathcal{D}$ implies $-q \in \mathcal{D}$). Second, the distribution of the crowd's trading volume can be non-normal. To establish a generalization of Proposition 1 the following definition is useful.

DEFINITION 5: A function $f : \mathcal{D} \to \mathbf{R}$ is *quasi-linear* if it has the representation $f(y) = \lambda y + R_f(y)$ on $\mathcal{D}$, $\lambda \geq 0$, $L(\mathcal{D})$-*a.e.*, where the $\mathcal{D}$-Borel-measurable function $R_f : \mathcal{D} \to \mathbf{R}$ satisfies

$$(6) \quad \mathbb{E}_{n,N}[R_f(\tilde{q}_{n,N} + \eta_{n,N})] = 0$$

for all $\mathcal{G}^{(N)}_{n\triangle_N}$-measurable random variables $\tilde{q}_{n,N} : \Omega \to \mathcal{D}_M$, $1 \leq n \leq N$, $N \geq 1$. We call R_f the *residual function* of f.

We will interpret this definition after Theorem 1. However, note that any linear function with non-negative slope is also quasi-linear and that equation (6) does not imply $R_f = 0$ (see Appendix C).

Having Definition 5, the proof of Proposition 1 can be easily modified to give the theorem below.

THEOREM 1: *The absence of price manipulation in $\mathcal{M}_1(\triangle_N)$ (NoPM) requires U to be quasi-linear, if $\triangle_N$ is sufficiently small. The quasi-linearity of U is also a consequence of each, the absence of unbounded price manipulation in $\mathcal{M}_1$ (NoUM), and the absence of quasi-arbitrage in $\mathcal{M}_2$ (NoQA).*

Theorem 1 says that each of (NoPM)-(NoQA) implies a price-update function that can be written as the sum of a linear function and its residual function R_U, which in conditional expected terms drops out. The latter holds regardless of what order the trader submits, because the trader's strategy set is identical to the set consisting of all $\mathcal{G}^{(N)}_{n\triangle_N}$-measurable random variables. Thus, traders always expect linear price updating.

One important formal feature of the price process (1) is that the price-impact function P can be chosen to include fixed per-share transaction costs. Therefore, Proposition 1 and Theorem 1 are also valid when commissions have to be paid per share.

Some empirical papers report evidence of asymmetric price impacts. For instance, Chan and Lakonishok (1995), Gemmill (1996), and Holthausen et al. (1987, 1990) find that block purchases have a larger price impact than block sales, whereas Keim and Madhavan (1996) and Scholes (1972) provide evidence that there are also markets with a stronger price impact of sales. A literal interpretation of Proposition 1 and Theorem 1 suggests quasi-arbitrage. Alternatively, this result suggests that a better specification should allow for time dependence of the price-impact function.

Proposition 1 and Theorem 1 above establish the linearity of the price-update function. Necessary conditions for the price-impact function are in Proposition 2. (Recall that the price-impact function is the sum of the permanent component - the price-update function - and the temporary component.)

PROPOSITION 2: *If either (NoPM) or (NoUM) holds, and if fixed costs $c(k)$ are proportional to k^a, $a < 1$, then the following two conditions must hold for all sufficiently small $\triangle_N$:*

(i) $\hat{P}(q) - \hat{P}(-q) \gtreqless \hat{U}(q)$ *for* $q \gtreqless 0$, $q \in \mathcal{D}_M$, *and*

(ii) $\hat{P} \neq 0$ *when* $\hat{U} \neq 0$.

If we interpret the left-hand side of condition *(i)* in Proposition 2 as the expected spread of the price-impact function, then condition *(i)* says that the expected spread *at any* trade size has to exceed the expected price update resulting from that trade. Were this not true, the trading strategy cited in the proof of Claim 1 (buying q shares in each of the first m periods and then selling q shares in each of the next m periods) would allow unbounded price manipulation (see

(5)). The same trading strategy also implies the second condition in Proposition 2. $\hat{P}$ always has to be a function of the trade size, unless $\hat{U} = 0$. The absence of quasi-arbitrage in general does not imply conditions *(i)* and *(ii)*.

3.2. *A Sufficient Condition for the Absence of Quasi-Arbitrage*

This subsection derives a sufficient condition for (NoPM)-(NoQA). With the aid of this condition we are able to establish that each (NoPM)-(NoQA) is equivalent to the linearity of the price-update function.

The main observation leading to this sufficient condition is the fact that $q^{0,N} = 0$ is the unique maximizer of $\sup_{q^{0,N}} \mathbb{E}[\pi(q^{0,N})]$ if $P(x) = U(x)/2 = \lambda x/2$, $\lambda \geq 0$, as is shown in Appendix B.

Proposition 3: *Let U be linear with non-negative slope, $L(\mathbf{R})$ – a.e. If $P(q) \geq U(q)/2$ and $P(-q) = -P(q)$ for non-negative $q \in \mathcal{D}$, then (NoPM)-(NoQA) are all satisfied for any $\triangle_N$.*

For $0 \leq \alpha < 1$, the price-impact function implied by the price processes (2) and (3) is $P = U/(1-\alpha)$. As a result, P meets the requirement on the price-impact function in Proposition 3. An important consequence of Proposition 3 is that nonlinear price-impact functions can lead to prices that preclude price manipulation or quasi-arbitrage. In fact, complicated price-impact functions are admissible.

Having Proposition 3, we are now able to characterize the absence of price manipulation and quasi-arbitrage. Indeed, Propositions 1 and 3 imply the following.

Proposition 4: *Suppose (i) P satisfies the condition given in Proposition 3 and (ii) the crowd's trades are normally distributed. Then, linearity of U (with non-negative slope, $L(\mathbf{R})$ –*

a.e.) is equivalent to the absence of price manipulation in $\mathcal{M}_1(\triangle_N)$, for all sufficiently small $\triangle_N$. The linearity of U (with non-negative slope, $L(\mathbf{R})-a.e.$) is also equivalent to each (NoUM) and (NoQA).

Proposition 4 connects (NoPM)-(NoQA) through one common characterizing property, namely, the linearity of the price-update function.

3.3. *Strong Viability*

Proposition 4 and the observation made before Proposition 3 enable us to determine when markets $\mathcal{M}_1$ and $\mathcal{M}_2$ are strongly or weakly viable.

Corollary 1: *Assume that the conditions (i)-(ii) in Proposition 4 are met. Then, the absence of quasi-arbitrage in market $\mathcal{M}_2$ is characterized by the strong viability of $\mathcal{M}_2$. If the trader is risk-neutral, the strong viability of market $\mathcal{M}_1$ is equivalent to the absence of unbounded price manipulation in $\mathcal{M}_1$.*

Proposition 4 and Corollary 1 can be directly applied to two problems which are often studied in the finance microstructure literature. The first involves insider trading, where a monopolistic insider solves $\sup_{\{q_n\}_{n=1}^N} \mathbb{E}[\sum_{n=1}^N (v-p_n)q_n]$ for a given N, after having received the value of the asset, v, in period 0 (e.g., see Dutta and Madhavan (1995)). The second problem is discretionary liquidity trading, as studied in Bertsimas and Lo (1998) and Huberman and Stanzl (2002). There, an uninformed trader faces the problem $\inf_{\{q_n\}_{n=1}^N} \mathbb{E}[\sum_{n=1}^N p_n q_n]$ subject to $\sum_{n=1}^N q_n = \overline{q} \neq 0$, given the number of trades, N. In other words, this trader wants to minimize the expected costs of trading a certain amount of shares, $\overline{q}$, over a certain time horizon. These studies employ a simpler version of the price process (3) with linear price-update functions. Proposition 4 and Corollary 1 justify their linearity assumptions.

4. TIME-DEPENDENT PRICE IMPACT

Until now, the price-impact and price-update functions have been time-independent, i.e., price reacts to traded quantity in the same manner in each period. *Liquidity*, which is represented by the first derivative of the price-impact and price-update functions (when they exist), is therefore constant through time. In what follows we relax this assumption and allow liquidity to vary across time.

4.1. *Absence of Price Manipulation*

One way to examine time-dependent liquidity is to consider linear price-impact and price-update functions that change over time. More specifically,

$$\tilde{p}_{n,N} = \tilde{p}_{n-1,N} + \lambda_{n-1,N}(q_{n-1,N} + \eta_{n-1,N}) + \varepsilon_{n,N}, \tag{7}$$

$$p_{n,N} = \tilde{p}_{n,N} + \mu_{n,N}(q_{n,N} + \eta_{n,N}),$$

for sequences $\{\lambda_{n,N}\}_{n=1}^{N}$ and $\{\mu_{n,N}\}_{n=1}^{N}$, where $\lambda_{1,N} = \mu_{1,N} \geq 0$. The price-impact and price-update functions are allowed to have any sign, that is, the slopes $\{\lambda_{n,N}\}_{n=2}^{N}$ and $\{\mu_{n,N}\}_{n=2}^{N}$ can either be positive or negative.

We proceed by establishing an equivalent condition for the absence of price manipulation. The main difference with the previous section is that unbounded price manipulation can usually be implemented with finitely many trades. In contrast to the last section, we can employ the global shape of the price-impact and price-update functions to seek for price manipulation.

To get an idea what kind of restrictions the absence of price manipulation imposes on the pair $(\{\lambda_{n,N}\}_{n=1}^{N}, \{\mu_{n,N}\}_{n=1}^{N})$, let us consider the simple case $N = 3$ where only three trades are feasible. Computing $\mathbb{E}[\pi(q^{0,3})]$ for any round-trip trade $q^{0,3}$ leads to

$$-\mathbb{E}\left[\sum_{n=1}^{3}[\tilde{p}_{0,3}+\sum_{j=1}^{n-1}\lambda_{j,3}(q_{j,3}+\eta_{j,3})+\mu_{n,3}(q_{n,3}+\eta_{n,3})+\sum_{j=1}^{n}\varepsilon_{j,3}]q_{n,3}\right]-c(T(q^{0,3}))$$

$$=-\mathbb{E}\left[\mu_{2,3}q_{2,3}^2+\lambda_{2,3}q_{2,3}q_{3,3}+\mu_{3,3}q_{3,3}^2\right]-c(T(q^{0,3}))$$

$$+\,\mathbb{E}\left[\sum_{n=1}^{2}\lambda_{n,3}\eta_{n,3}\sum_{j=1}^{n}q_{j,3}-\sum_{n=1}^{3}\mu_{n,3}q_{n,3}\eta_{n,3}+\sum_{n=2}^{3}\varepsilon_{n,3}\sum_{j=1}^{n-1}q_{j,3}\right],$$

which is $-\mathbb{E}\left[[q_{2,3}\ \ q_{3,3}]\Lambda_3[q_{2,3}\ \ q_{3,3}]^T\right]/2-c(T(q^{0,3}))$, where

$$\Lambda_3\equiv\begin{bmatrix}2\mu_{2,3} & \lambda_{2,3}\\ \lambda_{2,3} & 2\mu_{3,3}\end{bmatrix}.$$

The trades $q_{2,3}$ and $q_{3,3}$ in the above expression can take any value as long as the trades $q_{1,3}$, $q_{2,3}$, and $q_{3,3}$ do not cause negative expected prices. Hence, for all sufficiently large initial prices $\tilde{p}_{0,3}$, the absence of price manipulation is equivalent to Λ_3 being positive semidefinite.

The removal of $q_{1,3}$ is arbitrary. If we remove $q_{2,3}$ or $q_{3,3}$, we would find that

$$\begin{bmatrix}2\mu_{2,3} & 2\mu_{2,3}-\lambda_{2,3}\\ 2\mu_{2,3}-\lambda_{2,3} & 2(\mu_{2,3}-\lambda_{2,3}+\mu_{3,3})\end{bmatrix}\text{ and }\begin{bmatrix}2\mu_{3,3} & 2\mu_{3,3}-\lambda_{2,3}\\ 2\mu_{3,3}-\lambda_{2,3} & 2(\mu_{2,3}-\lambda_{2,3}+\mu_{3,3})\end{bmatrix},$$

respectively, have to be positive semidefinite to exclude price manipulation. Since all matrix representations employ the same parameters and each matrix is positive semidefinite if and only if the others are, either matrix can be used for the analysis. We work here only with Λ_3 henceforth.

For Λ_3 to be positive semidefinite, $\mu_{2,3}$ and $\mu_{3,3}$ must be non-negative and $\mu_{2,3}\mu_{3,3}\geq\lambda_{2,3}^2/4$.

These conditions, together with $\mu_{1,3} \geq 0$, say that the absence of price manipulation rules out negative price-impact sequences in all periods and that $\mu_{2,3}$ and $\mu_{3,3}$ have to be sufficiently large relative to $\lambda_{2,3}^2$.

The same method as above applied to the general case gives the following.

THEOREM 2: *Fix an arbitrary* $\triangle_N$*. For all sufficiently large initial prices* $\tilde{p}_{0,N}$*, no price manipulation in* $\mathcal{M}_1(\triangle_N)$ *is characterized by the positive semidefiniteness of the matrix*

$$(8) \quad \Lambda_N \equiv \begin{bmatrix} 2\mu_{2,N} & \lambda_{2,N} & \lambda_{2,N} & \dots & \lambda_{2,N} \\ \lambda_{2,N} & 2\mu_{3,N} & \lambda_{3,N} & \dots & \lambda_{3,N} \\ \lambda_{2,N} & \lambda_{3,N} & 2\mu_{4,N} & \dots & \lambda_{4,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{2,N} & \lambda_{3,N} & \lambda_{4,N} & \dots & 2\mu_{N,N} \end{bmatrix}.$$

Theorem 2 gives a specific computational criterion for the absence of price manipulation. Testing whether Λ_N is positive semidefinite can be easily done, unless N is too big. In any case, we suggest to first check whether Λ_N is positive definite. If $\Lambda_{j,N}$ denotes the j^{th}-order leading principal submatrix of Λ_N (delete the last $N-1-j$ rows and the last $N-1-j$ columns of Λ_N), then $\det \Lambda_{j,N} > 0$ for all $1 \leq j \leq N-1$ if and only if Λ_N is positive definite. Each $\det \Lambda_{j,N}$ can be obtained recursively from

$$(9) \quad \det \Lambda_{j,N} = 2(\mu_{j,N} + \mu_{j-1,N} - \lambda_{j-1,N}) \ \det \Lambda_{j-1,N} - (2\mu_{j-1,N} - \lambda_{j-1,N})^2 \ \det \Lambda_{j-2,N}$$

for $3 \leq j \leq N-1$, where $\Lambda_{N-1,N} = \Lambda_N$, with initial conditions $\det \Lambda_{1,N} = 2\mu_{2,N}$ and $\det \Lambda_{2,N} = 4\mu_{2,N}\mu_{3,N} - \lambda_{2,N}^2$. Hence, the complexity of testing for positive definiteness is only of linear order.

Another important implication of Theorem 2 is that for any given price-update sequence there exists a price-impact sequence that preserves the absence of price manipulation. This is a consequence of the price-update and price-impact slopes being different: the $\mu_{n,N}$'s only have to be chosen high enough, according to (9).

While all $\mu_{n,N}$'s must be non-negative, the signs of the $\lambda_{n,N}$'s are ambiguous. Negative $\lambda_{n,N}$'s are in discord with the interpretations that purchases either signal good news about the asset's value, or assert positive price pressure due to a declining inventory of the liquidity providers. But in this context negativity makes sense. The main mechanism that makes price manipulation successful is the positive relation between price update and trading volume. If the $\lambda_{n,N}$'s are negative, this mechanism would not work any more. For example, a purchase that drives up the price today but moves down future prices would erode the trader's ability to make money from trading.

The condition put forward in Theorem 2 is easiest interpreted when temporary price impacts are absent ($\lambda_{n,N} = \mu_{n,N}$ for $1 \leq n \leq N$). In this case, a positive semidefinite Λ_N implies that liquidity cannot increase over time by too high a rate. Otherwise, the trader could lock in expected profits from price manipulation: he begins pushing up the price in the early illiquid periods by consecutive purchases until the market becomes more liquid. He then sells the shares he is holding and makes profits since, due to the more liquid market, he can do the selling at an average price higher than the average purchase price. On the other hand, if liquidity decreases over time, price manipulation is infeasible. For instance, selling shares in the first periods and buying them back in later, less liquid periods would be unprofitable, because the purchase prices would exceed the selling prices on average.

Unfortunately, there is no handy characterizing condition for the absence of unbounded price manipulation or quasi-arbitrage. As a consequence, one has to verify whether quasi-arbitrage is possible for each individual case. However, this task is in general not too difficult, because the price-update and price-impact functions are linear. Here is an example. Consider $\tilde{p}_{0,N} = 10$, $\{\lambda_{n,3}\}_{n=1}^{3} = \{\mu_{n,3}\}_{n=1}^{3}$ with $\lambda_{1,3} = \lambda_{2,3} = 1$ and $\lambda_{3,3} = 0$. Buying q units of the asset in each of periods 1 and 2, and then selling the holdings in the third period results in expected profits of $\mathbb{E}[\pi(q^{0,3})] = -(10+q)q - (10+2q)q + 2q(10+2q) = q^2$, while $Std[\pi(q^{0,3})] = q\sqrt{\sigma_\eta^2(3) + 5\sigma_\varepsilon^2(3)}$. Hence, $\mathbb{E}[\pi(q^{0,3})] \to \infty$ and $\mathrm{SR}[\pi(q^{0,3})] \to \infty$ as $q \to \infty$. Such prices would not be weakly viable.

Remark: Breen et al. (2002), Chen et al. (2002), and Hasbrouck (1991) estimate the relation between prices and trading volume by employing price-impact functions which differ from ours. Theorems 1 and 2 are thus inapplicable to their models. Yet, a test for price manipulation can be easily conducted. Given their parameter estimates and reasonable assumptions on the levels of $\triangle_N$ and fixed transaction costs, price manipulation is feasible in each model. In such circumstances, empirical papers should either impose parameter conditions that exclude quasi-arbitrage or drop the assumption that the sequence of past trades does not affect the shape of the current price-impact function.

Remark: Nonlinear, time-dependent price-update functions may assume "chaotic" shapes without giving rise to price manipulation. Only when additional assumptions are made on the shapes of the price-update functions the analysis becomes meaningful. We focus here on linear price-impact and price-update functions, but one could impose other restrictions.

In the next subsection we will examine the insider trading problem based on the Kyle (1985) model in more detail.

4.2. *Discussion of the Kyle Model*

Black (1995) conjectures that the Kyle (1985) model allows "arbitrage opportunities" for uninformed agents if they pretend to be informed. We will prove that this conjecture is wrong and that the equilibrium price-update functions in Kyle must be linear asymptotically under certain conditions.

Kyle's model describes a game between competitive market-makers, who set the price in each period, and an individual risk-neutral insider trader, who has information on the liquidation value $v > 0$ of the single asset that is traded. The framework is as follows. All trades take place in the time interval $[0, 1]$ and the insider can choose any trade size. In each period, the market-makers observe only the aggregate trading volume, which is the sum of the insider's trading quantity and crowd's trades. They cannot observe the insider's amounts. Knowing the history of trades and the fact that there is one risk-neutral informed trader who maximizes his profits, they set the price equal to the conditional expected value, that is, $p_{n,N} = \mathbb{E}[v \mid \mathcal{F}^{(N)}_{n\triangle_N}]$ for $1 \leq n \leq N$. The insider, on the other hand, taking into account how the market-makers compute the price, submits in each round the quantity $q_{n,N}$ that maximizes his profits.

As Kyle shows for the case of normally and independently distributed trading volume of the crowd, this game has a unique linear equilibrium where the price evolves according to $p_{n,N} = p_{n-1,N} + \lambda_{n,N}(q_{n,N} + \eta_{n,N})$, the liquidity parameters $\lambda_{n,N}$ being endogenously (but deterministically) determined. But this price process is just a special case of (7) if the price impact has no temporary component and $\varepsilon_{n,N} = 0$. Thus, Theorem 2 can be applied to Kyle. *For any $\triangle_N$, Kyle's slopes are almost constant and hence price manipulation is infeasible. Consequently, Kyle's equilibrium is strongly viable.*

One can extend the Kyle model with one insider by allowing prices that incorporate more general nonlinear price-update functions, like $p_{n,N} = p_{n-1,N} + U_{n,N}(q_{n,N} + \eta_{n,N}) + \varepsilon_{n,N}$ (the price process given in (3)). Since the absence of unbounded price manipulation is necessary in equilibrium, it can be used to find the shape of equilibrium price-update functions.

PROPOSITION 5: *Suppose the price-update functions* $\{U_{n,N}\}_{n=1}^{N}$ *are symmetric, i.e.,* $U_{n,N}(x) = -U_{n,N}(-x) \geq 0$, $x \geq 0$, *and monotone in the sense that* $U_{n,N}(x) \leq U_{n-1,N}(x)$ *for* $x \geq 0$. *Then, given the price dynamics in (3), the absence of unbounded price manipulation in* $\mathcal{M}_1$ *implies that the expected price-update functions converge pointwise to a linear function on any interval* $(\tau, 1-\tau)$, $0 < \tau < 1/2$, *as* $N \to \infty$. *More formally, for any* $\varepsilon > 0$ *and* $x \in \mathbf{R}$ *there exists an index* N_0 *such that for* $N \geq N_0$ $\left|\hat{U}_{n,N}(x) - \lambda x\right| < \varepsilon$, *with* $\tau \leq n/N \leq 1-\tau$, $\lambda \geq 0$.

The monotonicity assumption implies that the insider's optimal policy causes the market maker to react less sensitively to trade size over time. This assumption is motivated by the properties of the linear equilibrium in Kyle.

Thus, as the number of auctions, N, becomes very large, only equilibria with approximately linear expected price-update functions are sustainable.

5. MULTIPLE ASSETS

So far we have discussed price manipulation and quasi-arbitrage only for one financial asset, but in typical applications investors trade many assets at the same time. In this section we extend our approach to the multivariate setting where a portfolio of $K > 1$ assets can be traded. The price-impact and price-update functions are assumed to be time-independent.

The multivariate case contains several interesting features not captured by the single-asset analysis. Presumably, the most important one is the ability to incorporate *cross-price impact* *functions*: the traded quantity of asset i affects not only the price of asset i but also the prices of other assets.

To extend formally the single-asset model to a multidimensional one, just replace all (single-valued) variables by K-dimensional vectors. In particular, the price-impact and price-update functions take each the traded quantities of the K assets as arguments, and map to a K vector, where component j gives the price-impact and price-update functions for asset j, respectively. Further, the non-negative slope λ in the definition of quasi-linearity (see Definition 5 above) now is a positive semidefinite matrix.

As the following example illustrates, cross-price-impact functions may lead to price manipulation. Suppose that there are two assets A and B, that only one share can be traded, and that price-impact functions are permanent. Table I shows a price-impact function for these two assets which exhibits asymmetric cross-price impacts. If one share of A is bought and none of B, then the price of A increases by one dollar and that of B by 50 cents. On the other hand, if one purchases one share of B and none of A, only the price of B rises (by one dollar). Now, buying one share of asset B in each of the first three periods, buying one share of A in each of the next eight periods, selling one share of B in each of the following three periods, and then selling one share of A in the subsequent eight periods, yields a profit of one dollar on average.

INSERT TABLE I ABOUT HERE

All results stated in Theorem 1 and Propositions 1 and 4 regarding the absence of price manipulation in $\mathcal{M}_1(\triangle_N)$ are literally true for the multi-asset case. Only Proposition 4 has to be adjusted slightly. Instead of condition (i) we impose $U = P$ (no temporary price impacts) as a sufficient condition.

Thus, the absence of price manipulation is equivalent to the price-impact function being linear, represented by a positive semidefinite matrix. In particular, all cross-price impact functions are linear and pairwise symmetric, that is, the price impact of asset i on asset j is the same as the price impact of asset j on asset i.

The concept of unbounded price manipulation or quasi-arbitrage is less applicable here. Trading infinite amounts introduces a technical difficulty of the following sort. Suppose there exists a quasi-arbitrage for asset A in the single-asset world. Introduce another asset B whose price goes down whenever asset A is bought. Since a quasi-arbitrage typically requires to buy a certain amount of shares of asset A infinitely many times, the price of B will eventually become negative, even if the negative price impact on B is tiny. To avoid such cases we would need to make additional assumptions about prices. We refrain from doing that in order not to lose generality. Anyway, price manipulation can become fairly large if the trade of an asset has a considerably larger impact on its own price than on the prices of other assets.

Black (1995) informally argues that the sum of the price update of individual trades must equal the price update of trading the "basket" containing these individual trades. In other words, the price update must be an additive function in the trading volume. Our results demonstrate that eliminating price manipulation requires more structure on the shape of the price-update function than Black claims.

6. PRICE MANIPULATION AND THE GAIN-LOSS RATIO

Bernardo and Ledoit (2000) propose to use the "gain-loss" ratio of an investment as a measure of its attractiveness. It is defined as the expectation of the investment's positive excess payoffs divided by the expectation of its negative excess payoffs. More formally, if z denotes the payoff of a zero-cost portfolio, then the gain-loss ratio equals $\mathrm{GLR}[z] \equiv \mathbb{E}[z^+]/\mathbb{E}[z^-]$, where

$z^+ = \max(z, 0)$ and $z^- = \max(-z, 0)$. One advantage of using the gain-loss ratio to detect attractive investment opportunities, rather than the Sharpe ratio, is that it recognizes a pure arbitrage with fat upper tails and flat lower tails as desirable, while the Sharpe ratio may not.

In the framework of Bernardo and Ledoit (2000) the absence of pure arbitrage is equivalent to the gain-loss ratio being finite. Thus, by imposing an upper bound on the gain-loss ratio arbitrage opportunities are ruled out. Even though our model generally does not exhibit this equivalence, one might want to exclude round-trip trades, $\{q_m^{0,N}\}_{m=1}^{\infty}$, with $\lim_{m\to\infty} \mathbb{E}[\pi(q_m^{0,N})] = \lim_{m\to\infty}\mathrm{GLR}[\pi(q_m^{0,N})] = \infty$. The existence of such *"great deals"* may threaten market viability.

To analyze great deals, we focus here only on time-independent price-impact and price-update functions. The market environment is described by $\mathcal{M}_1^*$ which is $\mathcal{M}_2$ with the difference that the growth conditions (as $\triangle_N$ becomes small) are imposed on the negative part of the random variables involved rather than on their variances (see Appendix A).

Everything derived in Section 3 for the absence of quasi-arbitrage in market $\mathcal{M}_2$ also applies to the absence of great deals in market $\mathcal{M}_1^*$. More precisely, all statements made in Theorem 1 and Propositions 1, 3, and 4 regarding the absence of quasi-arbitrage in $\mathcal{M}_2$ are also true for the absence of great deals in $\mathcal{M}_1^*$. Hence, linearity of the price-update function is necessary to rule out great deals. It is also sufficient for the absence of great deals if the conditions in Proposition 4 are met.

In order to examine market viability, suppose that the trader's utility is given by $u(\mathrm{GLR}[\pi(q^{0,N})])$, where u is strictly increasing and $\mathrm{GLR}[\pi(0)] \equiv 1$. The following result, which resembles Corollary 1, can be easily deduced when conditions (i)-(ii) in Proposition 4 are satisfied. The absence of great deals in $\mathcal{M}_1^*$ is equivalent to the strong viability of $\mathcal{M}_1^*$.

7. CONCLUDING REMARKS

This paper studies price manipulation and quasi-arbitrage for markets where trade size moves the price and prices are uncertain when trades are placed. A price manipulation and a quasi-arbitrage are both round-trip trading strategies, where the first creates a positive expected payoff, while the second produces an infinite expected payoff, as well as an infinite Sharpe ratio. Markets with time-independent price-impact functions are strongly viable if and only if there is no quasi-arbitrage, when agents' utility is measured by the Sharpe ratio of an investment opportunity.

We examine the conditions imposed by the absence of quasi-arbitrage on the functional shape of the temporary and permanent price effect of a trade. If the price-impact and price-update functions are time-independent and certain multiples of each other, then the absence of quasi-arbitrage is equivalent to the linearity of both functions. On the other hand, if the price-impact and price-update functions are independent, then only the price-update function must be linear in trading volume, while the temporary price impact can have various forms without offering quasi-arbitrage opportunities.

The theoretical microstructure literature usually assumes that the change in prices is time-independent and reacts linearly to trading volume. This paper demonstrates that the assumption of time independence of price changes already implies the linearity of the price-update function.

Linearity as a necessary condition for the absence of quasi-arbitrage calls for a careful examination of empirical estimations of price-update functions. To the extent that they detect deviations from linearity, one may wonder why some price manipulation possibilities had gone unexploited or one can suspect some misspecification (perhaps a time-dependent environment).

Postulating a finite gain-loss ratio instead of the absence of quasi-arbitrage does not change any of our conclusions. Also in this case the price-update function has to be linear, since otherwise the gain-loss ratio would become infinite.

The results of this paper call for one main extension, namely to permit the trading of market and limit orders at the same time. How do limit orders affect the market price? And what does a no-arbitrage condition look like if traders can submit market and limit orders simultaneously? Most important, we would like to examine how market and limit orders can coexist in an equilibrium exchange.

Columbia Business School, 3022 Broadway, New York, NY 10027,

and

Ziff Brothers Investments, L.L.C., 153 East 53rd Street, 43rd Floor, New York, NY 10022.

APPENDIX A: Details of the Market Classification

Consider two positive functions f and g, each defined on a set of the type $\mathbf{N}\backslash\{1, 2, ..., n-1\}$, $n \geq 1$. Following convention, write $f = O(g)$ to express the fact that f asymptotically evolves no faster than g, i.e., there exist constants $C > 0$ and $\bar{N}$ such that $f(N)/g(N) \leq C$ for all $N \geq \bar{N}$. Note that the domain of the functions $V_U(q, n, .)$ and $V_P(q, n, .)$ is $\mathbf{N}\backslash\{1, 2, ..., n-1\}$. In the following, the properties of the markets $\mathcal{M}_1(\triangle_N)$, $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_1^*$ are listed. In all markets, $\hat{P}_{n,N}(q) \geq 0$ and $\hat{U}_{n,N}(q) \geq 0$ if $q \geq 0$, and $\hat{P}_{n,N}(q) \leq 0$ and $\hat{U}_{n,N}(q) \leq 0$ if $q < 0$.

Market $\mathcal{M}_1(\triangle_N)$: (1) the exchange fixes the time between trades, $\triangle_N$; (2) the functions $\hat{P}_{n,N}$ and $\hat{U}_{n,N}$ exist for all $1 \leq n \leq N$; (3) $c(k) \propto k^e$, where $e < 2$.

Market $\mathcal{M}_1$: (1) the trader chooses the time between trades, $\triangle_N$; (2) the functions $\hat{P}_{n,N}$ and $\hat{U}_{n,N}$ exist for all $N \in \mathbf{N}$ and $1 \leq n \leq N$; (3) $\hat{U}_{n,N}(q) \geq -\hat{U}_{n,N}(-q)$ for all $q \in \mathcal{D}_M \cap \mathbf{R}^+$, $N \in \mathbf{N}$, and $1 \leq n \leq N$; (4) $c(k) = O(k^e)$, where $e < 2$.

Market $\mathcal{M}_2$: (1) the trader chooses the time between trades, $\triangle_N$; (2) the variances $V_P(q, n, N)$, $V_U(q, n, N)$, and $\sigma_\varepsilon^2(N)$ exist for all $q \in \mathcal{D}_M$, $N \in \mathbf{N}$, and $1 \leq n \leq N$; (3) $\hat{U}_{n,N}(q) \geq -\hat{U}_{n,N}(-q)$ for all $q \in \mathcal{D}_M \cap \mathbf{R}^+$, $N \in \mathbf{N}$, and $1 \leq n \leq N$; (4) $V_P(q, n, .) = O(N^{a(q,n)})$, $V_U(q, n, .) = O(N^{b(q,n)})$, and $\sigma_\varepsilon^2(N) = O(N^d)$, where $a(q, n) < 1$, $b(q, n) < 1$, and $d < 1$, for each $q \in \mathcal{D}_M$ and $n \in \mathbf{N}$; (5) $c(k) = O(k^e)$, where $e < 2$.

Market $\mathcal{M}_1^*$: (1) the trader chooses the time between trades, $\triangle_N$; (2) the expectations $\mathbb{E}[P_{n,N}(q + \eta_{n,N})^-]$, $\mathbb{E}[U_{n,N}(q + \eta_{n,N})^-]$, and $\mathbb{E}[\varepsilon_{1,N}^-]$ exist for all $q \in \mathcal{D}_M$, $N \in \mathbf{N}$, and $1 \leq n \leq N$; (3) $\hat{U}_{n,N}(q) \geq -\hat{U}_{n,N}(-q)$ for all $q \in \mathcal{D}_M \cap \mathbf{R}^+$, $N \in \mathbf{N}$, and $1 \leq n \leq N$; (4) $\mathbb{E}[P_{n,N}(q + \eta_{n,N})^-] = O(N^{a(q,n)})$, $\mathbb{E}[U_{n,N}(q + \eta_{n,N})^-] = O(N^{b(q,n)})$, and $\mathbb{E}[\varepsilon_{1,N}^-] = O(N^d)$, where $a(q, n) < 1$, $b(q, n) < 0$, and $d < 0$, for each $q \in \mathcal{D}_M$ and $n \in \mathbf{N}$; (5) $c(k) = O(k^e)$, where $e < 2$.

APPENDIX B: Proofs of the Results in Sections 3-5

Before proving Proposition 1 and Theorem 1, we derive two helpful results. To simplify the analysis, we make three technical assumptions:

(C1) If a sale is expected to induce a negative transaction price, then the expected revenue from this sale is set equal to zero;

(C2) If $q \in \mathcal{D}_M$ is irrational, then there exists at least one sequence $\{q_j\}_{j=1}^{\infty}$ such that all $q_j \in \mathcal{D}_M \cap \mathbf{Q}$ and $\lim_{j\to\infty} q_j = q$;

(C3) $\mathbb{P}[\varepsilon_{1,N} > a] > 0$ for all $a > 0$ and $N \in \mathbf{N}$.

Lemma 1: *Each of the conditions (NoPM)-(NoQA) in Theorem 1 implies:*

(i) $\hat{U}$ *is symmetric on* $\mathcal{D}_M$*, i.e.,* $\hat{U}(q) = -\hat{U}(-q)$ *for* $q \in \mathcal{D}_M$*; and*

(ii) $\lim_{j\to\infty} \hat{U}(q_j) = \hat{U}(q)$ *if* $\lim_{j\to\infty} q_j = q \neq 0$.

Proof: To verify *(i)* we start by proving that $\hat{U}(q) \leq -\hat{U}(-q)$ holds for all positive $q \in \mathcal{D}_M$. Suppose that this is not true, that is, there exists a $q > 0$ with $\hat{U}(q) > -\hat{U}(-q)$. Implement now the following trading strategy $q^{0,N}$: buy in each of the first m periods the volume q, and then sell the quantity q in each of the next m periods. This round-trip strategy implies non-negative expected prices and an expected profit given by (5). Therefore, if $\triangle_N$ is sufficiently small and $2m \leq N$, $\mathbb{E}[\pi(q^{0,N})] > 0$ and price manipulation is feasible.

If the variances exist, we can calculate

$$(10)\quad Var[\pi(q^{0,N})] \leq q^2 \left\{ \frac{m(m+1)(2m+1)}{6} V_U(q) + mV_P(q) + m(m+1)\sqrt{V_U(q)V_P(q)} \right.$$

$$\left. + \frac{(m-1)m(2m-1)}{6} V_U(-q) + mV_P(-q) + m(m-1)\sqrt{V_U(-q)V_P(-q)} + \frac{m(2m^2+1)}{3}\sigma_\varepsilon^2(N) \right\}$$

thanks to Minkowski's inequality, where $V_U(q) = V_U(q,n,N)$ and $V_P(q) = V_P(q,n,N)$. As $\mathbb{E}[\pi(q^{0,N})] = O(m^2)$ in $\mathcal{M}_1$ and $\mathcal{M}_2$, and $Std[\pi(q^{0,N})] = O(m^\theta)$, $\theta < 2$ in $\mathcal{M}_2$, the inequality $\hat{U}(q) > -\hat{U}(-q)$ also contradicts (NoUM) and (NoQA), for all sufficiently small $\triangle_N$.

Using even simpler arguments one can derive that $\hat{U}(0) = 0$.

Next, we show $\hat{U}(q) \geq -\hat{U}(-q)$ for all positive $q \in \mathcal{D}_M$, also by contradiction (only the case (NoPM) needs to be treated). For this purpose assume a $q > 0$ satisfying $\hat{U}(q) < -\hat{U}(-q)$. Now, conditional on $\tilde{p}_{0,N} + \varepsilon_{1,N} \geq \max\{-(m-1)\hat{U}(-q) - \hat{P}(-q), -m\hat{U}(-q)\}$, selling in each of the first m periods the quantity q and then buying the volume q in each of the following m periods results in $\mathbb{E}[\pi(q^{0,N})] > 0$ and $\mathbb{E}[p_{n,N}] \geq 0$ for all $1 \leq n \leq 2m \leq N$, if $\triangle_N$ is sufficiently small. As a consequence, (NoPM) is violated.

The second assertion, *(ii)*, is easiest shown by contradiction, too. Assume that $q \in \mathcal{D}_M \backslash \{0\}$ and $\lim_{j\to\infty} q_j = q$, and that *(ii)* does not hold, i.e., there exists a $q > 0$ (we can choose a positive q due to *(i)*) and $\varepsilon > 0$ such that one the following cases applies:

1. there exists a subsequence $q_{j'} \downarrow q$ with $\hat{U}(q_{j'}) \geq \hat{U}(q) + \varepsilon$,
2. there exists a subsequence $q_{j'} \downarrow q$ with $\hat{U}(q_{j'}) \leq \hat{U}(q) - \varepsilon$,
3. there exists a subsequence $q_{j'} \uparrow q$ with $\hat{U}(q_{j'}) \geq \hat{U}(q) + \varepsilon$,
4. there exists a subsequence $q_{j'} \uparrow q$ with $\hat{U}(q_{j'}) \leq \hat{U}(q) - \varepsilon$.

We shall show that U violates (NoPM)-(NoQA) in each case.

Case 1. Use the following strategy: buy $q_{j'}$ units of the asset in each of the first m periods, where j' is an arbitrary index of the subsequence; then sell the quantity q in the each of the following m periods and the remaining $m(q_{j'} - q)$ shares in period $2m+1$. Given *(i)*, the mean of these transactions' profit, $\mathbb{E}[\pi(q^{0,N})]$, is $O\left(m^2[(\hat{U}(q_{j'}) - \hat{U}(q))q + \hat{U}(q_{j'})(q - q_{j'})]/2\right)$, and the volatility, $Std[\pi(q^{0,N})]$, is $O(m^{\theta})$, $\theta < 2$. Since the coefficient in the former term is positive for sufficiently large j' (verify that the sequence $\{\hat{U}(q_{j'})\}$ must be bounded!), a contradiction to each of (NoPM)-(NoQA) is established, for all sufficiently small $\triangle_N$.

Case 2. Trading strategy: buy volume q in the each of the first m periods and then sell $q_{j'}$ units in each of the next $m-1$ periods and the remaining $mq-(m-1)q_{j'}$ shares in period $2m$ ($(m-1)\,q_{j'} \leq mq$ is met if j' is large enough). This implies $\mathbb{E}[\pi(q^{0,N})]$ to be $O\left(m^2[(\hat{U}(q)-\hat{U}(q_{j'}))q_{j'}+\hat{U}(q)(q_{j'}-q)]/2\right)$ and $Std[\pi(q^{0,N})] = O(m^{\theta})$, $\theta < 2$. But the coefficient in the first expression becomes positive if j' is sufficiently large. Again, (NoPM)-(NoQA) are all invalid, if $\triangle_N$ is sufficiently small.

The reader can easily check that for the remaining cases the following two trading strategies contradict each of (NoPM)-(NoQA): for case 3, buy $q_{j'}$ units in each of the first m periods and then sell quantity q in each of the following $m-1$ periods and the remaining shares in period $2m$; for case 4, buy q units in each of the first m periods and then sell the volume $q_{j'}$ in each of the next m periods and the residual shares in period $2m+1$. *Q.E.D.*

LEMMA 2: *Each of (NoPM)-(NoQA) requires that U satisfies the integral equation*

$$(11) \quad \int_{\Omega} R_U(q+\eta)d\mathbb{P} = 0 \quad \textit{for all } q \in \mathcal{D}_M,$$

where $R_U(x) \equiv U(x) - \lambda x$, $\lambda \geq 0$, $x \in \mathcal{D}$, and η assumes the distribution of the crowd's trades.

PROOF: Note that (11) is equivalent to $\hat{U}(q) = \lambda q$ for all $q \in \mathcal{D}_M$. To prove Lemma 2, suppose that $\hat{U}$ does not have the above property, i.e., there exists a $q > 0$, such that $\hat{U}(q) > \hat{U}(1)\,q$ or $\hat{U}(q) < \hat{U}(1)\,q$. Let us deal with the first case. Thanks to Lemma 1 *(ii)* and assumption (C2) we can choose q to be a rational number. Implement now the following trading strategy: buy q units of the asset in each of the first m periods such that mq is an integer, then sell one unit in each of the following mq periods ($m(1+q) \leq N$). It follows that $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}(q) - \hat{U}(1)q]/2)$ and $Std[\pi(q^{0,N})] = O(m^\theta)$, $\theta < 2$, contradicting each of (NoPM)-(NoQA), if $\triangle_N$ is sufficiently small.

The case $\hat{U}\,(q) < \hat{U}\,(1)\,q$ can be tackled similarly: it is easy to verify that the strategy of buying one unit in each of the first mq periods and then selling q units in each of the next m periods results in a violation of each (NoPM)-(NoQA). *Q.E.D.*

Proof of Theorem 1: Immediate consequence of Lemmas 1 and 2. *Q.E.D.*

Proof of Proposition 1: We only have to study here equation (11).

If $\mathbb{P}[\eta_{1,N} = 0] = 1$, then $U(q) = \lambda q$, $L(\mathbf{R}) - a.e.$, follows immediately from (11).

To simplify the analysis for case *(ii)*, we assume that there exists a number $a \in (0,1)$ (preferably close to one) such that the function $x \longmapsto U(x)e^{-ax^2/(2\sigma_\eta^2(N))}$ is $L(\mathbf{R})$-integrable. This is a mild assumption because $\mathbb{E}[U(\eta_{1,N})] < \infty$ in any case.

For normally distributed $\eta_{n,N}$'s the integral equation (11) becomes

$$(12) \quad \frac{1}{\sqrt{2\pi}\sigma_\eta(N)} \int_{\mathbf{R}} R_U(x) e^{-\frac{(x-q)^2}{2\sigma_\eta^2(N)}} dx = 0 \quad \text{for all } q \in \mathbf{R},\ \lambda \geq 0.$$

Using the above assumption, it is an easy exercise to verify that (12) can be reformulated as

$$(13) \quad \int_{\mathbf{R}} \left[R_U(x) e^{-a\frac{x^2}{2\sigma_\eta^2(N)}} \right] \left[\frac{1}{\sqrt{2\pi}\sigma_\eta(N)/\sqrt{1-a}} e^{-\frac{(x-q)^2}{2\sigma_\eta^2(N)/(1-a)}} \right] dx = 0$$

for all $q \in \mathbf{R}$, $\lambda \geq 0$.

Now, recall that the Fourier transform $F[f] : \mathbf{R} \to \mathbf{C}$ of a $L(\mathbf{R})$-integrable function $f : \mathbf{R} \to \mathbf{R}$ is defined by $F[f](x) \equiv \int_{\mathbf{R}} e^{ixy} f(y)dy$. Invoking the convolution theorem of Fourier transforms for (13) gives

$$F\left[y \mapsto R_U(y) e^{-a\frac{y^2}{2\sigma_\eta^2(N)}}\right](x)\; e^{-\frac{\sigma_\eta^2(N)}{2(1-a)}x^2} = 0 \quad \text{for all } x \in \mathbf{R},$$

which implies that $R_U = 0$, $L(\mathbf{R}) - a.e.$, since F is injective. So $U(q) = \lambda q$, $L(\mathbf{R}) - a.e.$, holds also for the case of normally-distributed $\eta_{n,N}$'s. *Q.E.D.*

Proof of Proposition 3: The result follows from the fact that, if $P(x) = U(x)/2 = \lambda x/2$, then $\mathbb{E}[\pi(q^{0,N})]$ equals

$$\mathbb{E}\Bigg[-\tfrac{\lambda}{2}\left(\textstyle\sum_{n=1}^{N} q_{n,N}\right)^2 + \lambda \textstyle\sum_{n=1}^{N-1} \eta_{n,N} \sum_{j=1}^{n} q_{j,N} - \tfrac{\lambda}{2}\sum_{n=1}^{N} \eta_{n,N} q_{n,N}$$
$$+ \textstyle\sum_{n=2}^{N} \varepsilon_{n,N} \sum_{j=1}^{n-1} q_{j,N}\Bigg] - c(T(q^{0,N})),$$

which is $-c(T(q^{0,N})) < 0$, for all round-trip trades $q^{0,N} \neq 0$. *Q.E.D.*

Proof of Proposition 5: Take any $\varepsilon > 0$ and define $n(\tau, N) \equiv \max\{j \in \mathbf{N} \mid j/N \leq \tau\}$, $\tau < 1/2$. First, we demonstrate that for an arbitrary $q \geq 0$, $\hat{U}_{n(\tau,N),N}(q) \leq \hat{U}_{N-n(\tau,N),N}(q) + \varepsilon$, if N is sufficiently large. Suppose not, i.e., there exists a sequence $N_m \to \infty$ such that $\hat{U}_{n(\tau,N_m),N_m}(q) > \hat{U}_{N_m-n(\tau,N_m),N_m}(q) + \varepsilon$. Then buying q shares in each of the first $n(\tau, N_m)$ periods, and then selling q shares in each of the periods $N_m - n(\tau, N_m) + 1$, $N_m - n(\tau, N_m) + 2, \ldots, N_m$, results in $\mathbb{E}[\pi(q^{0,N_m})] = O(n(\tau, N_m)^2 \varepsilon)$, which contradicts (NoUM).

Next, we verify that the inequalities

(14) $\quad q\hat{U}_{N-n(\tau,N),N}(1) - \varepsilon \leq \hat{U}_{N-n(\tau,N),N}(q) \leq q\hat{U}_{N-n(\tau,N),N}(1) + \varepsilon,$

$q \geq 0$, must hold for sufficiently large N. If the second inequality were false, then there would exist a $q > 0$ and a sequence $N_m \to \infty$ such that $\hat{U}_{N_m - n(\tau, N_m), N_m}(q) > q\hat{U}_{N_m - n(\tau, N_m), N_m}(1) + \varepsilon$. The strategy purchasing q shares in each of the first n'_m periods, and then selling one share in periods $N_m - n(\tau, N_m) + 1$, $N_m - n(\tau, N_m) + 2, \ldots$, $N_m - n(\tau, N_m) + qn'_m$, where n'_m and $qn'_m \in \mathbf{N}$ are smaller than $n(\tau, N_m)$, constitutes a price manipulation. The first inequality in (14) can be shown by a similar argument. The proposition follows then from the maintained assumptions. *Q.E.D.*

Proof of Multi-Asset Versions of Theorem 1, Propositions 1, and 4 (regarding price manipulation): We start by demonstrating the first part of Theorem 1. For this purpose we define $\hat{U}_{ij}(q) \in \mathbf{R}$ to be the expected price update of asset i when $q \in \mathcal{D}^j_M \subseteq \mathbf{R}$ shares of asset j and none of the other assets are traded. The proof is divided into five steps. If necessary, a round-trip trade below is conditioned on $\varepsilon_{1,N}$ to be sufficiently large so that it does not cause negative expected prices (recall assumption (C3)).

Step 1: $\hat{U}_{ij}$ is symmetric, i.e., $\hat{U}_{ij}(q) = -\hat{U}_{ij}(-q)$ on $\mathcal{D}^j_M$.

If not, then either (i) $\hat{U}_{ij}(q) > -\hat{U}_{ij}(-q)$ or (ii) $\hat{U}_{ij}(q) < -\hat{U}_{ij}(-q)$ for a $q > 0$, or (iii) $\hat{U}_{ij}(0) \neq 0$. For case (i) consider the strategy of buying q shares of asset i in each of the first m periods, buying q shares of asset j in each of the next m periods, selling q shares of asset j in each of the next m periods, and selling q shares of asset i in each of the following m periods. This implies $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}_{ij}(q) + \hat{U}_{ij}(-q)])$, given that $\triangle_N$ is sufficiently small. For case (ii) consider selling in each of the first m periods q shares of asset i, buying in each of the next m periods q shares of asset j, selling in each of the next m periods q shares of asset j, and buying in each of the subsequent m periods q shares of asset i. Then, $\mathbb{E}[\pi(q^{0,N})] = O(-m^2 q[\hat{U}_{ij}(q) + \hat{U}_{ij}(-q)])$. Both trading strategies contradict (NoPM). Case (iii) is easy to rebut and left to the reader.

Step 2: $\lim_{k\to\infty}\hat{U}_{ij}(q_k) = \hat{U}_{ij}(q)$ if $\lim_{k\to\infty} q_k = q$. In what follows, we verify that none of the four discontinuities stated in the proof of Lemma 1 can hold for $\hat{U}_{ij}$. For the first discontinuity take the strategy of buying q shares of asset i in each of the first m periods, buying $q_{k'}$ shares of asset j in each of the next m periods, selling q shares of asset j in each of the next m periods, selling q shares of asset i in each of the following m periods, and selling $m(q_{k'} - q)$ shares of asset j in period $4m+1$. This results in $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}_{ij}(q_{k'}) - \hat{U}_{ij}(q)])$. For the second discontinuity, consider buying q shares of asset i in each of the first m periods, buying q shares of asset j in each of the next m periods, selling $q_{k'}$ shares of asset j in each of the following $m-1$ periods, selling q shares of asset i in each of the next m periods, and selling $mq - (m-1)q_{k'}$ shares of asset j in period $4m$. As a consequence, $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}_{ij}(q) - \hat{U}_{ij}(q_{k'})])$. In the third case, take the strategy of buying q shares of asset i in each of the first m periods, buying $q_{k'}$ shares of asset j in each of the next m periods, selling q shares of asset j in each of the next $m-1$ periods, selling q shares of asset i in each of the following m periods, and selling $mq_{k'} - (m-1)q$ shares of asset j in period $4m$. We obtain $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}_{ij}(q_{k'}) - \hat{U}_{ij}(q)])$. For the last discontinuity, consider buying q shares of asset i in each of the first m periods, buying q shares of asset j in each of the next m periods, selling $q_{k'}$ shares of asset j in each of the following m periods, selling q shares of asset i in each of the next m periods, and selling $m(q - q_{k'})$ shares of asset j in period $4m+1$. This yields $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}_{ij}(q) - \hat{U}_{ij}(q_{k'})])$. All trading strategies contradict (NoPM).

Step 3: $\hat{U}_{ij}(q) = \hat{U}_{ij}(1)q$ on $\mathcal{D}_M^j$. If it were not, either $\hat{U}_{ij}(q) > \hat{U}_{ij}(1)q$ or $\hat{U}_{ij}(q) < \hat{U}_{ij}(1)q$ for a $q > 0$. In the first case, the trading strategy of buying q shares of asset i in the first m periods, buying q shares of asset j in the next m periods, selling one share of asset j in each of the next

mq periods, and selling q shares of asset i in each of the next m periods gives $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}_{ij}(q) - \hat{U}_{ij}(1)q])$. In the second case, we obtain $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}_{ij}(1)q - \hat{U}_{ij}(q)])$ from buying q shares of asset i in each of the first m periods, buying one share of asset j in each of the next mq periods, selling q shares of asset j in each of the next m periods, and selling q shares of asset i in each of the following m periods. Hence, both round-trip trades are at variance with (NoPM).

Step 4: $\hat{U}_{ij} = \hat{U}_{ji}$. Consider the strategy of buying q shares of asset i in each of the first m periods, buying q shares of asset j in each of the next m periods, selling q shares of asset i in each of the next m periods, and selling q shares of asset j in each of the next m periods. This implies $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q[\hat{U}_{ij}(q) - \hat{U}_{ji}(q)])$. Obviously, this is in discord with (NoPM) if $\hat{U}_{ij}(q) > \hat{U}_{ji}(q)$ for a $q > 0$. By symmetry, $\hat{U}_{ji}(q) \leq \hat{U}_{ij}(q)$, $q > 0$, and therefore $\hat{U}_{ij} = \hat{U}_{ji}$.

Last Step: $\hat{U}_i(q_1, q_2, \ldots, q_K) = \sum_{j=1}^{K} \hat{U}_{ij}(q_j)$. For brevity we prove the latter equality only for the case $K = 2$ here; the extension to arbitrary K is straightforward. Take m even and employ the following two strategies.

Strategy `X`: trade $-q_j$ shares of asset j in each of the first $m/2$ periods, trade (q_1, q_2) shares each of asset i and asset j in each of the next m periods, trade q_j shares of asset j in each of the next $m/2$ periods, trade $-q_j$ shares of asset j in each of the following m periods, and trade $-q_i$ shares of asset i in each of the next m periods;

Strategy `Y`: trade $-q_j$ shares of asset j in each of the first m periods, trade $-q_i$ shares of asset i in each of the next m periods, trade q_j shares of asset j in each of the following $m/2$ periods, trade (q_1, q_2) shares each of asset i and asset j in each of the next m periods, and trade $-q_j$ shares of asset j in each of the next $m/2$ periods.

Strategy `X` gives rise to $\mathbb{E}[\pi(q^{0,N})] = O(m^2 q_i[\hat{U}_i(q_1, q_2) - \hat{U}_{ii}(q_i) - \hat{U}_{ij}(q_j)])$, while strategy

Y has $\mathbb{E}[\pi(q^{0,N})] = O(-m^2 q_i[\hat{U}_i(q_1, q_2) - \hat{U}_{ii}(q_i) - \hat{U}_{ij}(q_j)])$ as a result. Thus, regardless of the value of q_i, (NoPM) implies $\hat{U}_i(q_1, q_2) = \hat{U}_{ii}(q_i) + \hat{U}_{ij}(q_j)$ and $\hat{U}_i$ is linear.

The slope λ of U has to be positive semidefinite: if there exists a q such that $q^T \lambda q < 0$, then consider the strategy of trading the vector q in the first period, and then trading $-q$ in second period. The result is $\mathbb{E}[\pi(q^{0,N})] = -q^T \lambda q$. Therefore, (NoPM) requires λ to be positive semidefinite. This completes the proof of the first part of Theorem 1.

The multi-asset version of Proposition 1 regarding (NoPM) is shown by using the same arguments as for the single-asset case. If $U = P$, then $\arg\sup_{q^{0,N}} \mathbb{E}[\pi(q^{0,N})] = 0$, as Huberman and Stanzl (2002) prove. Hence, the multiple-asset version of Proposition 4 for (NoPM) holds. *Q.E.D.*

APPENDIX C: Non-Zero Residual Functions

Example A *(Bernoulli distribution) Suppose $\mathcal{D}_M = \mathbf{R}$ and that the residual trades can only assume two values with positive probability, namely, $\mathbb{P}[\eta_{1,N} = -\eta_0] = \mathbb{P}[\eta_{1,N} = \eta_0] = 1/2 = 1/2 + \mathbb{P}[\eta_{1,N} = 0]$ for $n \in N$ and $\eta_0 > 0$. In this case, U is quasi-linear if and only if $U(x) = \lambda x + R_U(x)$, where $R_U(x) = -R_U(x - 2\eta_0)$ for all $x \in \mathbf{R}$.*

Example B *(Uniform distribution) Assume $\mathcal{D}_M = \mathbf{R}$ and that the $\eta_{n,N}$'s are uniformly distributed on $\mathbf{R}$, with compact support $[-s, s]$, $s > 0$, and that U is continuous and of bounded variation on $\mathbf{R}$. Then, U is quasi-linear if and only if $U(x) = \lambda x + R_U(x)$ where R_U is a $2s$-periodic trigonometric Fourier series satisfying $\int_0^{2s} R_U(x)dx = 0$. (For the precise form of R_U see below.)*

Observe that to derive the result in Example B, we need to impose smoothness assumptions on U, unlike the results in Propositions 1 and 4 and Example A.

In Examples A and B R_U can take on a variety of functional forms. For instance, any multiple of the sine function would be a possible candidate for the function R_U in Example B, if the periodicity is $s = 2\pi$. The reader is invited to construe candidate R_U-functions for Example A.

The precise shape of R_U is determined by the curvature of the density function of the crowd's trades and is therefore variable. For more complicated distributions, R_U can still be computed, albeit with much more intricate structure.

For the proofs note that $\int_\Omega R_U(q+\eta_{n,N})d\mathbb{P} = 0$ for all $q \in \mathcal{D}_M$ is equivalent to $\mathbb{E}_{n,N}[R_U(\tilde{q}_{n,N}+\eta_{n,N})] = 0$, for any $\mathcal{G}^{(N)}_{n\triangle_N}$-measurable random variable $\tilde{q}_{n,N}$.

Proof of Example A: $\int_\Omega R_U(q + \eta_{n,N})d\mathbb{P} = 0$ is just $R_U(q + \eta_0) + R_U(q - \eta_0) = 0$,

$q \in \mathbf{R}$. *Q.E.D.*

Proof of Example B: Under the assumption that U is quasi-linear, equation (6) becomes

$$(15)\quad \int_{q-s}^{q+s} R_U(x)dx = 0 \quad \text{for all } q \in \mathbf{R}.$$

By differentiating the above integral equation with respect to q, we obtain that R_U is $2s$-periodic on $\mathbf{R}$. Since, by assumption, R_U is continuous and of bounded variation, it has a trigonometric Fourier representation given by

$$(16)\quad R_U(q) = \sum_{n=1}^{\infty} \left[a_n \cos(\frac{\pi n}{s} q) + b_n \sin(\frac{\pi n}{s} q) \right],$$

$$a_n = \frac{1}{s} \int_{-s}^{s} R_U(x) \cos(\frac{\pi n}{s} x)dx, \;\; b_n = \frac{1}{s} \int_{-s}^{s} R_U(x) \sin(\frac{\pi n}{s} x)dx.$$

The above Fourier series does not have an intercept part, a_0, since $a_0 = \int_{-s}^{s} R_U(x)dx = 0$. Hence R_U possesses the claimed $2s$-periodic trigonometric Fourier series with $\int_0^{2s} R_U(x)dx = 0$.

Conversely, if $R_U(q) = U(q) - \lambda q$ is $2s$-periodic and satisfies $\int_0^{2s} R_U(x)dx = 0$, then it has the representation (16) and meets (15). *Q.E.D.*

REFERENCES

Allen, F., and D. Gale (1992): "Stock-Price Manipulation," *Review of Financial Studies*, 5, 503-529.

Amihud, Y., and H. Mendelson (1980): "Dealership Market: Market-Making with Inventory," *Journal of Financial Economics*, 8, 31-53.

Bernardo, A. E., and O. Ledoit (2000): "Gain, Loss, and Asset Pricing," *Journal of Political Economy*, 108, 144-172.

Bertsimas, D., and A. W. Lo (1998): "Optimal Control of Execution Costs," *Journal of Financial Markets* 1, 1-50.

Black, F. (1995): "Equilibrium Exchanges," *Financial Analysts Journal*, May-June, 23-29.

Breen, W. J., L. Hodrick, and R. A. Korajczyk (2002): "Prediciting Equity Liquidity," *Management Science*, 48, 470-483.

Chan, L. K., and J. Lakonishok (1995): "The Behavior of Stock Prices around Institutional Trades," *Journal of Finance*, 50, 1147-1174.

Chen, Z., W. Stanzl, and M. Watanabe (2002): "Price-Impact Costs and the Limit of Arbitrage," Unpublished Manuscript, Yale University.

Cochrane, J. H., and J. Saa-Requejo (2000): "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets," *Journal of Political Economy*, 108, 79-119.

Dutta, P. K., and A. Madhavan (1995): "Price Continuity Rules and Insider Trading," *Journal of Financial and Quantitative Analysis*, 30, 199-221.

Dybvig, P. H., and S. A. Ross (1987): "Arbitrage," in *The New Palgrave: Finance*, ed. by J. Eatwell, M. Milgate, and P. Newman, MacMillan Press, New York, 57-71.

Easley, D., and M. O'hara (1987): "Price, Trade Size, and Information in Securities

Markets," *Journal of Financial Economics*, 19, 69-90.

——— (1992), "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 577-605.

Gemmill, G. (1996): "Transparency and Liquidity: A Study of Block Trades on the London Stock Exchange under Different Publication Rules," *Journal of Finance*, 51, 1765-1790.

Glosten, L. R., and P. Milgrom (1985): "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents," *Journal of Financial Economics*, 14, 71-100.

Hasbrouck, J. (1991): "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46, 179-207.

Hasbrouck, J., and T. Ho (1987): "Order Arrival, Quote Behavior, and the Return Generating Process," *Journal of Finance*, 42, 1035-1048.

Hausman, J. A., A. W. Lo, and A. C. MacKinlay (1992): "An Ordered Probit Analysis of Transaction Stock Prices," *Journal of Financial Economics*, 31, 319-379.

Ho, T., and H. R. Stoll (1981): "Optimal Dealer Pricing Under Transactions and Return Uncertainty," *Journal of Financial Economics*, 9, 47-73.

Holthausen, R., R. Leftwich, and D. Mayers (1987): "The Effect of Large Block Transactions on Security Prices," *Journal of Financial Economics*, 19, 237-267.

——— (1990): "Large Block Transactions, the Speed of Response, and Temporary and Permanent Stock-Prices Effects," *Journal of Financial Economics*, 26, 71-95.

Huberman, G. (1982): "A Simple Approach to Arbitrage Pricing Theory," *Journal of Economic Theory*, 28, 183-191.

Huberman, G., and W. Stanzl (2002): "Optimal Liquidity Trading," Unpublished

Manuscript, Columbia University.

JARROW, R. A. (1992): "Market Manipulation, Bubbles, Corners, and Short Squeezes," *Journal of Financial and Quantitative Analysis*, 27, 311-336.

KEIM, D. B., AND A. MADHAVAN (1996): "The Upstairs Market for Large Block Transactions: Analysis and Measurement of Price Effects," *Review of Financial Studies*, 9, 1-36.

KEMPF, A., AND O. KORN (1999): "Market Depth and Order Size," *Journal of Financial Markets*, 2, 29-48.

KYLE, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315-1336.

LEDOIT, O. (1995): "Essays on Risk and Return in the Stock Market," Ph.D. dissertation, MIT.

MADHAVAN, A., M. RICHARDSON, AND M. ROMMANS (1997): "Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks," *Review of Financial Studies*, 10, 1035-1064.

MERTON, R. C. (1990): *Continuous-time Finance*, Blackwell, Cambridge, MA, & Oxford, UK.

SCHOLES, M. S. (1972): "The Market for Securities: Substitution versus Price Pressure and the Effects of Information on Share Prices," *Journal of Business*, 45, 179-211.

FOOTNOTES

[1]We are grateful to Prajit Dutta, Charles Jones, and Matthew Rhodes-Kropf for conversations and comments. Larry Glosten was especially and extremely helpful and supportive. We also thank the participants of the NASDAQ-Notre Dame Conference 2000, the Western Finance Association Meetings 2001, the European Finance Association Meetings 2001, and of the ESSFM Meetings 2001 in Gerzensee; in particular, Craig Holden, Eugene Kandel, Mark Loewenstein, Massimo Massa, and Anna Pavlova. An earlier version of this paper circulated under the title "Arbitrage-free Price-Update and Price-Impact Functions."

TABLES

TABLE I

PRICE-IMPACT FUNCTION WITH ASYMMETRIC CROSS-PRICE IMPACT

		Asset B		
		0	1	-1
Asset A	0	(0, 0)	(0, 1)	(0, -1)
	1	(1, ½)	(1, 1)	(0, 0)
	-1	(-1, -½)	(0, 0)	(-1, -1)

LIST OF SYMBOLS

1. Model related:

$\triangle_N$, $\tilde{p}_{n,N}$, $p_{n,N}$, $\tilde{q}_{n,N}$, $q_{n,N}$, $P_{n,N}$, $U_{n,N}$, $\varepsilon_{n,N}$, $\eta_{n,N}$,

$\hat{P}_{n,N}$, $\hat{U}_{n,N}$, $\hat{P}$, $\hat{U}$, $\hat{U}_{ij}$, $q^{0,N}$, $q_m^{0,N}$, V_P, V_U,

$T(q^{0,N})$, $\pi(q^{0,N})$, $\mathrm{SR}(q^{0,N})$, $\mathrm{GLR}[z]$, R_U, $\{q_j\}_{j=1}^{\infty}$, $c(.)$,

Ω, $\mathcal{F}$, $\mathbb{P}$, $\mathbb{E}$, Var, Std, $\mathbb{E}_{n,N}$, $\mathcal{F}_{n\triangle_N}^{(N)}$, $\mathcal{G}_{n\triangle_N}^{(N)}$,

$\mathcal{D}_M$, $\mathcal{D}_\eta$, $\mathcal{D}_\varepsilon$, $\mathcal{D}$, $\mathcal{D}_M^j$, $\mathcal{M}_1(\triangle_N)$, $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_1^*$,

$\lambda_{n,N}$, $\mu_{n,N}$, $\varepsilon_{1,N}^-$, z^+, z^-, u, σ_ε^2, σ_η^2, Λ_N, $\Lambda_{j,N}$,

α, λ, η, $N^{a(q,n)}$, $\mathtt{X}$, $\mathtt{Y}$.

2. Parameters & Constants:

θ, θ_m, δ, $\overline{q}$, v, n, N, N_0, N_m, x, τ, K,

i, j, π, n_m, n_m', η_0, a_0, ε, $\bar{N}$.

3. Sets:

$\mathbf{N}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{R}^+$, $L(\mathbf{R})$, $\mathbf{C}$.

4. Operators:

$\equiv$, $\subseteq$, $\propto$, $\neq$, $\gtreqless$, $\mp$, $\cap$, ∞, $\in$, $\downarrow$, $\uparrow$, $\rightarrow$,

$F[f]$, $O(g)$, $\sum_{n=1}^N$, $\sum_{n=1}^\infty$, $\int_\Omega$, $\int_{\mathbf{R}}$, $\int_a^b$, $\sqrt{x}$, $\sin$,

$\cos$, $\max$, $\min$, $\sup$, $\inf$, $\arg\max$, $\arg\sup$, $\lim$, $\det$.