Yale School of Management
Yale University

Working Paper No. ICF – 00-20

International Center for Finance
Yale University

Working Paper No. ICF – 00-20

Arbitrage-Free Price-Update and Price-Impact Functions

Werner Stanzl
Gur Huberman

ARBITRAGE-FREE PRICE-UPDATE AND PRICE-IMPACT FUNCTIONS

Gur Huberman
Columbia Business School

Werner Stanzl
Yale School of Management

Working Paper No. 00-20

OCTOBER 7, 2000

Arbitrage-free Price-Update and Price-Impact Functions*

Gur Huberman
Columbia Business School

Werner Stanzl
Yale School of Management

October 7, 2000

*We are grateful to Prajit Dutta and Larry Glosten for numerous conversations and comments. We also thank the participants of the NASDAQ-Notre Dame Conference 2000, where this paper was presented; in particular, Craig Holden and Eugene Kandel. Address for correspondence: Werner Stanzl. Yale School of Management, International Center for Finance, 135 Prospect Street, P.O. Box 208200, New Haven, CT 06520-8200, USA. Phone: (203) 436-0666, fax: (203) 436-0630, email: werner.stanzl@yale.edu.

Arbitrage-free Price-Update and Price-Impact Functions

ABSTRACT

Consider a trading environment where trading volume affects security prices. We show that when the price impact is time stationary, only linear price-impact functions rule out arbitrage. This is true whether a single asset or a portfolio of assets is traded. When the temporary and permanent effects of trades on prices are independent, only the permanent price impact must be linear while the temporary one can be of a more general form. We also examine what arbitrage-free temporary and permanent price impacts must look like in a nonstationary framework.

IN ANY MARKET, 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 with absolute precision, and is tempted to exploit this knowledge to his advantage. What are the possible relations between price changes and trades that rule out arbitrage for this market watcher?

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 traded 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 stationarity. Specifically, price-update functions that admit no arbitrage possibilities are linear in trade size.

Recent empirical papers assume in addition to stationarity 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 our work, these empirical results imply the feasibility of profitable manipulation. Alternatively, our work calls into question either the stationarity 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 latter distinction.

Black (1995) anticipates some of this paper's results. Black 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-stationary framework where trades have permanent price impacts only. Our 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 arbitrage. It thereby selects which price-update and price-impact functions qualify for equilibrium.

The rest of the paper is organized as follows. Section I presents a preview of our results. Section II introduces the model and characterizes the absence of arbitrage when the price-update and price-impact functions are stationary. Section III investigates nonstationary price-update and price-impact functions. Section IV treats multi-asset price dynamics, and Section V concludes. All proofs are relegated to the Appendix.

I A Preview of the Results

The results of this paper are formulated in ten propositions which offer conditions on the shape of the price impact functions that are necessary, sufficient, or equivalent to the absence of arbitrage. The first six propositions deal with time-stationary price-update and price-impact functions for one asset. Propositions 7 and 8 relax the assumption of stationarity and allow liquidity to vary across time, still in the single-asset framework. Propositions 9 and 10 tackle multi-asset, time-stationary price-update and price-impact functions. We omit nonstationary functions for the multiple-asset case because their analysis would add little to the findings in Propositions 7 and 8.

Propositions 1 and 2 assert that the price-update function must be linear in trade size to rule out arbitrage, regardless of the shape of the temporary price impact. Since Proposition 1 makes no assumptions about the distributions of the random elements of the price, such as news arrival and noise trades, the linearity of the price-update function holds only in expected terms. However, in Proposition 2, where all relevant random variables are normal, the linearity derived is exact.

As previously mentioned, this linearity contradicts empirical findings of nonlinear price updates. One argument often advanced to justify nonlinear price-update functions is that transaction costs outweigh the gains derived from exploiting nonlinear price updates. For example, if the price-update function is concave for purchases and convex for sales, as found in Hasbrouck (1991) and Kempf and Korn (1999), it would induce an arbitrageur to bid up the price by buying many small lots of shares over time and then to unwind his position by selling all the shares at once. Without transaction costs, such a strategy is profitable on average if sufficiently many trades are done. However, even in the presence of transaction costs, arbitrage is possible. This is the main implication of Propositions 1 and 2. Because transaction costs can be modeled by the price-impact function, Propositions 1 and 2 imply that fixed trading fees are insufficient to prohibit arbitrage if the price-update function is nonlinear.

Propositions 1 and 2 are useful for choosing equilibrium price-update and price-impact functions, for a general equilibrium should be arbitrage-free. Contrast this with Allen and Gale (1992) who employ a Glosten and Milgrom-type (1985) framework to construct equilibria in which uninformed agents profitably manipulate the security price. For this to happen, it is crucial to allow individuals to trade only one unit per period, and to forbid direct trading between the uninformed manipulator and informed agents. If the manipulator interacts only with uninformed agents, he can affect the price with his orders because he is considered to be possibly informed by those agents. This misperception by his trading partners allows the uninformed manipulator to move the price in the direction he needs for arbitrage profits. But if informed agents can choose when and with whom to trade, price manipulation is unlikely to happen. The insiders would on average countermatch the manipulator's trades, since they are not based on any information, and thus the price would not be updated by the uninformed traders in the market. In a free trading environment, equilibrium with arbitrage seems therefore unlikely.

Propositions 3 and 4 impose conditions on the random elements of the price and on the price-impact function, respectively. Proposition 3 contends that the absence of arbitrage rules out trends: the conditional expectation of noise trades and news must be zero. Otherwise, expected profits from arbitrage would be unbounded. Proposition 4 states two conditions that the price-impact function must respect in order to be arbitrage-free. First, the difference between the price impact of buying q shares and that of selling q shares must be no smaller than the price update of buying q shares. Second, the price-impact function cannot be constant unless the price-update function is zero. Both facts will become clearer after our model has been introduced below.

If the price-impact function is a multiple of the price-update function, then Proposition 5 shows that the absence of arbitrage is characterized by the linearity of both functions involved. Hence, linearity is also sufficient for the absence of arbitrage. For arbitrary price-impact functions and linear price-update functions, Proposition 6 provides a condition on the shape of the price-impact function that guarantees no arbitrage. This condition says that arbitrage is impossible whenever the price-impact function is large enough relative to the price-update function.

At this point a digression to the mathematical finance literature is in order. Arbitrage is studied there predominantly under the assumption that prices are purely stochastic, i.e., trading has no impact on prices. One exception is Jarrow (1992), who investigates whether a large trader whose trades move the price can make profits from price manipulation. Jarrow gives several examples of arbitrage and states a sufficient condition that rules out arbitrage, but is unable to characterize it. A characterization is quite difficult in his framework since very general price processes are permitted.

However, our Proposition 5 demonstrates that equivalent conditions that rule out arbitrage can be found if the price process has a more specific structure. Note that unlike Jarrow, we allow for price uncertainties that are realized only after an order has been submitted. This is why we define the absence of arbitrage in expected terms here.

Proposition 7 considers arbitrage-free, nonstationary price-update and price-impact functions when they are linear and when their slopes vary stochastically over time. It implies that expected market liquidity, measured by the slopes of both the price-update and price-impact functions, must not decrease too fast. If all functions are restricted to be deterministic, then the absence of arbitrage can even be characterized by a condition defined on the functions' slopes, as Proposition 8 claims.

Finally, Propositions 9 and 10 prove the multi-asset generalizations of the results in Propositions 1 and 2. The absence of arbitrage implies the price-update function to be represented by a positive semidefinite matrix, and the reverse is true if the price-impact function is positive semidefinite, too. Thus, the absence of arbitrage requires the price update of a portfolio to be the same as the sum of the price updates caused by trading all assets in the portfolio separately. In addition, all cross-price impacts must be symmetric.

II Single-asset Time-stationary Price Dynamics

Imagine one trader of a single asset over N periods. The asset can be bought or sold via market orders at any time. In each period n, the initial price of the asset is the last price update, $\tilde{p}_n$. In the absence of uncertainty a trader has to pay a total of $[\tilde{p}_n + P(q_n)]q_n$ if he buys the quantity q_n, and the initial price for the next period will be $\tilde{p}_{n+1} = \tilde{p}_n + U(q_n)$. (Interpret a negative q_n as a sale.) The *price-impact function* P measures the immediate price reaction to the traded volume q_n, including both the permanent and the temporary price impact. The *price-update function* U, on the other hand, describes the trade's permanent impact on future prices. Hence, the temporary price impact can be extracted from P and U by simply calculating $P - U$.

From the trader's perspective, other orders are random. In each period, all orders are submitted simultaneously. In addition, news that reveals value-relevant information arrives randomly. To incorporate both types of uncertainty, the price process is augmented with stochastic terms as follows. After the most recent trade q_{n-1} at the end of period $n-1$, the public news ε_n is revealed in the beginning of period n and the price is updated to $\tilde{p}_n$, taking into account both the last trade and the latest news. Since trading takes place only at the end of the period, the trader knows $\tilde{p}_n$ and ε_n before his trade in period n, but not the order sizes of the other market participants summarized in η_n. This structure gives rise to the following price dynamics:

$$\tilde{p}_n = \tilde{p}_{n-1} + U(q_{n-1} + \eta_{n-1}) + \varepsilon_n \tag{1}$$

$$p_n = \tilde{p}_n + P(q_n + \eta_n),$$

where p_n denotes the transaction price. The η_n's represent the residual trades over time, i.e., all orders other than that of the trader; they are *iid* with zero expected value. The ε_n's describe the disclosure of news through time, and are also *iid* random variables with zero mean, independent of the η_n's. Both stochastic processes are defined on the same probability space, $(\Omega, \mathcal{F}, \varphi)$. (Since the range of the random variables can cover $\mathbf{R}$, negative prices cannot be excluded.) The zero means of η_n and ε_n imply that the prices in (1) form a martingale if zero net total trading volume is expected.

In view of (1) buying q_n costs $p_n q_n$ and the initial price for the subsequent period is given by $\tilde{p}_{n+1}$. Moreover, note that the initial quote $\tilde{p}_n$ is the origin of the price-impact function in period n.

Since the trader knows ε_n but not η_n before his trade at time n, uncertainty over the current price is thus captured only by η_n, while uncertainty over subsequent prices is determined by the randomness of $\{\varepsilon_j\}_{j=n+1}^N$ and $\{\eta_j\}_{j=n}^N$. After the trade has occurred, the trader can extract η_n directly from the price p_n only when P is strictly monotonic; otherwise he must get the information on η_n from the publicly available records of trades at the exchange house. This environment should best capture real trading activity: while it is unlikely that new information occurs at the moment of submitting a trade, other trades not known to a trader are likely to happen.

We assume that in every period competitive liquidity providers stand ready to fill all the orders with a total volume of $q_n + \eta_n$. The prices given by (1) are thus set by those liquidity providers, with the price-update and price-impact functions representing their price reaction to trade size. Such providers resemble the market makers in Kyle (1985).

A relatively tractable special case of (1) is

$$\tilde{p}_n = \alpha \tilde{p}_{n-1} + (1-\alpha) p_{n-1} + \varepsilon_n \tag{2}$$

$$p_n = \tilde{p}_n + P(q_n + \eta_n)$$

which can be obtained by setting $U = (1-\alpha)P$, $\alpha \in [0,1]$. The individual faces an initial price that is a convex combination of the previous initial price and the 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 arbitrage. When $\alpha = 0$, i.e., $U = P$, then (2) simplifies to

$$p_n = p_{n-1} + U(q_n + \eta_n) + \varepsilon_n, \tag{3}$$

implying that the price change is a function of the current trade and randomness only, i.e., it does not depend on history. The recursion in (3) asserts that the transaction price and the price update coincide and that each trade has only a permanent impact on the security price.

Time stationarity underlies equations (1)-(3). Nonetheless, the definition of no arbitrage introduced next is independent of stationarity. Moreover, we later relax the stationarity assumption.

A A Definition of the Absence of Arbitrage

Arbitrage exists if one can start and end with no holdings of an asset but make money by trading it, i.e., if there is an integer N and a sequence $\{q_n\}_{n=1}^{N}$ of trades which sums to zero and for which (expected) profits, $-E[\sum_{n=1}^{N} p_n q_n]$, are positive. In the sequel, we write X, $\varphi - a.e.$ ($\varphi-$ almost everywhere), to indicate that the event X occurs with probability one.

Definition 1 *The price process (1) is admissible for a given initial price $p_0 \geq 0$ and number of trades N if*

$$\sum_{n=1}^{N} q_n = 0 \quad \varphi - a.e. \text{ implies that } E\left[\sum_{n=1}^{N} p_n q_n\right] \geq 0 \tag{4}$$

is satisfied. Condition (4) is the no-arbitrage condition.

Definition 2 *The pair (U, P) of price-update and price-impact functions is arbitrage-free if the associated price sequence is admissible for all $p_0 \geq 0$ and all integers N.*

Four comments about Definitions 1 and 2 are warranted. First, the integer N denotes the total number of trades within a fixed calendar time; a higher N is therefore equivalent to increasing the frequency of trading. Second, the trades q_n are stochastic because they are conditioned on history which includes the sequences $\{\varepsilon_j\}_{j=1}^{n}$ and $\{\eta_j\}_{j=1}^{n-1}$. Third, condition (4) rules out arbitrage during any subinterval of trading because zero trades are allowed.

Finally, our definition of arbitrage is “statistical arbitrage” and not “sure arbitrage.” Unlike in the mathematical finance literature (for instance, see Musiela and Rutkowski (1997)), arbitrage cannot be defined here for (almost) all states because the price is never known before the trades.

To illustrate Definitions 1 and 2, we provide two examples of price-update and price-impact functions that are not arbitrage-free.

Example 1. Suppose the price-update function is $U(q) = P(q) = \lambda q$ if $q \geq 0$ and $\frac{5}{16}\lambda q$ otherwise; $\lambda > 0$. In this case, purchases move the price more than sales. Certain empirical papers report such an asymmetric price impact in that block purchases have larger price impacts than block sales (see, e.g., Gemmill (1996) or Holthausen et al. (1987)). Chan and Lakonishok (1995) report the same for institutional trades. (However, Keim and Madhavan (1996) and Scholes (1972) provide evidence of a stronger permanent price impact of sales.). Such a price-update function is not admissible: the trading strategy of buying one unit in each of the first three periods and then selling the three purchased units in the fourth period yields profits of $-E[\sum_{n=1}^{4} p_n q_n] = -\sum_{n=1}^{3}(p_0 + n\lambda) + (p_0 + 3\lambda - 3 * \frac{5}{16}\lambda)3 = \frac{3}{16}\lambda > 0$, contradicting the no-arbitrage condition. It takes at least four trades to implement an arbitrage strategy.

Example 2. Suppose the price changes with trade sign and is not sensitive to trade size. A price-update function representing this situation is $U(q) = P(q) = \lambda\ sgn(q)$, $\lambda > 0$, where $sgn(q)$ is 1 if $q > 0$, 0 if $q = 0$, and -1 otherwise. But buying two units in the first period, buying one unit each in the second and third periods, and then selling everything in the fourth period renders profits of $-E[\sum_{n=1}^{4} p_n q_n] = -2(p_0 + \lambda) - (p_0 + 2\lambda) - (p_0 + 3\lambda) + 4(p_0 + 2\lambda) = \lambda > 0$, violating the no-arbitrage condition. Also, no arbitrage opportunities are available if no more than three trades are allowed.

B Necessary Conditions for the Absence of Arbitrage

For notational convenience, we introduce the vector of history

$$H_n \triangleq [\varepsilon_1, \ldots, \varepsilon_n, \eta_1, \ldots, \eta_{n-1}, q_1, \ldots, q_{n-1}]^T.$$

(In period zero only p_0 is known.) $L(\mathbf{R})$ is the Lebesque measure on $\mathbf{R}$, and $E_n[.] \triangleq E[. \mid H_n]$ is the conditional expectation given information H_n at time n (with some abuse of notation, we use H_n also to denote the sigma-algebra it generates). Let us also define the *expected price-update function* $\hat{U}(q) \triangleq E[U(q+\eta)]$, $q \in \mathbf{R}$,where η has the same distribution as the residual trades. The *expected price-impact function* $\hat{P}$ is defined in the same fashion.

Proposition 1 *If (U, P) is arbitrage-free, then U has the representation*

$$U(y) = \lambda y + S(y) \tag{5}$$

on $\mathbf{R}$, $L(\mathbf{R}) - a.e.$, $\lambda \geq 0$, where the $L(\mathbf{R})$-measurable function $S : \mathbf{R} \to \mathbf{R}$ satisfies

$$E_n[S(\tilde{q}_n + \eta_n)] = 0 \tag{6}$$

for all H_n-measurable random variables $\tilde{q}_n : \Omega \to \mathbf{R}$, $1 \leq n \leq N$.

Proposition 1 says that an arbitrage-free price-update function can always be written as the sum of a linear function (with nonnegative slope) and a *supplementary function* S, 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 H_n-measurable random variables. This insight implies that a risk-neutral trader ignores S when forming his trading strategy, since he always computes conditional expected prices to assess the profitability of his trades.

The intuition underlying Proposition 1 transpires from the outline of the proof offered below; the formal proof is in Appendix A. For convenience, we divide the outline of the proof into four steps.

Step 1: The expected price-update function is 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 offer arbitrage opportunities. 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. If the number of trades, $2m$, is large enough, the average sale price exceeds the average purchase price, i.e., an arbitrage opportunity exists, independent of what the price-update function looks like on $\mathbf{R}\backslash\{-q, q\}$. 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 arbitrage profits. It is straightforward to check that the absence of arbitrage also implies $\hat{U}(0) = 0$, and $\hat{U}(q) \geq 0$ for $q > 0$.

Step 2: $\hat{U}$ is continuous, with a possible exception at zero. 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, $\lim_{q' \to q+} \hat{U}(q') > \hat{U}(q)$. The strategy of buying $q' > q$ shares in each of the first m periods and selling q shares in each of the following m periods, where q' is chosen arbitrarily close to q, yields arbitrage profits for sufficiently large m. 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 A demonstrates that for any possible type of jump there exists an arbitrage trading strategy.

Step 3: $\hat{U}$ is linear, since the absence of arbitrage is incompatible with $\hat{U}(q) > \hat{U}(1)q$ or $\hat{U}(q) < \hat{U}(1)q$, 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) yields arbitrage profits for large m, since 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.

Step 4: Proving condition (6). Define $S(q) \triangleq U(q) - \hat{U}(q)$. Then, Step 3 implies $E[S(q+\eta_n)] = 0$ for all q, which in turn has (6) as a consequence.

Next, we propose two distributions of the residual trades, each of which causes the supplementary function in (6) to be zero. One possibility is that the residual trades are zero, and the second is that they are normally distributed.

Proposition 2 *Suppose that either*

i. $\varphi[\eta_n = 0] = 1$, *for* $1 \leq n \leq N$ *(zero residual trades) or*

ii. the residual trades are normally distributed.

Then the supplementary function S *equals zero,* $L(\mathbf{R}) -$ *a.e., in Proposition 1.*

Notice that case i describes the situation where only one trader affects the price in each period ($\eta_n = 0$ means that the total net trading volume of all the other traders is zero). Contrary to Proposition 1, the absence of arbitrage now requires the price-update function to be exactly linear and not only linear in expected terms. It can also be shown that Proposition 2 is true when the residual trades are a certain transform of a zero-mean normal random variable (for details see Remark 1 in Appendix A).

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. Hence, Propositions 1 and 2 are also valid when commissions have to be paid per share.

Note that the supplementary function need not be zero for other distributions, as three examples in Appendix B show. For empirical studies, then, nonlinear price-update functions can be used, but their conditional expectation must be linear in trade size.

Proposition 2 provides a theoretical justification for looking at linear additive price processes of the type $p_n = p_{n-1} + \lambda q_n + \varepsilon_n$ (i.e., setting $U = P$), which has been popular in the literature, with tractability being the main motivation (see Dutta and Madhavan (1995), Hausman et al. (1992), or Bertsimas and Lo (1998)). Note that this specification is also sufficient to rule out arbitrage. This is one of the main results in the next section.

To assume $E[\eta_n] \neq 0$ or $E[\varepsilon_n] \neq 0$ would be harmful in this context. For example, if $E[\eta_n] > 0$, then buying one share in the first period and selling this share many periods later would be profitable, because the price moves up between the purchase and the sale due to $E[\eta_n] > 0$. To exclude this kind of arbitrage, the price process (1) must not include trend components. This justifies our zero-mean assumptions, and is stated as the next proposition.

Proposition 3 *If either $E[\eta_n] \neq 0$ or $E[\varepsilon_n] \neq 0$, then the price process (1) will allow arbitrage.*

We can also derive necessary conditions for the price-impact function, although they have a less compact form than the conditions in Propositions 1 and 2. Thus, we are content here with giving only two of them.

Proposition 4 *If (U, P) is arbitrage-free, then the following two conditions must hold:*

i. $\hat{P}(q) - \hat{P}(-q) \gtreqless \hat{U}(q) \quad$ *for* $q \gtreqless 0$*, and*

ii. P *cannot be constant when* $U \neq 0$*.*

If we interpret the left-hand side of condition i in Proposition 4 as the spread of the price-impact function, then condition i says that the spread *at any* trade size has to exceed the price update resulting from that trade. Were this not true, the trading strategy cited in Step 1 of the proof following Proposition 1 (buying q shares in each of the first m periods and then selling q shares in each of the next m periods) would be profitable. The same trading strategy also implies the second condition in Proposition 4. P always has to be a function of the trade size, unless $U = 0$.

Proposition 2 rules out various tempting functional forms for the price-impact function. For instance, Breen et al. (1998) estimate a price-impact function where the inter-transaction return is linear in the traded quantity, i.e., $(p_n - p_{n-1})/p_{n-1} = \alpha + \lambda q_n + \varepsilon_n$. This regression equation implies $p_n = (1 + \alpha + \lambda q_n + \varepsilon_n)p_{n-1}$, giving rise to arbitrage. To see this, take $N = 2$ and $q_2 = -q_1$, and compute the costs, $E[\sum_{n=1}^{2} p_n q_n] = p_0 q_1(-\alpha + \lambda q_1)(1 + \alpha + \lambda q_1)$. This expression becomes negative if q_1 is negative and sufficiently large in absolute value. In fact, arbitrage profits are infinite when letting $q_1 \to -\infty$.

A second price process is $p_n = \sum_{i=1}^{m} \alpha_i p_{n-i} + \lambda q_n + \varepsilon_n$, $m > 0$ (see Hasbrouck (1991)). Here, the price law also offers room for arbitrage. Again take $N = 2$ and $q_2 = -q_1$, and compute expected costs $E[\sum_{n=1}^{2} p_n q_n] = q_1[p_0(\alpha_1 - \alpha_1^2 - \alpha_2) + (2 - \alpha_1)\lambda q_1)]$. The latter expression generally can be driven negatively by a proper choice of q_1, unless $\alpha_1 - \alpha_1^2 - \alpha_2 = 0$ and $\alpha_1 < 2$.

Finally, for price processes where the moving average is taken over the trading quantities, rather than over the prices, such as

$p_n = p_{n-1} + \lambda \sum_{i=0}^{m} \alpha_i q_{n-i} + \varepsilon_n$ (Dutta and Madhavan (1995) employ a special case of this price process), arbitrage opportunities can be found, too.

C A Sufficient Condition for the Absence of Arbitrage

This section derives a sufficient condition for the absence of arbitrage. With the aid of this condition we are able to establish a characterization of the absence of arbitrage for the case $U = (1 - \alpha)P$, where the price-update and price-impact functions are multiples of each other. No arbitrage is equivalent to the linearity of both the price-update and price-impact functions. If U and P are independent, our sufficient condition will serve us to find some interesting examples of price-impact functions that give rise to arbitrage-free (U, P)'s.

The main observation leading to this sufficient condition is the fact that the no-arbitrage condition in Definition (1) is tantamount to the cost-minimization problem

$$\min_{\{q_n,\ H_n-measurable\}_{n=1}^N} E[\sum_{n=1}^N p_n q_n] \quad \text{subject to} \quad \sum_{n=1}^N q_n = 0 \tag{7}$$

having nonnegative costs as its solution. The optimization problem given in (7) can be approached by standard dynamic programming arguments if the price-update and price-impact functions are both assumed to be linear and $P > \frac{1}{2}U$ for $q > 0$. In this case, Appendix A shows that the minimal costs are always zero. Hence, we can state the following auxiliary result.

Lemma 1 *If U and P are both linear and $P \geq \frac{1}{2}U$ for $q \geq 0$, then (U, P) is arbitrage-free.*

Suppose now that the price-update and the price-impact functions satisfy $U = (1 - \alpha)P$, and that both have the representation given in Proposition (1). In this case, the trader ignores the supplementary function when he computes his optimal trading strategy and treats both U and P as linear, as was pointed out in the comments below Proposition 1. Then, Lemma 1 and Propositions 1 and 2 imply the following.

Proposition 5 *Suppose that $U = (1 - \alpha)P$. Then the pair (U, P) is arbitrage-free if and only if U and P have the representation given in Proposition (1). If the residual trades assume one of the distributions stated in Proposition 2, then we obtain the stronger result that the absence of arbitrage is characterized by the linearity of U and P.*

Proposition 5 says that the set of arbitrage-free price-update and price-impact functions coincides with the set of linear functions if the price evolves according to (2) or (3).

If P is not a multiple of U, then nonlinear price-impact functions can also lead to arbitrage-free (U, P)'s. With the help of the proposition below, which follows directly from Lemma 1, we can construct examples illustrating this point.

Proposition 6 *Let U be linear. If $P \geq \frac{1}{2}U$ for $q \geq 0$ and $P \leq \frac{1}{2}U$ for $q < 0$, then (U, P) is arbitrage-free.*

Consider the price-impact function $P(q) = \frac{1}{2}[A\ sgn(q) + U(q)]$, where $A > 0$ is a constant. This function exhibits a discontinuity at zero (with jump size A) and intersects the price-update function twice. From Proposition 6, the pair (U, P) is arbitrage-free if U is linear. Situations like this, where price revision intersects the actual price schedule, are described in Glosten (1994), where an equilibrium in an open limit order book is constructed.

The empirical studies in Hasbrouck (1991) give evidence that security prices are concave for purchases and convex for sales. This relation can be modeled here by taking a symmetric price-impact function that is concave in some positive range without violating the conditions stated in Proposition 6, which imply that P has to grow (decline) at least linearly eventually.

Evidently, many more arbitrage-free (U, P)'s with very complicated price-impact functions can be found here. This suggests that in the case $U \neq (1 - \alpha)P$, sufficient conditions for no arbitrage that are also necessary may be very hard to derive. We refrain from further examination.

III Nonstationary Price Dynamics

Until now, the price-update and price-impact functions have been time-stationary, 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-update and price-impact functions (when they exist), is therefore constant through time. In what follows we relax this assumption and allow liquidity to vary across time.

A Analysis

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

$$\tilde{p}_n = \tilde{p}_{n-1} + \lambda_{n-1}(q_{n-1} + \eta_{n-1}) + \varepsilon_n \tag{8}$$

$$p_n = \tilde{p}_n + \mu_n(q_n + \eta_n),$$

where the sequences of random variables $\{\lambda_n : \Omega \to \mathbf{R}\}_{n=1}^{\infty}$ and $\{\mu_n : \Omega \to \mathbf{R}\}_{n=1}^{\infty}$ are assumed to be independent across time as well as independent of each other and all other uncertainty in this model; in addition, each λ_n and μ_n is H_n-measurable, $\mu_1 \geq 0$ $\varphi - a.e.$, and for convenience we set $\hat{\lambda}_1 = E[\lambda_1] = E[\mu_1] = \hat{\mu}_1 \geq 0$.

The nonstationary framework requires some adaptation of our terminology. In the spirit of the above analysis, we call a pair $(\{\lambda_n\}_{n=1}^{\infty}, \{\mu_n\}_{n=1}^{\infty})$ of *price-update sequence* and *price-impact sequence* arbitrage-free if it does not offer any arbitrage opportunities, in other words, if the price process (8) is admissible with respect to all $p_0 \geq 0$ and $N \in \mathbf{N}$. (The definition also includes finite sequences where $\lambda_n = \infty$ for all $n > N$. We distinguish them from infinite sequences by writing $\{\lambda_n\}_{n=1}^{N}$ instead of $\{\lambda_n\}_{n=1}^{\infty}$.)

We proceed as follows. First, we establish a necessary and a sufficient condition for no arbitrage. These conditions are then used to characterize the absence of arbitrage for the case of deterministic price-update and price-impact functions. At the end of this section we relate our results to the extant literature.

Before we work through the above agenda, let us provide an example of sequences that permit arbitrage. Consider $p_0 = 10$, $\{\lambda_n\}_{n=1}^{3} = \{\mu_n\}_{n=1}^{3}$ with $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 0$, $\varphi - a.e.$ Buying one unit of the asset in each of periods 1 and 2 (paying $1 \times 11 + 1 \times 12 = 23$) and then selling the holdings (two units) in the third period (collecting $2 \times 12 = 24$) results in expected profits of one. Therefore, arbitrage is possible. The arbitrage profits are actually unbounded in this case: buy q units in the first and second period and sell the holdings in the third period. This trading strategy yields profits of $-(10+q)q-(10+2q)q+2q(10+2q) = q^2$ which grow without bound when $q \to \infty$.

To get an idea what kind of restrictions the absence of arbitrage imposes on the pair $(\{\lambda_n\}_{n=1}^{\infty}, \{\mu_n\}_{n=1}^{\infty})$, let us consider the simple case $N = 3$ where only three trades are permitted. For the sequences to be arbitrage-free, all trade sequences $\{q_n\}_{n=1}^{3}$that sum to zero must result in nonnegative (expected) costs.

Computing $E[\sum_{n=1}^{3} p_n q_n]$ under the constraint $\sum_{n=1}^{3} q_n = 0$ leads to the quadratic form $E[\sum_{n=1}^{3} p_n q_n] = \hat{\mu}_2 q_2^2 + \hat{\lambda}_2 q_2 q_3 + \hat{\mu}_3 q_3^2$ if only deterministic trades are chosen, or in matrix notation,

$$E[\sum_{n=1}^{3} p_n q_n] = \frac{1}{2}\,[q_2, q_3] \begin{bmatrix} 2\hat{\mu}_2 & \hat{\lambda}_2 \\ \hat{\lambda}_2 & 2\hat{\mu}_3 \end{bmatrix} \begin{bmatrix} q_2 \\ q_3 \end{bmatrix}.$$

(The term $p_0 + \hat{\lambda}_1 q_1$ drops out because it is a price component in each period and is thus canceled by $\sum_{n=1}^{3} q_n = 0$.) The trades q_2 and q_3 in the above expression can take any value since q_1 can always be chosen such that $q_1 = -q_2 - q_3$. We therefore draw the simple conclusion that the absence of arbitrage requires the positive semidefiniteness of the matrix $\Lambda_{3,-1} \triangleq \begin{bmatrix} 2\hat{\mu}_2 & \hat{\lambda}_2 \\ \hat{\lambda}_2 & 2\hat{\mu}_3 \end{bmatrix}$. (The first subindex of Λ refers to the total number of periods, while the second indicates that the matrix does not depend on first-period variables.)

The removal of q_1 is arbitrary. If we remove q_2 or q_3 we obtain expected costs of

$$E[\sum_{n=1}^{3} p_n q_n] = \frac{1}{2}\,[q_1, q_3]\,\Lambda_{3,-2} \begin{bmatrix} q_1 \\ q_3 \end{bmatrix} = \frac{1}{2}\,[q_1, q_2]\,\Lambda_{3,-3} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$$

where $\Lambda_{3,-2} \triangleq \begin{bmatrix} 2\hat{\mu}_2 & \hat{\lambda}_2 \\ \hat{\lambda}_2 & 2\hat{\mu}_3 \end{bmatrix}$ and $\Lambda_{3,-3} \triangleq \begin{bmatrix} 2\hat{\mu}_3 & 2\hat{\mu}_3 - \hat{\lambda}_2 \\ 2\hat{\mu}_3 - \hat{\lambda}_2 & 2\hat{\mu}_3 \end{bmatrix}$. The first quadratic form is obtained by substituting $q_2 = -q_1 - q_3$ in $E[\sum_{n=1}^{3} p_n q_n] = \hat{\mu}_2 q_2^2 + \hat{\lambda}_2 q_2 q_3 + \hat{\mu}_3 q_3^2$, and the second is the result of plugging in $q_3 = -q_1 - q_2$. Consequently, the absence of arbitrage also implies that the matrix $\Lambda_{3,-3}$ is positive semidefinite. It is straightforward to show that positive semidefiniteness of one of these matrices implies the other two to be positive semidefinite.

For $\Lambda_{3,-1}$ to be positive semidefinite, $\hat{\mu}_2$ and $\hat{\mu}_3$ must be nonnegative and $\hat{\mu}_2\hat{\mu}_3 \geq \hat{\lambda}_2^2/4$. These conditions, together with $\hat{\mu}_1 \geq 0$, say that the absence of arbitrage rules out negative (expected) price-impact sequences in all periods and that $\hat{\mu}_2$ and $\hat{\mu}_3$ have to be sufficiently large relative to $\hat{\lambda}_2^2$. Notice that $\hat{\lambda}_2$ can be negative, conflicting with the interpretation that purchases convey positive news about the asset's value and push the price up. We will discuss this issue below when we have at our disposal a sufficient condition for no arbitrage.

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

Proposition 7 *If the pair $(\{\lambda_n\}_{n=1}^{\infty}, \{\mu_n\}_{n=1}^{\infty})$ is arbitrage-free, then the expected value $\hat{\Lambda}_n$ of the matrix Λ_n defined by*

$$\Lambda_n \triangleq \begin{bmatrix} 2\mu_2 & \lambda_2 & \lambda_2 & \dots & \lambda_2 \\ \lambda_2 & 2\mu_3 & \lambda_3 & \dots & \lambda_3 \\ \lambda_2 & \lambda_3 & 2\mu_4 & \dots & \lambda_4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_2 & \lambda_3 & \lambda_4 & \dots & 2\mu_n \end{bmatrix} \tag{9}$$

must be positive semidefinite for all $n \geq 2$.

Unfortunately, the reverse is not true. To understand this, note that expected costs can be written as

$$E[\sum_{n=1}^{N} p_n q_n] = E\left[q_{N,-1}^T[(\mu_1 - \lambda_1)1_{N-1\times N-1} + \frac{1}{2}\Lambda_N]q_{N,-1}\right] \tag{10}$$

when $\sum_{n=1}^{N} q_n = 0$, where $q_{N,-1} \triangleq (q_2, q_3, \ldots, q_N)$ and $1_{N-1\times N-1}$ is the $N-1$ dimensional square matrix containing only ones. We are going to show that profits can be made here even when the condition in Proposition 7 is met. To this end, consider the following example: $N = 2$, Ω can be partitioned into Ω_1, Ω_2, and Ω_3, and λ_1 and μ_1 are discrete with $\lambda_1(\Omega_1) = \{4\}$, $\mu_1(\Omega_1) = \mu_1(\Omega_2) = \lambda_1(\Omega_2) = \{2\}$, $\mu_1(\Omega_3) = \{1\}$, and $\lambda_1(\Omega_3) = \{0\}$. In addition, the probabilities are $\varphi(\Omega_1) = \varphi(\Omega_2) = 1/4$, $\varphi(\Omega_3) = 1/2$, and $\hat{\mu}_2 = 2/3$. In this case, the trading strategy $q_1 = -q_2 = \mu_1 - \lambda_1$ would render, due to (10), an expected profit of $-E[\sum_{n=1}^{2} p_n q_n] = 1$, without violating the condition in Proposition 7.

A sufficient condition for the absence of arbitrage thus has to be stronger than the condition in Proposition 7. One that derives immediately from (10) is $\mu_1 \geq \lambda_1 \geq 0$, $\varphi - a.e.$, and Λ_n being positive semidefinite, $\varphi - a.e.$, for all $n \geq 2$. Since $\mu_1 \geq 0$ forbids arbitrage with one trade, positive semidefinite Λ_n's together with $\mu_1 - \lambda_1 \geq 0$ guarantee the absence of arbitrage.

However, when the $\{\lambda_n\}_{n=1}^{\infty}$ and $\{\mu_n\}_{n=1}^{\infty}$ are both restricted to be deterministic, the necessary condition stated in Proposition 7 is also sufficient for no arbitrage. The crucial fact giving rise to this result is that the optimal trading strategy of (7) is deterministic (see Appendix A). Hence, from (10) and $\mu_1 = \lambda_1$ we have the following.

Proposition 8 *A deterministic pair $(\{\lambda_n\}_{n=1}^{\infty}, \{\mu_n\}_{n=1}^{\infty})$ is arbitrage-free if and only if Λ_n is positive semidefinite for all $n \geq 2$. The sum of two arbitrage-free (deterministic) pairs of price-update and price-impact sequences is again arbitrage-free, as well as any positive multiple of an arbitrage-free pair. If $\lambda_n = \mu_n = \infty$ for $n > N$, then a positive semidefinite Λ_N is already sufficient and necessary for the absence of arbitrage.*

For the remainder of this section we confine our attention to the deterministic case. Proposition 8 contains two important facts. The first is that for any given price-update sequence there exists a price-impact sequence that preserves the absence of arbitrage. This is a consequence of the price updates and immediate price impacts being different: the μ_n's only have to be chosen high enough. The second fact is that it gives a specific computational criterion for testing for no arbitrage.

While all μ_n's must be nonnegative, the signs of the λ_n's are ambiguous. Negative λ_n's are in discord with the interpretation that purchases signal good news about the asset's value. But in a pure arbitrage framework, negativity makes perfect sense. The main mechanism that makes manipulation successful is the positive relation between price update and trading volume. If the λ_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.

This problem does not appear, of course, if $\{\lambda_n\}_{n=1}^{\infty} = \{\mu_n\}_{n=1}^{\infty}$ (no temporary price impact) is assumed. Then, all liquidity parameters are nonnegative. In this case, arbitrage-free sequences can be found quite easily. Since $det\ \Lambda_n > 0$ for all $n \geq 2$ implies the absence of arbitrage and since $det\ \Lambda_n$ can be computed recursively by

$$det\ \Lambda_n = 2\lambda_n\ det\ \Lambda_{n-1} - \lambda_{n-1}^2\ det\ \Lambda_{n-2} \quad \text{for } n \geq 3, \tag{11}$$

with initial conditions $det\Lambda_2 = 2\lambda_2$ and $det\Lambda_3 = \lambda_2(4\lambda_3 - \lambda_2)$, the complexity of constructing an arbitrage-free price-update sequence is only of linear order.

By virtue of (11), sequences like $\{1+\frac{1}{2^n}\}_{n=1}^{\infty}$ or $\{1+e^{-n}\}_{n=1}^{\infty}$ satisfy $det\ \Lambda_n > 0$ for all $n \geq 2$ and are therefore arbitrage-free. Another interesting example is $\{1+\frac{(-1)^n}{2^n}\}_{n=1}^{\infty}$, which is nonmonotonic; it is arbitrage-free, too. Even alternating sequences are possible candidates for arbitrage-free price-impact sequences. If they converge with "sufficient speed" they offer no arbitrage opportunities.

When a price-update sequence does not satisfy $det\ \Lambda_n > 0$ for all $n \geq 2$, one has no alternative but Proposition 8 to verify whether it offers profits from arbitrage. Two examples of sequences where (11) cannot be employed are $\{\frac{1}{n}\}_{n=1}^{\infty}$ and $\{1 + \frac{1}{n}\}_{n=1}^{\infty}$. The criterion in Proposition 8 tells us that neither is arbitrage-free. These decreasing price-update sequences converge too slowly. A price-update sequence that fails to meet $det\ \Lambda_n > 0$ for all $n \geq 2$ but is arbitrage-free is $\{\lambda_n\}_{n=1}^{3} = \{1, 1, 1/4\}$ (recall our introductory example where $\{\lambda_n\}_{n=1}^{3} = \{1, 1, 0\}$ leads to arbitrage opportunities).

For nondecreasing and recurrent price-update sequences (a recurrent sequence is the infinite repetition of the same finite series of real numbers), Proposition 8 and (11) imply that *nondecreasing price-update sequences are necessarily arbitrage-free, and recurrent price-update sequences must be constant if they are arbitrage-free.*

Summarizing, the results of this section tell us several things. First, all price-impact sequences must be nonnegative, while price-update sequences need not be so, when permanent and immediate price impacts are different. Second, for each price-update sequence there exists a price-impact sequence that makes the pair of both arbitrage-free. For this to happen, the price impacts have to be sufficiently large relative to the price updates. Third, if price impacts and updates are the same, arbitrage becomes more likely if liquidity increases over time. Too high a rise in liquidity enables the trader to lock in arbitrage profits: he begins pushing up the price 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.

B Discussion

That liquidity is not allowed to rise too rapidly is also an implicit result in Kyle (1985). His model describes a game between a competitive market-maker, who sets the price in each period, and an individual risk-neutral insider trader, who has information on the liquidation value of the single asset that is traded. The framework is as follows. In each period, the market-maker observes only the aggregate trading volume, which is the sum of the insider's trading quantity and residual trades. He cannot observe the insider's amounts. Knowing the history of trades and the fact that there is one informed trader who maximizes his profits, he sets the price equal to the conditional expected value. The insider, on the other hand, taking into account how the market-maker computes the price, submits in each round the quantity that maximizes his profits. As Kyle shows, this game has a unique linear equilibrium where the price evolves according to $p_n = p_{n-1} + \lambda_n q_n + \varepsilon_n$, the liquidity parameters λ_n being endogenously (but deterministically) determined. (The residual trades in Kyle are represented here by a stochastic term $\varepsilon_n \sim NID(0, \sigma^2)$.) For small and large N, Kyle's slopes are almost constant and hence arbitrage-free. Consequently, in Kyle's equilibrium arbitrage is impossible. If arbitrage were possible, the insider's optimum would be undefined.

If multiple (equally informed) insiders are introduced into the Kyle model, the picture changes dramatically. In this case, arbitrage opportunities arise when the number of insiders is sufficiently large. As Holden and Subrahmanyam (1992) prove, insiders trade very aggressively in early periods resulting in very low liquidity in the beginning. Only after a few periods almost all insider information is impounded in the price. At the point where this happens, liquidity alters abruptly to higher levels. In the presence of sufficiently many insiders, this change in liquidity occurs too fast in the sense that it violates the conditions that are put forward in Proposition 8. (Actually, some of the numerical examples presented in Holden and Subrahmanyam imply arbitrage opportunities. We thank Craig Holden who made us aware of this fact.)

This finding raises some doubt about the equilibrium concept proposed in Holden and Subrahmanyam for the Kyle framework with many informed traders. If the market makers and insiders took into account manipulative trades by uninformed market watchers, their optimal trading behavior would look different than described in Holden and Subrahmanyam. We conjecture that, in equilibrium, the possibility of manipulative trades causes the insiders to trade less aggressively in the initial periods.

After the analysis of this section one may be tempted to study more general price processes, such as $p_n = p_{n-1} + U_n(q_n + \eta_n) + \varepsilon_n$, and try to characterize the arbitrage-free price-update functions $\{U_n(.)\}_{n=1}^{\infty}$. This is very difficult unless one puts restrictions on the functions U_n. Furthermore, the arbitrage-free U_n's may no longer have a nice shape. Consider, for example, $U_n(q) = (n-1)q$ if $q = 1$ and nq otherwise. Using (11) it can be easily shown that these price-update functions are arbitrage-free. However, these functions are not smooth: they are neither continuous, symmetric, nor increasing.

An even more drastic example that illustrates how arbitrage-free price-update functions can have very arbitrary shapes is the following: for a given monotonically strictly increasing nonnegative sequence $\{\lambda_n\}_{n=1}^{\infty}$, any sequence of functions $\{U_n(.)\}_{n=1}^{\infty}$ that satisfies $U_1 \geq 0$ and $\lambda_{n-1}q \leq U_n(q) \leq \lambda_n q$ if $q \geq 0$ and $\lambda_n q \leq U_n(q) \leq \lambda_{n-1}q$ if $q < 0$, for all $n \geq 2$, is arbitrage-free. This follows also directly from (11). Since we want to work with smooth functions, we abstain from investigating the general case.

IV Multi-asset Price Dynamics and Arbitrage

So far we have discussed arbitrage 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 $M \geq 1$ assets can be traded. In particular, consider the price process (1) and replace all scalar variables there with M-dimensional vectors. The price-update and price-impact functions are then functions of the form $U : \mathbf{R}^M \to \mathbf{R}^M$ and $P : \mathbf{R}^M \to \mathbf{R}^M$, respectively. Moreover, the absence of arbitrage is defined as in Definitions 1 and 2, though the no-arbitrage condition reads

$$\sum_{n=1}^{N} q_n = 0 \ \ \varphi - a.e. \text{ implies that } E\left[\sum_{n=1}^{N} p_n^T q_n\right] \geq 0. \tag{12}$$

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 impacts.* If the traded quantity of asset i has an impact not only on the price of asset i but also on the prices of all other assets, then many more arbitrage opportunities may exist. The task of this section is to find necessary and sufficient conditions for the shape of the price-update and price-impact functions that rule out arbitrage.

Fortunately, our results from the single-asset case extend to the multi-asset case in a natural way. The absence of arbitrage requires the price-update function to be represented by a positive semidefinite matrix, by which we mean that $U(q) = \Lambda q$ on $\mathbf{R}^M$, Λ being positive semidefinite. We state and prove the multi-asset analogues to Propositions 1 and 2 and then establish a characterization of no arbitrage as in Proposition 5 for the case $U = P$.

Let us begin with the generalizations of Propositions 1 and 2.

Proposition 9 *If (U, P) is arbitrage-free, then $U : \mathbf{R}^M \to \mathbf{R}^M$ has the representation*

$$U(y) = \Lambda y + S(y) \tag{13}$$

on $\mathbf{R}^M$, $L(\mathbf{R}^M) - a.e.$, Λ positive semidefinite, where $S : \mathbf{R}^M \to \mathbf{R}^M$ is a $L(\mathbf{R}^M)$-measurable function that satisfies

$$E_n[S(\tilde{q}_n + \eta_n)] = 0 \tag{14}$$

for all H_n-measurable random variables $\tilde{q}_n : \Omega \to \mathbf{R}^M$, $1 \leq n \leq N$. If the residual trades are multivariate normal, then $S = 0$, $L(\mathbf{R}^M) - a.e.$

To provide the intuition behind this result we present an outline of the proof in what follows. For this purpose we define $\hat{U}_{ij}(q) \in \mathbf{R}$ to be the expected price update of asset i when $q \in \mathbf{R}$ shares of asset j and none of the other assets are traded.

Step 1: $\hat{U}_{ij}$ is linear. As in the single-asset case, we prove this after we have shown that $\hat{U}_{ij}$ is symmetric on $\mathbf{R}$ and continuous on $\mathbf{R}\backslash\{0\}$. Note that we have established already the linearity of $\hat{U}_{ii}$ in Proposition 1. Thus we only need to consider the case $i \neq j$ here.

Suppose $\hat{U}_{ij}(q) > -\hat{U}_{ij}(-q)$ for a $q > 0$, contradicting symmetry. Then the trading 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 following m periods, and selling q shares of asset i in each of the next m periods would render profits from arbitrage if m is big enough. Similarly, $\hat{U}_{ij}(q) < -\hat{U}_{ij}(-q)$ for a $q > 0$ and $\hat{U}_{ij}(0) \neq 0$ can be rejected.

We skip the arguments for the continuity of $\hat{U}_{ij}$ (see Appendix A) and go directly to the linearity of $\hat{U}_{ij}$. Again, by way of contradiction, assume that there exists a $q \in \mathbf{Q}_+$ such that $\hat{U}_{ij}(q) > \hat{U}_{ij}(1)q$. Then buying q shares of asset i in each the first m periods, buying q shares of asset j in each of 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 period is obviously profitable for large m. Analogously, $\hat{U}_{ij}(q) < \hat{U}_{ij}(1)q$ can be excluded.

Step 2: $\hat{U}_{ij} = \hat{U}_{ji}$, i.e., cross-price updates are symmetrical. Suppose, on the contrary, $\hat{U}_{ij}(q) > \hat{U}_{ji}(q)$ for a $q > 0$. This says that trading asset j has a stronger impact on the price of asset i than the other way round. Then, the trading 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 violates the no-arbitrage condition if m is large enough. The mechanism of this strategy is clear: the profit that arises from putting the purchase of asset j shares between the purchase and the sale of asset i shares outweighs the losses that derive from selling asset i shares between the purchase and the sale of asset j shares. The reverse inequality, $\hat{U}_{ij}(q) < \hat{U}_{ji}(q)$ for a $q > 0$, cannot be true either, as similar arguments show.

Step 3: $\hat{U}_i$ is additive separable, that is,

$$\hat{U}_i(q_1, q_2, \ldots, q_M) = \sum_{j=1}^{M} \hat{U}_{ij}(q_j).$$

The trading strategies verifying the last equality are a bit more involved and are relegated to Appendix A. Hence there exists a symmetric matrix Λ such that $\hat{U}(q) = \Lambda q$, for all $q \in \mathbf{R}$. That Λ has to be positive semidefinite follows from $E[\sum_{n=1}^{2} p_n^T q_n] = q^T \Lambda q$ if $q_1 = -q_2 = q$ is chosen, which has to be nonnegative for all $q \in \mathbf{R}^M$ to rule out arbitrage. The remaining claims in Proposition 9 can be shown as in the single-asset case.

If the price-update and price-impact functions are both represented by a positive semidefinite matrix, say $U(q) = \Lambda q$ and $P(q) = \Gamma q$ on R^M, then we can use the same approach as in the single-asset case to find a sufficient condition for the absence of arbitrage. Indeed, the solution to the multidimensional version of the constrained cost-minimization problem (7) yields that (U, P) *is arbitrage-free if* $\Gamma - \frac{1}{2}\Lambda$ *is positive semidefinite.* (For a detailed analysis of multidimensional optimal trading problems of this kind see Huberman and Stanzl (2000).)

Combining the last result with Proposition 9 yields the following for the case $U = P$.

Proposition 10 *The pair* (U, U) *is arbitrage-free if and only if* U *has the representation given in Proposition 9. If the residual trades are multivariate normal, then the absence of arbitrage is characterized by* U *being represented by a positive semidefinite matrix.*

Proposition 10 has one important consequence, which is that a multi-asset environment with nonzero cross-price effects can always be reduced to one that exhibits *no* cross-price impacts. To understand this, note that any positive semidefinite matrix Λ can be written as the product $C^T \Psi C$ of a diagonal matrix Ψ, which diagonal is formed by the nonnegative eigenvalues of Λ, and a matrix C constructed by the eigenvectors of Λ. If we interpret the entries of C as portfolio weights of the underlying assets, then C is a collection of M portfolios. If we replace the original assets with these portfolios, the relevant price-update function becomes $U_C(q) = \Psi q$ which incorporates no cross-price impacts.

In his paper, Black (1995) informally argues that the sum of the price updates 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 arbitrage requires more structure on the shape of the price-update function than Black claims.

V Concluding Remarks

This paper examines the conditions imposed by the absence of arbitrage on the functional shape of the temporary and permanent price effect of a trade. If the price-update and price-impact functions are stationary and multiples of each other, then the absence of arbitrage is equivalent to the linearity of both functions. On the other hand, if the price-update and price-impact 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 violating no-arbitrage requirements.

The theoretical micro-structure literature usually assumes that the change in prices is time-independent and reacts linearly to trading volume. This paper demonstrates that the assumption of stationarity of the price change makes the assumption of linear price-update functions redundant, as the latter is implied by the former.

Linearity as a necessary condition for the absence of arbitrage calls for a careful examination of empirical estimations of price-update functions. To the extent that they detect deviations from linearity, one can suspect some misspecification (perhaps a nonstationary environment) or wonder if indeed some arbitrage possibilities had gone unexhausted.

Our results also hold in a strategic environment where arbitrageurs are uninformed. Whenever the price-update function deviates from linearity, an oligopolist could increase his profits from trading by using one of the arbitrage strategies of the individual trader considered in this text. This clearly contradicts equilibrium. Hence the price impact must be linear if the price change is only a function of the current aggregate trading volume and stochastic terms. But linearity is also sufficient for the absence of arbitrage if temporary price impacts are absent, for in equilibrium the arbitrageur must have positive profits from trading, otherwise he stays out of the market. If those arbitrageurs form a monopoly, they will make profits, contradicting the fact that an individual cannot benefit from arbitrage.

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 whether market and limit orders can coexist in an equilibrium exchange.

Appendix A. Proofs

Before we prove Propositions 1 and 2, we derive two very useful results. To this end let us introduce the following convenient notation: Two real functions f and g, both defined on the integers, are said to be *asymptotically equivalent*, in sign $f \simeq g$, if $lim_{n\to\infty}\frac{f(n)}{g(n)}$ is a positive constant.

Lemma 2 *The absence of arbitrage requires the expected price-update function $\hat{U}$ to satisfy the following conditions:*

i. $\hat{U}$ is symmetric on $\mathbf{R}$, i.e., $\hat{U}(q) = -\hat{U}(-q)$ for $q \in \mathbf{R}$; and

ii. $\hat{U}$ is continuous on $\mathbf{R}\backslash\{0\}$.

Proof. To verify i we start by proving that $\hat{U}(q) \leq -\hat{U}(-q)$ holds for $q > 0$. 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: buy in each of the first m periods the volume q, and then sell the quantity q in each of the next m periods. The costs of this strategy are

$$E[\sum_{n=1}^{2m} p_n q_n] = \sum_{n=1}^{m}\left[p_0 + (n-1)\hat{U}(q) + \hat{P}(q)\right]q$$

$$-\sum_{n=1}^{m}\left[p_0 + m\hat{U}(q) + (n-1)\hat{U}(-q) + \hat{P}(-q)\right]q$$

$$= -\frac{m^2}{2}q[\hat{U}(q) + \hat{U}(-q)] + \frac{m}{2}q[\hat{U}(-q) - \hat{U}(q) + 2(\hat{P}(q) - \hat{P}(-q))],$$

which implies $E[\sum_{n=1}^{2m} p_n q_n] \simeq -\frac{m^2}{2}q[\hat{U}(q) + \hat{U}(-q)]$. But this expression is negative, contradicting the hypothesis of no arbitrage (note that we assume here $N \geq 2m$ to make this trading strategy happen).

Next, we show $\hat{U}(q) \geq -\hat{U}(-q)$ for $q > 0$, also by contradiction. For this purpose assume a $q > 0$ satisfying $\hat{U}(q) < -\hat{U}(-q)$. Now, 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 costs asymptotically equivalent to $\frac{m^2}{2}q[\hat{U}(q)+\hat{U}(-q)]$. Again this term is negative, implying arbitrage (notice that p_0 must be large enough to avoid negative expected prices).

The second assertion, *ii*, is easiest shown by contradiction, too. Assume 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 $q_{n'} \to q+$ with $\hat{U}(q_{n'}) \geq \hat{U}(q)+\varepsilon$,
2. there exists a subsequence $q_{n'} \to q+$ with $\hat{U}(q_{n'}) \leq \hat{U}(q)-\varepsilon$,
3. there exists a subsequence $q_{n'} \to q-$ with $\hat{U}(q_{n'}) \geq \hat{U}(q)+\varepsilon$,
4. there exists a subsequence $q_{n'} \to q-$ with $\hat{U}(q_{n'}) \leq \hat{U}(q)-\varepsilon$.

We shall show that U violates the no-arbitrage condition in each case. N is assumed large enough so that the trading strategies below can be implemented.

Case 1. Use the following strategy: buy $q_{n'}$ units of the asset in each of the first m periods, where n' is an arbitrary index of the subsequence; then sell the quantity q in the each of the following m periods. Given i, the costs of these transactions are $E[\sum_{l=1}^{2m} p_l q_l] \simeq -\frac{m^2}{2}[(\hat{U}(q_{n'})-\hat{U}(q))q+\hat{U}(q_{n'})(q-q_{n'})]$. Since the latter term is negative for sufficiently large n' (verify that the sequence $\left\{\hat{U}(q_{n'})\right\}$ must be bounded!), profits from arbitrage can be realized, in contradiction to the hypothesis.

Case 2. Trading strategy: buy volume q in the each of the first m periods and then sell $q_{n'}$ units in each of the next $m-1$ periods. This implies $E[\sum_{l=1}^{2m-1} p_l q_l] \simeq -\frac{m^2}{2}[(\hat{U}(q)-\hat{U}(q_{n'}))q_{n'}-\hat{U}(q)(q_{n'}-q)]$. But the last expression becomes negative if n' is sufficiently large (note that $(m-1)\,q_{n'} \leq mq$ is met if n' is large enough). Again, arbitrage is possible.

The reader can easily check that for the remaining cases the following two trading strategies contradict the no-arbitrage condition: for case 3, buy $q_{n'}$ units in each of the first m periods and then sell quantity q in each of the following $m-1$ periods; for case 4, buy q units in each of the first m periods and then sell the volume $q_{n'}$ in each of the next m periods. □

Lemma 3 *If U is arbitrage-free, then U satisfies the linear integral equation*

$$E\left[U(q+\eta)\right] = \int_{\Omega} U(q+\eta)d\varphi = \lambda q \qquad \textit{for all } q \in \mathbf{R} \tag{15}$$

where η is as in Lemma 2, and $\lambda \geq 0$.

Proof. Note that (15) is equivalent to $\hat{U}(q) = \lambda q$ for all $q \in \mathbf{R}$. To prove Lemma 3, 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 2 *ii* 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. It follows that $E[\sum_{n=1}^{m(1+q)} p_n q_n] \simeq -\frac{m^2}{2} q[\hat{U}(q) - \hat{U}(1)q]$, contradicting the no-arbitrage condition. Note again that N is chosen large enough to make the strategy above realizable.

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 negative costs for an appropriate m. This completes the proof. □

Proof of Proposition 1. Thanks to Lemma 3 we only need to show conditions (5) and (6).

From Lemma 3 we know that there exists a $\lambda \geq 0$ such that $E[U(q+\eta)] = \lambda q$ for all $q \in \mathbf{R}$, provided that U is arbitrage-free. Take this λ and define the supplementary function S on $\mathbf{R}$ by

$S(q) \triangleq U(q) - \lambda q$. The integral equation (15) can now be restated as

$$\int_{\Omega} S(q+\eta)d\varphi = 0 \qquad \text{for all } q \in \mathbf{R}. \tag{16}$$

Having (16) at hand, we are ready to show that S satisfies the integral equation, $E_n[S(\tilde{q}_n+\eta_n)] = 0$, for any H_n-measurable random variable $\tilde{q}_n$. First, note that we are allowed to write

$$E[S(\tilde{q}_n + \eta_n) \mid H_n] = E[S(g(H_n) + \eta_n) \mid H_n = .] \circ H_n,$$

since, due to the Doob-Dynkin lemma (see, e.g., Rao (1984)), there exists a Borel measurable function $g : \mathbf{R}^{3n-2} \rightarrow \mathbf{R}$ such that $\tilde{q}_n = g(H_n)$. Then, using the notation $\varphi_{\eta_n|H_n=x}$ for the distribution of η_n given the event $\{H_n = x\}$, we obtain

$$E[S(g(H_n) + \eta_n) \mid H_n = x] = \int_{\mathbf{R}} S(g(x) + y)d\varphi_{\eta_n|H_n=x}(y)$$

$$= \int_{\mathbf{R}} S(g(x) + y)d\varphi_{\eta_n}(y) = 0$$

for all $x \in \mathbf{R}^{3n-2}$, thanks to (16) and the independence of η_n. Therefore, $E[S(\tilde{q}_n + \eta_n) \mid H_n] = 0$, ending the proof. $\square$

Proof of Proposition 2. We only have to study here equation (16).

If $\varphi[\eta_n = 0] = 1$, then $U(q) = \lambda q$, $L(\mathbf{R}) - a.e.$, follows immediately from (16).

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 $x \longmapsto U(x)e^{-a\frac{x^2}{2\sigma_\eta^2}}$ is $L(\mathbf{R})$-integrable. This is a mild assumption because $E[U(\eta_n)] < \infty$ holds in any case, as the price process (1) is assumed to be integrable.

For normally distributed η_n's the integral equation (16) becomes

$$\frac{1}{\sqrt{2\pi}\sigma_\eta}\int_{\mathbf{R}} S(x)e^{-\frac{(x-q)^2}{2\sigma_\eta^2}}\,dx = 0 \quad \text{for all } q \in \mathbf{R},\ \lambda \geq 0. \tag{17}$$

Using the above assumption, it is an easy exercise to verify that (17) can be reformulated as

$$\int_{\mathbf{R}} \left[S(x)e^{-a\frac{x^2}{2\sigma_\eta^2}}\right]\left[\frac{1}{\sqrt{2\pi}\sigma_\eta/\sqrt{1-a}}e^{-\frac{(x-q)^2}{2\sigma_\eta^2/(1-a)}}\right]dx = 0 \tag{18}$$

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 $F[f](x) \triangleq \int_{\mathbf{R}} e^{ixy} f(y)dy$. Invoking the convolution theorem of Fourier transforms for (18) gives

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

which implies that $S = 0$, $L(\mathbf{R}) - a.e.$, since F is injective. So $U(q) = \lambda q$, $L(\mathbf{R}) - a.e.$, holds also for the case of normal-distributed η_n's. $\square$

Remark 1 *The supplementary function is also zero when the residual trades are a transform* $W : \mathbf{R} \to \mathbf{R}$ *of a zero-mean normal random variable, where* W *satisfies* $W(x) = -W(-x)$ *and* $\frac{dW}{dx}(x) > 0$ *for all* $x \geq 0$.

Proof. In this case, (16) has the form

$$\frac{1}{\sqrt{2\pi}\sigma_\eta}\int_{\mathbf{R}} S(q + W(x))e^{-\frac{x^2}{2\sigma_\eta^2}}dx = 0 \quad \text{for all } q \in \mathbf{R}. \tag{19}$$

But this is evidently equivalent to

$$\frac{1}{\sqrt{2\pi}\sigma_\eta}\int_{\mathbf{R}}\frac{S(x)}{W'(x)}\exp\left[-\frac{1}{2\sigma_\eta^2}\left(W^{-1}\right)^2(x-q)\right]dx=0 \quad \text{for all } q\in\mathbf{R}. \tag{20}$$

By applying Fourier transforms to this equation we obtain

$$F\left[y\mapsto\frac{S(y)}{W'(y)}\right](x)\ F\left[y\mapsto\exp\left[-\frac{1}{2\sigma_\eta^2}\left(W^{-1}\right)^2(y)\right]\right](x)=0$$

for all $x\in\mathbf{R}$, from which $S=0$, $L(\mathbf{R})-a.e.$, follows, because W' and the modulus of $F[y\mapsto\exp\left[-\frac{1}{2\sigma_\eta^2}\left(W^{-1}\right)^2(y)\right]]$ are both positive on $\mathbf{R}$. We have therefore established the validity of this remark. □

Proof of Lemma 1. Only the case $P>\frac{1}{2}U$ has to be considered here since $E[\sum_{n=1}^N p_n q_n]=0$ when $\sum_{n=1}^N q_n=0$ and $P=\frac{1}{2}U$. It is convenient here to modify (7) slightly by replacing the constraint $\sum_{n=1}^N q_n=0$ with $\sum_{n=1}^N q_n=Q\geq 0$. The associated Bellman equation of (7) with the more general constraint is

$$C_n=\min_{q_n,\ H_n\text{-measurable}} E_n[p_n q_n+C_{n+1}] \tag{21}$$

$$\text{subject to}\quad Q_n=Q_{n-1}-q_{n-1},$$

$$Q_0\triangleq 0,\ Q_1\triangleq Q\geq 0,\ \text{and } Q_{N+1}=0,$$

where the Q_n's denote the remaining shares to be traded, and C_n represents the remaining costs of trading. Standard computations show that optimal trades and cost function have the form

$$q_n=\frac{Q}{N}\geq 0,\quad \text{for } 1\leq n\leq N,\ \text{and}$$

$$E[C_1] = p_0 Q + \frac{N+1}{2N}[2P(1) - U(1)]Q^2 \geq 0,$$

proving the lemma. □

Proof of Proposition 8. Due to (10) the only thing we have to demonstrate here is that the optimal trading sequence of (7) is deterministic if $\{\lambda_n\}_{n=1}^{\infty}$ and $\{\mu_n\}_{n=1}^{\infty}$ are deterministic. Using the Bellman equation (21), a little algebra shows that its solution is given by

$$q_n = (1 - \frac{2\mu_n - \lambda_n}{2\gamma_{n+1}})Q_n, \;\; q_N = Q_N, \text{ and}$$

$$C_n = [p_{n-1} + \varepsilon_n + (\lambda_{n-1} - \mu_{n-1})(Q_{n-1} + \eta_{n-1})]Q_n + \gamma_n Q_n^2$$

for $1 \leq n \leq N$, where

$$\gamma_n = [\lambda_n - \lambda_{n-1} - (\mu_n - \mu_{n-1}) + (2\mu_n - \lambda_n)(1 - \frac{2\mu_n - \lambda_n}{4\gamma_{n+1}})]$$

and $\gamma_N = \mu_N + \mu_{N-1} - \lambda_{N-1}$, confirming our claim. □

Proof of Proposition 9. The proof is divided into 5 steps. $\hat{U}_{ij}$, $i \neq j$, is as defined in the main text.

Step 1: $\hat{U}_{ij}$ is symmetric, i.e., $\hat{U}_{ij}(q) = -\hat{U}_{ij}(-q)$ on $\mathbf{R}$.

If not, then (i) there exists $q > 0$ with $\hat{U}_{ij}(q) > -\hat{U}_{ij}(-q)$, or (ii) there exists $q > 0$ with $\hat{U}_{ij}(q) < -\hat{U}_{ij}(-q)$, 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 $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq -m^2 q[\hat{U}_{ij}(q) + \hat{U}_{ij}(-q)]$. 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, $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq m^2 q[\hat{U}_{ij}(q) + \hat{U}_{ij}(-q)]$. Both trading strategies offer arbitrage opportunities for appropriate choice of m. Case (iii) is easy to rebut and left to the reader.

Step 2: $\hat{U}_{ij}$ is continuous on $\mathbf{R}\backslash\{0\}$. In what follows, we verify that none of the four cases stated in Lemma 2 can hold for $\hat{U}_{ij}$. For the first case take the strategy of buying q shares of asset i in each of the first m periods, buying $q_{n'}$ 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 results in $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq -m^2 q[\hat{U}_{ij}(q_{n'}) - \hat{U}_{ij}(q)]$. For the second case, consider selling q shares of asset i in each of the first m periods, buying $q_{n'}$ shares of asset j in each of the next m periods, selling q shares of asset j in each of the following m periods, and buying q shares of asset i in each of the next m periods. Costs are thus $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq m^2[\hat{U}_{ji}(q)(q-q_{n'}) + \hat{U}_{jj}(q_{n'})(\frac{1}{2}q_{n'} - q) + \frac{1}{2}q\hat{U}_{jj}(q) + q[\hat{U}_{ij}(q_{n'}) - \hat{U}_{ij}(q)]]$. In the third case, take the strategy of selling 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 $q_{n'}$ shares of asset j in each of the next m periods, and buying q shares of asset i in each of the following m periods. We obtain $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq m^2[\hat{U}_{ji}(q)(q_{n'} - q) + \hat{U}_{jj}(q)(\frac{1}{2}q - q_{n'}) + \frac{1}{2}q_{n'}\hat{U}_{jj}(q_{n'}) - q[\hat{U}_{ij}(q_{n'}) - \hat{U}_{ij}(q)]]$. For the last case, 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 $q_{n'}$ shares of asset j in each of the following m periods, and selling q shares of asset i in each of the next m periods. This yields $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq m^2[\hat{U}_{ji}(q)(q - q_{n'}) + \hat{U}_{jj}(q)(\frac{1}{2}q - q_{n'}) + \frac{1}{2}q_{n'}\hat{U}_{jj}(q_{n'}) - q[\hat{U}_{ij}(q) - \hat{U}_{ij}(q_{n'})]]$. All trading strategies render profits from arbitrage when m and the index n' are chosen appropriately.

Step 3: $\hat{U}_{ij}(q) = \hat{U}_{ij}(1)q$ on $\mathbf{R}$. If it were not, either $\hat{U}_{ij}(q) > \hat{U}_{ij}(1)q$ for a $q > 0$ 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 $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq -m^2 q[\hat{U}_{ij}(q) - \hat{U}_{ij}(1)q]$. In the second case, we obtain $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq m^2 q[\hat{U}_{ij}(q) - \hat{U}_{ij}(1)q]$ from selling 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 one share of asset j in each of the next mq periods, and buying q shares of asset i in each of the following m periods. Both are at variance with the no-arbitrage condition if m is large enough.

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 costs of $E[\sum_{n=1}^{4m} p_n^T q_n] \simeq -m^2 q[\hat{U}_{ij}(q) - \hat{U}_{ji}(q)]$. Obviously, this is in discord with the absence of arbitrage if $\hat{U}_{ij}(q) > \hat{U}_{ji}(q)$ for a $q > 0$. The reader is invited to falsify the opposite inequality, i.e., $\hat{U}_{ij}(q) < \hat{U}_{ji}(q)$ for a $q > 0$.

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

1. Strategy A: sell q shares of asset j in each of the first $\frac{m}{2}$ periods, buy q shares each of asset i and asset j in each of the next m periods, buy q shares of asset j in each of the next $\frac{m}{2}$ periods, sell q shares of asset j in each of the following m periods, and sell q shares of asset i in each of the next m periods;

2. Strategy B: sell q shares of asset j in each of the first m periods, sell q shares of asset i in each of the next m periods, buy q shares of asset j in each of the following $\frac{m}{2}$ periods, buy q shares each of asset i and asset j in each of the next m periods, and sell q shares of asset j in each of the next $\frac{m}{2}$ periods.

Strategy A costs are asymptotically equivalent to $-\frac{m^2}{2}q_i[\hat{U}_i(q_1,q_2)-\hat{U}_{ii}(q_i)-\hat{U}_{ij}(q_j)]$, while strategy B's are asymptotically equivalent to $\frac{m^2}{2}q_i[\hat{U}_i(q_1,q_2)-\hat{U}_{ii}(q_i)-\hat{U}_{ij}(q_j)]$. Hence, regardless of the value of q_i, the absence of arbitrage implies $\hat{U}_i(q_1,q_2)=\hat{U}_{ii}(q_i)+\hat{U}_{ij}(q_j)$. The rest follows from the arguments presented in the main text. □

Appendix B. Examples of Nonzero Supplementary Functions

We give here three examples where the supplementary function, as stated in Proposition 1, fails to be zero. The proofs are presented after a brief discussion of these examples.

Example A (Bernoulli distribution) Suppose that the residual trades can only assume two values, namely, $\varphi[\eta_n=-\eta_0]=\varphi[\eta_n=\eta_0]=\frac{1}{2}$ *for* $n\in N$ *and* $\eta_0>0$*. In this case,* $U=(1-\alpha)P$ *being arbitrage-free is equivalent to* S *satisfying* $S(x)=-S(x-2\eta_0)$ *for all* $x\in\mathbf{R}$*.*

Example B (Uniform distribution) Assume that the η_n*'s are uniformly distributed on* $\mathbf{R}$*, with compact support* $[-s,s]$*,* $s>0$*, and that* U *is either continuous and of bounded variation or piece-wise continuously differentiable. Then,* S *is a* $2s$*-periodic trigonometric Fourier series satisfying* $\int_0^{2s}S(x)dx=0$ *if and only if* $U=(1-\alpha)P$ *is arbitrage-free. (For the precise form of* S *see below.)*

Example C (Triangle distribution) Let the residual trades have the "triangle density"

$$f_\eta(x)=\begin{cases}(1+\frac{x}{s})/s & x\in[-s,0]\\ (1-\frac{x}{s})/s & x\in(0,s]\end{cases},\ s>0,$$

on $\mathbf{R}$*. In this case,* $U=(1-\alpha)P$ *is arbitrage-free if and only if* S *is given by*

$$S(q)=S_1(q)+S_2(q)q, \tag{22}$$

where $S_1 : \mathbf{R} \to \mathbf{R}$ *and* $S_2 : \mathbf{R} \to \mathbf{R}$ *are s-periodic functions satisfying* $\int_0^s S_1(q)dq = \int_0^s S_2(q)dq = 0$.

Observe that to derive the result in Example B, we need to impose smoothness assumptions on U, unlike the results in Propositions 2 and 5 and Examples A and C.

In Examples A, B, and C, S can take on a variety of functional forms. What they have in common is that they are periodic and that either S or its components integrate to zero over any interval with length equal to their periodicity. For instance, any multiple of the sine function would be a possible candidate for the functions S, S_1, and S_2 in Examples B and C, if the periodicity is $s = \pi$ and $s = 2\pi$, respectively. The reader is invited to construe candidate S-functions for Example A.

If the supplementary functions are $2s$ (or s)-periodic, then the supplementary price updates of the total trading volume, $S(q+\eta_n)$, are $2s$ (or s)-periodic, too. Further, the conditional expectation $E_n[S(q + \eta_n)]$ vanishes if S integrates to zero over all intervals with length $2s$ (or s). As these conclusions are straightforward, the main job of verifying Examples B and C is to prove also the converse, i.e., that no arbitrage requires the supplementary functions to have the two properties stated above.

Note that the precise shape of S is determined by the curvature of the residual trades' density function and is therefore variable. For more complicated distributions, S can still be computed, albeit with much more intricate structure.

Proof of Example A. We have seen above that the absence of arbitrage requires the supplementary function to meet (16), which in this case becomes

$$S(q + \eta_0) + S(q - \eta_0) = 0 \quad \text{for all } q \in \mathbf{R}. \tag{23}$$

On the other hand, (23) implies (16). This together with Proposition 5 has the absence of arbitrage as consequence. □

Proof of Example B. Under the maintained assumptions of this example, equation (16) becomes

$$\int_{q-s}^{q+s} S(x)dx = 0 \quad \text{for all } q \in \mathbf{R}. \tag{24}$$

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

$$S(q) = \sum_{n=1}^{\infty} \left[a_n \cos(\frac{\pi s}{n} q) + b_n \sin(\frac{\pi s}{n} q) \right], \tag{25}$$

where

$$a_n = \frac{1}{s} \int_{-s}^{s} S(x) \cos(\frac{\pi s}{n} x) dx \tag{26}$$

and

$$b_n = \frac{1}{s} \int_{-s}^{s} S(x) \sin(\frac{\pi s}{n} x) dx. \tag{27}$$

Note that the above Fourier series does not have an intercept part, a_0, since $a_0 = \int_{-s}^{s} S(x)dx = 0$. Hence S has the representation (25)-(27) if $U = (1-\alpha)P$ is to be arbitrage-free.

Conversely, if $S(q) = U(q) - \lambda q$ is $2s$-periodic and satisfies $\int_0^{2s} S(x)dx = 0$, then it has the representation (25) and meets (24). Then, from Proposition (5) the absence of arbitrage follows. Thus the proof is complete. □

Proof of Example C. We first show that in case of the absence of arbitrage S is given by (22), where $S_1 : \mathbf{R} \to \mathbf{R}$ and $S_2 : \mathbf{R} \to \mathbf{R}$ are s-periodic functions.

If the density of the residual trades is f_η, equation (16) has the form

$$\int_{q-s}^{q+s} S(x) f_\eta(x-q)dx = 0 \quad \text{for all } q \in \mathbf{R}. \tag{28}$$

Now, differentiating this integral equation with respect to q yields

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

By differentiating again, we obtain that S satisfies the difference equation

$$S(q+2s) - 2S(q+s) + S(q) = 0 \quad \text{for all } q \in \mathbf{R}.$$

But the general solution of it is just given by (22), where S_1 and S_2 are both s-periodic.

That S_1 and S_2 satisfy the two integral conditions stated in Example C follows from the identity

$$\int_{q-s}^{q+s} S(x) f_\eta(x-q)dx =$$

$$\frac{1}{s}\int_{q}^{q+s} S_1(x)dx + \frac{1}{s}\int_{q}^{q+s} S_2(x)dx * q = 0$$

for all $q \in \mathbf{R}$.

Alternately, the last equation implies that S satisfies (28) if it has the representation (22) with $\int_0^s S_1(q)dq = \int_0^s S_2(q)dq = 0$. Again, by invoking Proposition (5) we obtain that $U = (1-\alpha)P$ is arbitrage-free. This finishes the proof. $\square$

REFERENCES

Allen, Franklin, and Douglas Gale, 1992, Stock-Price Manipulation, *Review of Financial Studies, Vol.5, No.3,* 503-529.

Allen, Franklin, and Gary Gorton, 1992, Stock-Price Manipulation, Market Microstructure and Asymmetric Information, *European Economic Review, Vol.36,* 624-630.

Back, Kerry., 1992, Insider Trading in Continuous Time*, Review of Financial Studies, Vol.5, No.3,* 387-409.

Bertsimas, Dimitris, and Andrew W. Lo, 1998, Optimal Control of Execution Costs, *Journal of Financial Markets 1*, 1-50.

Black, Fischer, 1995, Equilibrium Exchanges, *Financial Analysts Journal*, *May-June*, 23-29.

Breen, William J., Laurie Hodrick, and Robert A. Korajczyk, 1998, Equity Price Concession, *Working Paper #205*, Kellog Graduate School of Management, Northwestern University.

Chan, Louis K., and Josef Lakonishok, 1995, The Behavior of Stock Prices around Institutional Trades*, Journal of Finance, Vol.50, No.4,* 1147-1174.

Dutta, Prajit K., and Ananth Madhavan, 1995, Price Continuity Rules and Insider Trading, *Journal of Financial and Quantitative Analysis, Vol.30, No.2,* 199-221.

Easley, David, and Maureen O'Hara, 1992, Time and the Process of Security Price Adjustment, *Journal of Finance, Vol.47, No.2,* 577-605.

Gemmill, Gordon, 1996, Transparency and Liquidity: A Study of Block Trades on the London Stock Exchange under Different Publication Rules*, Journal of Finance, Vol.51, No.5,* 1765-1790.

Glosten, Lawrence R., and Paul Milgrom, 1985, Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents, *Journal of Financial Economics, Vol.14,* 71-100.

Glosten, Lawrence R., 1994, Is the Electronic Open Limit Order Book Inevitable?, *Journal of Finance, Vol.49, No.4*, 1127-1161.

Hasbrouck, Joel, 1991, Measuring the Information Content of Stock Trades, *Journal of Finance, Vol.46, No.1,* 179-207.

Hausman, Jerry A., Andrew W. Lo, and A. Craig MacKinlay, 1992, An Ordered Probit Analysis of Transaction Stock Prices, *Journal of Financial Economics, Vol.31,* 319-379.

Holden, Craig W., and Avanidhar Subrahmanyam, 1992, Long-Lived Private Information and Imperfect Competition, *Journal of Finance, Vol.47, No.1,* 247-270.

Holthausen, Robert, Richard Leftwich, and David Mayers, 1987, The Effect of Large Block Transactions on Security Prices, *Journal of Financial Economics, Vol.19,* 237-267.

Holthausen, Robert, Richard Leftwich, and David Mayers, 1990, Large Block Transactions, the Speed of Response, and Temporary and Permanent Stock-Prices Effects, *Journal of Financial Economics, Vol.26,* 71-95.

Huberman, Gur, and Werner Stanzl, 2000, Optimal Liquidity Trading, *Mimeo,* Columbia University.

Jarrow, Robert A., 1992, Market Manipulation, Bubbles, Corners, and Short Squeezes, *Journal of Financial and Quantitative Analysis, Vol.27, No.3,* 311-336.

Keim, Donald B., and Ananth Madhavan, 1996, The Upstairs Market for Large Block Transactions: Analysis and Measurement of Price Effects, *Review of Financial Studies, Vol.9, No.1,* 1-36.

Kempf, Alexander, and Olaf Korn, 1999, Market Depth and Order Size, *Journal of Financial Markets, Vol.2, No.1,* 29-48.

Kyle, Albert S., 1985, Continuous Auctions and Insider Trading, *Econometrica, Vol.53, No.6,* 1315-1336.

Medrano, Luis A., and Xavier Vives, 1997, Strategic Behavior and Price Discovery, *Unpublished Paper*, Harvard University.

Musiela, Marek, and Marek Rutkowsi, 1997, *Martingale Methods in Financial Modelling*, Applications of Mathematics, Springer Verlag.

Rao, M.M., 1984, *Probability Theory with Applications*, Academic Press.

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