

NBER WORKING PAPER SERIES

NOTES FOR A CONTINGENT CLAIMS
THEORY OF LIMIT ORDER MARKETS

Bruce N. Lehmann

Working Paper 11533

<http://www.nber.org/papers/w11533>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

July 2005

This research is part of the NBER's program on Asset Pricing and was supported in part by a Morgan Stanley Equity Microstructure Research Grant for which I am quite grateful. The views expressed herein are solely my own and not those of any other person or entity including Morgan Stanley and the National Bureau of Economic Research. It has also benefitted from the comments of participants at the conferences in *Honor of David K. Whitcomb* at Rutgers University (where participants found a dreadful blunder), on the *Econometrics of the Microstructure of Financial Markets* at Tilburg University, on the *Analysis of High-Frequency Financial Data and Market Microstructure* at Academia Sinica, and on the *Workshop on Microstructure of Financial Markets* at the Bolsa Madrid. All errors that remain are my own responsibility. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2005 by Bruce N. Lehmann. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Notes for a Contingent Claims Theory of Limit Order Markets

Bruce N. Lehmann

NBER Working Paper No. 11533

July 2005

JEL No. G10, G13, G19

ABSTRACT

This paper provides a road map for building a contingent claims theory of limit order markets grounded in a simple observation: limit orders are equivalent to a portfolio of cash-or-nothing and asset-or-nothing digital options on market order flow. However, limit orders are not conventional derivative securities: order flow is an endogenous, non-price state variable; the underlying asset value is a construct, the value of the security in different order flow states; and arbitrage trading or hedging of limit orders is not feasible. Fortunately, none of these problems is fatal since options on order flow can be conceptualized as bets implicit in limit orders, arbitrage trading can be replaced by limit order substitution, and plausible assumptions can be made about the endogeneity of order flow states and their associated asset values. The analysis yields two main results: Arrow-Debreu prices for order flow “states” are proportional to the slope of the limit order book and the limit order book at one time proves to be identical to that at an earlier time adjusted for the net order flow since that time when all information arrives via trades.

Bruce Lehmann

University of California, San Diego

IR/PS

1415 Robinson Building Complex

La Jolla, CA 92093-0519

and NBER

blehmann@ucsd.edu

I. Introduction

A dozen or so years ago, I visited David Whitcomb in his New York apartment to view his limit order trading system. He invited me to do so under the misapprehension that I might be able to help him find a buyer, a notion of which he was rapidly disabused.¹ I came away remembering one detail of the system: limit orders were always canceled if they were not executed within two minutes because the specialist's commission – a number on the order of a mil at the time as I recall – was only paid on orders that resided on the book more than two minutes. The idea that stuck with me is that one could learn a great deal about optimal trading strategies by paying close attention to the minutiae of the microstructure of a market. Anyone who feels the need to verify the claim that I see value in paying attention to such minutiae should examine the description of the Tokyo Stock Exchange in Lehmann and Modest (1995).

The research reported here marries this observation with a simple idea: the apparatus of state pricing can be applied profitably in microstructure. Just as Hicks (1939) noted that conventional microeconomic theory could be applied to economic dynamics by treating the same goods on different dates as different commodities and Arrow (1964) and Debreu (1959) showed that dynamics under uncertainty could be analyzed similarly by treating the same good on different dates in different states of nature as different commodities, it has long seemed to me that one could exploit this strategy in microstructure by defining states in terms of trade prices, quotes, and quantities. Limit order markets are a natural environment within which to apply this idea precisely because the mechanical nature of the order execution process makes for a clear definition of order flow states.

The idea of using contingent claims analysis in a microstructure setting is hardly new: since at least Copeland and Galai (1983), it has been commonplace to view quotes or limit orders

as free options given by limit order traders or market makers to market order traders who may possess superior information regarding asset values. On this view, a limit buy (sell) order is a call (put) option with a strike price equal to the limit price with an implicit option premium — actually a contingent premium since it is not paid unless the option is exercised — given by the spread between the limit price and the midpoint of the bid-ask spread, albeit one with an uncertain expiration date since limit orders are typically good until cancelled. This spread reflects the by now classic balance between losses to informed investors and profits from uninformed liquidity traders along with traders who falsely believe they possess value-relevant private information.

Examining this implicit option from the perspective of limit order traders sheds quite different light on the sources of value in limit orders. A limit buy (sell) order for X_t shares involves the receipt (expenditure) of $P_t X_t$ dollars in exchange for (delivery of) X_t shares. In the language of exotic options markets, the first payoff is that of X_t cash-or-nothing digital call options struck at P_t and the second is the same as that on X_t asset-or-nothing digital call options struck at P_t as well. Since limit order markets only result in transactions at a given price $P_t(q)$ when a market order of size Q_t is large enough to hit the limit orders posted at the price (i.e., when $Q_t \geq q$), one can view these implicit derivative securities as options on order flow.

The chapter is laid out as follows. The next section describes both my assumptions about limit order markets and the digital options implicit in limit orders. The third section is devoted primarily to the identification of the Arrow-Debreu prices for order flow ‘states’ based on the implicit options embedded in the limit order book. Care is needed since these implicit digital options differ in important ways from their analogues in conventional derivative asset markets since the nature of the underlying asset, the definition of states of nature, and the notion of

¹ As I recall, the asking price was \$10 million and the proposed commission was 10 per cent.

arbitrage are not entirely straightforward in this context. The penultimate section derives some perhaps surprising implications for limit order book dynamics and the final section provides brief concluding remarks.

2. Limit Orders as Order Flow Derivatives

Consider a limit order book market for a single security with a marginal price schedule $P_t(q)$ where q is positive for buy orders and negative for sell orders. Limit orders are placed prior to time t and so $P_t(q) \in \mathcal{F}_{t-1}$ where \mathcal{F}_{t-1} is public information available before time t including, at minimum, past quotes along with transactions prices, quantities, and times. Assume for now that $P_t(q)$ is continuous; that is, there is no minimum price variation or minimum tick.

Since incoming market orders are executed against standing limit orders, the overall cost of a market order for Q_t shares is $\int_0^{Q_t} P_t(q) dq$ where Q_t and q are positive for buy orders and negative for sell orders – that is, the market order walks up or down the book until it is filled. Note that all orders at the same limit price are treated symmetrically in this limit order book — that is, price priority is strictly maintained while both size and time priority are ignored. Neither omission should be of great importance in this market since one can *always* step ahead of an existing order by posting one at an infinitesimally better price.

This market should be thought of as one that operates without frictions, save for enough slippage to permit limit order traders to cancel old orders and/or submit new ones before the next market order arrives. There are no dealers like NYSE specialist who can see the book before taking a position or stepping in front of an existing limit order in the book. There are no hidden limit orders that market order traders might “ping” to discover how much immediacy resides in the undisclosed portion of the book. There are no limits on how many tiers of the book a single

market order can clear out as there are on the Tokyo Stock Exchange. Some of these frictions can be handled by careful definition of the “state” of the book and some cannot. I will not deal with them in what follows.

No special assumptions are made about order flows at this point. There can be limit or market orders that follow arbitrary stochastic processes that depend in arbitrary ways on the slope, depth, and history of the limit order book as well as other public information. To be sure, most of the heavy lifting in microstructure models involves the determination of the optimal order placements of informed investors and/or optimizing liquidity traders, all of which requires numerous assumptions about the economic setting. The present exercise is much easier since it only involves the *mechanics* of order execution in a stylized limit order market. Some modest assumptions about order flows and execution will be made in the next two sections.

The analysis of contingent claims written on order flows begins with the three possible events that might transpire at time t : the arrival of a market buy order (i.e., $Q_t > 0$), of a market sell order (i.e., $Q_t < 0$), or of information without an intervening trade including, perhaps, that the clock has ticked without the arrival of a trade (i.e., $Q_t = 0$). This last prospect will often lead to a revision of the limit order book, an instance of the good until cancelled feature of limit orders. In this context, my assumption that there is a tatōnnement-like process permitting limit order traders to freely cancel and submit orders prior to the next market order arrival means that limit order traders, by assumption, have placed the limit orders they think appropriate given all of the information available to them at time t .

Now suppose a market buy order arrives for $Q_t > q$ shares at time t and consider a limit order that offered one share at price $P_t(q)$. Since this limit order is executed for sure (i.e., $Q_t > q$), the limit order trader exchanges one share of stock for $P_t(q)$ dollars. In the language of exotic options, a contingent claim that pays a fixed cash flow in a given eventuality is called a cash-or-

nothing digital call option and so the cash flow from a limit sell order is the payoff on a cash-or-nothing digital call option struck at $P_t(q)$. Similarly, a contingent claim that pays a share in a given state is called an asset-or-nothing digital call option and so the share transfer in this transaction is the payoff of an asset-or-nothing digital call option struck at $P_t(q)$ as well. I will follow Ingersoll's (2000) nomenclature and refer to cash-or-nothing digital calls and puts as digital options and to asset-or-nothing digital calls and puts as digital shares.

Hence, limit orders implicitly bundle digital option and share positions. A limit buy order is equivalent to a long position in a digital option and a short position in a digital share, each struck at $P_t(q)$. Each implicit digital option pays $P_t(q)$ dollars if $Q_t \geq q$ and expires worthless if $Q_t < q$. Similarly, each implicit digital share converts into a share if $Q_t \geq q$ and expires worthless otherwise. By the same token, a limit sell order is equivalent to a short position in a digital option and a long position in a digital share. The former requires the payment of $P_t(q)$ dollars when $Q_t \leq q$ and the latter involves the receipt of a share in these circumstances. Both expire worthless if $Q_t > q$.

Since a limit order involves no up front cash flows, the value of these two implicit options must be the same. Accordingly, any analysis that delivers the value of one option position implicitly determines that of the other position as well. In addition, any such valuation must take account of the fact that these derivatives are implicitly written on order flow, a non-price state variable.

This is a weaker restriction than that derived in Ingersoll (2000) for conventional digital options or shares. After all, the underlying asset price, a quantity measured in dollars, is always available to compute the relative value of digital options and shares in any state of nature. However, market order size is a non-price state variable and there is no observable that corresponds to the value of a share in different order flow states. This inability to mechanically

assign a value to the digital share on expiration or to a claim to a dollar in the same state of nature is the main obstacle to the straightforward computation of the value of the contingent claims implicit in limit order books.

A digital call (put) can be replicated by a bull (bear) spread in conventional call (put) options with an infinitesimal spread between the two strike prices. This observation is of limited utility for limit order valuation purposes since it is hard to imagine a market in which it is easy to form a bull spread in conventional options at the appropriate strikes just before a market order arrives that expires on the arrival of the market order. It may be of some use in reverse: inferences regarding the prices of order flow states implicit in the limit order book can provide insight into microstructural effects on the value of very short expected maturity options.

However, this fact does yield one modest insight concerning the riskiness of the derivatives implicit in limit orders. As is well-known, at-the-money digital options with short time to expiration are highly sensitive to changes in the underlying asset value because the absolute values of their deltas and gammas – that is, the sensitivity of the bull spread to changes in the underlying asset price and its square, respectively – can both be large in absolute value and fluctuate a great deal. Moreover, the gamma of such a digital call can change signs and its delta is unbounded. Accordingly, the prices of limit orders should be quite sensitive to changes in both order flow dynamics and the elasticity of limit order prices with respect to order flow. Perhaps this sensitivity explains why changes in the market for liquidity in a stock are associated with volatility in limit order cancellations and submissions.

More revealing for valuation purposes, however, is a spread in digital calls with an infinitesimal spread between the two strike prices. Since a digital call (put) is equivalent to a spread in conventional calls (puts), a spread in digital call (put) options can be replicated by a spread in bull (bear) spreads in conventional options or, in more common parlance, by a butterfly

spread in three conventional calls (puts). As shown in Breeden and Litzenberger (1978), a butterfly spread in conventional options with infinitesimal differences across the three strike prices has a payoff proportional to that of a pure Arrow-Debreu claim that pays off only when the stock price on expiration equals the intermediate strike price. Similarly, the payoff of a bull spread in digital calls (puts) on order flow pays off only in the intermediate order flow state when there is an infinitesimal spread in the order flow states. Hence, the analysis of the contingent claims implicit in limit orders yields a new interpretation of the *slope* of the limit order book.

3. Limit Order Valuation and Order Flow Bets

Hence, the contingent claims implicit in limit orders can be viewed as the payoffs on two American, deferred premium, one-touch, binary or digital options. They are American options because limit orders are generally good until cancelled and they are one-touch options because they pay a fixed quantity of shares or dollars if the market order exceeds the tier of the book on which the limit order resides. They are also very short maturity options since they last, at most, one trading day. We need only value the digital options or digital shares implicit in limit orders to provide a contingent claims interpretation of the sources of value in limit order books.

There are nontrivial challenges associated with exploiting this observation. The “underlying asset” is order flow – a non-price state variable – and there is no observable or risk premium for order flow risk. Moreover, order flow is not an exogenous set of outcomes but rather is the endogenous product of the trading decisions of informed and uninformed investors. In addition, the application of arbitrage reasoning – the natural starting point for the valuation of contingent claims – is hampered by the presence of short sales restrictions and the absence of natural hedging instruments. Finally, any such analysis must account for the American nature of these implicit order flow derivatives.

There is a simple but useful device for thinking through these issues: imagining a parallel but separate market for wagering on order flow in which participants bet by buying and selling order flow contingent claims. In such a market, claim q pays \$1 if the next order is for exactly q shares – i.e., when $Q_t = q$ – and zero otherwise, making it an Arrow-Debreu claim on order flow. As is readily apparent, the cash flows associated with any trading strategy in the actual limit order market can be perfectly replicated in the betting market. When this fictitious betting market is isomorphic to general arbitrage-free financial markets, we can freely import analysis from that setting to this one when it is convenient to do so. Additional insights into the contingent claims interpretation can be gleaned by being precise about when the betting market produces exactly the same state prices as those implicit in the limit order market.

Before proceeding, it is worthwhile to consider the role of interest rates in the analysis. As is commonplace in the analysis of generic derivatives, the time value of money drops out of the relevant pricing relations if the numeraire is a savings account that accrues at the relevant risk-free interest rate. There is an even better reason to do so in the present setting: interest rates are a second order consideration over (most) market microstructure time scales and so most researchers are content to simply set the riskless rate to zero. On this hypothesis, state prices are also risk neutral probabilities for order flow states.

One nettlesome problem associated with the betting market analogy concerns the endogeneity of order flow. Since order flow states represent the transactions of market order traders, such traders have an incentive to place bets on market orders of given sizes and then to submit the corresponding market orders in the limit order market. Accordingly, there can, in general, be no trade in the betting market in equilibrium since trades in the actual limit order market can be used to perfectly manipulate the betting market. The simplest solution is to assume that market order traders cannot wager in the betting market prior to the placement of

market orders. That is, the natural solution to the problem of endogeneity of order flow is to assume there is no intermarket front-running.²

The limit order market analogue is Assumption 3 of Glosten (1994), which requires traders to use only public information when submitting limit orders. That is, informed and uninformed traders interpret the information content of posted limit orders in the same way under this assumption. This circumstance will arise when informed investors cannot use limit orders by assumption, which is the usual interpretation of Glosten (1994). However, this assumption will also hold when informed traders have no incentive to post limit orders at prices different from those that would be submitted by uninformed traders. If the supply schedule of limit orders that would be submitted by uninformed traders is common knowledge, informed traders would have no incentive to post orders off the supply schedule.³ They will have the incentive to post orders on the schedule if it lies above their valuations for limit sell orders and below their valuations for limit buy orders.

Accordingly, suppose the set of possible market order sizes is given by $\mathbb{Q}_t \subseteq \mathbb{R}$ and that market participants agree that all order sizes in \mathbb{Q}_t are possible.⁴ Note that \mathbb{Q}_t is likely to be bounded since the market will not provide bets for all order sizes if the adverse selection problem is severe enough. In the absence of frictions in the betting market, the usual arbitrage argument insures the existence of a set of (not necessarily unique) strictly positive state prices $\{\psi_t(q), q \in \mathbb{Q}_t\}$, where $\psi_t(q) = E_{\nu}[1_{Q_t=q} | \mathcal{F}_{t-1}]$ with $E_{\nu}[\bullet | \bullet]$ denoting the associated risk neutral

² See the bluffing subsection of Lehmann (2005) for a more detailed analysis. I don't mean to suggest that this issue is easy to deal with in substance but rather that it is easy to find assumptions under which order flow endogeneity is not a problem.

³ By informed trader, I mean those who actually possess value relevant private information as well as noise traders who think they are informed but who are not.

expectation.]

While nothing in what follow depends on it, it is natural to ask whether spanning would naturally arise in the betting market.⁵ That is, would bettors naturally write claims on all of \mathbb{Q}_t ? The answer to this question is yes because the size of the subsequent market order is observable and its endogeneity has been assumed away. If there were at least two investors with different shadow prices for some non-traded claim for q^* shares, some investor could earn an arbitrage profit by buying a claim that pays \$1 when the next market order is for q^* shares from the low valuation investor and selling it to the high valuation investor. Such riskless profit opportunities will not arise if this betting market is arbitrage-free. Accordingly, it is natural to assume that state prices $\{\psi_t(q), q \in \mathbb{Q}_t\}$ are unique because the menu of claims that will be offered in arbitrage-free betting markets spans \mathbb{Q}_t . The analogue in the actual limit order market is that its book should have orders posted at all feasible trade sizes, a prediction clearly at variance with the evidence from actual limit order markets.]

In any event, the question at hand is whether these state prices can be used to value limit orders. What is missing is $V_t(q)$, the value of the asset if a market order of size q arrives at time t . If $V_t(q)$ is common knowledge among market participants and the state prices in the betting and limit order markets are identical, the value of a claim to a share in order flow state q is simply $\psi_t(q)V_t(q)$. The existence of such values crops up in the analysis of generic financial markets in the definition of states of nature: the typical assumption is that there is a deterministic mapping between cash flows and asset values on the one hand and their associated states of

⁴ Agreement on the possible is necessary for equilibrium to exist in frictionless markets because any market participant who believed a given state is impossible will be happy to sell an infinite quantity of bets that pay off in that state.]

⁵ Spanning isn't necessary since order flow is a non-price state variable and \mathbb{Q}_t can simply be defined to be the set of trade sizes offered, which is trivially spanned by construction.]

nature on the other. The problem in the present setting, of course, is that there is no observable mapping between the value of the asset being traded and order flow states. The absence of such a mapping motivates Assumption 2 of Glosten (1994): the existence of a function $V_t(q)$ that is both nondecreasing in q and common knowledge among market participants, although I will assume it is strictly increasing for simplicity.

Limit order valuation is straightforward when state prices from the betting market are identical to those implicit in the book. Consider a market buy order since the sell side is symmetric. The digital option implicit in the corresponding limit sell order pays $P_t(q)$ when $Q_t > q$. The implicit digital share value is given by the value of a claim to a share in each order flow state integrated over order flow states. Since the values of the implicit digital option and digital share are equal:

$$E_{\psi} [1_{Q_t \geq q} P_t(q) | \mathcal{F}_{t-1}] = E_{\psi} \left[\int_q^{\infty} 1_{Q_t=u} V_t(u) du | \mathcal{F}_{t-1} \right] \Rightarrow P_t(q) \int_q^{\infty} \psi_t(u) du = \int_q^{\infty} \psi_t(u) V_t(u) du \quad (1)$$

the limit price is given by:

$$P_t(q) = \frac{\int_{u \geq q} \psi_t(u) V_t(u) du}{\int_{u \geq q} \psi_t(u) du} = \frac{E_{\psi} [V_t(Q_t) 1_{Q_t \geq q} | \mathcal{F}_{t-1}]}{E_{\psi} [1_{Q_t \geq q} | \mathcal{F}_{t-1}]} \quad (2)$$

which is the risk neutral analogue of the upper tail expectation formula in Glosten (1994).⁶

The only remaining question is whether state prices from the betting market are identical

⁶ As Glosten emphasizes, the fact that it is an upper tail expectation reflects the discriminatory nature of the limit order book. As (2) clearly shows, this observation is independent of the probabilities used in the calculation. In contrast, the conditional mean $P_t(q) = E_{\psi} [V_t(Q_t) | Q_t = q, \mathcal{F}_{t-1}]$ is the limit order schedule in a nondiscriminatory book. Most limit order markets open with a single price auction in which all market orders are consolidated into an aggregate net order and executed against the book at the same price. As noted above, this price can be treated as the “liquidating” asset value embedded in the expectations $V_t(q)$. Note also that these single price auctions are isomorphic to the batch auctions studied by Huberman and Stanzi (2004). It may be possible use their results to place restrictions on the “terminal asset

to those implicit in the book. This brings us to the third basic problem with the betting market analogy: the absence of short sales restrictions. In a betting market, all gambles are in zero net supply and, thus, there are obviously no short sales restrictions. However, the limit order market permits no short selling of limit orders and so the frictionless markets assumption cannot apply. Since state prices implicit in the book are given by bull spreads in the relevant limit orders, this absence of short sales in the limit order book would appear to be an insuperable barrier to the application of the betting market analogy.

The way out of this conundrum is to recall that a zero net investment portfolio always has two interpretations. The standard one is as an arbitrage portfolio in which any long positions are financed by short positions. However, a zero net investment portfolio also represents a feasible change in an existing portfolio with no short positions in which security purchases are financed by security sales. A feasible change is one in which none of the asset sales is so large as to create a short position in any asset and an arbitrage portfolio can always be scaled so that no long position is exhausted in the absence of indivisibilities.

It is this second interpretation that is the right one to apply in this context. A limit order trader can always contemplate a swap of all or part of an existing limit order for another at a different price point or, for that matter, for none at all. If one limit order trader with no private information perceives that a limit order is mispriced, all such traders will cancel their orders at this price point if it is overpriced and they will submit new orders at this price point if it is underpriced. Limit order substitution by sufficiently many patient or value traders is a perfect substitute for explicit arbitrage in these circumstances.⁷

values” determined in opening call auctions which might, in turn, deliver additional intertemporal restrictions on order flow state prices. See the end of Section 4.

⁷ This version of the *Fundamental Theorem of Asset Pricing* corresponds to one in which the portfolio choice problem of an investor with (possibly state dependent) von Neumann-Morgenstern preferences will not have an interior maximum if arbitrage opportunities are

The mathematics of risk neutral upper tail expectations naturally dovetails with the observation that bull spreads in limit orders reveal the associated implicit state prices. The slope of the limit order book at q is proportional to the risk neutral probability $\psi_t(q)$ since:

$$P'_t(q) = \lambda_t(q)[P_t(q) - V_t(q)]; \quad \lambda_t(q) = -\frac{\psi_t(q)}{\int_q^\infty \psi_t(u) du} \quad (3)$$

by Leibniz's rule where $\lambda_t(q)$ is the hazard function familiar from survival analysis. Accordingly, the risk neutral probabilities implicit in $P_t(q)$ can be extracted from $P'_t(q)$ given $V_t(q)$ via the recursion:⁸

$$\psi_t(q) = \lambda_t(q) \exp\left\{-\int_q^\infty \lambda_t(u) du\right\} \quad (4)$$

Of course, $V_t(q)$ can be uncovered from $P'_t(q)$ given $\lambda_t(q)$ and, thus, $\psi_t(q)$ as well.

There is one state that cannot be valued by limit order substitution alone: the null "trade" $Q_t = 0$ that corresponds to changes in the state of the book that arise when non-trade-related information arrives in the market. Recall that limit orders were predetermined with respect to information arrives in the market. Recall that limit orders were predetermined with respect to market orders conditional on public information by assumption through the tatōnnement-like process by which traders could freely cancel and replace their limit orders before the next market order arrived. In this model, limit order traders can revise their limit orders in light of any non-trade-related information as well before exposing them to a market order when $0 \in \mathbb{Q}_t$

present. Alternatively, optimizing behavior on the part of market order traders along the lines of Assumption 1 of Glosten (1994) along with uninformed risk neutral limit order traders is sufficient as well.

⁸ After making the relevant substitutions, the discrete version of this formula:

$$\psi_t(q_k) = \frac{\sum_{j=k+1}^{\infty} q_t(q_j) [V_t(q_j) - P_t(q_k)]}{P_t(q_k) - V_t(q_k)}; \quad \sum_{j=0}^{\infty} q_t(q_j) = 1$$

is identical to equation 5 of Banz and Miller (1978) with the option price set to zero, $P_t(q)$ equal to the underlying asset price, and $V_t(q)$ equal to the strike price. The formal similarity ends there

The analogue in the betting market is a provision for the cancellation of all existing bets when the null “trade” arrives. The wagers on this revised claim are refunded to all market participants and new bets are placed, which results in the discovery of new order flow state prices during the tâtonnement-like process. That is, these bets are not pure Arrow-Debreu securities because they pay a dollar if state $q \in \mathbb{Q}_t$ occurs, zero if a market order for any other size $Q_t \neq 0$ arrives, and $-\phi_t(q)$ if $Q_t = 0$, where $1 > \phi_t(q) > 0$ is the initial price paid for this claim.⁹

It is a straightforward matter to construct pure contingent claims by augmenting the asset menu – that is, the set of trade size bets – with one additional claim: a savings account that pays a dollar irrespective of order flow outcomes since the interest rate has been normalized to be

zero. For each $q \neq 0$, a portfolio that is long $\frac{\phi_t(q)}{1 + \int_{\mathbb{Q}_t} \phi_t(u) du}$ units of bet q , short $\frac{\phi_t(q)}{1 + \int_{\mathbb{Q}_t} \phi_t(u) du}$ units of each bet $u \neq q$, and long $\frac{\phi_t(q)}{1 + \int_{\mathbb{Q}_t} \phi_t(u) du}$ units of the riskless savings

account pays one dollar in state $q \neq 0$ and zero otherwise.¹⁰ Similarly, a portfolio that is short

as Banz and Miller are not concerned with bull and bear spreads in digital options as opposed to Breeden and Litzenberger (1978), who (implicitly) construct state prices from such spreads.

⁹ Of course, the bets could have been designed so that there was no need to refund the wager. That is, bets that pay $1 - \psi_t(q)$ if state $q \in \mathbb{Q}_t$ occurs, $-\psi_t(q)$ if a market order for any other size $Q_t \neq 0$ arrives, and nothing if $Q_t = 0$ have no up-front costs and, hence, require no refunds if the expected risk neutral payoff is zero as in $\psi_t(q)[1 - \psi_t(q)] + [1 - \psi_t(q)][-\psi_t(q)] = 0$. The formulation in the text has the virtue of showing how the valuation of Arrow-Debreu securities for all of \mathbb{Q}_t save for 0 – that is, $q \in \mathbb{Q}_t \setminus 0$ – is affected by the presence of the null trade.

¹⁰ This portfolio can be formed by any limit order trader who is long more than $\frac{\phi_t(q)}{1 + \sum_{u \in \mathbb{Q}_t} \phi_t(u)}$ units of each bet. Since $\phi_t(q) < 1$, one unit of the q^{th} bet suffices in the absence of indivisibilities.

$$\frac{1}{1 + \sum_{u=0}^q \phi_t(u)} \quad \text{units of each bet } q \text{ and long} \quad \frac{1}{1 + \sum_{u=0}^q \phi_t(u)} \quad \text{units of the riskless savings}$$

account will yield a dollar when $q = 0$ and zero otherwise. Hence, state prices are given by the prices of these portfolios:

$$\begin{aligned} \psi_t(q) &\equiv \phi_t(q) = \frac{1}{1 + \sum_{u=0}^q \phi_t(u)} \\ \psi_t(0) &\equiv \frac{1}{1 + \sum_{u=0}^0 \phi_t(u)} \end{aligned} \quad (5)$$

which sum to one because these are risk neutral probabilities due to the normalization of the interest rate to zero.

Hence, a betting market coupled with access to riskless savings is isomorphic to the actual limit order market and so state prices calculated in one setting can be freely transferred to the other. All that was necessary was care in constructing the claims in the betting market so that they possessed the relevant features of the contingent claims implicit in limit orders. These characteristics include arbitrage via limit order substitution, the existence of asset values for each order flow state, and enough time for the limit order book to be refreshed – after either a market order has executed or the arrival of information via the “null trade” – before the next market order arrives so that market order traders cannot “front run” the book. Informed investors will have no incentive to post limit orders off of the supply schedule because the schedule that would be posted by uninformed traders is common knowledge in these circumstances.

4. Limit Order Book Dynamics

What happens to state prices if information arrives in the market via a market order at time $t-1$? The answer lies in the change in the information available in the market between time $t-1$ and time t . Since the information state of the market was \mathcal{F}_{t-2} at time $t-1$ and no non-trade-

related information hit the market, the new information state is $\mathcal{F}_{t+1} = \{Q_{t+1}, \mathcal{F}_{t+2}\}$. As a result, the state of the world in which the trades at times t and $t-1$ are q and Q_{t-1} , respectively, is identical to that in which the trade at time $t-1$ is $\bar{Q}_{t-1} = q + Q_{t-1}$ due to the structure of information flows when $Q_{t-1} \neq 0$. Hence, state prices at times t and $t-1$ must be linked by:

$$\psi_t(q) = E_{\psi} [1_{Q_t=q} | \mathcal{F}_{t-1}] = E_{\psi} [1_{Q_t=q} | Q_{t-1}, \mathcal{F}_{t-2}] = E_{\psi} [1_{\bar{Q}_{t-1}=q+Q_{t-1}} | \mathcal{F}_{t-2}] = \psi_{t-1}(q + Q_{t-1}) \quad (6)$$

when information arrives only via market orders.

This restriction generalizes in an obvious way for longer intervals during which all information arrival is trade-related. If the null trade does not occur between times I and T , the recursion suggested by (6) reveals:

$$\begin{aligned} \psi_T(q) &= E_{\psi} [1_{Q_T=q} | \mathcal{F}_{T-1}] = E_{\psi} [1_{Q_T=q} | \mathcal{F}_{T-2}] = \dots = E_{\psi} [1_{Q_T=q} | \mathcal{F}_I] \\ &= E_{\psi} [1_{Q_T=q} | Q_I, \mathcal{F}_I] = E_{\psi} [1_{Q_T=q} | Q_I, \mathcal{F}_I] = E_{\psi} [1_{Q_T=q} | Q_I, \mathcal{F}_I] \\ &= E_{\psi} [1_{Q_T=q} | Q_I, \mathcal{F}_I] = E_{\psi} [1_{Q_T=q} | Q_I, \mathcal{F}_I] = E_{\psi} [1_{Q_T=q} | Q_I, \mathcal{F}_I] \\ &\equiv \psi_I(q + \sum_{i=I}^{T-1} Q_i) \equiv \psi_I(q + Q_I^{T-1}) \end{aligned} \quad (7)$$

where Q_I^{T-1} is cumulative signed volume between times I and $T-1$. It is as though all bets are placed at time zero and the bet on the value of cumulative net order flow given by Q_I^{T-1} is the one that paid off.

What is the analog of these relations in the actual limit order market? The answer is concealed in the equivalence between trades of q and Q_{t-1} at times t and $t-1$, respectively, and one of $\bar{Q}_{t-1} = q + Q_{t-1}$ at time $t-1$ when $\{q, Q_{t-1}, \bar{Q}_{t-1}\} \neq 0$. The limit order book is discriminatory and so limit order traders place orders as though they might confront a single order of \bar{Q}_{t-1} at time $t-1$ or an order of size Q_{t-1} at time $t-1$ and another of size q at time t before limit order traders have a chance to cancel and replace their limit orders. When limit orders do so, they need not fear any such order splitting on the demand side. This is the essence of part ii of Proposition 3 of Gilosten

(1994). Continuation of the recursion in (6) yields the required relation $\psi_T(q) = \psi_1(q + Q_1^{T-1})$ between the time t and time 0 limit order books when all information arrives via trades.

The discriminatory nature of the book coupled with information arrivals that are entirely driven by order flow has a surprising implication for limit order book dynamics: moving up or down the limit order book at a point in time is isomorphic to moving across limit order books over time in these circumstances. It is as though there is a single market of size Q_1^T arrives at time one and walks up or down the book to the marginal price $P_1(Q_1^T)$. Put differently, the marginal price at time T is path independent and Markovian, depending only on Q_1^T if the “null trade” representing value-relevant, non-trade-related information arrives at time $T+1$ or later.

Note that I am not assuming that the actual time zero limit order book has orders posted at all of the possible values of cumulative signed order flow Q_1^T that can occur given the stochastic process generating market order flow. As noted earlier, we would expect the actual book to have finite depth at time zero such that sufficiently large values of Q_1^T are not bid or offered. Put differently, actual markets can force informed market order traders to break up their orders over time within the implicit constraints associated with free entry and competition in limit order placement.

However, limit order markets that routinely remain in information epochs – that is, periods in which all information arrives via market orders – are more likely to satisfy the assumption that limit order traders can freely cancel and replace limit orders before the next market order arrives. Limit order traders will post orders at time $t-1$ on the price schedule $P_1(Q_1^{t-1} + q)$ if no value-relevant, non-trade-related information arrives between times one and $t-1$. If a market order arrives at time t , limit order traders will post limit orders on the marginal

price schedule $P_t(Q_t^T + q)$. If the null trade arrives at time t , limit order traders can cancel all of their orders and repost them when they have assimilated the new information. In fact, this process can be automated if limit order traders know how the set of feasible trades evolve over time as well as the hypothetical marginal price at time zero for each feasible future level of cumulative net order flow. It is certainly common for microstructure theorists to assume that markets experience long information regimes of this sort.

Most markets open (sometimes more than once per day) and some markets close with a single price auction. While the marginal price schedule is given by the upper tail expectation $P_t(q) = E_{\nu}[V_t(Q_t) | Q_t \geq q, \mathcal{F}_{t-1}]$ due to the discriminatory nature of the book, the uniform nature of single price auctions produces prices that are simple expectations $P_t^A(Q_t^A) = E_{\nu}[V_t(Q_t^A) | \mathcal{F}_{t-1}]$ where T is the time of the single price auction and A denotes auction. This completes the picture of dynamics in such markets: the upper tail expectation is also given by $P_t(q) = E_{\nu}[P_t^A(Q_t^A) | Q_t \geq q, \mathcal{F}_{t-1}]$ due to the law of iterated expectations and the expected midquote of the post-auction limit order book is given by $P_t^A(Q_t^A)$. It may prove useful to connect this analysis with Huberman and Stanzl's (2004) study of sequences of batch auction markets that are isomorphic to single price auctions.

5. Conclusion

This paper provides a road map for building a contingent claims theory of limit order markets. It is grounded in a simple observation: limit orders are equivalent to a portfolio of cash-or-nothing and asset-or-nothing digital options on order flow. However, limit orders are not conventional derivative securities: order flow is an endogenous, non-price state variable; the underlying asset value is a construct, the value of the security in different order flow states; and arbitrage trading or hedging of limit orders is not feasible. Fortunately, none of these problems

is fatal since options on order flow can be conceptualized as bets implicit in limit orders, arbitrage trading can be replaced by limit order substitution, and plausible assumptions can be made about the endogeneity of order flow states and their associated asset values.

Perhaps surprisingly, both limit order books and their associated implicit state prices have a simple dynamic stochastic structure under plausible conditions. A special feature of limit order books is their discriminatory nature and a market order that walks up or down the book as it executes is identical to an order that executes one share at a time *over time* when the only new information is the arrival of yet another market order. Hence, prices at time t can be read off of the time one limit order book if the only source of new information arriving between those two times is trade-related.

While I have said nothing about it in this paper, the most interesting aspect of this analysis is its implications for the empirical analysis of limit order markets. Implicit state prices can shed light on basic microstructure questions such as the role of risk aversion, if any, in the provision of immediacy in limit order markets. The successful identification of episodes in which information arrives only via trade can serve to better measure information flows and to provide new insight into the process of price formation. The expectational linkages between periodic single price auctions and continuous trading in limit order markets and even between these markets and options markets should shed considerable light on the economics of the associated intertemporal and cross-market comparisons.

I have called the paper “Notes...” because the theory is as yet incomplete. It seems to me to be worthwhile to present it in this somewhat embryonic state since both the broad outline of the theory and the work needed to complete it seem reasonably clear. And I thought that its link to my glimpse of David Whitcomb’s limit order trading system, the progenitor of his extremely successful company Automated Trading Desk, made it fitting as well.

References

- Arrow, K., 1964, "The Role of Securities in the Optimal Allocation of Risk Bearing," *Review of Economic Studies*, 31, pp. 91-96.
- Banz, R. and M. H. Miller, 1978, "Prices for State-Contingent Claims: Some Estimates and Applications," *Journal of Business* 51, pp. 653-672.
- Breeden, D. T., and R. H. Litzenberger, 1978, "The Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business*, 51, pp. 621-651.
- Copeland, T., and D. Galai, 1983, "Information Effects on the Bid-Ask Spread," *Journal of Finance*, 38, pp. 1457-1469.
- Debreu, G., 1959, *Theory of Value* (New York: John Wiley & Sons).
- Glosten, Lawrence, 1994, "Is the Electronic Limit Order Book Inevitable?" *Journal of Finance* 49, pp. 1127-1161.
- Hicks, J. R., 1939, *Value and Capital*, (London: Oxford University Press).
- Huberman, G. and W. Stanzl, 2004, "Quasi-arbitrage and Price Manipulation," *Econometrica* 72, pp. 1247-1275.
- Ingersoll, J. E., Jr., 1987, "Digital Contracts: Simple Tools for Pricing Complex Derivatives," *Journal of Business* 73, pp. 67-88.
- Lehmann, B. N., 2005, "Arbitrage-free Limit Order Books and the Pricing of Order Flow Risk," (University of California at San Diego, Unpublished Working Paper).
- Lehmann, B. N. and D. Modest, 1995, "Market Structure and Liquidity on the Tokyo Stock Exchange" in Andrew W. Lo (ed.), *The Industrial Organization and Regulation of Securities Markets* (Chicago: University of Chicago Press).
- Parlour, C., 1998, "Price Dynamics in Limit Order Markets," *Review of Financial Studies* 11, pp. 789-816.
- Sandas, P., 2001, "Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market," *Review of Financial Studies* 14, pp. 705-734.
- Seppi, D., 1997, "Liquidity Provision with Limit Orders and a Strategic Specialist," *Review of Financial Studies* 10, pp. 103-150.
-