Optimal Algorithmic Trading with Limit Orders

Craig W. Holden
Indiana University

May 2012

Abstract

I determine the optimal trading strategy for an institutional trader who wants to purchase a large number of shares over a fixed time horizon. First, I consider the case when limit orders can be used as well as market orders. I develop a simple binomial model where limit orders either execute or not. I find that the optimal sequence of limit orders involves small changes in price aggressiveness from node-to-node over a binomial tree. That is, if a given limit order executes (or not), then the next limit order optimally has a slightly less (more) aggressive price. I find that this trading strategy beats the benchmark trading strategy from the existing literature nearly all of the time and ties it in one special case. Second, I consider the case where trading algorithms can depend on a rich set of state variables. I develop a rich simulation model of trader who tries to satisfy the trading request of a fund manager. I model a pure limit order book exchange and allow the trader to select from a wide range of trading algorithms. I calibrate the simulation to real-world summary statistics based on order data. I find that if the fund manager is opportunistic, then the optimal algorithm involves only limit orders with low price aggressiveness. Conversely if the fund manager is committed, then limit orders should be followed by market orders at the end. I find that if the fund manager is informed and not using effective spread to measure the cost of trading, then market orders should be front-loaded in time. Conversely, if effective spread is used or if the fund manager is uninformed, then less aggressive orders should be spread evenly over time.

Keywords: Limit order, market order, liquidity, order flow.
JEL classification: G14, D44.

Addresses for correspondence: Craig Holden, Kelley School of Business, Indiana University, 1309 E. Tenth St., Bloomington, IN 47405-1701. Email: cholden@indiana.edu

Acknowledgements: I thank Noah Stoffman, Charles Trzcinka, Scott Yonker, and seminar participants at Indiana University. I am solely responsible for any errors.

Consider an institutional trader who wishes to buy 100,000 shares of a given stock over a trading day. What orders should the trader submit? There are many choices. First is the choice of order type: market order versus limit order.[1] A market order that will get the shares requested, but at a higher cost. A limit order will cost less if it executes, but it may not execute. Second, if a limit order is used, then there is a choice of price. A higher limit buy price will be more likely to execute, but a lower limit buy price is a better price. Third is the choice of order size. Should the trader submit a smaller number of big orders or a larger number of small orders? Fourth is the choice of dynamic strategies. If new information arrives, should the unexecuted portion of a limit order be cancelled and resubmitted at an updated price? Since market orders pay the spread, should one strategically wait for moments when the spread is relatively small? And there are many more dynamic strategies that one could imagine.

In a seminal article that began the optimal execution literature, Bertsimas and Lo (1998) consider a risk-neutral institutional trader who wishes to buys $\overline{S}$ shares in T periods. The trader is allowed to submit market orders. Trades have a temporary price impact, meaning that a larger trade causes the price to deviate more from its fundamental value. The trader's objective function is to minimize total trading costs. They show that the "naïve strategy," trading an equal amount per period $\left(\overline{S}/T\right)$, is optimal under the special circumstances that price changes have a zero mean and are normally distributed. More generally they find the analytical solution for the optimal sequence of trades: (1) when price changes are lognormally distributed, (2) when there is an autoregressive state variable for "market conditions," and (3) for the natural generalization to trading portfolios.

[1] A market order is a request to buy or sell a specific quantity of shares at available market price(s). A limit order is a request to buy or sell at a specified price as many shares as possible up to desired quantity.

Gatheral and Schied (2012) provides an excellent review of the optimal execution literature.[2] Almgren and Chriss (2001) consider an institutional trader with a mean-variance objective function whose trades cause both temporary and permanent price impacts. They find that the trader should frontload trading in earlier periods to reduce risk. Harris (1998) is closest to this paper. He considers a trader who can submit a limit or market order at multiple points in time. The trader's goal is to purchase a single unit of the security. Harris finds that it is optimal for a trader to start with a limit order, then if it fails to execute become more aggressive over time, and then if it still fails to execute by the end submit a market order to guarantee execution.[3] The trading strategy of becoming increasingly aggressive in price over time is taken as the benchmark from the existing literature.

I consider two research questions. First, when limit orders can be used as well as market orders, is it possible to beat the benchmark trading strategy from the existing literature? Second, when a trading algorithm can depend on a rich set of state variables, is it possible to beat existing trading algorithms in the literature?

To answer these questions, I develop two models. First, I develop a simple binomial model. In this model trading is done with unit-sized limit orders and market orders. The model is binomial in the sense that orders either execute or don't. Since all orders are unit-sized, there is no case of partial execution. I find the analytic solutions for the optimal trading strategy for both single-period and multiple-period problems. Importantly, I find that the optimal sequence of limit

[2] Important papers in optimal execution literature include Harris (1998), Huberman and Stanzl (2000), Almgren and Chriss (2001), He and Mamaysky (2001), Dubil (2002), Almgren (2003), He and Mamaysky (2005), Obizhaeva and Wang (2005), Ly, Mnif, and Pham (2007), Rogers and Singh (2008), Forsyth (2009), Kharroubi and Pham (2009), Schied and Schoneborn (2009), Rosenthal (2010), Alfonsi, Fruth, and Schied (2010), Guilbaud, Mnif, and Pham (2010), Gatheral and Schied (2010), Ankirchner and Kruse (2011a), and Ankirchner and Kruse (2011b).

[3] Harris and Hasbrouck (1996) discuss the idea that a patient liquidity trader might begin by submitting a limit order, and if need be, switch to a market order at the end. Handa and Schwartz (1996) analyze a limit order first and market order last strategy compared to a market order first strategy.

orders involves small changes in price aggressiveness[4] from node-to-node over a binomial tree. That is, if a given limit order executes (or doesn't), then the next limit order optimally has a slightly less (more) aggressive price. This trading strategy beats the benchmark trading strategy nearly all of the time and ties it in one special case.

Secondly, I develop a rich simulation model of a trader who tries to satisfy the trading request of a fund manager. I model of a pure limit order book exchange and allow the trader to select from a wide range of trading algorithms. The model is wide-open. Orders of any type and size can be submitted. Prices are on a discrete penny grid. Unexecuted limit orders can be cancelled at any time. Trading algorithms can depend on a wide variety of state variables such as bid, ask, midpoint, bid depth, ask depth, order arrival from other traders, order type from other trades, order size from other traders, own quantity executed so far, number of periods remaining, lagged values of everything, etc. A given trading algorithm could potentially depend on dozens of state variables.[5] I calibrate the simulation to real-world summary statistics for order arrival, order size, and intraday price volatility. For each trading problem, I test a wide variety of trading algorithms with a large number of iterations for each algorithm. I determine the optimal trading algorithm for a wide variety of trading problems.

I find that if the fund manager is opportunistic, then the optimal algorithm should use only limit orders with low price aggressiveness, because these are the only trades that will immediately be profitable. Conversely, if the fund manager is committed, then the optimal algorithm uses limit orders followed by market orders at the end, because the market orders will guarantee purchasing the request amount. I find that if the fund manager is informed and performance is not measured using effective spread, then the optimal algorithm uses market

[4] For limit buy (sell) orders, a higher (lower) price is more aggressive.
[5] For comparison, Bertsimas and Lo (1998) obtain an analytic solution based on two state variables.

orders that are front-loaded in time so as to trade before prices move in the predicted direction. Conversely, if performance is measured using effective spread or if the fund manager is uninformed, then the optimal algorithm uses less aggressive orders spread evenly over time. The reason is that under effective spread the benchmark is the contemporaneous quote midpoint, which moves up or down with the price, so there is no penalty to trading later in the day. Also, if the fund manager is uninformed, then there isn't any predicted direction of prices to avoid.

This paper is also related to two other literatures. First, it is related to the empirical literature on how institutional traders implement trading requests.[6] Second, it is related to the theoretical literature on limit order book exchanges in equilibrium.[7]

The paper is organized as follows. Section 1 develops a simple binomial model of trading using limit orders and market orders, solves it analytically, and determines if the optimal trading strategy beat existing trading strategies from the academic or practitioner literature. Section 2 develops a rich simulation of model of trading on a pure limit order book exchange, calibrates it to real-world summary statistics based on order data, and numerically determines if it is possible to beat existing trading algorithms from the academic or practitioner literature. Section 3 concludes. The appendix contains all proofs.

1. A Simple Binominal Model

1.1 The Single-Period Version

I develop a simple binomial model of a pure, open limit order market. This type of market structure is used by NYSE-ARCA, BATS, Direct Edge, the Tokyo Stock Exchange, Euronext, and many other exchanges around the world. This simple model provides insight and

---

[6] See Irvine, Lipson, and Puckett (2007), Conrad, Johnson, and Wahal (2003), Domowitz, Glen, and Madhavan (2001), Keim and Madhavan (1995, 1996, 1997), Handa and Schwartz (1996), Harris and Hasbrouck (1996), Angel (1995), and Lakonishok, Schleifer, and Vishny (1992).

[7] See Rosu (2009), Goettler, Parlour, and Rajan (2005, 2009), Parlour and Seppi (2008), Foucault, Kadan, and Kandel (2005), Foucault (1999), Chakravarty and Holden (1995), and Glosten (1994).

intuition into optimal trading strategy that carries over to the rich simulation model in the next section of the paper.

I begin with a single-period version. The initial limit order book for a given stock is empty. A single seller wishes to trade with a single buyer. Both traders need to set a limit order price. By convention, the seller is exogenous. The seller submits a limit sell order at a limit sell price s for one unit of the asset. The limit sell price s is drawn from an uniform distribution over the continuous interval $[v-d, v+d]$, where v is the public value of the stock and d specifies the dispersion of potential prices. The buyer is endogenous. The buyer submits a limit buy order at a limit buy price b for one unit of the asset. He chooses the limit buy price b from the same the continuous interval $[v-d, v+d]$. Define limit buy price aggressiveness a, as the excess of the limit price above the public value, $a = b - v$.

Both orders are submitted simultaneously. Both orders execute in full when $s \le b$ and both fail to execute when $s > b$. It immediately follows that the limit buy's probability of execution p is the probability that $s \le b$, which is

$$p = \frac{b-(v-d)}{(v+d)-(v-d)} = \frac{b-v+d}{2d}. \quad (1)$$

This way of modeling a limit buy captures the key property that a higher buy price b increases the probability of execution p. At the extreme, a buy price of $v+d$ yields a 100% probability of execution. I interpret this case as a market buy order.[8] For simplicity, I adopt the convention that the two orders execute at the buy price b.[9]

[8] Equivalently, this can be interpreted as a marketable limit buy. The simple binomial model does not distinguish between a market buy and a marketable limit buy.
[9] It is easy to show that the results of the model are qualitatively similar for the alternative convention that they execute half the time at the buy price and half the time at the sell price.

The buyer wishes to minimize his expected disutility $E[D]$ at the end of the period. If the limit buy executes, then he receives disutility from the cost of trading $b-v$, the excess price paid relative to the public value v.[10] If the limit buy fails to execute, then he receives a disutility penalty k, which represents the buyer's degree of unhappiness for failing to obtain the desired amount. The buyer's decision problem is to choose the limit buy price b so as to minimize his expected disutility

$$\underset{b}{Min}\, E[D] = p(b-v) + (1-p)k \,. \tag{2}$$

A higher b increases p which reduces the chance of receiving the disutility penalty k for failing to execute. The tradeoff is that a higher b increases the cost of trading $b-v$ when execution happens. The proposition below gives the analytic, single-period solution.

Proposition 1. If $k < 3d$, then the optimal limit price aggressiveness is

$$a = \frac{1}{2}(k-d) < d \,, \tag{3}$$

the corresponding probability of execution is

$$p = \frac{k+d}{4d} < 1 \,, \tag{4}$$

and the corresponding expected disutility is

$$E[D] = \frac{6dk - d^2 - k^2}{8d} < d \,. \tag{5}$$

If $k \geq 3d$, then the optimal limit price aggressiveness is $a = d$, the corresponding probability of execution is $p = 1$, and the corresponding expected disutility is $E[D] = d$.

[10] This measure of the cost of trading is very similar to the effective half spread, which for buy trades is defined as the trade price minus the quote midpoint. In this simple model, the cost of trading is equal to price aggressiveness, $b-v=a$. In the rich simulation model in the next section, we will measure the cost of trading with four different metrics.

The key driver of the optimal trading strategy is the failure penalty k. If the failure penalty is low enough $(k > 3d)$, then the optimal order submission is a limit buy with an interior optimum price aggressiveness $a < d$. The price aggressiveness is increasing in the failure penalty *k* and decreasing in the dispersion *d*. Conversely, if the failure penalty is high enough $(k \geq 3d)$, then the optimal order submission is a market buy at the corner solution price aggressiveness $a = d$.

1.2 The Multi-Period Version

Now I extend the simple binomial model to a multi-period version. There are T order submission times $t = 1, 2, \ldots, T$ and a terminal time $T + 1$.[11] Let v_t be the public value of the stock at time *t*. This public value is conditional on all public information at time *t*, including the history of all prior trades and quotes in the stock. Let next time's public value be given by $v_{t+1} = v_t + \varepsilon_{t+1}$, where ε_{t+1} is next time's innovation the pubic value and is common knowledge.

At any node on the tree, a single exogenous seller wishes to trade immediately. At time t, the seller submits a limit sell at a sell price s_t for one unit that expires after one time period. The limit sell price s_t is drawn from an uniform distribution over the continuous interval $[v_t - d, v_t + d]$. Again, the limit buy price aggressiveness $a_{t,n}$ is defined as the excess of the limit price above the public value, $a_{t,n} = b_{t,n} - v_t$.

A single endogenous long-lived buyer wishes to purchase N units by the terminal time $T + 1$. Since there is only one unit available for sale at a given time, this will require the execution of N limit or market buy orders of one unit each. We constrain $N \leq T$ so that it is

[11] In principle, *T* could be arbitrarily large. There could be so many periods that one period could represent one second or even one millisecond.

feasible in some state of nature to purchase the desired quantity. We also impose the economically sensible non-negativity constraints $d > 0,\ N > 0,$ and $v_t - d > 0$ at all times.

Consider a buyer who at time t has already succeeded in purchasing n units. At the (t, n) node of a binomial tree, the buyer cancels any unexecuted limit buy from the previous time[12] and submits a new limit buy order at a buy price $b_{t,n}$ for one unit. He chooses $b_{t,n}$ from the same interval $[v_t - d, v_t + d]$ as the contemporaneous seller.

Both orders are submitted simultaneously. Both orders execute in full when $s_t \le b_{t,n}$ and both fail to execute otherwise. The limit buy's probability of execution $p_{t,n}$ is the probability that $s_t \le b_{t,n}$, which is given by

$$p_{t,n} = \frac{b_{t,n} - (v_t - d)}{(v_t + d) - (v \ - d)} = \frac{b_{t,n} - v_t + d}{2d}\,. \tag{6}$$

The long-lived buyer's disutility function at the terminal $(T+1, n)$ node is

$$D_{T+1,n} = TCT_{T+1} + k(N - n) \qquad \text{for } n = 0, 1, \ldots, N\,, \tag{7}$$

where TCT_{T+1} is the total cost of trading prior to time $T+1$ (which includes the final trading time T) and the second term is the trader's degree of unhappiness due to the quantity shortfall $N - n$.

Let $J_{t,n}$ be the long-lived buyer's derived disutility function at the (t, n) node. By dynamic programming, the decision problem at the (t, n) node is given recursively by

[12] In the simple binomial model, canceling any unexecuted limit buy always leaves the buyer either better off or indifferent compared to having two limit buys outstanding at the same time. To see this, recall that only one unit is available for sale in any period and thus, it is not possible for two buy orders to execute in the same period. Further, if the prior buy price is higher than the current period optimal buy price, then the prior buy is too high and the buyer is better off by canceling it. If the prior buy price is lower than the current optimal buy price, then the prior buy price is dominated and the buyer is indifferent to its existence. So in the latter case, we adopt the tie-breaking convention that he cancels it.

$$J_{t,n} = \underset{b_{t,n}}{Min}\, E_{t,n}\left[J_{t+1,\tilde{n}} \right]. \tag{8}$$

where $\tilde{n}$ is the random number of units purchased through the next time $t+1$ and $J_{T+1,n} = D_{T+1,n}$ for all values of *n*. At any given (t,n) node, if $n < N$, then an order will be submitted that will either execute or not. However, if $n = N$, then no order will be submitted since the full desired quantity has already been purchased. Thus, derived disutility $J_{t,n}$ at the (t,n) node can be written as

$$J_{t,n} = \underset{b_{t,n}}{Min}\, E_{t,n}\left[J_{t+1,\tilde{n}} \right] = \begin{cases} \underset{b_{t,n}}{Min}\ p_{t,n} J_{t+1,n+1} + \left(1 - p_{t,n}\right) J_{t+1,n} & \text{if } n < N \\ TCT_t & \text{if } n = N. \end{cases} \tag{9}$$

In the case that $n = N$, the absence of further orders means that the current derived utility equals the terminal disutility, which also equals the total cost of trading that has already been realized prior to the current date TCT_t and no penalty due to quantity shortfall.

In the case that $n < N$, we can write the derived disutility as the expected value of terminal disutility under the optimal trading strategy as determined recursively backwards by dynamic programming. In the event that the time *t* order executes, then the updated expected value of terminal disutility is

$$J_{t+1,n+1} = TCT_t + b_{t,n} - v_t + f_{t+1,n+1} + k\left(N - (n+1) - h_{t+1,n+1}\right), \tag{10}$$

where TCT_t is the total cost of trading prior to time *t*, $b_{t,n} - v_t$ is the increment to the total cost of trading caused by the time *t* order executing, $f_{t+1,n+1}$ is defined as the expected future addition to the total cost of trading starting from the $(t+1, n+1)$ node, and $h_{t+1,n+1}$ is defined as the expected number of units to be purchased in future trades starting from the $(t+1, n+1)$ node. Both $f_{t+1,n+1}$ and $h_{t+1,n+1}$ are the forecasted outcome of future trades under the optimal trading strategy.

In the event that the time t order doesn't execute, then the updated expected value of terminal disutility is

$$J_{t+1,n} = TCT_t + f_{t+1,n} + k\left(N - n - h_{t+1,n}\right), \tag{11}$$

where $f_{t+1,n}$ and $h_{t+1,n}$ are defined analogously starting from the $(t+1,n)$ node.

Again there is a trade-off. A higher buy price $b_{t,n}$ increases the probability of execution $p_{t,n}$ which reduces the chance of receiving the disutility penalty for a quantity shortfall, but it increases the increment to the total cost of trading $b_{t,n} - v_t$ when execution does happen. The proposition below gives the analytic, multi-period solution.

Proposition 2. The solution is specified by binomial trees for a, p, f, and *h. At the* (t,n) *nodes of these trees where the trader is not done* $(n < N)$*, the optimal limit price aggressiveness is*

$$a_{t,n} = Min\left[d, \frac{1}{2}\left(f_{t+1,n} - f_{t+1,n+1} + k\left(1 + h_{t+1,n+1} - h_{t+1,n}\right) - d\right)\right], \tag{12}$$

the probability of execution at the optimal limit buy price is

$$p_{t,n} = \frac{a_{t,n} + d}{2d}, \tag{13}$$

the expected total cost of trading on future trades is

$$f_{t,n} = p_{t,n}\left(b_{t,n} - v_t + f_{t+1,n+1}\right) + \left(1 - p_{t,n}\right) f_{t+1,n}, \tag{14}$$

and the expected number of units to be purchased in future trades is

$$h_{t,n} = p_{t,n}\left(1 + h_{t+1,n+1}\right) + \left(1 - p_{t,n}\right) h_{t+1,n}. \tag{15}$$

At the (t,n) *nodes of these trees where the trader is done* $(n \geq N)$*, no further orders will be submitted and so* $f_{t,n} = 0$, $h_{t,n} = 0$, $p_{t,n} = 0$, *and* $a_{t,n}$ *is undefined. For all terminal nodes,*

$f_{T+1,n} = 0,$ and $h_{T+1,n} = 0.$ *The* a *and* p *trees do not include the terminal date* $T+1$. *Further, a binomial tree for the derived disutility minus the total cost of trading prior to time t* $\left(J_{t,n} - TCT_t\right)$ *is given by equations (9), (10), and (11). On the initial date, the total cost of trading before the first trade* $TCT_1 = 0$ *and so the first node (1,0) of this binomial tree yields the ex-ante derived disutility* $J_{1,0}$.

The binomial trees are calculated backwards from the last date to the first. Start at the date $T+1$ nodes for the f and h binomial trees (which are equal to zero), next calculate the date T nodes for the a and p binomial trees using (12) and (13), then calculate the date T nodes for the f and h trees using (14) and (15), …, keep calculating the a and p trees first and then the f and h trees second on each date all the way back to date 1. Finally, the $J - TCT$ binomial tree can be calculated from the a, p, f, and h trees using (9), (10), and (11) starting at date T and working back to date 1.

1.3 The Character of the Optimal Trading Strategy

A numerical example will help illustrate the character of the optimal trading strategy. Suppose that the long-live buyer wishes to purchase 4 units $\left(N = 4\right)$ has 8 time periods $\left(T = 8\right)$ to do so. Further suppose that the remaining parameters are $v_0 = \$30.00,$ $d = \$0.10,$ and $k = 0.32.$ In this example, the failure penalty is high $\left(k > 3d\right)$.

Figure 1 shows a binomial tree for the optimal limit price aggressiveness a and a binomial tree for the probability of execution p. Starting at the (1,0) node of the upper tree, the optimal limit price aggressiveness on date 1 is $a_{1,0} = \$0.00,$ meaning the date 1 limit buy price is set equal to the date 1 true value $b_{1,0} = v_1$. Looking at the corresponding (1,0) node of the lower tree, this order has a probability of execution $p_{1,0} = 50\%$. If this order *fails* to execute, then you go to the *up*-node (2,0) in both trees. At this node, you cancel the time 1 limit order and submit an optimal time 2 limit order with a slightly more aggressive price $a_{2,0} = \$0.01$ (meaning a limit buy price $b_{2,0} = v_2 + .01)$, which corresponds to a slightly higher probability of execution $p_{2,0} = 56\%$. However, if the time 1 order *succeeds* in executing, then you go to the *down*-node (2,1) in both trees. At this node, you submit an optimal time 2 limit order with a slightly less aggressive price $a_{2,1} = -\$0.01$ (meaning a limit buy price $b_{2,1} = v_2 - .01$), which corresponds to a slightly lower probability of execution $p_{2,1} = 44\%$.

I call this pattern of slightly higher price aggression on up-steps and slightly lower price aggression on down-steps "step-by-step aggressiveness." That is, from any node if the current limit order fails to execute, then the next step at the up-node will optimally be at a slightly more aggressive price with a slightly higher probability of execution. Intuitively, the failure to execute makes the remaining problem *more* difficult (you have one less period to purchase the same number of units as before) and so you act *more* aggressively. Conversely, if the current limit order succeeds in executing, then the next step at the down-node will optimally be at a slightly less aggressive price with a slightly lower probability of execution. Intuitively, the success at execution makes the remaining problem *less* difficult (you have one less unit to purchase in the remaining time) and so you act *less* aggressively.

Step-by-step aggressiveness characterizes nearly all of the 2^T possible paths that you might take over the binomial tree. For example, suppose your first four steps happen to be up, down, down, and up. Then, your price aggressiveness and probability of execution will slightly increase, decrease, decrease, and increase.

More specifically, step-by-step aggressiveness characterizes any path over the entire diamond-shaped zone with the yellow shading (light grey shading with solid borders in black and white) on the binomial trees. I call this area the Limit Order zone, because it turns out that it is always optimal to submit non-marketable limit orders $\left(p_{t,n} < 1\right)$ in this zone (see Proposition 4 below). Specifically, the Limit Order Zone contains all $\left(t,n\right)$ nodes where the trader is not done $\left(n < N\right)$ and where there is a positive amount of slack time (i.e., more time periods left than remaining units to be purchased $T+1-t > N-n$). The following proposition formally specifies this characterization.

Proposition 3. In the Limit Order Zone, the optimal strategy exhibits step-by-step aggressiveness. That is, from any $\left(t,n\right)$ *node in this zone:*

- *If the order* <u>*fails*</u> *to execute, then you go to the up-node* $\left(t+1,n\right)$ *where the optimal strategy is to cancel the unexecuted order and submit a new limit buy at a more aggressive price* $a_{t+1,n} > a_{t,n}$ *with a higher probability of execution* $p_{t+1,n} > p_{t,n}$.
- *If the order* <u>*succeeds*</u> *in executing, then you go to the down-node* $\left(t+1,n+1\right)$ *where the optimal strategy is to submit a limit buy at a less aggressive price* $a_{t+1,n+1} < a_{t,n}$ *with a lower probability of execution* $p_{t+1,n+1} < p_{t,n}$.

To get a feel for step-by-step aggressiveness, take a look at the lower binomial tree for the probability of execution in Figure 1. You can readily see the probability of execution at any node in the Limit Order Zone is an intermediate value between the following two connected nodes. Formally, one can make a stronger statement. Part of the proof of Proposition 3 is to prove that the probability of execution at any node where the trader is not done $\left(n < N\right)$ is a weighted average of the following two connected nodes

$$p_{t,n} = w_{t,n}\, p_{t+1,n+1} + \left(1 - w_{t,n}\right) p_{t+1,n}\,, \tag{16}$$

where the weight $w_{t,n} = \left(p_{t+1,n+1} + p_{t+1,n}\right)/2$. Logically, if the $\left(t,n\right)$ node is in the middle, then one of the following connected nodes will have a higher probability of execution and the other will have a lower probability of execution.

An interesting case is when there have been four failures to execute in a row (the up-up-up-up-node). This is the (5,0) node, where there are only four time periods left and there remains four units to be purchased. In other words, there is zero slack time. The only way to guarantee execution is with the “corner solution" strategy of submitting four market buys over the next four time periods. The optimal limit price aggression for these four orders is the maximum possible $a_{5,0} = a_{6,1} = a_{7,2} = a_{8,3} = d = \$0.10$ (meaning the most aggressive possible sequence of bid prices $b_{5,0} = v_5 + d$, $b_{6,1} = v_6 + d$, $b_{7,2} = v_7 + d$, and $b_{8,3} = v_8 + d$). These market buys will execute with certainty $p_{5,0} = p_{6,1} = p_{7,2} = p_{8,3} = 100\%$. I call the upper right triangle with green shading (medium grey shading in black and white) of the binominal trees the “Market Order / Maximum Aggressiveness Zone.” The name describes the optimal trading strategy in this zone (see Proposition 4 below).[13]

Another interesting case is when there have been four successes in executing in a row (the down-down-down-down-node). This is the (5,4) node, where the entire desired quantity has already been purchased. Hence, no more orders are submitted and the probability that an order

---

[13] When the failure penalty is high $\left(k > 3d\right)$, the nodes *above* the SouthWest edge of the Market Order / Maximum Aggressiveness Zone are unreachable. To see this, consider the (5,0) node which is *on* the SouthWest edge. Here it is optimal to submit a market buy that is certain to execute $p_{5,0} = 100\%$. Thus, it is certain that the trader will go to the down-node (6,1) next and cannot reach the up-node (6,0).

will be executed is zero $p_{5,4} = p_{6,4} = p_{7,4} = p_{8,4} = 0\%.$ I call the lower right triangle with red shading (dark grey shading in black and white) of the binominal trees the "Done" zone.[14]

Now consider an example where the failure penalty is low $(k < 3d)$. Reduce the failure penalty to $k = 0.28$ and keep the rest of the parameters the same: $N = 4,$ $T = 8,$ $v_0 = \$30.00,$ and $d = \$0.10.$ Figure 2 shows the binomial trees in this case. The main difference in this case is in the Market Order / Maximum Aggressiveness Zone. Here the optimal limit price aggression is less than the maximum possible $a_{t,n} < d$ (meaning the bid prices are less than the maximum possible $b_{t,n} < v_t + d$). These non-marketable limit orders have a probability of execution less than 100%, and specifically in this case, $p_{t,n} = 95\%.$ This limit order exhibits the Maximum Aggressiveness that is acceptable to this trader and it is greater than the aggressiveness of any order in the Limit Order Zone.[15]

Indeed, it is interesting that all of the nodes in this zone have a probability of execution which is *identical* to probability of execution for the same parameter values in single-period case. This makes sense when you consider that all of the nodes on the last trading date, (8,0), (8,1), (8,2), and (8,3), in fact only have one trading opportunity left and so logically the probability of execution $(p_{8,0} = p_{8,1} = p_{8,2} = p_{8,3} = 95\%)$ must be the identical to the single-period solution. Further, the remainder of nodes in this zone have two following connected nodes

[14] The nodes *below* the NorthWest edge of the Done Zone are unreachable. To see this, consider the (5,4) node which is *on* the NorthWest edge. Here no more orders are submitted and so the probability that an order will be executed is zero $p_{5,4} = 0\%.$ Thus, it is certain that the trader will go to the up-node (6,4) next and cannot reach the down-node (6,5).

[15] When the failure penalty is low $(k < 3d),$ the optimal limit orders in the Market Order / Maximum Aggressiveness Zone have a probability of execution of less than 100%. Hence, all of the nodes can be reached.

with a 95% probability of execution. Since all nodes are a weighted average of the following two connected nodes by equation (16), they must have the same 95% probability of execution.[16]

The proposition below specifies that the character of the optimal solutions that we have discussed for these particular numerical examples holds in general.

Proposition 4. The binomial trees for the optimal limit price aggressiveness and the probability of execution have three zones:

1. *A Limit Order Zone, where it is optimal to submit a non-marketable limit order.*
2. *A Market Order / Maximum Aggressiveness Zone. In this zone, if the failure penalty is high enough* $(k \geq 3d)$*, then it is optimal to submit a market order. Conversely, if the failure penalty is low enough* $(k < 3d)$*, then it is optimal to submit a non-marketable limit order with higher price aggressiveness than any node in the limit order zone.*
3. *A Done Zone where the trader is done* $(n \geq N)$ *and so no further orders are submitted.*

Under step-by-step aggressiveness, if the (t,n) node order fails to execute, then the remaining trading problem is more difficult and thus requires a more aggressive strategy. Pursuing this reasoning, multiple failures to execute cause the trader to cross-over into a *more aggressive zone*, namely, the Market Order / Maximum Aggressiveness Zone. Intuitively, all of the slack time is gone and so the trader must act very aggressively. Conversely, N successes in executing cause in the trader to cross-over into a *less aggressive zone*, namely, the Done Zone. Here the trader adopts the least aggressive strategy possible: stopping trading.

[16] The situation changes when we get to the limit order zone, because at least one of the following connected nodes has a lower probability of execution. For example, the (7,4) node in the limit order zone is connected to the (8,4) node with a 95% probability and to the (8,5) node with a 0% probability. The weighted average of these two following connected nodes yields a 50% probability for the (7,4) node.

Recall that the benchmark from the existing literature is becoming increasingly aggressive in price over time. Since this strategy is different that step-by-step aggressiveness strategy that Proposition 3 proves is optimal strategy, it must yield strictly higher (worse) ex-ante expected disutility in nearly all cases. The one exception is the one that Harris (1998) considers, which can be viewed as a special case of the binomial model when only a single order is requested $\left(N = 1\right)$ and the failure penalty is high $\left(k \geq 3d\right)$. Figure 3 shows the binomial trees in this case. Here the optimal strategy begins by submitting a limit order at a low -\$0.06 aggressiveness with an 18% chance of execution. If it executes, you are done. If not, the next limit order will be more aggressive with a 20% chance of execution. The pattern continues at each node going along the upper edge of the binomial tree. From each node, you will be done if your order executes or you will submit a more aggressive next time if it fails. In the last period, you submit a market order. In this special case, the optimal strategy *exactly matches* the benchmark from the existing literature, which is to become increasingly aggressive in price over time. Thus, it ties in ex-ante expected disutility in this one case.

2. A Rich Simulation Model

2.1 Model Setup

Many institutional traders separate the “selection” task and the “implementation” task. For each fund (e.g., mutual fund, pension fund, endowment fund, etc.), there is a fund manager who decides what securities to buy and sell in what overall amounts at what times, which is called selection. The fund manager’s trading requests are sent to an “order desk” employee. This employee decides what order types at what order prices in what order sizes should be sent to what exchanges or broker-dealers at what times, which is called implementation.

I develop a rich simulation model from the point of view of an order desk employee. This employee takes the fund manager's request as given and tries to figure out the best way to implement the request so as to minimize the fund manager's unhappiness.

The fund manager's request is a request to buy specific quantity of a particular security expressed as a percentage of average daily volume.[17] For example, a request to purchase 1% of average daily volume would be an easy request, whereas a request to purchase 30% of average daily volume would be a difficult request. Multiplying (Requested Buy as a percentage of Average Daily Volume) times (Average Daily Volume measured in round lots) yields the requested number of round lots N.[18]

There are two types of agents. One agent is the endogenous, long-lived order desk employee. The second class is exogenous, other traders who can submit a variety of orders.[19]

The trading day is divided into 20 order submission times $t = 1, 2, \ldots, 20$ that are 20 minutes apart. Let T be the fund manager's order submission deadline to fulfill the purchase request and $T+1$ be a terminal valuation time after the deadline. For example, a fund manager's deadline of $T = 1$ (corresponding to 9:40 a.m.) would be very impatient, whereas a fund manager's deadline of $T = 20$ (corresponding to 4:00 p.m.) would be very patient.

The two classes of agents alternate submitting orders. At time $t = 1$, other traders submit an order and then the order desk employee submits an order. At time $t = 2$, other traders submit an order and then the order desk employee submits an order. And so on until the deadline time

---

[17] I only consider buy requests. Sell requests are incorporated implicitly as the mirror image of buy requests. I only consider requests to trade in a single security. Bertsimas and Lo (1998) show that portfolio requests are readily accommodated by accounting for the multivariate stochastic process across securities.

[18] One round lot is equal to 100 shares.

[19] Like the rest of the optimal execution literature, both of my models are based on individual optimality (that is, partial equilibrium). Compared to general equilibrium, the strength of this approach is its enormous generality. By calibrating the actions of other agents using real data, it does not impose strong assumptions about all agents in the economy, such as rationality, full understanding of distributions, sophisticated computational ability, etc. The inherent weakness of this approach is that it cannot address the interaction of multiple strategic agents.

$t = T$, when other traders submit their final order by the deadline and then the order desk employee submits a final order by the deadline.

The fund manager's disutility function D_{T+1} at the terminal valuation time $T+1$ is

$$D_{T+1} = TCT^{m}_{T+1} + k\left|N - n_{T+1}\right| + A\sigma, \tag{17}$$

where TCT^{m}_{T+1} is the total cost of trading at the terminal valuation time based on metric *m* (explained below), n_{T+1} is the realized number of round lots purchased, A is the fund manager's risk aversion, and σ is the standard deviation of trade prices. The second term is the fund manager's degree of unhappiness due to a quantity shortfall (or overrun) $\left|N - n_{T+1}\right|$ and the third term captures the fund manager's disutility due to execution price risk.

I incorporate four alternative metrics that are widely-used in practice for measuring the cost of trading C^{m}_{k} due to the k^{th} trade under metric *m*:

$$\textit{Effective Spread}: \quad C^{ES}_{k} = 2I_k\left(P_k - M_k\right), \tag{18}$$

$$\textit{Implementation Shortfall}: \quad C^{IS}_{k} = 2I_k\left(P_k - M_{IS}\right), \tag{19}$$

$$\textit{Volume-Weighted Average Price}: \quad C^{VWAP}_{k} = 2I_k\left(P_k - P_{VWAP}\right), \tag{20}$$

$$\textit{Closing Price}: \quad C^{C}_{k} = 2I_k\left(P_k - P_C\right), \tag{21}$$

where I_k is an indicator variable that equals +1 if the k^{th} trade is a buy and -1 if the k^{th} trade is a sell, M_k is the quote midpoint immediately before the k^{th} trade, M_{IS} is the quote midpoint at the time of fund manager's request (interpreted as before the start of the trading day), P_{VWAP} is the weighted average trade price over the trading day as weighted by the volume of each trade, and P_C is the last trade price of the trading day. The total cost of trading over all trades $k = 1, 2, \ldots, K$ that take place before the deadline is

$$TCT_{T+1}^{m} = \sum_{k=1}^{K} C_k^m + I_{IS}\left(M_C - M_{IS}\right)\left(N - n_{T+1}\right), \tag{22}$$

where I_{IS} is an indicator variable that equals +1 under Implementation Shortfall and 0 otherwise and M_C is quote midpoint at the close of the trading day. The second term is an extra opportunity cost component that is part of Implementation Shortfall (but not the other metrics), which measures the foregone profits (or losses) on the quantity shortfall (or overrun) $\left(N - n_{T+1}\right)$.

I incorporate the possibility that the fund manager may be informed. Let s be the fund manager's private signal prior to the trading day about the terminal value of the security and let $s - v_1$ be the difference between that signal and the time 1 public value of the security. Let $v_{T+1} - v_1$ be the cumulative innovation in the pubic value of the security from time 1 to time T+1. Define the fund manager's information ρ as the correlation between signal difference $s - v_1$ and the cumulative innovation $v_{T+1} - v_1$. For example, a fund manager's information $\rho = 0$ would be uninformed, whereas a fund manager's information $\rho = 1$ would be perfectly informed.

When the fund manager is unformed, the public value innovations are equally likely to be positive or negative. Conversely, if the fund manager is truly informed and choose to request buying, then it must been because the fund manager's signal was of good news (i.e., the signal difference is positive $s - v_1 > 0$). Under perfect information $\left(\rho = 1\right)$ the fund manager's good news signal must have been correct, and so in this case, the public value innovations are strictly positive.[20]

[20] A fund manager who is truly informed and gets a signal of bad news which leads to a sell request is incorporated implicitly as the mirror image of a signal of good news leading to a buy request. In the mirror image case under perfect information $\left(\rho = 1\right)$ the fund manager's signal must have been correct and so public value innovations would be strictly negative.

Like the simple binomial model, the rich simulation is a model of a pure, open limit order market. Unlike the previous model, the initial limit order book is *not* empty.

The order desk employee has a rich set of orders to choose from. Each period either a market order or a limit order may be submitted. A limit order may have *any* limit price as selected from a discrete price grid with penny increments. A market or limit order may be for *any* integer amount of round lots. At any time you may simultaneously cancel the unexecuted limit orders and submit an updated market or limit order.

2.2 Model Calibration

In the model, I have calibrated the actions of other traders to real-world summary statistics based on order data. On Nov 15, 2000, the Securities and Exchange Commission (SEC) adopted Rule 605[21] that mandated the disclosure monthly summary statistics by each exchange (or other market center) by each individual stock based on order data. I selected the Better Alternative Trading System (BATS) exchange to calibrate to because its market structure is a pure, open limit order book just like my model. As a starting point, I selected Microsoft stock to calibrate to. I intend to investigate a variety of other stocks as well. I selected data for the month of December 2011.

Table I shows the calibrated inputs to the simulation. Panel A shows the probability of various order types. The Rule 605 data shows that market orders and marketable limit orders represent 61.5% of all Microsoft orders on BATS. So I assigned 30.75% to market buys and 30.75% market sells. Conversely, non-marketable limit orders represent 38.5% of all Microsoft orders on BATS. So I assigned 19.25% for (non-marketable) limit buys and 19.25% for (non-marketable) limit sells.

---

[21] It was formerly named Rule 11Ac1-5.

Panel B shows the probability of order size by order type. The Rule 605 data reports the total shares for market orders and marketable limit orders. I divided the monthly figure by 21 trading days in December 2011 and by 20 intervals per day to obtain 230.4 round lots per 20-minute period. I assigned probabilities to various market order sizes to obtain an average size of 230.4 round lots. Similarly, the total share for non-marketable limit orders were divided by 21 and then by 20 to obtain 276.7 round lots per 20-minute period. I assigned probabilities to various (non-marketable) limit order sizes to obtain an average size of 276.7 round lots. Implicitly, the average calibrated order in the simulation represents the *aggregate* quantity over a 20-minute interval.

Panel C shows the probability of price changes. I selected New York Stock Exchange (NYSE) Trade and Quote (TAQ) data for Microsoft in December 2011 and sampled trade prices every 20 minutes. To determine the intraday volatility, I computing the mean absolute deviation of the 20-minute prices as 0.033. I assigned probabilities to public value innovations such that the mean absolute deviation was 0.033. Finally, out of all Microsoft non-marketable limit orders on BATS, 0.0% were inside-the-quote, 87.4% were at-the-quote, and 12.6% were behind-the-quote. To approximately capture this distribution, I assigned a probability of non-negative limit price deviation to be 0.0%, the probability of -$0.01 limit price deviation to be 87.4%, and the probability of a -$0.02 and -$0.03 limit price deviation to add up to 12.6%.

2.3 Model Testing

The rich simulation model provides a wide variety of state variables that may be relevant to trading algorithms. I test a wide range of trading algorithms. Table II describes the 46 algorithmic trading strategies tested. They fall into groups as follows:

- **Strategies 1-3, 7-9, 31-33, 37-39.** Limit buys are used at varying limit prices relative to the bid (bid-$0.01, bid, bid+$0.01) and with varying cancellation (none, every 12 periods, every 4 periods)
- **Strategies 4-6.** Market buys are used in varying sizes (few large, many small, lots of tiny)
- **Strategies 10-12, 16-18.** Market buys in fixed amounts or tied to the ask depth are only submitted when the spread is less than various cutoffs ($0.08, $0.06, $0.04) and no conditions are imposed in the last 5 periods
- **Strategies 13-15, 34, 40.** Medium limit buys are used with varying time-patterns (front-loaded, evenly spread, back-loaded) and varying limit prices relative to the bid (bid-$0.01, bid, bid+$0.01)
- **Strategies 19-30.** Limit orders are submitted with step-by-step varying price aggressiveness relative to the bid or relative to the midpoint ("absolute price aggressiveness") with size variations (1/3, 1/4, 1/5 of requested lots) and zone variations (normal or large); all end with a market buy
- **Strategies 35, 36, 41-46.** Medium limit buys with a more aggressive price relative to the bid or relative to the midpoint ("absolute price aggressiveness") when there are various number of periods left (4, 6, 8); all end with a market buy.

The strategies above were tested on a wide range problems. I varied the requested buy amount as a percentage of daily volume (10%, 50%, 100%), the deadline (2 periods, 10 periods, and 20 periods), the failure penalty (none = 0.00, medium = 0.50, and high = 1.00), fund manager risk aversion (none = 0, medium = 100, high = 1,000), fund manager information (none = 0, medium = 0.50, and high = 1.00), and performance metric (effective spread, implementation shortfall, volume-weighted average price, and closing price). Combining all of the variations, I tested 972 problems.

For each iteration of the simulation, there are four random variables that must be updated for each of the 20 time periods, plus a few more one-time random variables, for a total of 87 random variables. For each problem, I tested each strategy with 100 iterations of the simulation and computed the average disutility over all iterations. I repeated this process for all 46 strategies and identified the strategy with the lowest average disutility. I continued this process for all 972 problems.

2.4 Results

Table III shows the optimal trading algorithm by failure penalty, risk aversion, and performance metric. Panel A shows when the failure penalty is zero, which represents the most purely opportunistic fund manager. Strikingly, the optimal strategy in each case involves limit buys only – no market buys. If there is no penalty for failing to get the requested quantity, then the only trades that you want to do are those that incur a *negative* cost of trading, meaning a profit. In other words, you want every single trade to make a profit.

Panels B and C show when the failure penalty is medium (0.50) and high (1.00), respectively. In every case, the optimal strategy involves market buys. One set of strategies mainly use limit buys, but are followed by market buys at the end. Another set of strategies uses market buys only. In either case, it is worth incurring the extra cost of trading due to market buys in order to avoid the failure penalty for a quantity shortfall. Given the other parameters (requested buy = 10% of daily volume, deadline = 20 periods, and fund manager’s information = 0), the optimal strategy doesn’t vary very much by risk aversion and by performance metric.

Table IV shows the optimal trading algorithm by fund manager’s information, deadline, and performance metric. Panel A shows when the fund manager’s information is zero. When the deadline is very short (2 periods), then the optimal trading algorithm is very aggressive with large market buys. When the deadline is medium or long (10 or 20 periods), then the optimal trading algorithm is mainly limit buys, but with market buys at the end to get the requested amount.

Panels B and C show when the fund manager's information is medium (0.50) and high (1.00), respectively. As before when the deadline is very short, it is best to go with large market buys. However, when the deadline is medium or long and the performance metric is *not* effective spread, then it is also best to submit large market buys. The intuitive explanation for this has three steps: (1) the fund manager has a private signal of good news leading to a buy request, (2) statistically the fund manager is correct leading to a price increase over the trading day on average, and (3) anticipating this, it is best to purchase the security early in the day before the price has risen very much.

An interesting exception to this logic occurs when the performance metric *is* effective spread. In this case, it is best to submit mainly limit buys, but with market buys. The benchmark under the effective spread metric is the contemporaneous midpoint. This has the unique property that the benchmark moves upward if trades are later in the day at higher prices. So under effective spread, there is no penalty for trading later in the day and it is best to spread out the trades over time. By contrast, the benchmark in the other three performance metrics (the request midpoint, the volume-weighted average price, and the closing price) are unaffected by the timing of trades. So under these three metrics, it is best to aggressively front-load the trading when you anticipate that the price rise on average.

3. Conclusion

I determine the optimal trading strategy for an institutional trader who wants to purchase a large number of shares over a fixed time horizon. First, I consider the case when limit orders can be used as well as market orders. I develop a simple binomial model where limit orders either execute or not. I find that the optimal sequence of limit orders involves small changes in price aggressiveness from node-to-node over a binomial tree. That is, if a given limit order executes (or not), then the next limit order optimally has a slightly less (more) aggressive price. I find that this trading strategy beats the benchmark trading strategy from the existing literature nearly all of the time and ties it in one special case. Second, I consider the case where trading algorithms can depend on a rich set of state variables. I develop a rich simulation model of trader who tries to satisfy the trading request of a fund manager. I model a pure limit order book exchange and allow the trader to select from a wide range of trading algorithms. I calibrate the simulation to real-world summary statistics based on order data. I find that if the fund manager is opportunistic, then the optimal algorithm involves only limit orders with low price aggressiveness. Conversely if the fund manager is committed, then limit orders should be followed by market orders at the end. I find that if the fund manager is informed and not using effective spread to measure the cost of trading, then market orders should be front-loaded in time. Conversely, if effective spread is used or if the fund manager is uninformed, then less aggressive orders should be spread evenly over time.

Appendix

Proof of Proposition 1. When $k < 3d$, the optimal limit price aggressiveness is obtained by substituting (1) into (2) and taking the derivative with respect to b. The second order condition is positive and thus the objective function is minimized. The probability of execution is obtained by substituting (3) in place of $b - v$ in (1). The expected disutility is obtained by substituting (3) and (4) into (2). When $k = 3d$, substitute $3d$ in place of k in (3), (4), and (5) to obtain $a = d$,

$p=1,$ and $E[D]=d.$ By definition the price aggressiveness can't go above d, so for any value of $k>3d$ the solution is constrained to the corner solution $a=d,$ $p=1,$ and $E[D]=d.$ *Q.E.D.*

Proof of Proposition 2. The optimal limit price aggressiveness is obtained by substituting (6), (10), and (11) into (9) and taking the derivative with respect to b. The second order condition is positive and thus the objective function is minimized. The probability of execution is obtained by substituting (12) in place of $b_{t,n}-v_t$ in (6). Equation (14) for the expected total cost of trading on future trades at the (t,n) node is the probability weighted average of the expected total cost of trading on future trades at $(t+1,n+1)$ node and the $(t+1,n)$ node, where the incremental cost of trading $b_{t,n}-v_t$ is added when the current order executes. Equation (15) for the expected number of units to be purchased in future trades at the (t,n) node is the probability weighted average of the expected number of units to be purchased in future trades at $(t+1,n+1)$ node and the $(t+1,n)$ node, where the number of units is incremented by one unit when the current order executes. *Q.E.D.*

Combined Proof of Propositions 3 and 4. The first step of the combined proof is to show that the probability of execution at any node where the trader is not done $(n<N)$ is a weighted average of the following two connected nodes. The proof is based on the connection between the *f* and *h* variables at a given node and the *f* and *h* variables at the following connected notes. Substituting (12) into (13) and then twice substituting (14) and (15) into the resulting equation yields

$$p_{t,n} = \left(\frac{1}{4d}\right)\begin{pmatrix} \left[p_{t+1,n}\left(b_{t+1,n} - v_{t+1} + f_{t+2,n+1}\right) + \left(1 - p_{t+1,n}\right) f_{t+2,n} \right] \\ -\left[p_{t+1,n+1}\left(b_{t+1,n+1} - v_{t+1} + f_{t+2,n+2}\right) + \left(1 - p_{t+1,n+1}\right) f_{t+2,n+1} \right] \\ +k\begin{pmatrix} 1 + \left[p_{t+1,n+1}\left(1 + h_{t+2,n+2}\right) + \left(1 - p_{t+1,n+1}\right) h_{t+2,n+1} \right] \\ -\left[p_{t+1,n}\left(1 + h_{t+2,n+1}\right) + \left(1 - p_{t+1,n}\right) h_{t+2,n} \right] \end{pmatrix} + d \end{pmatrix}. \tag{17}$$

Rearranging (17) yields

$$p_{t,n} = \left(\frac{1}{4d}\right)\begin{pmatrix} p_{t+1,n}\left(b_{t+1,n} - v_{t+1} - \left[f_{t+2,n} - f_{t+2,n+1} + k\left(\left(1 + h_{t+2,n+1} - h_{t+2,n}\right)\right)\right]\right) \\ -p_{t+1,n+1}\left(b_{t+1,n+1} - v_{t+1} - \left[f_{t+2,n+1} - f_{t+2,n+2} + k\left(1 + h_{t+2,n+2} - h_{t+2,n+1}\right)\right]\right) \\ +\left[f_{t+2,n} - f_{t+2,n+1} + k\left(\left(1 + h_{t+2,n+1} - h_{t+2,n}\right)\right)\right] + d \end{pmatrix}. \tag{18}$$

Substituting from (12) into (18) yields

$$p_{t,n} = \left(\frac{1}{4d}\right)\begin{pmatrix} p_{t+1,n}\left(a_{t+1,n} - \left[2a_{t+1,n} + d\right]\right) \\ -p_{t+1,n+1}\left(a_{t+1,n+1} - \left[2a_{t+1,n+1} + d\right]\right) \\ +\left[2a_{t+1,n} + d\right] + d \end{pmatrix}. \tag{19}$$

Substituting from (13) into (19) yields

$$p_{t,n} = \left(\frac{1}{4d}\right)\begin{pmatrix} p_{t+1,n}\left(d\left(2p_{t+1,n} - 1\right) - d\left(4p_{t+1,n} - 1\right)\right) \\ -p_{t+1,n+1}\left(d\left(2p_{t+1,n+1} - 1\right) - d\left(4p_{t+1,n+1} - 1\right)\right) \\ +d\left(4p_{t+1,n} - 1\right) + d \end{pmatrix}. \tag{20}$$

Rearranging (20) yields equation (16), in which the probability of execution at any node where the trader is not done $\left(n < N\right)$ is a weighted average of the following two connected nodes. Equation (16) will yield step-by-step aggressiveness in the Limit Order Zone once it is proven that up-nodes have higher probabilities and down-nodes have lower probabilities.

The next step is to define the Market Order / Maximum Aggressiveness Zone for both the *a* and *p* binomial trees at the set of all nodes where there is non-positive slack time (i.e., the same or fewer time periods left than remaining units to be purchased $T+1-t \leq N-n$). At all terminal nodes, there are no future trades, so $f_{T+1,n}=0,$ and $h_{T+1,n}=0.$ Substitute these equations into (12) for all of the nodes on the last trading date $\left(t=T\right)$ in the Market Order / Maximum Aggressiveness Zone to get

$$a_{T,n}=\begin{cases}\frac{1}{2}\left(k-d\right) & \text{when } k<3d \\ d & \text{when } k>3d\end{cases}, \tag{21}$$

and substitute this into (13) to get

$$p_{T,n}=\begin{cases}\frac{k+d}{4d} & \text{when } k<3d \\ 1 & \text{when } k>3d\end{cases}. \tag{22}$$

(21) and (22) are identical to the single-period solution. All of the nodes in this zone on earlier trading dates $\left(t<T\right)$ have two following connected nodes with a probability of execution given by (22). Since all nodes are a weighted average of the following two connected nodes by equation (16), their probability of execution must also be given by (22). Inverting (13) yields

$$a_{t,n}=d\left(2p_{t,n}-1\right). \tag{23}$$

Substituting (22) into (23) yields (21).

The next step is to define the Limit Order Zone for both the *a* and *p* binomial trees at the set of all nodes where the trader is not done $\left(n<N\right)$ and there is positive slack time (i.e., more time periods left than remaining units to be purchased $T+1-t>N-n$). Consider all of the nodes in the Limit Order Zone that border the Market Order / Maximum Aggressiveness Zone (i.e., have one following connection in the Market Order / Maximum Aggressiveness Zone and one following connection *not* in the Market Order / Maximum Aggressiveness Zone). Since all nodes are a weighted average of the following two connected nodes by equation (16), these nodes must have a probability of execution that is strictly less than (22) and substituting into (23) yields an optimal limit order aggressiveness strictly less than (21).

The next step is to define the Done Zone for both the a and p binomial trees at the set of all nodes where the trader is done $(n \geq N)$. For all of the nodes in the Done Zone, no more orders will be submitted and so the probability that an order will execute is zero $p_{t,n} = 0$. Consider all of the nodes in the Limit Order Zone that border the Done Zone (i.e., have one following connection in the Done Zone and one following connection *not* in the Done Zone). Since all nodes are a weighted average of the following two connected nodes by equation (16), these nodes must have a probability of execution that is strictly greater than zero and substituting into (23) must have an optimal limit order aggressiveness strictly greater than $-d$.

The final step is to consider all nodes in the Limit Order Zone. Since all nodes are a weighted average of the following two connected nodes by equation (16), then one of the following connected nodes will have a higher probability of execution and the other will have a lower probability of execution. Since the nodes in the Limit Order Zone that border the Market Order / Maximum Aggressiveness Zone must connect to the *higher* probability of execution of the Market Order / Maximum Aggressiveness Zone and since the nodes in the Limit Order Zone that border the Done Zone must connect to the *lower (zero)* probability of execution of the Done Zone, then up-nodes must have a higher probability of execution and down-nodes must have a lower probability of execution. This proof that up-nodes have higher probabilities and down-nodes have lower probabilities when combined with (16) yield that the probabilities exhibit step-by-step aggressiveness in the Limit Order Zone, and by substituting into (23), their optimal limit order aggressiveness must also exhibit step-by-step aggressiveness in the Limit Order Zone. *Q.E.D.*

References

Alfonsi, A., Fruth, A. & Schied, A., 2010, Optimal execution strategies in limit order books with general shape functions, *Quantitative Finance* 10, 143-157.

Almgren, R., 2003, Optimal execution with nonlinear impact functions and trading enhanced risk, *Applied Mathematical Finance* 10, 1-18.

Almgren, R. & Chriss, N., 2000, Optimal execution of portfolio transactions, *Journal of Risk* 3, 5-39.

Angel, J., 1995, Limit Versus Market Orders, Unpublished paper, Georgetown University.

Ankirchner, S. and T. Kruse, 2011a, Optimal trade execution under price-sensitive risk preferences, working paper, University of Bonn.

Ankirchner, S. and T. Kruse, 2011b, Price sensitive liquidation in continuous-time, working paper, University of Bonn.

Bae, K-H., H. Jang, and K. Park, 2002, Traders' Choice between Limit and Market Orders: Evidence from NYSE Stocks, *Journal of Financial Markets*, forthcoming.

Bertsimas, D. & Lo, A., 1998, Optimal control of execution costs, *Journal of Financial Markets* 1, 1-50.

Biais, B., P. Hillion, and C. Spatt, 1995, An Empirical Analysis of the Limit-order Book and the Order Flow in the Paris Bourse, *Journal of Finance* 50, 1655-89.

Bloomfield, R., M. O'Hara, and G. Saar, 2005, The 'Make or Take' Decision in an Electronic Market: Evidence on the Evolution of Liquidity, *Journal of Financial Economics* 75, 165-199.

Chakravarty, S. and C. Holden, 1995 An Integrated Model Of Market And Limit Orders, *Journal of Financial Intermediation* 4, 213-241

Chung, K., B. Van Ness, and R. Van Ness, 1999, Limit Orders and the Bid-Ask Spread, *Journal of Financial Economics* 53, 255-87.

Cohen, K., S. Maier, R. Schwartz, and D. Witcomb, 1981, "Transactions Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread," *Journal of Political Economy* 89, 287-305.

Conrad, J., Johnson, K., and S. Wahal, 2003, Institutional Trading and Alternative Trading

Systems, *Journal of Financial Economics* 70, 99-134.

Cushing, D., and A. Madhavan, 2000, Stock Returns and Trading at the Close, *Journal of Financial Markets* 3, 45-62.

Domowitz, I., Glen, J. and A. Madhavan, 2001, Liquidity, Volatility, and Equity Trading Costs Across Countries and Over Time, *International Finance* 4, 221-255.

Dubil, 2002, Optimal liquidation of large security holdings in thin markets, Research in International Entrepreneurial Finance and Business Ventures at the Turn of the Third Millennium, Editor: Yochanan Shachmurove, Academy of Entrepreneurial Finance.

Forsyth, P., 2009 A Hamilton Jacobi Bellman approach to optimal trade execution. working paper, University of Waterloo.

Foucault, T., 1999, Order Flow Composition and Trading Costs in a Dynamic Limit Order Market, *Journal of Financial Markets* 2, 99-134.

Foucault, T., Kadan, O., and E. Kandel, 2005. Limit Order Book as a Market for Liquidity, *Review of Financial Studies* 18, 1171-1217.

Gatheral, J. & Schied, A., 2011, Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework, *International Journal of Theoretical and Applied Finance* 14, 353-0368.

Gatheral, J. & Schied, A., 2012, Dynamical models of market impact and algorithms for order execution, working paper, City University of New York.

Glosten, L., 1994. Is the Electronic Open Limit Order Book Inevitable? *Journal of Finance* 49, 1127-61.

Goettler, R., C. Parlour, and U. Rajan, 2005, Equilibrium in a Dynamic Limit Order Book, *Journal of Finance*, 60, 2149-2192.

Goettler, R., Christine P., and U. Rajan, 2009. Informed traders and limit order markets, *Journal of Financial Economics* 93, 67-87.

Guilbaud, F., Mnif, M., and H. Pham, 2010, Numerical methods for an optimal order execution problem, working paper, University of Paris.

Handa, P., and R. Schwartz, 1996, Limit Order Trading, *Journal of Finance* 51, 1835-61.

Harris, L., 1998, Optimal Dynamic Order Submission Strategies in Some Stylized Trading Problems, *Financial Markets, Institutions, and Instruments* 7, 1-75.

Harris, L., and J. Hasbrouck, 1996, Market versus Limit Orders: the SuperDOT Evidence on Order Submission Strategy, *Journal of Financial and Quantitative Analysis* 31, 212-31.

Hasbrouck, J., and G. Saar, 2002, Limit Orders and Volatility in a Hybrid market: The Island ECN, unpublished paper, New York University.

He, H. and H. Mamaysky, 2001, Dynamic trading policies with price impact, *Journal of Economic Dynamics and Control* 29, 891-930.

Hollifield, B., R. Miller, and P. Sandas, 1999, An Empirical Analysis of Limit Order Markets, Unpublished paper, Rodney White Center for Financial Research Working Paper 029-99, University of Pennsylvania.

Hollifield, B., R. Miller, P. Sandas, and J. Slive, 2002, Liquidity Supply and Demand in Limit Order Markets, Unpublished paper, Rodney White Center for Financial Research Working Paper, University of Pennsylvania.

Huberman, G. and Stanzl, W., 2004, Price manipulation and quasi-arbitrage, *Econometrica* 72, 1247-1275.

Irvine, P, Lipson, M. and A. Puckett, 2007, Tipping, *Review of Financial Studies* 20, 741-768.

Kavajecz, K., 1999, A Specialist's Quoted Depth and the Limit Order Book, *Journal of Finance* 54, 747-72.

Keim, D. and A. Madhavan, 1995, Anatomy of the Trading Process: Empirical Evidence on the behavior of institutional traders, *Journal of Financial Economics* 37, 371-398.

Keim, D. and A. Madhavan, 1996, The Upstairs Market for Large-Block Transactions: Analysis and Measurement of Price Effects, *Review of Financial Studies* 9, 1-36.

Keim, D. and A. Madhavan, 1997, Transaction Costs and Investment Style: An Inter-exchange Analysis of Institutional Equity Trades, *Journal of Financial Economics* 46, 265-292.

Kharroubi, I. and Pham, H., 2010, Optimal portfolio liquidation with execution cost and risk, *Journal on Financial Mathematics* 1, 897-931.

Lakonishok, J., Schleifer, A., and R. Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23-43.

Ly Vath, V., Mnif, M., and H. Pham, 2007, A model of optimal portfolio selection under liquidity risk and price impact, *Finance and Stochastics* 11, 51-90.

Obizhaeva, A. & Wang, J., 2005, Optimal trading strategy and supply/demand dynamics, Working paper.

Parlour, C., 1998, Price Dynamics in Limit Order Markets, *Review of Financial Studies* 11, 789-816.

Parlour, C. and D. Seppi, 2008, Limit Order Markets: A Survey, *Handbook of Financial Intermediation and Banking*, Boot and Thakor (Editors), North-Holland, Amsterdam, The Netherlands.

Peterson, M., and E. Sirri, 2000, Order Submission Strategy and the Curious Case of Marketable Limit Orders, *Journal of Financial and Quantitative Analysis* 37, 221-41.

Ranaldo, A., 2002, Order Aggressiveness in Limit Order Book Markets, *Journal of Financial Markets*, forthcoming.

Rogers L. and S. Singh (2008) The cost of illiquidity and its effects on hedging, *Mathematical Finance* 20, 597-615.

Rosenthal, D., 2010, Performance metrics for algorithmic traders, working paper, University of Illinois.

Rosu, I., 2009, A Dynamic Model of the Limit Order Book, *Review of Financial Studies* 22, 4601-4641.

Sandas, P., 2001, Adverse Selection and Competitive Market Making: Empirical Evidence from a Pure Limit Order Market, *Review of Financial Studies* 14, 705-34.

Schied, A. and Schoneborn, T., 2009, Risk aversion and the dynamics of optimal trading strategies in illiquid markets, *Finance and Stochastics* 13, 181-204.

Seppi, D., 1997, Liquidity Provision with Limit Orders and a Strategic Specialist, *Review of Financial Studies* 10, 103-50.

Smith, J., 2000, Market versus Limit Order Submission Behavior at a Nasdaq Market Maker, Unpublished paper, Nasdaq Economic Research.

United States Securities and Exchange Commission, 2001, Report on Comparison of Order Executions Across Equity Market Structures, Office of Economic Analysis.

United States Securities and Exchange Commission, 2001, Disclosure of Order Execution and Routing Practices.

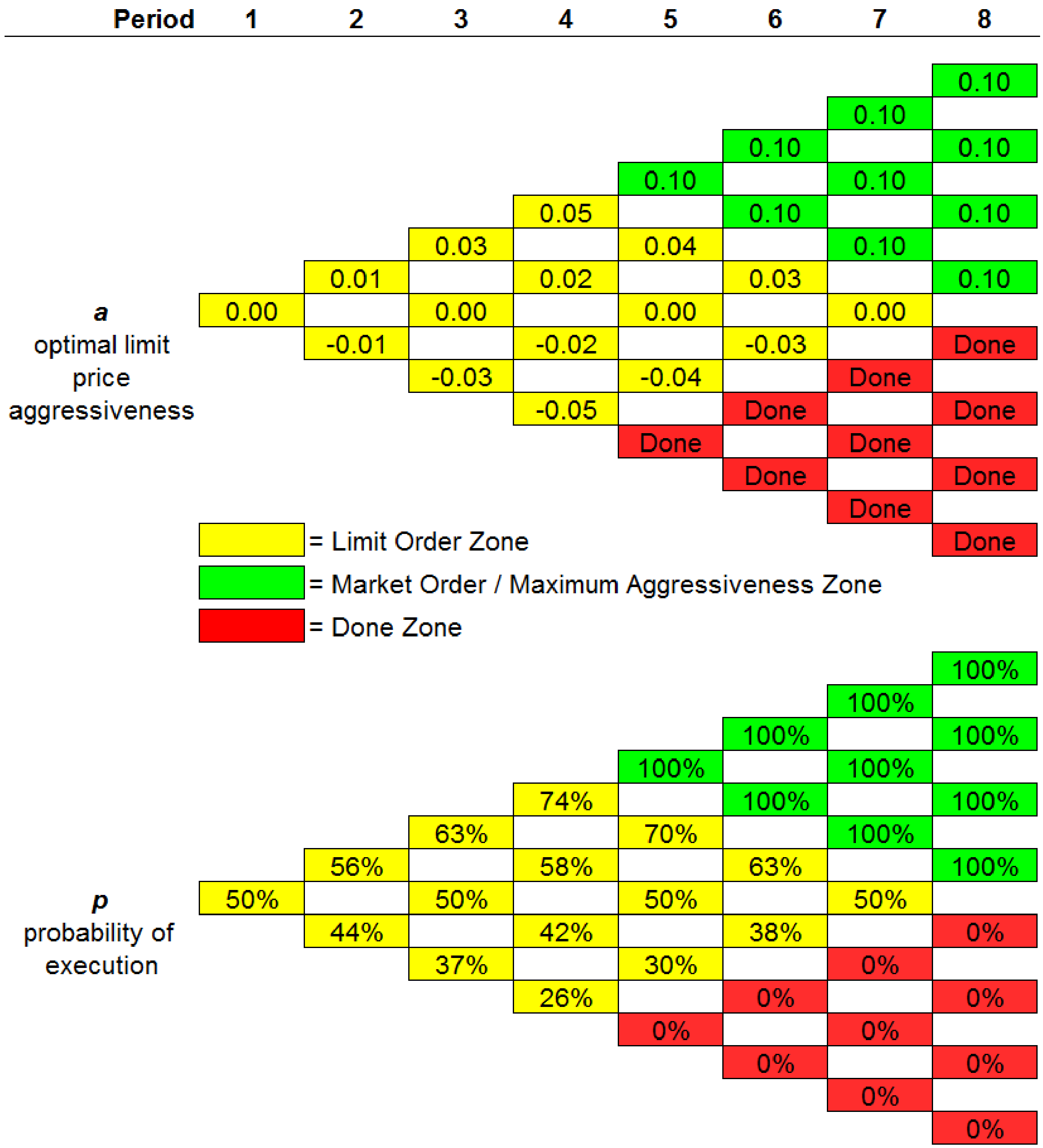

Figure 1. Binomial Trees for *a* and *p* when the Failure Penalty is High.

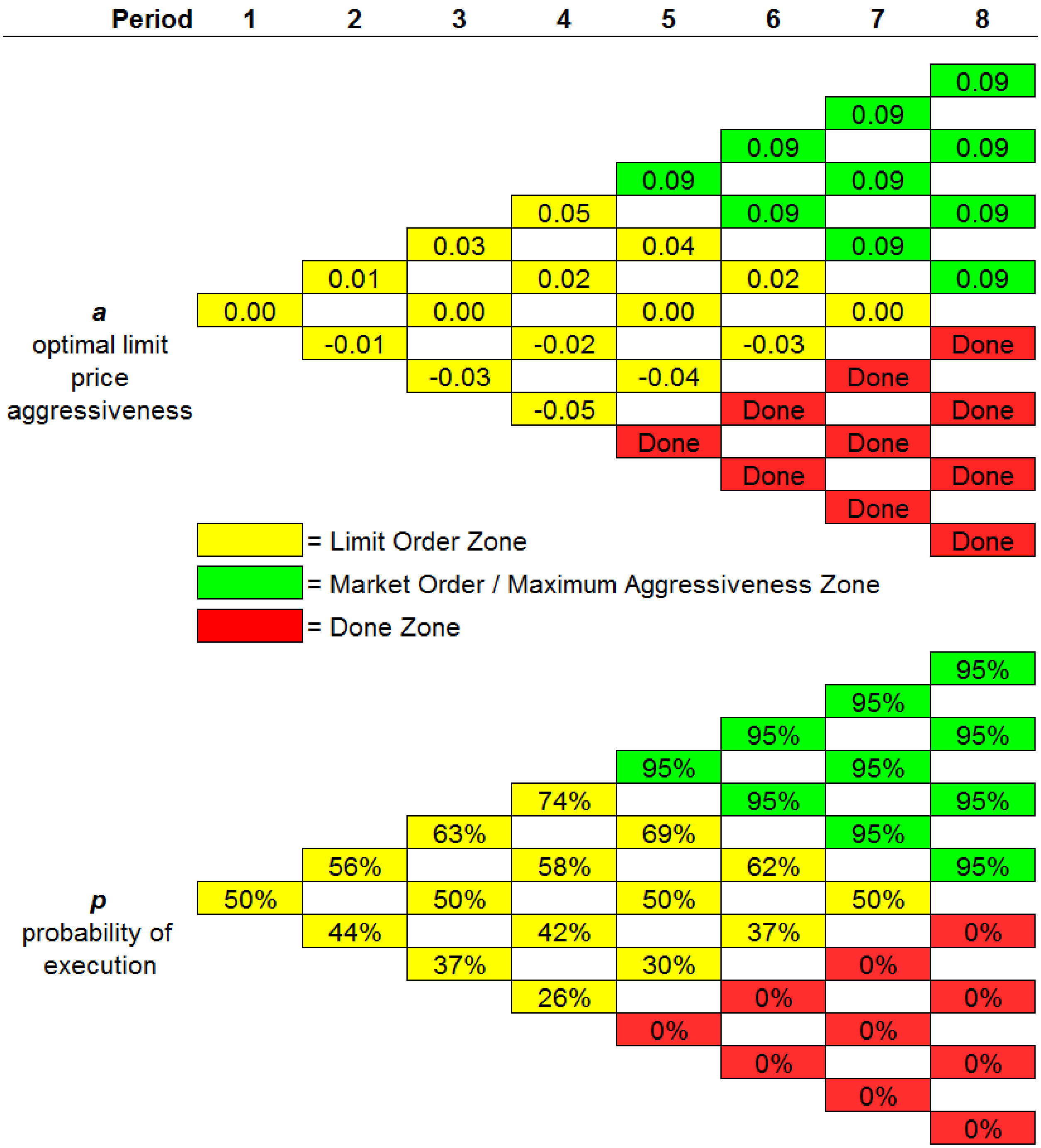

Figure 2. Binomial Trees for *a* and *p* when the Failure Penalty is Low.

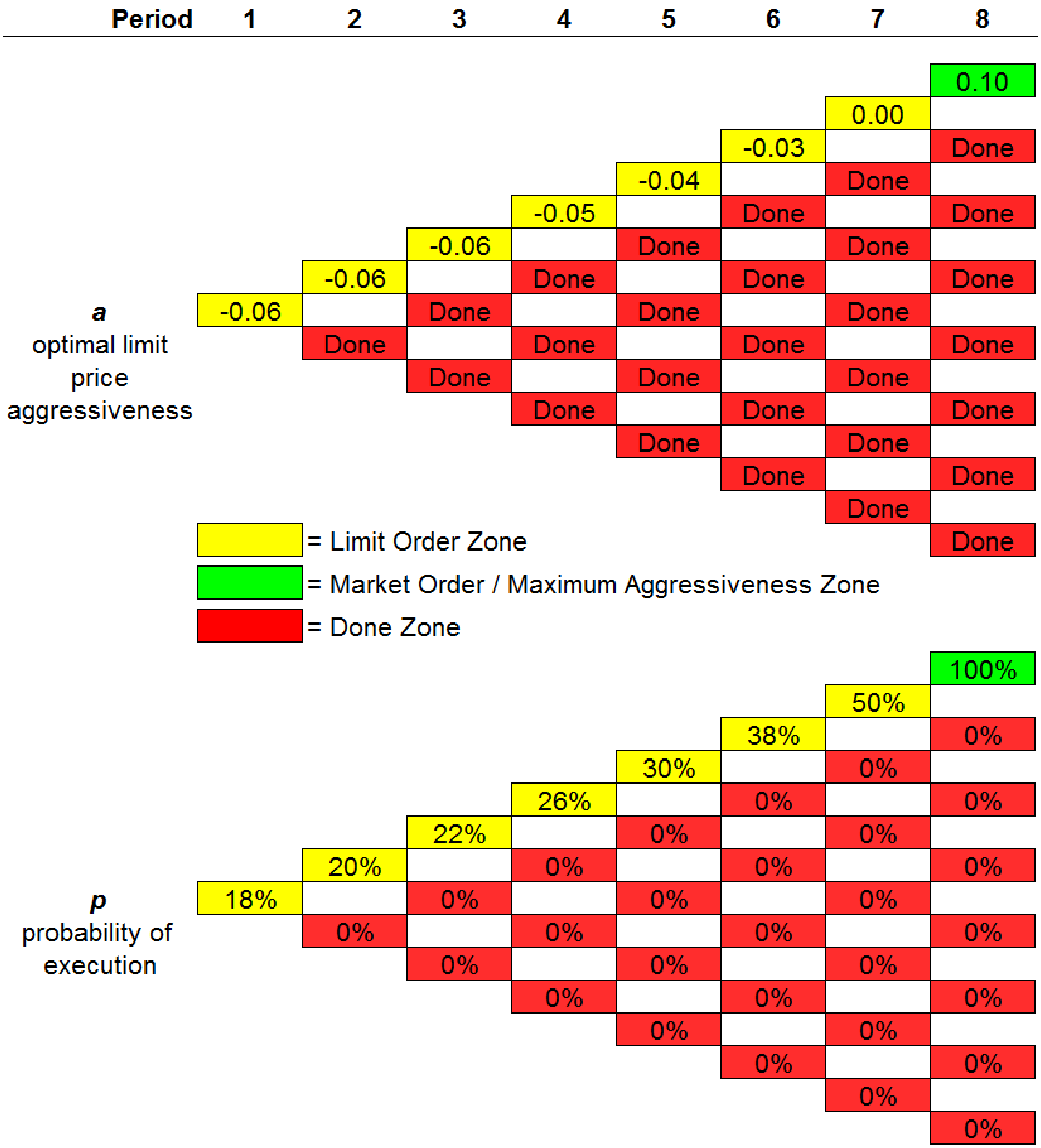

Figure 3. Binomial Trees for *a* and *p* when $N = 1$ and the Failure Penalty is High.

Table I
Calibration of Simulation Input Parameters

The simulation input parameters are calibrated to SEC Rule 605 mandated disclosure of summary statistics by exchange by stock. Summary statistics for Dec 2011 are divided by 21 trading days and by 20 intervals to scale to 20 minute intervals. NYSE TAQ data is sampled every 20 minutes to determine price volatility at 20 minute intervals.

Panel A: Probability of order types

Order Type	Probability
Market buy	30.75%
Market sell	30.75%
Limit buy	19.25%
Limit sell	19.25%
No order	0.00%
Total	100.00%

Panel B: Probability of Order Size by Order Type

Order Size	Probability of Market Order	Probability of Limit Order
10	3.0%	0.5%
20	3.0%	1.0%
30	3.0%	1.0%
40	4.0%	1.0%
50	4.0%	1.5%
100	19.0%	20.0%
200	20.0%	20.0%
300	20.0%	20.0%
400	14.0%	20.0%
500	10.0%	15.0%
Total	100.0%	100.0%
Average Size	230.4	276.7

Panel C: Probability of Price Changes

Price Changes	Probability of Public Value Innovations	Probability of Limit Price Deviations
-$0.05	13.0%	0.0%
-$0.04	12.0%	0.0%
-$0.03	10.0%	4.6%
-$0.02	8.0%	8.0%
-$0.01	6.0%	87.4%
$0.00	2.0%	0.0%
$0.01	6.0%	0.0%
$0.02	8.0%	0.0%
$0.03	10.0%	0.0%
$0.04	12.0%	0.0%
$0.05	13.0%	0.0%
Total	100.0%	100.0%
Mean Absolute Deviation	0.033	

Table II
Trading Algorithms Tested

Algorithm Description
1. LBs at bid -.01
2. Strat 1, except LB price = bid
3. Strat 1, except LB price = bid + .01
4. Large MBs as needed
5. Small MBs every other period
6. Tiny MBs every period
7. LBs with no cancel & resub
8. Strat 7, except cancel & resub every 12 periods
9. Strat 7, except cancel & resub every 4 periods
10. Medium MB when spread <.08; no condition when < 5 per
11. Medium MB when spread <.06; no condition when < 5 per
12. Medium MB when spread <.04; no condition when < 5 per
13. Medium LBs; More front-loaded; MB at end
14. Medium LBs; MB at end
15. Medium LBs; more agg prices below 6; MB at end
16. MB Ask Depth when spr <.08; no condition when < 5 per
17. MB Ask Depth when spr <.06; no condition when < 5 per
18. MB Ask Depth when spr <.04; no condition when < 5 per
19. 1/4 LBs; Step-by-step price agg; End MB
20. 1/4 LBs; Step-by-step price agg; End MB; Large At zone
21. 1/3 LBs; Step-by-step price agg; End MB
22. 1/3 LBs; Step-by-step price agg; End MB; Large At zone
23. 1/5 LBs; Step-by-step price agg; End MB
24. 1/5 LBs; Step-by-step price agg; End MB; Large At zone
25. 1/4 LBs; Step-by-step abs price agg (1,2,3); End MB
26. 1/4 LBs; Step-by-step abs price agg (2,3,4); End MB
27. 1/3 LBs; Step-by-step abs price agg (1,2,3); End MB
28. 1/3 LBs; Step-by-step abs price agg (2,3,4); End MB
29. 1/5 LBs; Step-by-step abs price agg (1,2,3); End MB
30. 1/5 LBs; Step-by-step abs price agg (2,3,4); End MB
31. Strat 7, except LB price = bid
32. Strat 8, except LB price = bid
33. Strat 9, except LB price = bid
34. Strat 13, except LB price = bid
35. Medium LBs; more abs agg prices (2,1) below 6; MB end
36. Medium LBs; more abs agg prices (3,2) below 6; MB end
37. Strat 7, except LB price = bid-.01
38. Strat 8, except LB price = bid-.01
39. Strat 9, except LB price = bid-.01
40. Strat 13, except LB price = bid-.01
41. Medium LBs; more agg prices below 8; MB at end
42. Medium LBs; more agg prices below 4; MB at end
43. Medium LBs; more abs agg prices (2,1) below 8; MB end
44. Medium LBs; more abs agg prices (2,1) below 4; MB end
45. Medium LBs; more abs agg prices (3,2) below 8; MB end
46. Medium LBs; more abs agg prices (3,2) below 4; MB end

Table III

Optimal Trading Algorithm By Failure Penalty, Risk Aversion, and Performance Metric

The optimal trading algorithm is shown by failure penalty, risk aversion, and performance metric. In all cases below, the requested buy is 10% of daily volume, the deadline is 20, and the fund manager's information is zero.

Risk Aversion	Performance Metric	Optimal Trader Strategy Description
Panel A Failure Penalty is Zero		
0	Effective Spread	Limit Buys only, Bid - .01, no cancellation
0	Imple. Shortfall	Limit Buys only, Bid - .01, no cancellation
0	VWAP	Limit Buys only, Bid - .01, no cancellation
0	Closing Price	Limit Buys; Midpoint-.02; in last 4 rounds Midpoint-.03; MB end
100	Effective Spread	Limit Buys only, Bid - .01, no cancellation
100	Imple. Shortfall	Limit Buys only, Bid - .01, no cancellation
100	VWAP	Limit Buys only, Bid - .01, no cancellation
100	Closing Price	Limit Buys; Midpoint-.01; in last 4 rounds Midpoint-.02; MB end
1000	Effective Spread	Limit Buys only, Bid - .01, no cancellation
1000	Imple. Shortfall	Limit Buys only, Bid - .01, no cancellation
1000	VWAP	Limit Buys only, Bid - .01, no cancellation
1000	Closing Price	Limit Buys only, Bid - .01, no cancellation
Panel B Failure Penalty is Medium (0.50)		
0	Effective Spread	Limit Buys mainly; Market Buy at end
0	Imple. Shortfall	Limit Buys mainly; Market Buy at end
0	VWAP	Limit Buys mainly; Market Buy at end
0	Closing Price	Limit Buys mainly; Market Buy at end
100	Effective Spread	Limit Buys; Midpoint-.01; in last 4 rounds Midpoint-.02; MB end
100	Imple. Shortfall	Limit Buys; Midpoint-.01; in last 4 rounds Midpoint-.02; MB end
100	VWAP	Limit Buys; Midpoint-.01; in last 4 rounds Midpoint-.02; MB end
100	Closing Price	Limit Buys; Midpoint-.01; in last 8 rounds Midpoint-.02; MB end
1000	Effective Spread	Limit Buys; Midpoint-.01; in last 8 rounds Midpoint-.02; MB end
1000	Imple. Shortfall	Limit Buys; Midpoint-.01; in last 4 rounds Midpoint-.02; MB end
1000	VWAP	Limit Buys; Midpoint-.01; in last 4 rounds Midpoint-.02; MB end
1000	Closing Price	Limit Buys; Midpoint-.01; in last 4 rounds Midpoint-.02; MB end
Panel C Failure Penalty is High (1.00)		
0	Effective Spread	Limit Buys mainly; Market Buy at end
0	Imple. Shortfall	Limit Buys mainly; Market Buy at end
0	VWAP	Limit Buys mainly; Market Buy at end
0	Closing Price	Limit Buys mainly; Market Buy at end
100	Effective Spread	Limit Buys; Midpoint-.01; in last 4 rounds Midpoint-.02; MB end
100	Imple. Shortfall	Large Market Buys
100	VWAP	Large Market Buys
100	Closing Price	Limit Buys; Midpoint-.01; in last 8 rounds Midpoint-.02; MB end
1000	Effective Spread	Limit Buys; Midpoint-.01; in last 8 rounds Midpoint-.02; MB end
1000	Imple. Shortfall	Limit Buys; Midpoint-.01; in last 8 rounds Midpoint-.02; MB end
1000	VWAP	Limit Buys; Midpoint-.01; in last 8 rounds Midpoint-.02; MB end
1000	Closing Price	Limit Buys; Midpoint-.01; in last 8 rounds Midpoint-.02; MB end

Table IV
Optimal Trading Algorithm By Fund Manager's Information, Deadline, and Performance Metric

The optimal trading algorithm is shown by fund manager's information, deadline, and performance metric. In all cases below, the requested buy is 10% of daily volume, the failure penalty is 0.50, and the fund manager's risk aversion is zero.

Deadline	Performance Metric	Optimal Trader Strategy Description
Panel A Fund Manager's Information is Zero		
2	Effective Spread	Large Market Buys
2	Imple. Shortfall	Large Market Buys
2	VWAP	Large Market Buys
2	Closing Price	Large Market Buys
10	Effective Spread	Limit Buys mainly; Market Buy at end
10	Imple. Shortfall	Limit Buys mainly; Market Buy at end
10	VWAP	Limit Buys mainly; Market Buy at end
10	Closing Price	Limit Buys mainly; Market Buy at end
20	Effective Spread	Limit Buys mainly; Market Buy at end
20	Imple. Shortfall	Limit Buys mainly; Market Buy at end
20	VWAP	Limit Buys mainly; Market Buy at end
20	Closing Price	Limit Buys mainly; Market Buy at end
Panel B Fund Manager's Information is Medium (0.50)		
2	Effective Spread	Large Market Buys
2	Imple. Shortfall	Large Market Buys
2	VWAP	Large Market Buys
2	Closing Price	Large Market Buys
10	Effective Spread	Limit Buys mainly; Market Buy at end
10	Imple. Shortfall	Large Market Buys
10	VWAP	Large Market Buys
10	Closing Price	Large Market Buys
20	Effective Spread	Limit Buys; Midpoint-.01; in last 6 rounds Midpoint-.02; MB end
20	Imple. Shortfall	Large Market Buys
20	VWAP	Large Market Buys
20	Closing Price	Limit Buys mainly; Market Buy at end
Panel C Fund Manager's Information is High (1.00)		
2	Effective Spread	Large Market Buys
2	Imple. Shortfall	Large Market Buys
2	VWAP	Large Market Buys
2	Closing Price	Limit Buys Mainly; More front-loaded; MB at end
10	Effective Spread	Limit Buys mainly; Market Buy at end
10	Imple. Shortfall	Large Market Buys
10	VWAP	Large Market Buys
10	Closing Price	Large Market Buys
20	Effective Spread	Limit Buys; Midpoint-.01; in last 8 rounds Midpoint-.02; Market Buy at end
20	Imple. Shortfall	Large Market Buys
20	VWAP	Large Market Buys
20	Closing Price	Large Market Buys