Statistical Arbitrage with Limit Order Book Imbalance

Anton Rubisov

University of Toronto Institute for Aerospace Studies
Faculty of Applied Science and Engineering
University of Toronto

A thesis submitted in conformity with the requirements
for the degree of *Master of Applied Science*

Statistical Arbitrage with Order Imbalance

Anton Rubisov

University of Toronto Institute for Aerospace Studies
Faculty of Applied Science and Engineering
University of Toronto

2015

Abstract

This is your abstract. It should be a summary and should include motivation, thesis objective and anticipated significance.

Acknowledgements

And I would like to acknowledge ...

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Algorithmic trading is the process of running computer algorithms to execute orders on an electronic exchange such as NASDAQ. Speed of execution is typically paramount (and hence the alternate moniker of high-frequency trading), often requiring running the algorithms on servers directly wired to the exchange (known as co-location) which comes at a steep price of X. Theoretically, high-frequency trading is encompassed by algorithmic trading, but not all algorithmic trading need be high-frequency; however, the two terms are often used interchangeably.

High-frequency trading can be proprietary or non-proprietary, and varies greatly across the different types of strategies employed. For example, low-latency strategies utilize naive algorithms to profit from price discrepancies for a single stock cross-listed on multiple exchanges, while statistical arbitrage strategies use complex algorithms to profit from observed statistical patterns of a single stock on a single exchange.

Proprietary trading refers to trading for-profit with an institution's money, e.g. the livelihood of hedge funds. Cross-exchange arbitrage uses primitive algorithms and emphasizes speed: the server is co-located at one of the exchanges, algorithm latency is on the order of microseconds, and the limiting reagent is the time taken for information to travel to and from the other exchange. In the case of Chicago and New York, for example, information can make the trip in 6ms via fibreoptics that send information at half the speed of light. For this reason, agents are now paying for access to a system of ground satellites that has been set up to relay information between the two exchanges via microwaves, shaving latency down to 4ms [Laughlin et al., 2014].

Non-proprietary algorithmic trading is generally aimed at transaction cost reduction, with the primary theoretical paper on the subject being [Almgren and Chriss, 2001]. When an institutional investor wishes to buy or sell a large quantity of shares, the aim of the trader is to obtain the best possible price compared with some benchmark (often taken to be the midprice at the time of initiating the strategy). Here the definition of 'large' is determined relative to the liquidity of the stock - either in comparison to the average size of trades for the given stock, or to the available quantity to be bought/sold at the best listed price. The goal of the algorithmic trading is then to split up the large order into smaller pieces and execute them on an algorithmically-determined schedule, balancing the total time for execution with the volatility of the price the trader will receive.

In statistical arbitrage, the aim is to exploit predictable statistical patterns in the available data provided by the exchange, e.g. predicting stock price movements from prices observed thus far. This method too requires co-location, and operates on the scale of milliseconds. It is this type of high-frequency trading that I will explore in this dissertation.

As part of the Dodd-Frank Act of 2010, the Volcker Rule has banned US banks from making certain speculative investments, and in so doing effectively curbed their proprietary trading activity. Nevertheless, as they are still required to provide liquidity to markets via market-making, banks utilize algorithmic trading to determine the bid/ask bands they will send to exchanges. Thus it remains feasible for banks to simply shift the bands they send in order to comply with market-making requirements and simultaneously execute a given strategy.

1.1 Limit Order Book Dynamics

A *limit order* is an instruction submitted by an agent to buy or sell up to a specified quantity or volume of a financial instrument, and at a specified price. A *limit order book (LOB)* is the accumulated list of such orders sent to a given exchange, where each order is accompanied by a timestamp and an anonymous unique key for identifying the agent. The exchange runs a trade matching engine that utilizes the LOB to pair buy and sell requests that concur on price, even if only partially on volume. Orders remain in effect until they are modified, canceled, or fully filled [Kyle, 1989].

The unfilled or partially filled orders accumulate in the limit order book and provide liquidity to the market. At any given time, the structure of the LOB can be visualized

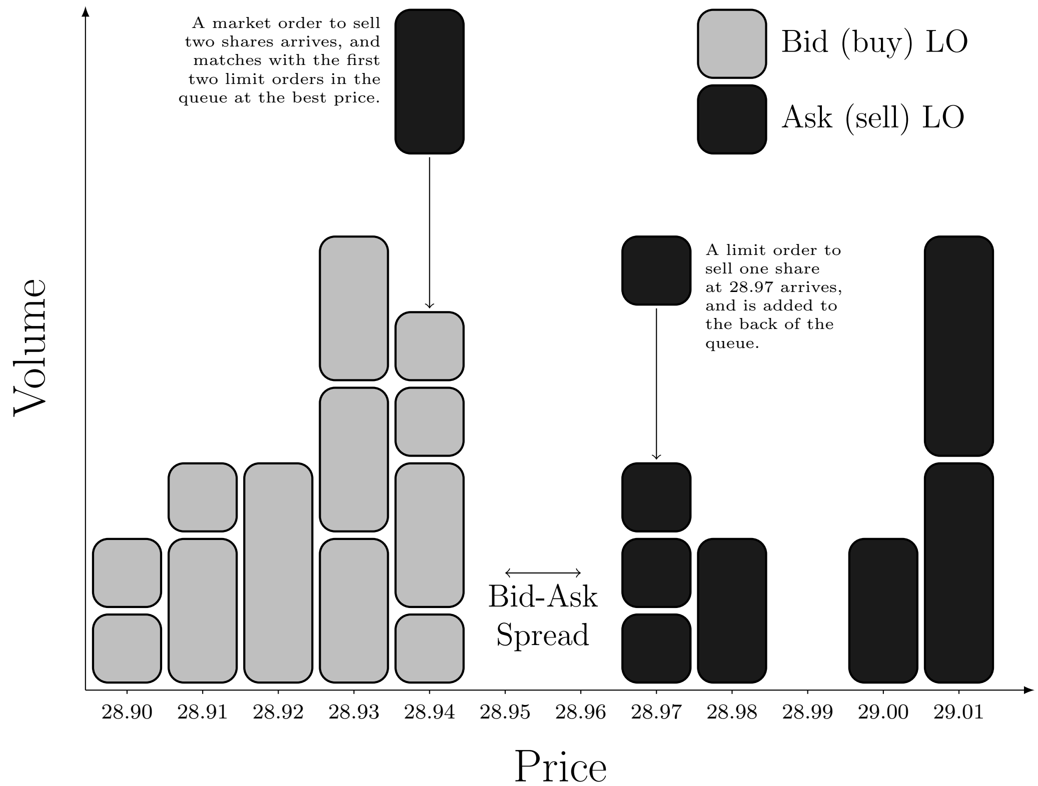


Figure 1.1: Structure and mechanics of the limit order book, adapted from [Booth, 2015]. Each block represents an order, of varying volumes, submitted by an agent.

as in Figure 1.1. As new limit orders arrive, they are compared with existing opposing orders in the book in search of a match - and if so, existing orders are *filled* or *lifted* according to a first-in-first-out priority queue for each price level. *Market orders* extend the idea of limit orders by specifying only the volume, and accept the best possible price currently available in the LOB; whereas limit orders are free to post, modify, and cancel (as an incentive for providing liquidity), market orders are charged a fee.

In the literature, LOBs are generally modelled in one of two ways: either by an economics-based approach, or a physics-based approach [Gould et al., 2013]. The economics based approaches are trader-centric, assume perfect rationality, view order flow as static, and seek to understand trader strategies, in particular through game-style theories. By contrast, the physics approach, with which we are more concerned here, assumes zero-intelligence, provides conceptual toy models of the evolution of the book, and is concerned with the search for statistical regularity. The dynamics of the book, namely order arrivals and cancellations, are governed by stochastic processes of varying complexity, from particles on a 1-D price lattice [Bak et al., 1997] to independent Poisson processes governing the arrival, modification, and cancellation of orders [Cont et al., 2010]. See [Gould et al., 2013] for an excellent literature survey on LOB modelling.

In this thesis, I will be concerned primarily with limit order book order imbalance. *Imbalance* is a ratio of limit order volumes between the bid and ask side, and in this work is calculated as

$$I(t) = \frac{V_{bid}(t) - V_{ask}(t)}{V_{bid}(t) + V_{ask}(t)} \in [-1, 1] \tag{1.1}$$

where both V_{bid} and V_{ask} are computed as the weighted average volumes at the best three prices, with exponentially decreasing weights.

1.2 ITCH Data Set

Chapter 2

Exploratory Data Analysis

2.1 Modelling Imbalance: Continuous Time Markov Chain

The aim of this research project is to utilize the LOB volume imbalance $I(t)$ in an algorithmic trading application; hence, a suitable choice of model for $I(t)$ must be made. Rather than modelling imbalance directly as a real-valued process, an alternative approach, and that which is utilized herein, is to discretize the imbalance value $I(t)$ into subintervals, or bins, and fit the resulting process to a continuous-time Markov chain.

The following definitions and properties are adapted from [Takahara, 2014]:

Definition 1. *A continuous-time stochastic process* $\{X(t) \mid t \geq 0\}$ *with state space* S *is called a* continuous-time Markov chain (CTMC) *if it has the Markov property; namely, that*

$$\mathbb{P}\left[X(t) = j \mid X(s) = i, X(t_{n-1} = i_{n-1}, \ldots, X(t_1) = i_1\right] = \mathbb{P}\left[X(t) = j \mid X(s) = i\right] \quad (2.1)$$

where for any integer $n \geq 1$, $0 \leq t_1 \leq \cdots \leq t_{n-1} \leq s \leq t$ *is any non-decreasing sequence of* $n+1$ *times, and* $i_1, \ldots, i_{n-1}, i, j \in S$ *are any* $n+1$ *states.*

Definition 2. *A CTMC* $X(t)$ *is* time homogeneous *if for any* $s \leq t$ *and any states* $i, j \in S$,

$$\mathbb{P}\left[X(t) = j \mid X(s) = i\right] = \mathbb{P}\left[X(t-s) = j \mid X(0) = i\right] \quad (2.2)$$

Definition 3. *The key quantities that determine a CTMC $X(t)$ are the* transition rates q_{ij}*, which specify the rate at which X jumps from state i to j. Conditional on leaving state i, X transitions to state j with* conditional transition probability p_{ij}*. The amount of time that X spends in state i, called the* holding time*, is exponentially distributed with rate v_i. These quantities are related by:*

$$v_i = \sum_{\substack{j \in S \\ j \neq i}} q_{ij} \tag{2.3}$$

$$q_{ij} = v_i \cdot p_{ij} \tag{2.4}$$

$$p_{ij} = \frac{q_{ij}}{v_i} \tag{2.5}$$

Definition 4. *A CTMC $X(t)$ has an* infinitesimal generator matrix $\boldsymbol{G}$*, whose entries are*

$$g_{ij} = q_{ij}, \qquad i \neq j \tag{2.6}$$

$$g_{ii} = -v_i \tag{2.7}$$

If $X(t)$ has transition probabilities $P_{ij}(t) = \mathbb{P}[X(t) = j \mid X(0) = i]$ and matrix $\boldsymbol{P}(t) = \{P_{ij}(t)\}$, then $\boldsymbol{P}(t)$ and $\boldsymbol{G}$ are related by

$$\dot{\boldsymbol{P}}(t) = \boldsymbol{G} \cdot \boldsymbol{P}(t) \tag{2.8}$$

$$\boldsymbol{P}(t) = e^{\boldsymbol{G}t} \tag{2.9}$$

Conditional on $Z(t) = k$, we assume the arrival of buy and sell market orders follow independent Poisson processes with intensities $\lambda_k^{\pm}$, where λ_k^{+} (λ_k^{-}) is the rate of arrivals of market buys (resp. sells). Such processes are called *Markov-modulated Poisson processes*, as the Poisson intensities are themselves stochastic processes determined by the state of the Markov chain. Thus, a timeline of observations of arrivals of buy/sell market orders and of regime switches might look like:

In the sections that follow, I derive maximum likelihood estimations for the parameters of the CTMC, and evaluate the fit of the model to the data.

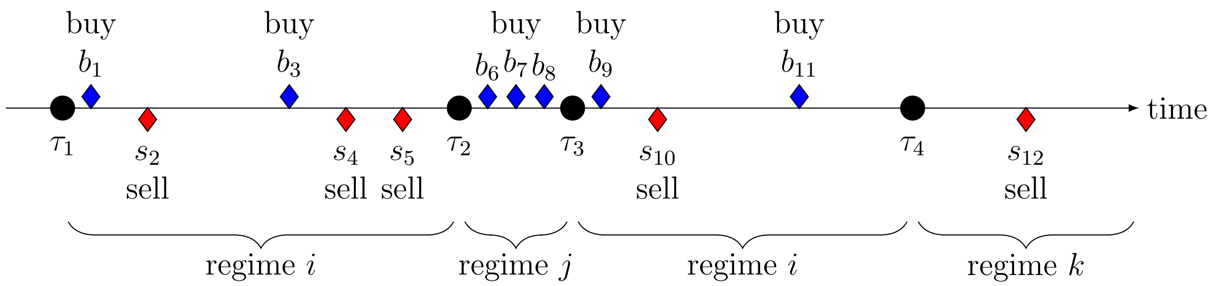


Figure 2.1: Hypothetical timeline of market orders arriving during changing order imbalance regimes.

2.2 Maximum Likelihood Estimate of a Markov-modulated Poisson Process

2.2.1 Maximum Likelihood Estimation of G

Let $\boldsymbol{G}$ be the generator matrix for a CTMC $Z(t)$ with state space K. From observations, e.g. the fictional events in the timeline given in Figure 2.1, we want to estimate the entries of $\boldsymbol{G}$. Since the holding time in a given state i has pdf $f(t; v_i) = v_i e^{-v_i t}$, the likelihood function (allowing for repetition of terms) would therefore be:

$$\mathcal{L}(\boldsymbol{G}) = (v_i e^{-v_i(\tau_2-\tau_1)} p_{ij})(v_j e^{-v_j(\tau_3-\tau_2)} p_{ji})(v_i e^{-v_i(\tau_4-\tau_3)} p_{ik}) \dots \tag{2.10}$$

$$= \prod_{i=1}^{K} \prod_{i \neq j} (v_i p_{ij})^{N_{ij}(T)} e^{-v_i H_i(T)} \tag{2.11}$$

$$= \prod_{i=1}^{K} \prod_{i \neq j} (q_{ij})^{N_{ij}(T)} e^{-v_i H_i(T)} \tag{2.12}$$

where:

$$N_{ij}(T) \equiv \text{number of transitions from regime } i \text{ to } j \text{ up to time } T$$

$$H_i(T) \equiv \text{holding time in regime } i \text{ up to time } T$$

So that the log-likelihood becomes:

$$\ln \mathcal{L}(\boldsymbol{G}) = \sum_{i=1}^{K} \sum_{i \neq j} \left[N_{ij}(T) \ln(q_{ij}) - v_i H_i(T) \right] \tag{2.13}$$

$$= \sum_{i=1}^{K} \sum_{i \neq j} \left[N_{ij}(T) \ln(q_{ij}) - \left(\sum_{i \neq k} q_{ik} H_i(T) \right) \right] \tag{2.14}$$

To get a maximum likelihood estimate $\hat{q}_{ij}$ for transition rates and therefore the matrix $\boldsymbol{G}$, we take the partial derivative of $\ln \mathcal{L}(\boldsymbol{G})$ and set it equal to zero:

$$\frac{\partial \ln \mathcal{L}(\boldsymbol{G})}{\partial q_{ij}} = \frac{N_{ij}(T)}{q_{ij}} - H_i(T) = 0 \tag{2.15}$$

$$\Rightarrow \hat{q}_{ij} = \frac{N_{ij}(T)}{H_i(T)} \tag{2.16}$$

2.2.2 Maximum Likelihood Estimation of $\lambda_k^{\pm}$

Now we want to derive an estimate for the intensity of the Poisson process of market order arrivals conditional on being in state k. We'll look first at just the buy market orders for some regime k, as the sell case is identical. Let the buy market order arrival times be indexed by b_i. Since we're assuming that the arrival process is Poisson with the same intensity throughout trials, we can consider the inter-arrival time of events conditional on being in state k. Then the MLE derivation follows just as for the generator matrix:

$$\mathcal{L}(\lambda_k^+; b_1, \ldots, b_N) = \prod_{i=2}^{N} \lambda_k^+ e^{-\lambda_k^+ (b_i - b_{i-1})} \tag{2.17}$$

$$= (\lambda_k^+)^{N_k^+(T)} e^{-\lambda_k^+ H_k(T)} \tag{2.18}$$

where:

$$N_k^+(T) \equiv \text{number of market order arrivals in regime } k \text{ up to time } T$$
$$H_k(T) \equiv \text{holding time in regime } k \text{ up to time } T$$

So that the log-likelihood becomes:

$$\ln \mathcal{L}(\lambda_k^+) = N_k^+(T) \ln(\lambda_k^+) - \lambda_k^+ H_k(T) \tag{2.19}$$

And the ML estimate for $\hat{\lambda}_k^+$ is:

$$\frac{\partial \ln \mathcal{L}}{\partial \lambda_k^+} = \frac{N_k^+(T)}{\lambda_k^+} - H_k(T) = 0 \tag{2.20}$$

$$\Rightarrow \hat{\lambda}_k^+ = \frac{N_k^+(T)}{H_k(T)} \tag{2.21}$$

2.3 2-Dimensional CTMC

Next we consider a CTMC that jointly models the imbalance bin and the price change over a subsequent interval. That is, the CTMC models the joint distribution $(I(t), \Delta S(t))$ where $I(t) \in \{1, 2, \ldots, \#_{bins}\}$ is the bin corresponding to imbalance averaged over the interval $[t - \Delta t_I, t]$, and $\Delta S(t) = \text{sgn}(S(t + \Delta t_S) - S(t)) \in \{-1, 0, 1\}$. The pair $(I(t), \Delta S(t))$ is then reduced into one dimension with a simple encoding which we will denote $\varphi(I(t), S(t))$; for example, using 3 bins:

$Z(t)$	Bin $I(t)$	$\Delta S(t)$	$Z(t)$	Bin $I(t)$	$\Delta S(t)$	$Z(t)$	Bin $I(t)$	$\Delta S(t)$
1	Bin 1	< 0	4	Bin 1	0	7	Bin 1	> 0
2	Bin 2	< 0	5	Bin 2	0	8	Bin 2	> 0
3	Bin 3	< 0	6	Bin 3	0	9	Bin 3	> 0

Table 2.1: $\varphi(I(t), S(t))$: 1-Dimensional Encoding of 2-Dimensional CTMC

2.3.1 Cross-Validation

We cross-validate the CTMC calibration by means of a time-homogeneity test similar to that done in [Tan and Ylmaz, 2002]. The null hypothesis is given by [Weißbach and Walter, 2010]:

$$H_0 = \forall i, j \in S \;:\; \exists q_{ij} \in \mathbb{R}^+ \;:\; q_{ij}(t) \equiv q_{ij} \forall t \in [0, T] \tag{2.22}$$

whereas the alternative hypothesis states that transition rates/probabilities are time-dependent. To test the hypothesis, we fix an imbalance averaging time Δt_I, number of imbalance bins, and calculate the MLE estimate of the infinitesimal generator matrix $\boldsymbol{G}$ on the full timeseries. For a chosen error threshold ϵ, we use the relationship in Equation (2.9) to calculate the number of timesteps n_{conv} of size Δt_I such that

$$||\boldsymbol{P}\left((n_{conv} + 1)\Delta t_I\right) - \boldsymbol{P}\left(n_{conv} \Delta t_I\right)|| < \epsilon \tag{2.23}$$

This value n_{conv} determines the size of the cross-validation timewindow into which to partition the full timeseries, yielding K equal subintervals of length n. For comparison, we also partitioned the timeseries into 8, 4, and 2 equal intervals. For each "removed series" $k \in \{1, \ldots, K\}$, we recalibrate a CTMC generator matrix $\boldsymbol{G}_k$. Finally, we test whether the one-step transition probabilities p_{ij}^k contained in $\boldsymbol{P}_k(\Delta t_I)$ are statistically different from those of the full period. The asymptotically equivalent test statistic to the likelihood ratio test statistic is:

$$D = -2\ln(\mathcal{L}) = 2\sum_k \sum_{i,j} n_{i,j}^k \left[\ln(p_{ij}^k) - \ln(p_{ij})\right] \tag{2.24}$$

where n_{ij}^k is the number of observed transitions from state i to j in subinterval k. This test statistic has a χ^2 distribution with $(K-1)(3 \cdot \#_{bins})(3 \cdot \#_{bins} - 1)$. The tests were run for each ticker for each trading day of 2013, and averaged over the year. Table 2.2 summarizes the p-value scores for the tests. Considering the standard cutoff p-value of 0.05, the cross-validation results show a strong case for the rejection of the homogeneity hypothesis. However, utilizing a non-homogeneous model falls outside of the scope of this research project, and instead suggests possible extensions to this research wherein the trading day is broken down into subintervals to better account for fluctuations and patterns in trading activity - perhaps early morning, mid-day, and final hour of trading.

2.4 Predicting Future Price Changes

It is crucial to note that the value $\Delta S(t)$ contains the price change from time t over the *future* Δt_S seconds - hence in real-time one cannot know the state of the Markov Chain. However, the analytic results do prove enlightening: from the resulting timeseries we estimate a generator matrix $\boldsymbol{G}$, and transform it into a one-step transition probability matrix $\boldsymbol{P} = e^{\boldsymbol{G}\Delta t_I}$. The entries of $\boldsymbol{P}$ are the conditional probabilities

$$\boldsymbol{P}_{ij} = \mathbb{P}\left[\varphi(I_{[t-\Delta t_I, t]}, \Delta S_{[t, t+\Delta t_S]}) = j \mid \varphi(I_{[t-2\Delta t_I, t-\Delta t_I]}, \Delta S_{[t-\Delta t_I, t]}) = i\right] \tag{2.25}$$

which can be expressed semantically as

$$= \mathbb{P}\left[\varphi(\rho_{curr}, \Delta S_{future}) = j \mid \varphi(\rho_{prev}, \Delta S_{curr}) = i\right] \tag{2.26}$$

		subintervals			
Δt_I	n_{conv}	n_{conv}	8	4	2
		FARO			
$\#_{bins} = 3$					
100ms	4933	0.000	0.000	0.000	0.003
1000ms	727	0.000	0.002	0.001	0.005
10000ms	149	0.000	0.005	0.010	0.017
$\#_{bins} = 5$					
100ms	6450	0.000	0.001	0.002	0.004
1000ms	941	0.000	0.001	0.003	0.006
10000ms	187	0.000	0.000	0.000	0.005
		NTAP			
$\#_{bins} = 3$					
100ms	1320	0.000	0.000	0.000	0.000
1000ms	237	0.000	0.000	0.000	0.000
10000ms	72	0.000	0.006	0.003	0.007
$\#_{bins} = 5$					
100ms	1777	0.000	0.000	0.000	0.000
1000ms	308	0.000	0.001	0.000	0.001
10000ms	87	0.000	0.000	0.002	0.010

		subintervals			
Δt_I	n_{conv}	n_{conv}	2	4	8
		ORCL			
$\#_{bins} = 3$					
100ms	1803	0.000	0.000	0.000	0.000
1000ms	303	0.000	0.000	0.000	0.001
10000ms	84	0.000	0.007	0.005	0.010
$\#_{bins} = 5$					
100ms	2503	0.000	0.000	0.000	0.000
1000ms	404	0.000	0.001	0.002	0.003
10000ms	103	0.000	0.000	0.001	0.009
		INTC			
$\#_{bins} = 3$					
100ms	2545	0.000	0.000	0.000	0.001
1000ms	408	0.000	0.001	0.001	0.002
10000ms	105	0.000	0.004	0.006	0.009
$\#_{bins} = 5$					
100ms	3498	0.000	0.001	0.001	0.001
1000ms	771	0.000	0.001	0.002	0.002
10000ms	133	0.000	0.000	0.000	0.007

Table 2.2: χ^2-test p-values for testing the time homogeneity hypothesis. Tests were run for each ticker for each trading day of 2013, and averaged over the year. For calculating n_{conv}, the converge error threshold was $\epsilon = 1 \times 10^{-10}$.

Since we can easily decode the 1-dimensional Markov state back into two dimensions, we can think of $\boldsymbol{P}$ as being four-dimensional and re-write its entries as

$$
\begin{align}
&= \mathbb{P}\left[\rho_{curr} = i, \Delta S_{future} = j \mid \rho_{prev} = k, \Delta S_{curr} = m\right] \tag{2.27}\\
&= \mathbb{P}\left[\rho_{curr} = i, \Delta S_{future} = j \mid B\right] \tag{2.28}
\end{align}
$$

where we're using the shorthand $B = (\rho_{prev} \in k, \Delta S_{curr} \in m)$ to represent the states in the previous timestep. Applying Bayes' Rule:

$$
\mathbb{P}\left[\Delta S_{future} \in j \mid B, \rho_{curr} \in i\right] = \frac{\mathbb{P}\left[\rho_{curr} \in i, \Delta S_{future} \in j \mid B\right]}{\mathbb{P}\left[\rho_{curr} \in i \mid B\right]} \tag{2.29}
$$

where the right-hand-side numerator is each individual entry of the one-step probability matrix $\boldsymbol{P}$, and the denominator can be computed from $\boldsymbol{P}$ by:

$$
\mathbb{P}\left[\rho_{curr} \in i \mid B\right] = \sum_j \mathbb{P}\left[\rho_{curr} \in i, \Delta S_{future} \in j \mid B\right] \tag{2.30}
$$

This result is of great interest to us: the left-hand-side value is the probability of seeing a given price change over the immediate future time interval conditional on past imbalances and the most recent price change, and therefore allows us to predict future price moves. We'll denote by $\boldsymbol{Q}$ the matrix containing all values given by Equation (2.29).

The following $\boldsymbol{Q}$ matrix was obtained using data for `MMM` from 2013-05-15, averaging imbalance timewindow $t_I = 1000$ms, $K = 3$ imbalance bins, and price change timewindow $t_S = 1000$ms:

	$\Delta S_{curr} < 0$			$\Delta S_{curr} = 0$			$\Delta S_{curr} > 0$		
	$\rho_n = 1$	$\rho_n = 2$	$\rho_n = 3$	$\rho_n = 1$	$\rho_n = 2$	$\rho_n = 3$	$\rho_n = 1$	$\rho_n = 2$	$\rho_n = 3$
$\Delta S_{future} < 0$									
$\rho_{n-1} = 1$	**0.53**	0.15	0.12	0.05	0.10	0.14	0.08	0.13	0.14
$\rho_{n-1} = 2$	0.10	**0.58**	0.14	0.07	0.04	0.10	0.13	0.06	0.12
$\rho_{n-1} = 3$	0.08	0.12	**0.52**	0.09	0.06	0.03	0.11	0.10	0.05
$\Delta S_{future} = 0$									
$\rho_{n-1} = 1$	0.41	0.75	0.78	**0.91**	0.84	0.79	0.42	0.79	0.77
$\rho_{n-1} = 2$	0.79	0.36	0.71	0.83	**0.92**	0.82	0.75	0.37	0.78
$\rho_{n-1} = 3$	0.79	0.74	0.40	0.81	0.83	**0.91**	0.70	0.76	0.39
$\Delta S_{future} > 0$									
$\rho_{n-1} = 1$	0.06	0.10	0.09	0.04	0.06	0.07	**0.50**	0.09	0.09
$\rho_{n-1} = 2$	0.10	0.06	0.15	0.10	0.04	0.08	0.12	**0.57**	0.10
$\rho_{n-1} = 3$	0.13	0.14	0.08	0.10	0.11	0.05	0.19	0.14	**0.56**

Table 2.3: The $\boldsymbol{Q}$ matrix: conditional probabilities of future price changes, conditioned on current imbalance, current price change, and previous imbalance.

Immediately evident from Table 2.3 is that in most cases we are expecting no price change. In fact, the only cases in which the probability of a price change is > 0.5 show evidence of *momentum*; for example, the way to interpret the value in row 1, column 1 is: if $\rho_{prev} = \rho_{curr} = 1$ and previously we saw a downward price change, then we expect to again see a downward price change. The bolded diagonal values in the table lend themselves to the empirical conlusion:

$$\mathbb{P}\left[\Delta S_{future} = \Delta S_{curr} \mid \rho_{curr} = \rho_{prev}\right] > 0.5 \tag{2.31}$$

2.5 Naive Trading Strategies

Utilizing the key insight drawn from Equation (2.31), we implemented several naive trading strategies, descriptions of which follow:

Naive Trading Strategy Using the conditional probabilities obtained from $\boldsymbol{Q}$, we will execute a buy (resp. sell) market order if the probability of an upward (resp. downward) price change is > 0.5. Pseudocode for this strategy is given in 1.

Algorithm 1 Naive Trading Strategy

```
cash = 0
asset = 0
for t = 2 : length(timeseries) do
    if P[ΔS_future < 0 | ρ_curr, ρ_prev, ΔS_curr] > 0.5 then
        cash += data.BuyPrice(t)
        asset −= 1
    else if P[ΔS_future > 0 | ρ_curr, ρ_prev, ΔS_curr] > 0.5 then
        cash −= data.SellPrice(t)
        asset += 1
    end if
end for
if asset > 0 then
    cash += asset × data.BuyPrice(t)
else if asset < 0 then
    cash += asset × data.SellPrice(t)
end if
```

Naive+ Trading Strategy If no change in midprice is expected then we'll post buy and sell limit orders at the touch, front of the queue. We'll track MO arrival between timesteps, assume we always get executed, and immediately repost the limit orders.

Naive++ Trading Strategy As a reformulation of the Naive strategy to use limit orders instead of market orders, if we expect a downward (resp. upward) price change then we'll post an at-the-touch sell (resp. buy) limit order, which may be lifted by an agent who is executing a market order going against the price change momentum.

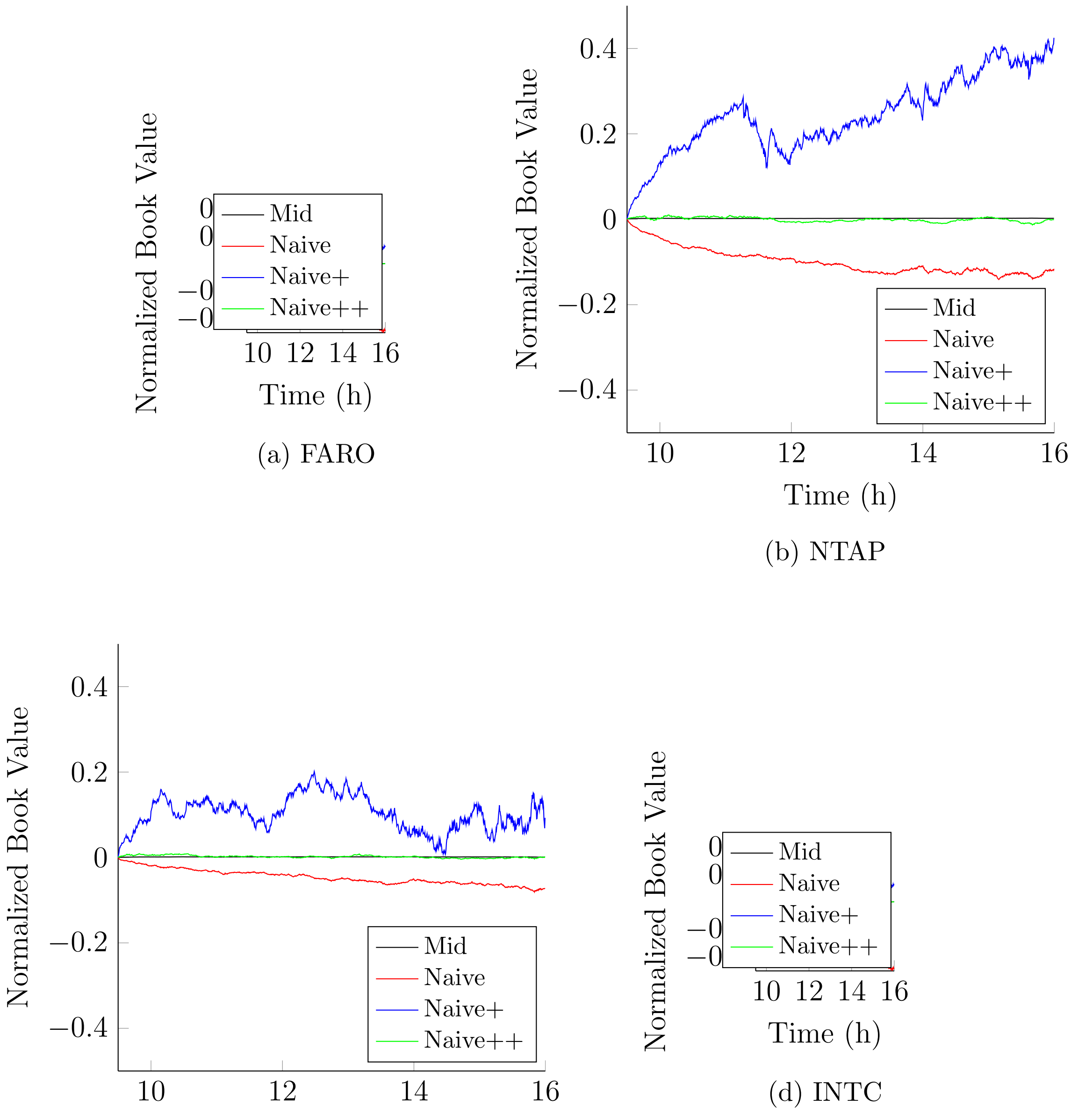
2.6 Calibration and Backtesting

Backtesting these naive trading strategies required a choice of parameters for the price change observation period Δt_S, the imbalance averaging period Δt_I, and the number of imbalance bins $\#_{bins}$. We used a brute force calibration technique that, for each ticker and each day, traversed the potential parameter space searching for the highest number of timesteps at which Equation (2.31) could be utilized. We found that $\#_{bins} = 4$ provided the highest expected number of trades for most tickers. However, as we were utilizing percentile bins symmetric around zero, we wanted to have $\#_{bins}$ as an odd number such that all behaviour around zero imbalance was treated equally; thus all backtesting was done with either $\#_{bins} = 3$ or $\#_{bins} = 5$. Additionally, we found empirically that calibration always yielded $\Delta t_S = \Delta t_I$, so this was taken as a given. The backtest for each ticker then consisted of first calibrating the value Δt_I from the first day of data by maximizing the intra-day Sharpe ratio, then using the calibrated parameters to backtest the entire year.

2.7 Conclusions from the Naive Trading Strategies

As seen in Figure 2.2, the Naive strategy on average underperformed the average mid price, while the Naive+ (at-the-touch limit orders when no change was expected) and Naive++ (adding limit orders to adversely select agents that traded against the price change momentum) strategies both on average generated revenue.

Question 1 Why is the Naive strategy producing, on average, normalized losses? On calibration, we see that our intra-day Sharpe ratio is around 0.01 or 0.02 when we choose our optimal parameters, so at the very least on the calibration date the strategy produces positive returns. The remainder of the calendar days are out-of-sample, and the parameters are (likely) not optimal. This suggests non-stationary data, and in particular not every day can be modelled by the same Markov chain. The problem may be exaggerated by the fact that we're calibrating on the first trading day of the calendar year, when we might expect reduced, or at least non-representative, trading activity. Further, we are using midprices to obtain the $\boldsymbol{Q}$ probability matrix while ignoring the bid-ask spread. Thus predicting a "price change" may be insufficient when considering a monetizable opportunity, as we won't be able to profit off a predicted increase followed by a predicted decrease unless the interim mid-price move is greater than the bid-ask spread (assuming constant spread); this flaw affects trading on `FARO` in particular, which has a spread of about 15 cents.



(a) FARO

(b) NTAP

(c) ORCL

(d) INTC

Figure 2.2: Comparison of Naive (red), Naive+ (blue), and Naive++ (green) trading strategies, with benchmark Midprice (black). Plotted are normalized book values against time of trading day, averaged across the trading year.

t	$I(t)$	Bid/Ask	Prediction	Action	Inv	PnL
0	1	9.99/10.01	$\mathbb{P}[\Delta S_{future} = 0] > 0.5$	None	0	0
1	1	10.00/10.02	$\mathbb{P}[\Delta S_{future} > 0] > 0.5$	BUY @ 10.02	1	-0.02
2	0	10.01/10.03	$\mathbb{P}[\Delta S_{future} = 0] > 0.5$	None	1	-0.01
3	-1	10.01/10.03	$\mathbb{P}[\Delta S_{future} = 0] > 0.5$	None	1	-0.01
4	-1	10.00/10.02	$\mathbb{P}[\Delta S_{future} < 0] > 0.5$	SELL @ 10.00	0	-0.02

Table 2.4: Hypothetical timeline of adverse selection with market orders.

Question 2 Why do the Naive+ and Naive++ strategies outperform the Naive strategy? This is particularly interesting since the probabilities are being obtained from the same matrix. The obvious difference between the successful and unsuccessful strategies is that the former (a) uses limit orders, and (b) executes when we predict a zero change, whereas the latter uses (a) market orders, and (b) executes when we predict non-zero changes.

(a) obviously leads to a different transaction price being used: a stock purchase with a limit order is executed at the bid price, while a purchase with a market order is at the ask price. Since the asset is marked-to-market at the more conservative price, and the mid price doesn't move as a result of my transaction, then a limit order purchases the share for the same value at which it is marked-to-market, whereas a market order 'crosses the spread' and loses value.

(b) seems to be the largest flaw in the Naive strategy, to which there are two factors. One, we are not predicting the magnitude of the price change, only whether it is zero or non-zero. Two, from the probabilities presented above, *we will only predict a price change if we've already seen a price change*. Thus we're effectively reacting too late. Table 2.4 presents a hypothetical series of events demonstrating the adverse effects of this flaw. Since the strategy is reacting to an already observed price change, the adverse effect would be exacerbated if the initial price change at timestep 4 were larger. All these considerations suggest potential modifications to the strategies.

Chapter 3

Stochastic Optimal Control

Hello and welcome to this limited edition chapter on stochastic optimal control. Read on if you dare subject yourself to the infinite wisdom contained herein.

3.1 Continuous Time

Below we list the processes involved in the optimization problem:

Imbalance & Midprice Change	$\boldsymbol{Z}_t = (\rho_t, \Delta_t)$	CTMC with generator G
Imbalance	$\rho_t = \boldsymbol{Z}_t^{(1)}$	LOB imbalance at time t
Midprice	S_t	evolves according to CTMC
Midprice Change	$\Delta_t = \boldsymbol{Z}_t^{(2)} = S_t - S_{t-s}$	s a pre-determined interval
Bid-Ask half-spread	ξ_t	constant?
LOB Shuffling	N_t	Poisson with rate $\lambda(\boldsymbol{Z}_t)$
ΔPrice: LOB shuffled	$\{\eta_{0,z}, \eta_{1,z}, \dots\} \sim F_z$	i.i.d. with $z = (k, l)$, where
		$k \in \{\#\text{bins}\},\ l \in \{\Delta\$\}$
Other Agent MOs	$K_t^{\pm}$	Poisson with rate $\mu^{\pm}(\boldsymbol{Z}_t)$
LO posted depth	$\delta_t^{\pm}$	our $\mathcal{F}$-predictable controlled processes
Our LO fill count	$L_t^{\pm}$	$\mathcal{F}$-predictable, non-Poisson
Our MOs	$M_t^{\pm}$	our controlled counting process
Our MO execution times	$\boldsymbol{\tau}^{\pm} = \{\tau_k^{\pm} : k = 1, \dots\}$	increasing sequence of $\mathcal{F}$-stopping times
Cash	$X_t^{\boldsymbol{\tau},\delta}$	depends on our processes M and δ
Inventory	$Q_t^{\boldsymbol{\tau},\delta}$	depends on our processes M and δ

$L_t^{\pm}$ are counting processes (not Poisson) satisfying the relationship that if at time t we have a sell limit order posted at a depth δ_t^-, then our fill probability is $e^{-\kappa\delta_t^-}$ conditional on a buy market order arriving; namely:

$$\mathbb{P}[\,\mathrm{d}L_t^- = 1 \,|\, \mathrm{d}K_t^+ = 1] = e^{-\kappa\delta_t^-} \tag{3.1}$$

$$\mathbb{P}[\,\mathrm{d}L_t^+ = 1 \,|\, \mathrm{d}K_t^- = 1] = e^{-\kappa\delta_t^+} \tag{3.2}$$

The midprice S_t evolves according to the Markov chain and hence is Poisson with rate λ and jump size η, both of which depend on the state of the Markov chain. This Poisson process is all-inclusive in the sense that it accounts for any midprice change, be it from executions, cancellations, or order modifications with the LOB. Thus, the stock midprice S_t evolves according to the SDE:

$$\mathrm{d}S_t = \eta_{N_{t^-},Z_{t^-}}\,\mathrm{d}N_t \tag{3.3}$$

and additionally satisfies:

$$S_t = S_{t_0} + \int_{t_0+s}^{t} \Delta_u \,\mathrm{d}u \tag{3.4}$$

In executing market orders, we assume that the size of the MOs is small enough to achieve the best bid/ask price, and not walk the book. Hence, our cash process evolves according to:

$$\begin{aligned}\mathrm{d}X_t^{\boldsymbol{\tau},\delta} = &\underbrace{(S_t + \xi_t + \delta_t^-)\,\mathrm{d}L_t^-}_{\text{sell limit order}} - \underbrace{(S_t - \xi_t - \delta_t^+)\,\mathrm{d}L_t^+}_{\text{buy limit order}} \\ &+ \underbrace{(S_t - \xi_t)\,\mathrm{d}M_t^-}_{\text{sell market order}} - \underbrace{(S_t + \xi_t)\,\mathrm{d}M_t^+}_{\text{buy market order}}\end{aligned} \tag{3.5}$$

Based on our execution of limit and market orders, our inventory satisfies:

$$Q_0^{\boldsymbol{\tau},\delta} = 0, \qquad Q_t^{\boldsymbol{\tau},\delta} = L_t^+ + M_t^+ - L_t^- - M_t^- \tag{3.6}$$

We define a new variable for our net present value (NPV) at time t, call it $W_t^{\boldsymbol{\tau},\delta}$, and hence $W_T^{\boldsymbol{\tau},\delta}$ at terminal time T is our 'terminal wealth'. In algorithmic trading, we want to finish the trading day with zero inventory, and assume that at the terminal time T we will submit a market order (of a possibly large volume) to liquidate remaining stock. Here we do not assume that we can receive the best bid/ask price - instead, the price achieved will be $(S - \text{sgn}(Q)\xi - \alpha Q)$, where $\text{sgn}(Q)\xi$ represents crossing the spread in the direction of trading, and αQ represents receiving a worse price linearly in Q due to walking the book. Hence, $W_t^{\boldsymbol{\tau},\delta}$ satisfies:

$$W_t^{\boldsymbol{\tau},\delta} = \underbrace{X_t^{\boldsymbol{\tau},\delta}}_{\text{cash}} + \underbrace{Q_t^{\boldsymbol{\tau},\delta}\left(S_t - \text{sgn}(Q_t^{\boldsymbol{\tau},\delta})\xi_t\right)}_{\text{book value of assets}} - \underbrace{\alpha\left(Q_t^{\boldsymbol{\tau},\delta}\right)^2}_{\text{liquidation penalty}} \tag{3.7}$$

The set of admissible trading strategies is the product of the sets $\mathcal{T}$, the set of all $\mathcal{F}$-stopping times, and $\mathcal{A}$, the set of all $\mathcal{F}$-predictable, bounded-from-below depths δ. We only consider $\delta^{\pm} \geq 0$, since at $\delta = 0$ our fill probability is $e^{-\kappa\delta} = 1$, so we cannot increase the chance of our limit order being filled by posting any lower than at-the-touch; doing so would only diminish our profit.

For deriving an optimal trading strategy via dynamic programming, I will consider the performance criteria that maximizes terminal wealth. With the above notation, the performance criteria function can be written

$$H^{\boldsymbol{\tau},\delta}(t, x, s, \boldsymbol{z}, q) = \mathbb{E}\left[W_T^{\boldsymbol{\tau},\delta}\right] \tag{3.8}$$

And the value function, in turn, is given by

$$H(t, x, s, \boldsymbol{z}, q) = \sup_{\boldsymbol{\tau}\in\mathcal{T}_{[t,T]}} \sup_{\delta\in\mathcal{A}_{[t,T]}} H^{\boldsymbol{\tau},\delta}(t, x, s, \boldsymbol{z}, q) \tag{3.9}$$

3.1.1 Dynamic Programming

The following theorems establish the dynamic programming method we will utilize to solve this type of problem:

Theorem 5 ([Cartea et al., 2015]). ***Dynamic Programming Principle for Optimal Stopping and Control.*** *If an agent's performance criteria for a given admissible control* $\boldsymbol{u}$ *and admissible stopping time* τ *are given by*

$$H^{\tau,\boldsymbol{u}}(t, \boldsymbol{x}) = \mathbb{E}_{t,\boldsymbol{x}}[G(X_\tau^{\boldsymbol{u}})]$$

and the value function is

$$H(t, \boldsymbol{x}) = \sup_{\tau \in \mathcal{T}_{[t,T]}} \sup_{\boldsymbol{u} \in \mathcal{A}_{[t,T]}} H^{\tau, \boldsymbol{u}}(t, \boldsymbol{x})$$

then the value function satisfies the Dynamic Programming Principle

$$H(t, \boldsymbol{x}) = \sup_{\tau \in \mathcal{T}_{[t,T]}} \sup_{\boldsymbol{u} \in \mathcal{A}_{[t,T]}} \mathbb{E}_{t,\boldsymbol{x}} \left[G(X_\tau^{\boldsymbol{u}}) \mathbb{1}_{\tau < \theta} + H(\theta, X_\theta^{\boldsymbol{u}}) \mathbb{1}_{\tau \geq \theta} \right] \tag{3.10}$$

for all $(t, \boldsymbol{x}) \in [0, T] \times \mathbb{R}^m$ and all stopping times $\theta \leq T$.

Theorem 6 ([Cartea et al., 2015])**.** ***Dynamic Programming Equation for Optimal Stopping and Control.*** *Assume that the value function $H(t, \boldsymbol{x})$ is once differentiable in t, all second-order derivatives in $\boldsymbol{x}$ exist, and that $G : \mathbb{R}^m \to \mathbb{R}$ is continuous. Then H solves the quasi-variational inequality*

$$0 = \max \left\{ \partial_t H + \sup_{\boldsymbol{u} \in \mathcal{A}_t} \mathcal{L}_t^{\boldsymbol{u}} H \; ; \; G - H \right\} \tag{3.11}$$

on $\mathcal{D}$, where $\mathcal{D} = [0, T] \times \mathbb{R}^m$.

3.1.2 Maximizing Terminal Wealth (Continuous)

In this section we solve the DPE that results from using the maximal terminal wealth performance criteria. The quasi-variational inequality in equation 3.11 can be interpreted as follows: the max operator is choosing between posting limit orders or executing market orders; the second term, $G - H$, is the stopping region and represents the value derived from executing a market order; and the first term is the continuation region, representing the value of posting limit orders. We'll use the shorthand $H(\cdot) = H(t, x, s, \boldsymbol{z}, q)$ and solve for $\mathrm{d}H$ inside the continuation region, hence $\mathrm{d}M^{\pm} = 0$, in order to then extract out the infinitesimal generator.

$$\begin{aligned} \mathrm{d}H(t, x, s, \boldsymbol{z}, q) &= \sum_i \partial_{x_i} H \, \mathrm{d}x_i & (3.12) \\ &= \partial_t H \, \mathrm{d}t + \partial_{K^\pm} H \, \mathrm{d}K^\pm + \partial_{\boldsymbol{Z}} H \, \mathrm{d}\boldsymbol{Z} & (3.13) \end{aligned}$$

$$
\begin{aligned}
&= \partial_t H \,\mathrm{d}t + \left\{ e^{-\kappa\delta^-} \mathbb{E}\big[H(t, x + (s + \xi + \delta^-), s, \boldsymbol{z}, q-1) - H(\cdot)\big] \right\} \mathrm{d}K^+ \\
&\qquad + \left\{ e^{-\kappa\delta^+} \mathbb{E}\big[H(t, x - (s - \xi - \delta^+), s, \boldsymbol{z}, q+1) - H(\cdot)\big] \right\} \mathrm{d}K^- \\
&\qquad + \sum_{\mathbf{j}} \mathbb{E}\left[H(t, x, s + \eta_{0,\mathbf{j}}, \mathbf{j}, q) - H(\cdot)\right] \mathrm{d}Z_{\boldsymbol{z},\mathbf{j}}
\end{aligned}
\tag{3.14}
$$

We substitute in the following identities for the compensated processes, which are continuous-time martingales. For Poisson processes this is given by [Cartea et al., 2015]

$$
\mathrm{d}K^{\pm} = \mathrm{d}\tilde{K}^{\pm} + \mu^{\pm}(\boldsymbol{z})\,\mathrm{d}t \tag{3.15}
$$

while for the Markov Chain, this is [Kurtz, 2004]

$$
\mathrm{d}Z_{\boldsymbol{z},\mathbf{j}} = \mathrm{d}\tilde{Z}_{\boldsymbol{z},\mathbf{j}} + G_{\boldsymbol{z},\mathbf{j}}\,\mathrm{d}t \tag{3.16}
$$

$$
\begin{aligned}
&= \partial_t H \,\mathrm{d}t + \Bigg\{ \mu^+(\boldsymbol{z}) e^{-\kappa\delta^-} \mathbb{E}\big[H(t, x + (s + \xi + \delta^-), s, \boldsymbol{z}, q-1) - H(\cdot)\big] \\
&\qquad\qquad + \mu^-(\boldsymbol{z}) e^{-\kappa\delta^+} \mathbb{E}\big[H(t, x - (s - \xi - \delta^+), s, \boldsymbol{z}, q+1) - H(\cdot)\big] \\
&\qquad\qquad + \sum_{\mathbf{j}} G_{\boldsymbol{z},\mathbf{j}} \mathbb{E}\left[H(t, x, s + \eta_{0,\mathbf{j}}, \mathbf{j}, q) - H(\cdot)\right] \Bigg\} \mathrm{d}t \\
&\qquad + \left\{ e^{-\kappa\delta^-} \mathbb{E}\big[H(t, x + (s + \xi + \delta^-), s, \boldsymbol{z}, q-1) - H(\cdot)\big] \right\} \mathrm{d}\tilde{K}^+ \\
&\qquad + \left\{ e^{-\kappa\delta^+} \mathbb{E}\big[H(t, x - (s - \xi - \delta^+), s, \boldsymbol{z}, q+1) - H(\cdot)\big] \right\} \mathrm{d}\tilde{K}^- \\
&\qquad + \sum_{\mathbf{j}} \mathbb{E}\left[H(t, x, s + \eta_{0,\mathbf{j}}, \mathbf{j}, q) - H(\cdot)\right] \mathrm{d}\tilde{Z}_{\boldsymbol{z},\mathbf{j}}
\end{aligned}
\tag{3.17}
$$

From which we can see that the infinitesimal generator is given by

$$
\begin{aligned}
\mathcal{L}_t^{\delta} H = {} & \mu^+(\boldsymbol{z}) e^{-\kappa\delta^-} \mathbb{E}\big[H(t, x + (s + \xi + \delta^-), s, \boldsymbol{z}, q-1) - H(\cdot)\big] \\
& + \mu^-(\boldsymbol{z}) e^{-\kappa\delta^+} \mathbb{E}\big[H(t, x - (s - \xi - \delta^+), s, \boldsymbol{z}, q+1) - H(\cdot)\big] \\
& + \sum_{\mathbf{j}} G_{\boldsymbol{z},\mathbf{j}} \mathbb{E}\left[H(t, x, s + \eta_{0,\mathbf{j}}, \mathbf{j}, q) - H(\cdot)\right]
\end{aligned}
\tag{3.18}
$$

Now, our DPE has the form

$$
\begin{aligned}
0 = \max\Big\{ \partial_t H + \sup_{\boldsymbol{u}\in\mathcal{A}_t} \mathcal{L}_t^{\boldsymbol{u}} H \; ; \; H(t, x-(s+\xi), s, \boldsymbol{z}, q+1) - H(\cdot) \; ; \\
H(t, x+(s-\xi), s, \boldsymbol{z}, q-1) - H(\cdot) \Big\}
\end{aligned}
\tag{3.19}
$$

with boundary conditions

$$
H(T, x, s, \boldsymbol{z}, q) = x + q(s - \operatorname{sgn}(q)\xi) - \alpha q^2 \tag{3.20}
$$

$$
H(t, x, s, \boldsymbol{z}, 0) = x \tag{3.21}
$$

The three terms over which we are maximizing represent the continuation regions and stopping regions of the optimization problem. The first term, the continuation region, represents the limit order controls; the second and third terms, each a stopping region, represent the value gain from executing a buy market order and a sell market order, respectively.

Let's introduce the ansatz $H(\cdot) = x + q(s - \operatorname{sgn}(q)\xi) + h(t, \boldsymbol{z}, q)$. The first two terms are the wealth plus book value of assets, hence a mark-to-market of the current position, whereas the $h(t, \boldsymbol{z}, q)$ captures value due to the optimal trading strategy. The corresponding boundary conditions on h are

$$
h(T, \boldsymbol{z}, q) = -\alpha q^2 \tag{3.22}
$$

$$
h(t, \boldsymbol{z}, 0) = 0 \tag{3.23}
$$

Substituting this ansatz into equation 3.18, we get:

$$
\begin{aligned}
\mathcal{L}_t^{\delta} H = \mu^+(\boldsymbol{z}) e^{-\kappa\delta^-} \big[\delta^- + \xi[1 + \operatorname{sgn}(q-1) + q(\operatorname{sgn}(q) - \operatorname{sgn}(q-1))] \\
+ h(t, \boldsymbol{z}, q-1) - h(t, \boldsymbol{z}, q)\big] \\
+ \mu^-(\boldsymbol{z}) e^{-\kappa\delta^+} \big[\delta^+ + \xi[1 - \operatorname{sgn}(q+1) + q(\operatorname{sgn}(q) - \operatorname{sgn}(q+1))] \\
+ h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q)\big] \\
+ \sum_{\mathbf{j}} G_{\boldsymbol{z},\mathbf{j}} \left[q\mathbb{E}[\eta_{0,\mathbf{j}}] + h(t, \mathbf{j}, q) - h(t, \boldsymbol{z}, q)\right]
\end{aligned}
\tag{3.24}
$$

We can further simplify the factors of ξ; for example, in the case of the δ^+ term, we can write

$$
\begin{aligned}
1 - \mathrm{sgn}(q+1) + q(\mathrm{sgn}(q) - \mathrm{sgn}(q+1)) &= 1 - (-\mathbb{1}_{q\le -2} + \mathbb{1}_{q\ge 0}) + \mathbb{1}_{q=-1} \\
&= 1 + (\mathbb{1}_{q\le -1} - \mathbb{1}_{q\ge 0}) \\
&= 2 \cdot \mathbb{1}_{q\le -1}
\end{aligned}
$$

This gives us the simplified infinitesimal generator term

$$
\begin{aligned}
\mathcal{L}_t^{\delta} H = \mu^{+}(\boldsymbol{z})e^{-\kappa\delta^{-}} &\big[\delta^{-} + 2\xi \cdot \mathbb{1}_{q\ge 1} + h(t,\boldsymbol{z},q-1) - h(t,\boldsymbol{z},q)\big] \\
&+ \mu^{-}(\boldsymbol{z})e^{-\kappa\delta^{+}}\big[\delta^{+} + 2\xi \cdot \mathbb{1}_{q\le -1} + h(t,\boldsymbol{z},q+1) - h(t,\boldsymbol{z},q)\big] \\
&+ \sum_{\mathbf{j}} G_{\boldsymbol{z},\mathbf{j}}\left[q\mathbb{E}[\eta_{0,\mathbf{j}}] + h(t,\mathbf{j},q) - h(t,\boldsymbol{z},q)\right]
\end{aligned}
\tag{3.25}
$$

In the DPE, the first term requires finding the supremum over all $\delta^{\pm}$ of the infinitesimal generator. For this we can set the partial derivatives with respect to both δ^{+} and δ^{-} equal to zero to solve for the optimal posting depth, which we denote with a superscript asterisk. For δ^{+} we get:

$$
0 = \partial_{\delta^{+}}\left[e^{-\kappa\delta^{+*}}\big[\delta^{+*} + 2\xi \cdot \mathbb{1}_{q\le -1} + h(t,\boldsymbol{z},q+1) - h(t,\boldsymbol{z},q)\big]\right] \tag{3.26}
$$

$$
= -\kappa e^{-\kappa\delta^{+*}}\big[\delta^{+*} + 2\xi \cdot \mathbb{1}_{q\le -1} + h(t,\boldsymbol{z},q+1) - h(t,\boldsymbol{z},q)\big] + e^{-\kappa\delta^{+*}} \tag{3.27}
$$

$$
= e^{-\kappa\delta^{+}}\big[-\kappa\big(\delta^{+*} + 2\xi \cdot \mathbb{1}_{q\le -1} + h(t,\boldsymbol{z},q+1) - h(t,\boldsymbol{z},q)\big) + 1\big] \tag{3.28}
$$

Since $e^{-\kappa\delta^{+*}} > 0$, the term inside the square braces must be equal to zero:

$$
0 = -\kappa\big(\delta^{+*} + 2\xi \cdot \mathbb{1}_{q\le -1} + h(t,\boldsymbol{z},q+1) - h(t,\boldsymbol{z},q)\big) + 1 \tag{3.29}
$$

$$
\delta^{+*} = \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q\le -1} - h(t,\boldsymbol{z},q+1) + h(t,\boldsymbol{z},q) \tag{3.30}
$$

Recalling that our optimal posting depths are to be non-negative, we thus find that the optimal buy limit order posting depth can be written in feedback form as

$$
\delta^{+*} = \max\left\{0\ ;\ \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q\le -1} - h(t,\boldsymbol{z},q+1) + h(t,\boldsymbol{z},q)\right\} \tag{3.31}
$$

We can follow similar steps to solve for the optimal sell limit order posting depth

$$
\delta^{-*} = \max\left\{0\ ;\ \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q\ge 1} - h(t,\boldsymbol{z},q-1) + h(t,\boldsymbol{z},q)\right\} \tag{3.32}
$$

Turning our attention to the stopping regions of the DPE, we can use the ansatz to simplify the expressions:

$$
\begin{aligned}
&H(t, x-(s+\xi), s, \boldsymbol{z}, q+1) - H(\cdot) \\
&\qquad = x - s - \xi + (q+1)(s - \operatorname{sgn}(q+1)\xi) + h(t, \boldsymbol{z}, q+1) && (3.33)\\
&\qquad\quad - \big[x + q(s - \operatorname{sgn}(q)\xi) + h(t, \boldsymbol{z}, q)\big] \\
&\qquad = -\xi\big[(q+1)\operatorname{sgn}(q+1) - q\operatorname{sgn}(q) + 1\big] + h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q) && (3.34)\\
&\qquad = -2\xi \cdot \mathbb{1}_{q\geq 0} + h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q) && (3.35)
\end{aligned}
$$

and similarly,

$$
H(t, x+(s-\xi), s, \boldsymbol{z}, q-1) - H(\cdot) = -2\xi \cdot \mathbb{1}_{q\leq 0} + h(t, \boldsymbol{z}, q-1) - h(t, \boldsymbol{z}, q) \tag{3.36}
$$

Substituting all these results and simplifications into the DPE, we find that h satisfies

$$
\begin{aligned}
0 = \max\Big\{ &\partial_t h + \mu^+(\boldsymbol{z}) e^{-\kappa\delta^{-*}} \left(\delta^{-*} + 2\xi\mathbb{1}_{q\geq 1} + h(t, \boldsymbol{z}, q-1) - h(t, \boldsymbol{z}, q)\right) \\
&\qquad + \mu^-(\boldsymbol{z}) e^{-\kappa\delta^{+*}} \left(\delta^{+*} + 2\xi \cdot \mathbb{1}_{q\leq -1} + h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q)\right) \\
&\qquad + \sum_{\mathbf{j}} G_{\boldsymbol{z},\mathbf{j}} \left[q l \mathbb{E}\left[\eta_{0,\mathbf{j}}\right] + h(t, \mathbf{j}, q) - h(t, \boldsymbol{z}, q)\right] \; ; \\
&\quad - 2\xi \cdot \mathbb{1}_{q\geq 0} + h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q) \; ; \\
&\quad - 2\xi \cdot \mathbb{1}_{q\leq 0} + h(t, \boldsymbol{z}, q-1) - h(t, \boldsymbol{z}, q) \Big\}
\end{aligned} \tag{3.37}
$$

Looking at the simplified feedback form in the stopping region, we see that a buy market order will be executed at time τ_q^+ whenever

$$
h(\tau_q^+, \boldsymbol{z}, q+1) - h(\tau_q^+, \boldsymbol{z}, q) = 2\xi \cdot \mathbb{1}_{q\geq 0} \tag{3.38}
$$

and a sell market order whenever

$$
h(\tau_q^+, \boldsymbol{z}, q-1) - h(\tau_q^+, \boldsymbol{z}, q) = 2\xi \cdot \mathbb{1}_{q\leq 0} \tag{3.39}
$$

Consider than when our inventory is positive, we can purchase a stock at $s+\xi$, but it will be marked-to-market at $s-\xi$, resulting in a value difference of 2ξ. With negative inventory, we will still purchase at s_ξ, but will now also value at $s+\xi$ because our overall position is still negative, producing no value difference. In particular, with negative inventory, we will execute a buy market order so long as it does not change our value function; and with zero or positive inventory, only if it increases the value function by the value of the spread. The opposite holds for sell market orders. Together, these indicate a penchant for using market orders to drive inventory levels back toward zero when it has no effect on value, and using them to gain extra value only when the expected gain is equal to the size of the spread. This is reminiscent of what we saw in the exploratory data analysis: if a stock is worth S, we can purchase it at $S + \xi$ and immediately be able to sell it at $S - \xi$, at a loss of 2ξ; this was the most significant source of loss in the naive trading market order strategy. Hence we need to expect our value to increase by at least 2ξ when executing market orders for gain.

The variational inequality in Equation (3.37) yields that whilst in the continuation region, we instead have

$$h(\tau_q^+, \boldsymbol{z}, q+1) - h(\tau_q^+, \boldsymbol{z}, q) \leq 2\xi \cdot \mathbb{1}_{q \geq 0} \tag{3.40}$$
$$h(\tau_q^+, \boldsymbol{z}, q-1) - h(\tau_q^+, \boldsymbol{z}, q) \leq 2\xi \cdot \mathbb{1}_{q \leq 0} \tag{3.41}$$

Taken together, these inequalities yield

$$-2\xi \cdot \mathbb{1}_{q \geq 0} \leq h(t, \boldsymbol{z}, q) - h(t, \boldsymbol{z}, q+1) \leq 2\xi \cdot \mathbb{1}_{q \leq -1} \tag{3.42}$$
$$-2\xi \cdot \mathbb{1}_{q \leq 0} \leq h(t, \boldsymbol{z}, q) - h(t, \boldsymbol{z}, q-1) \leq 2\xi \cdot \mathbb{1}_{q \geq 1} \tag{3.43}$$

or alternatively,

$$h(t, \boldsymbol{z}, q) \overset{\text{sell if }=}{\overset{\downarrow}{\leq}} h(t, \boldsymbol{z}, q+1) \overset{\text{buy if }=}{\overset{\downarrow}{\leq}} h(t, \boldsymbol{z}, q) + 2\xi, \qquad q \geq 0 \tag{3.44}$$
$$h(t, \boldsymbol{z}, q) \underset{\text{buy if }=}{\underset{\uparrow}{\leq}} h(t, \boldsymbol{z}, q-1) \underset{\text{sell if }=}{\underset{\uparrow}{\leq}} h(t, \boldsymbol{z}, q) + 2\xi, \qquad q \leq 0 \tag{3.45}$$

Recalling the boundary condition $h(t, \boldsymbol{z}, 0) = 0$, Equation (3.44) and Equation (3.45) tell us that the function h is non-negative everywhere. Furthermore, noting the feedback form of our optimal buy limit order depth given in equation Equation (3.31), together with the inequalities in Equation (3.42) and Equation (3.42), we obtain bounds on our posting depths given by

$$\begin{aligned} \delta^{+^*} &= \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q \leq -1} - h(t, \boldsymbol{z}, q+1) + h(t, \boldsymbol{z}, q) && (3.46) \\ &\geq \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q \leq -1} - 2\xi \cdot \mathbb{1}_{q \geq 0} && (3.47) \end{aligned}$$

$$= \frac{1}{\kappa} - 2\xi \tag{3.48}$$

$$\delta^{+*} \leq \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q \leq -1} + 2\xi \cdot \mathbb{1}_{q \leq -1} \tag{3.49}$$

$$= \frac{1}{\kappa} \tag{3.50}$$

Combined with the identical conditions on the sell depth, we have the conditions

$$\boxed{\frac{1}{\kappa} - 2\xi \leq \delta^{\pm *} \leq \frac{1}{\kappa}} \tag{3.51}$$

A possible interpretation of the unexpected upper bound on the posting depth is that if the calculated buy (resp. sell) depth is 'sufficiently' large so as to indicate a disposition against buying (resp. selling), then it is actually optimal to sell (resp. buy) instead.

3.2 Discrete Time

Reminder of our processes (a little bit of abuse of notation going on):

$\boldsymbol{z}_k = (\rho_k, \Delta_k)$ - 2-D time-homogenous Markov Chain with transition probabilities $\boldsymbol{P}_{ij}$, where $\rho_k \in \Gamma$ and Γ represents the set of imbalance bins, and $\Delta_k = \operatorname{sgn}(s_k - s_{k-1}) \in \{-1, 0, 1\}$.

$$\text{State } \vec{x}_k = \begin{pmatrix} x_k \\ s_k \\ \boldsymbol{z}_k \\ q_k \end{pmatrix} \qquad \begin{matrix} \text{cash} \\ \text{stock price} \\ \text{Markov chain state, as above} \\ \text{inventory} \end{matrix}$$

$$\text{Control } \vec{u}_k = \begin{pmatrix} \delta_k^+ \\ \delta_k^- \\ M_k^+ \\ M_k^- \end{pmatrix} \qquad \begin{matrix} \text{bid posting depth} \\ \text{ask posting depth} \\ \text{buy MO - binary control} \\ \text{sell MO - binary control} \end{matrix}$$

$$\text{Random } \vec{w}_k = \begin{pmatrix} K_k^+ \\ K_k^- \\ \omega_k \end{pmatrix} \qquad \begin{matrix} \text{other agent buy MOs - binary} \\ \text{other agent sell MOs - binary} \\ \text{random variable uniformly distributed on [0,1]} \end{matrix}$$

Following [Kwong, 2015], we'll write the evolution of the Markov chain as a function of the current state and a uniformly distributed random variable ω:

$$\boldsymbol{z}_{k+1} = T(\boldsymbol{z}_k, \omega_k) = \sum_{i=0}^{|\Gamma|} i \cdot \mathbb{1}_{\left(\sum_{j=0}^{i-1} \boldsymbol{P}_{\boldsymbol{z}_k, j}, \sum_{j=0}^{i} \boldsymbol{P}_{\boldsymbol{z}_k, j}\right]}(\omega_k) \tag{3.52}$$

Here $\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$, and hence Z_{k+1} is assigned to the value i for which ω_k is in the indicated interval of probabilities.

Our Markovian state evolution function f, given by $\vec{x}_{k+1} = f(\vec{x}_k, \vec{u}_k, \vec{w}_k)$, can be written explicitly as

$$\begin{pmatrix} x_{k+1} \\ s_{k+1} \\ \boldsymbol{z}_{k+1} \\ q_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ s_k + \eta_{k+1,T(\boldsymbol{z}_k,\omega_k)} \\ T(\boldsymbol{z}_k, \omega_k) \\ q_k \end{pmatrix} + \begin{pmatrix} s_k + \xi + \delta_k^- \\ 0 \\ 0 \\ -1 \end{pmatrix} L_k^- + \begin{pmatrix} -(s_k - \xi - \delta_k^+) \\ 0 \\ 0 \\ 1 \end{pmatrix} L_k^+ \tag{3.53}$$

The cash process at a subsequent timestep is equal to the cash at the previous step, plus the profits and costs of executing market and/or limit orders. At time k, if the agent posts a sell limit order that gets filled "between timesteps" k and $k+1$ (depending on the binary random variable L_k^-, itself depending on the binary random variable K_k^+), the revenue depends on the stock price at k. This is consistent with reality as with backtesting: while we are choosing to model the posting *depth*, in reality a submitted limit order has a specific price specified - thus once the order is submitted at k, the potential cash received is fixed.

Our impulse control at every time step is given by

$$\begin{pmatrix} x_k \\ s_k \\ \boldsymbol{z}_k \\ q_k \end{pmatrix} = \begin{pmatrix} x_k \\ s_k \\ \boldsymbol{z}_k \\ q_k \end{pmatrix} + \begin{pmatrix} s_k - \xi \\ 0 \\ 0 \\ -1 \end{pmatrix} M_k^- + \begin{pmatrix} -(s_k + \xi) \\ 0 \\ 0 \\ 1 \end{pmatrix} M_k^+ \tag{3.54}$$

Our market orders assume immediate execution, and are assumed to be sufficiently small in volume so as to not affect order imbalance or the midprice.

3.2.1 Dynamic Programming

The system formulation allows both continuous and impulse control to mimic what was done in the continuous time section, though in discrete time there is no *a priori* distinction between the two [Bensoussan, 2008]. The following theorem shows that in this case a quasi-variational inequality formulation does exist, and that it is equivalent to the standard dynamic programming formulation. The result is a simplified expression that mirrors the continuous time analysis.

Theorem 7 ([Bensoussan, 2008])**. *Dynamic Programming with Impulse Control in Discrete Time.*** *Consider a controlled Markov Chain with state space $X = \mathbb{R}^d$, tran-*

sition probability $\pi(x, v, \mathrm{d}\eta)$, and positive, bounded, uniformly continuous cost function $l(x, v)$.

Introduce an impulse control w. Define the extended cost function by $l(x, v, w) = l(x + w, v) + c(w)$, the extended transition probability by $\pi(x, v, w, \mathrm{d}\eta) = \pi(x + w, v, \mathrm{d}\eta)$ with the associated operator $\Phi^{v,w} f(x) = \int_{\mathbb{R}^d} f(\eta)\pi(x, v, w, \mathrm{d}\eta) = \Phi^v f(x + w)$.

Consider a decision rule V, W with associated probability $\mathbb{P}^{V,W,x}$ on $\Omega, \mathcal{A}$ for which $y_1 = x$ a.s. Consider the pay-off function

$$J_x(V, W) = \mathbb{E}^{V,W,x}\left[\sum_{n=1}^{\infty} \alpha^{n-1} l(y_n, v_n, w_n)\right] \tag{3.55}$$

and the corresponding Bellman equation

$$u(x) = \inf_{\substack{v \in U \\ w \geq 0}} [l(x + w, v) + c(w) + \alpha \Phi^v u(x + w)] \tag{3.56}$$

Assume:

1. *$\Phi^V \phi_v(x)$ is continuous in v, x if $\phi_v(x) = \phi(x, v)$ is uniformly continuous and bounded in x, v;*
2. *$c(w) = K\mathbb{1}_{w=0} + c_0(w), \quad c_0(0) = 0, \qquad c_0(w) \to \infty$ as $|w| \to \infty$, $c_0(w)$ is sub-linear positive continuous;*
3. *U is compact.*

Then there exists a unique, positive, bounded solution of Equation (3.56) belonging to the space of uniformly continuous and bounded functions. Further, this solution is identical to that of

$$u(x) = \min\left\{K + \inf_{w \geq 0}[c_0(w) + u(x + w)]\ ;\ \inf_{v \in U}[l(x, v) + \alpha \Phi^v u(x)]\right\} \tag{3.57}$$

3.2.2 Maximizing Terminal Wealth (Discrete)

Following the dynamic programming with impulse control programme, we introduce the value function $V_k^{\delta^{\pm}}$. Here, as in the continuous-time formulation, our objective is to maximize the terminal wealth performance criteria given by

$$V_k^{\delta\pm}(x,s,\boldsymbol{z},q) = \mathbb{E}\left[W_T^{\delta\pm}\right] = \mathbb{E}_{k,x,s,\boldsymbol{z},q}\left[X_T^{\delta\pm} + Q_T^{\delta\pm}(S_T - \operatorname{sgn}(Q_T^{\delta\pm})\xi) - \alpha(Q_T^{\delta\pm})^2\right] \tag{3.58}$$

where, as before, the notation $\mathbb{E}_{k,x,s,\boldsymbol{z},q}[\,\cdot\,]$ represents the conditional expectation

$$\mathbb{E}[\,\cdot \mid X_k = x, S_k = s, \boldsymbol{Z}_k = \boldsymbol{z}, Q_k = q]$$

In this case, our dynamic programming equations (DPEs) are given by

$$V_T(x,s,\boldsymbol{z},q) = x + q(s - \operatorname{sgn}(q)\xi) - \alpha q^2 \tag{3.59}$$

$$\begin{aligned} V_k(x,s,\boldsymbol{z},q) = \max\Big\{ &\sup_{\delta\pm}\big\{\mathbb{E}_{\mathbf{w}}\left[V_{k+1}(f((x,s,\boldsymbol{z},q),\boldsymbol{u},\mathbf{w}_k)\right]\big\} \ ; \\ &V_k(x + s_k - \xi, s_k, \boldsymbol{z}_k, q_k - 1) \ ; \\ &V_k(x - s_k - \xi, s_k, \boldsymbol{z}_k, q_k + 1)\Big\} \end{aligned} \tag{3.60}$$

where expectation is with respect to the random vector $\mathbf{w}_k$. Note that in this formulation we do not have per stage costs, as the cost of execution is bundled into the state x. Nevertheless, it is rather immediate that the execution costs could be disentangled from the system state and seen to satisfy the theorem assumptions. Hypothetically we could add the fourth case where $M^+ = M^- = 1$, though a quick substitution shows that it is always strictly 2ξ less in value than the case of only limit orders, where $M^+ = M^- = 0$. This should be evident, as buying and selling with market orders in a single timestep yields a guaranteed loss as the agent is forced to cross the spread.

To simplify the DPEs, we introduce a now familiar ansatz:

$$V_k(x,s,\boldsymbol{z},q) = x + q(s - \operatorname{sgn}(q)\xi) + h_k(\boldsymbol{z},q) \tag{3.61}$$

with boundary condition $h_k(\boldsymbol{z},0) = 0$ and terminal condition $h_T(\boldsymbol{z},q) = -\alpha q^2$. Substituting this ansatz into the Equation (3.60), we obtain

$$\begin{aligned} 0 = \max\Big\{ &\sup_{\delta\pm}\big\{\mathbb{E}_{\mathbf{w}}\left[V_{k+1}(f((x,s,\boldsymbol{z},q),\boldsymbol{u},\mathbf{w}_k)\right] - V_k(x,s,\boldsymbol{z},q)\big\} \ ; \\ &V_k(x + s_k - \xi, s_k, \boldsymbol{z}_k, q_k - 1) - V_k(x,s,\boldsymbol{z},q) \ ; \\ &V_k(x - s_k - \xi, s_k, \boldsymbol{z}_k, q_k + 1) - V_k(x,s,\boldsymbol{z},q)\Big\} \end{aligned} \tag{3.62}$$

$$
\begin{aligned}
0 = \max\Bigg\{ & \sup_{\delta^\pm}\Big\{\mathbb{E}_{\mathbf{w}}\big[(s+\xi+\delta^-)L_k^- - (s-\xi-\delta^+)L_k^+ \\
&\qquad + (L_k^+ - L_k^-)\big(s + \eta_{0,T(\boldsymbol{z},\omega)} - \operatorname{sgn}(q + L_k^+ - L_k^-)\xi\big) \\
&\qquad + q\left(\eta_{0,T(\boldsymbol{z},\omega)} - \big(\operatorname{sgn}(q + L_k^+ - L_k^-) - \operatorname{sgn}(q)\big)\,\xi\right) \\
&\qquad + h_{k+1}(T(\boldsymbol{z},\omega), q + L_k^+ - L_k^-) - h_k(\boldsymbol{z},q)\big]\Big\}\ ; \\
& -2\xi\cdot\mathbb{1}_{q\geq 0} + h_k(\boldsymbol{z},q+1)\ ; \\
& -2\xi\cdot\mathbb{1}_{q\leq 0} + h_k(\boldsymbol{z},q-1)\Bigg\}
\end{aligned} \tag{3.63}
$$

We'll begin by concentrating on the first term in the quasi-variational inequality. Thus, we want to solve

$$
\begin{aligned}
\sup_{\delta^\pm}\Bigg\{\mathbb{E}_{\mathbf{w}}\Bigg[& (s+\xi+\delta^-)L_k^- - (s-\xi-\delta^+)L_k^+ \\
&\qquad + (L_k^+ - L_k^-)\big(s + \eta_{0,T(\boldsymbol{z},\omega)} - \operatorname{sgn}(q + L_k^+ - L_k^-)\xi\big) \\
&\qquad + q\left(\eta_{0,T(\boldsymbol{z},\omega)} - \big(\operatorname{sgn}(q + L_k^+ - L_k^-) - \operatorname{sgn}(q)\big)\,\xi\right) \\
&\qquad + h_{k+1}(T(\boldsymbol{z},\omega), q + L_k^+ - L_k^-) - h_k(\boldsymbol{z},q)\Bigg]\Bigg\}
\end{aligned} \tag{3.64}
$$

As other agents' market orders as Poisson distributed, we have that

$$
[K_k^+ = 0] = \frac{e^{-\mu^+(\boldsymbol{z})\Delta t}(\mu^+(\boldsymbol{z})\Delta t)^0}{0!} = e^{-\mu^+(\boldsymbol{z})\Delta t} \tag{3.65}
$$

and so the probability of seeing some positive number of market orders is

$$
\mathbb{P}[K_k^+ > 0] = 1 - e^{-\mu^+(\boldsymbol{z})\Delta t} \tag{3.66}
$$

Now we make the simplified assumption that the *aggregate* of the orders walks the limit order book to a depth of p_k, and if $p_k > \delta^-$, then our sell limit order is lifted. As in the continuous time section, we will assume that the probability of our order being lifted is $e^{-\kappa\delta^-}$. Thus we have the following preliminary results:

$$
\mathbb{P}[L_k^- = 1 | K_k^+ > 0] = e^{-\kappa\delta^-} \tag{3.67}
$$

$$
\mathbb{P}[L_k^- = 0 | K_k^+ > 0] = 1 - e^{-\kappa\delta^-} \tag{3.68}
$$

$$
\begin{aligned}
\mathbb{E}[L_k^-] &= \mathbb{P}[L_k^- = 1 | K_k^+ > 0]\cdot\mathbb{P}[K_k^+ > 0] && (3.69)\\
&= (1 - e^{-\mu^+(\boldsymbol{z})\Delta t})e^{-\kappa\delta^-} && (3.70)
\end{aligned}
$$

For ease of notation, we'll write the probability of the $L_k^- = 1$ event as $p(\delta^-)$. This gives us the additional results:

$$\begin{aligned}
\mathbb{P}[L_k^- = 1] &= p(\delta^-) = \mathbb{E}[L_k^-] && (3.71)\\
\mathbb{P}[L_k^- = 0] &= 1 - p(\delta^-) && (3.72)\\
\partial_{\delta^-}\mathbb{P}[L_k^- = 1] &= -\kappa p(\delta^-) && (3.73)\\
\partial_{\delta^-}\mathbb{P}[L_k^- = 0] &= \kappa p(\delta^-) && (3.74)
\end{aligned}$$

Let's pre-compute some of the terms that we'll encounter in the supremum, namely the expectations of the random variables. To each we will assign an uppercase Greek letter as shorthand, as will be evident from the analysis.

$$\begin{aligned}
\mathbb{E}[\operatorname{sgn}(q + L_k^+ - L_k^-)] &= \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 1] \cdot \operatorname{sgn}(q)\\
&\quad + \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 0] \cdot \operatorname{sgn}(q-1)\\
&\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 1] \cdot \operatorname{sgn}(q+1)\\
&\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 0] \cdot \operatorname{sgn}(q) && (3.75)\\
&= p(\delta^-)p(\delta^+)\operatorname{sgn}(q)\\
&\quad + p(\delta^-)(1 - p(\delta^+))\operatorname{sgn}(q-1)\\
&\quad + (1 - p(\delta^-))p(\delta^+)\operatorname{sgn}(q+1)\\
&\quad + (1 - p(\delta^-))(1 - p(\delta^+))\operatorname{sgn}(q) && (3.76)\\
&= \operatorname{sgn}(q)\big[1 - p(\delta^+) - p(\delta^-) + 2p(\delta^+)p(\delta^-)\big]\\
&\quad + \operatorname{sgn}(q-1)\big[p(\delta^-) - p(\delta^+)p(\delta^-)\big]\\
&\quad + \operatorname{sgn}(q+1)\big[p(\delta^+) - p(\delta^+)p(\delta^-)\big] && (3.77)\\
&= \begin{cases} 1 & q \geq 2\\ 1 - p(\delta^-)(1 - p(\delta^+)) & q = 1\\ p(\delta^+) - p(\delta^-) & q = 0\\ -\big[1 - p(\delta^+)(1 - p(\delta^-))\big] & q = -1\\ -1 & q \leq -2 \end{cases} && (3.78)\\
&= \Phi(q, \delta^+, \delta^-) && (3.79)
\end{aligned}$$

Similarly:

$$
\begin{aligned}
\mathbb{E}[L_k^+ \operatorname{sgn}(q + L_k^+ - L_k^-)] &= \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 1] \cdot \operatorname{sgn}(q) \\
&\quad + \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 0] \cdot 0 \operatorname{sgn}(q-1) \\
&\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 1] \cdot \operatorname{sgn}(q+1) \\
&\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 0] \cdot 0 \operatorname{sgn}(q) && (3.80) \\
&= p(\delta^+)\big[p(\delta^-) \operatorname{sgn}(q) + (1 - p(\delta^-) \operatorname{sgn}(q+1)\big] && (3.81) \\
&= p(\delta^+) \begin{cases} 1 & q \geq 2 \\ 1 & q = 1 \\ (1 - p(\delta^-)) & q = 0 \\ -p(\delta^-) & q = -1 \\ -1 & q \leq -2 \end{cases} && (3.82) \\
&= p(\delta^+)\Psi(q, \delta^-) && (3.83)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}[L_k^- \operatorname{sgn}(q + L_k^+ - L_k^-)] &= p(\delta^-)\big[p(\delta^+) \operatorname{sgn}(q) + (1 - p(\delta^+)) \operatorname{sgn}(q-1)\big] && (3.84) \\
&= p(\delta^-) \begin{cases} 1 & q \geq 2 \\ p(\delta^+) & q = 1 \\ -(1 - p(\delta^+)) & q = 0 \\ -1 & q = -1 \\ -1 & q \leq -2 \end{cases} && (3.85) \\
&= p(\delta^-)\Upsilon(q, \delta^+) && (3.86)
\end{aligned}
$$

We'll also require the partial derivatives of these expectations, which we can easily compute. Below we'll use the simplified notation Φ_+ to denote the function closely associated with the partial derivative of Φ with respect to δ^+.

$$\partial_{\delta^-}\mathbb{E}[\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^-}\Phi(q,\delta^+,\delta^-) = \kappa p(\delta^-)\begin{cases} 0 & q\geq 2 \\ (1-p(\delta^+)) & q=1 \\ 1 & q=0 \\ p(\delta^+) & q=-1 \\ 0 & q\leq -2 \end{cases} \tag{3.87}$$

$$= \kappa p(\delta^-)\Phi_-(q,\delta^+) \tag{3.88}$$

$$\partial_{\delta^+}\mathbb{E}[\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^+}\Phi(q,\delta^+,\delta^-) = \kappa p(\delta^+)\begin{cases} 0 & q\geq 2 \\ -p(\delta^-) & q=1 \\ -1 & q=0 \\ -(1-p(\delta^-)) & q=-1 \\ 0 & q\leq -2 \end{cases} \tag{3.89}$$

$$= \kappa p(\delta^+)\Phi_+(q,\delta^-) \tag{3.90}$$

$$\partial_{\delta^-}\mathbb{E}[L_k^+\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^-}p(\delta^+)\Psi(q,\delta^-) = \kappa p(\delta^+)p(\delta^-)\begin{cases} 0 & q\geq 2 \\ 0 & q=1 \\ 1 & q=0 \\ 1 & q=-1 \\ 0 & q\leq -2 \end{cases} \tag{3.91}$$

$$= \kappa p(\delta^+)p(\delta^-)\Psi_-(q) \tag{3.92}$$

$$\partial_{\delta^+}\mathbb{E}[L_k^+\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^+}p(\delta^+)\Psi(q,\delta^-) = -\kappa p(\delta^+)\Psi(q,\delta^-) \tag{3.93}$$

$$\partial_{\delta^-}\mathbb{E}[L_k^-\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^-}p(\delta^-)\Upsilon(q,\delta^+) = -\kappa p(\delta^-)\Upsilon(q,\delta^+) \tag{3.94}$$

$$\partial_{\delta^+}\mathbb{E}[L_k^-\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^+}p(\delta^-)\Upsilon(q,\delta^+) = \kappa p(\delta^+)p(\delta^-)\begin{cases} 0 & q\geq 2 \\ -1 & q=1 \\ -1 & q=0 \\ 0 & q=-1 \\ 0 & q\leq -2 \end{cases} \tag{3.95}$$

$$= \kappa p(\delta^+)p(\delta^-)\Upsilon_+(q) \tag{3.96}$$

Recalling that we have $\boldsymbol{P}$ the transition matrix for the Markov Chain $\boldsymbol{Z}$, with $\boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} = \mathbb{P}[\boldsymbol{Z}_{k+1} = \mathbf{j} | \boldsymbol{Z}_k = \boldsymbol{z}]$, then we can also write:

$$\begin{aligned} \mathbb{E}[h_{k+1}(T(\boldsymbol{z},\omega), q + L_k^+ - L_k^-)] = \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \Big[& h_{k+1}(\mathbf{j}, q) \big[1 - p(\delta^+) - p(\delta^-) + 2p(\delta^+)p(\delta^-)\big] \\ & + h_{k+1}(\mathbf{j}, q-1) \big[p(\delta^-) - p(\delta^+)p(\delta^-)\big] \\ & + h_{k+1}(\mathbf{j}, q+1) \big[p(\delta^+) - p(\delta^+)p(\delta^-)\big] \Big] \end{aligned} \tag{3.97}$$

and its partial derivatives as

$$\begin{aligned} \partial_{\delta^-} \mathbb{E}[h_{k+1}(T(\boldsymbol{z},\omega), q + L_k^+ - L_k^-)] = \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \Big[& h_{k+1}(\mathbf{j}, q) \big[\kappa p(\delta^-) - 2\kappa p(\delta^+)p(\delta^-)\big] \\ & + h_{k+1}(\mathbf{j}, q-1) \big[-\kappa p(\delta^-) + \kappa p(\delta^+)p(\delta^-)\big] \\ & + h_{k+1}(\mathbf{j}, q+1) \big[\kappa p(\delta^+)p(\delta^-)\big] \Big] \end{aligned} \tag{3.98}$$

$$\begin{aligned} = \kappa p(\delta^-) \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \Big[& h_{k+1}(\mathbf{j}, q) \big[1 - 2p(\delta^+)\big] \\ & + h_{k+1}(\mathbf{j}, q-1) \big[-1 + p(\delta^+)\big] \\ & + h_{k+1}(\mathbf{j}, q+1) \big[p(\delta^+)\big] \Big] \end{aligned} \tag{3.99}$$

$$\begin{aligned} \partial_{\delta^+} \mathbb{E}[h_{k+1}(T(\boldsymbol{z},\omega), q + L_k^+ - L_k^-)] = \kappa p(\delta^+) \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \Big[& h_{k+1}(\mathbf{j}, q) \big[1 - 2p(\delta^-)\big] \\ & + h_{k+1}(\mathbf{j}, q-1) \big[p(\delta^-)\big] \\ & + h_{k+1}(\mathbf{j}, q+1) \big[-1 + p(\delta^-)\big] \Big] \end{aligned} \tag{3.100}$$

Now we tackle solving the supremum in equation 3.64 and thus finding the optimal posting depths, again denoted by a subscript asterisk. First we consider the first-order condition on δ^-, namely that the partial derivative with respect to it must be equal to zero.

$$\begin{aligned}0 = \partial_{\delta^-} \Big\{ &(s + \xi + \delta^{-*})\mathbb{E}[L_k^-] - (s - \xi - \delta^+)\mathbb{E}[L_k^+] \\ &+ \mathbb{E}[L_k^+] \left(s + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]\right) - \xi\mathbb{E}\left[L_k^+ \operatorname{sgn}(q + L_k^+ - L_k^-)\right] \\ &- \mathbb{E}[L_k^-] \left(s + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]\right) + \xi\mathbb{E}\left[L_k^- \operatorname{sgn}(q + L_k^+ - L_k^-)\right] \\ &+ q\mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - q\xi\mathbb{E}[\operatorname{sgn}(q + L_k^+ - L_k^-)] + q\xi \operatorname{sgn}(q) \\ &+ \mathbb{E}\left[h_{k+1}(T(\boldsymbol{z},\omega), q + L_k^+ - L_k^-)\right] - h_k(\boldsymbol{z}, q)\Big\}\end{aligned} \tag{3.101}$$

$$\begin{aligned}= \partial_{\delta^-} \Big\{ &(s + \xi + \delta^{-*})\mathbb{E}[L_k^-] - \xi\mathbb{E}\left[L_k^+ \operatorname{sgn}(q + L_k^+ - L_k^-)\right] \\ &- \mathbb{E}[L_k^-] \left(s + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]\right) + \xi\mathbb{E}\left[L_k^- \operatorname{sgn}(q + L_k^+ - L_k^-)\right] \\ &- q\xi\mathbb{E}[\operatorname{sgn}(q + L_k^+ - L_k^-)] + \mathbb{E}\left[h_{k+1}(T(\boldsymbol{z},\omega), q + L_k^+ - L_k^-)\right]\Big\}\end{aligned} \tag{3.102}$$

$$\begin{aligned}= &p(\delta^{-*}) - \kappa p(\delta^{-*})(s + \xi + \delta^{-*}) - \xi\kappa p(\delta^+)p(\delta^{-*})\Psi_-(q) \\ &+ \kappa p(\delta^{-*}) \left(s + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]\right) - \xi\kappa p(\delta^{-*})\Upsilon(q, \delta^+) - q\xi\kappa p(\delta^{-*})\Phi_-(q, \delta^+) \\ &+ \kappa p(\delta^{-*}) \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \Big[h_{k+1}(\mathbf{j}, q)\big[1 - 2p(\delta^+)\big] + h_{k+1}(\mathbf{j}, q - 1)\big[-1 + p(\delta^+)\big] \\ &\qquad + h_{k+1}(\mathbf{j}, q + 1)\big[p(\delta^+)\big]\Big]\end{aligned} \tag{3.103}$$

Dividing through by $\kappa p(\delta^{-*})$, which is nonzero, and re-arranging, we find that the optimal sell posting depth is given by

$$\begin{aligned}\delta^{-*} = &\frac{1}{\kappa} + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - \xi\left(1 + p(\delta^+)\Psi_-(q) + \Upsilon(q, \delta^+) + q\Phi_-(q, \delta^+)\right) \\ &+ \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \Big[h_{k+1}(\mathbf{j}, q)\big[1 - 2p(\delta^+)\big] + h_{k+1}(\mathbf{j}, q - 1)\big[-1 + p(\delta^+)\big] + h_{k+1}(\mathbf{j}, q + 1)\big[p(\delta^+)\big]\Big]\end{aligned} \tag{3.104}$$

$$\begin{aligned}= &\frac{1}{\kappa} + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\left(\mathbb{1}_{q\geq 1} + p(\delta^+)\mathbb{1}_{q=0}\right) \\ &+ \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \Big[h_{k+1}(\mathbf{j}, q)\big[1 - 2p(\delta^+)\big] + h_{k+1}(\mathbf{j}, q - 1)\big[-1 + p(\delta^+)\big] + h_{k+1}(\mathbf{j}, q + 1)\big[p(\delta^+)\big]\Big]\end{aligned} \tag{3.105}$$

Recalling that we want $\delta^{\pm} \geq 0$, we find:

$$\begin{aligned}\delta^{-*} = \max\Big\{0\ ;\ &\frac{1}{\kappa} + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\mathbb{1}_{q\geq 1} + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\big[h_{k+1}(\mathbf{j},q) - h_{k+1}(\mathbf{j},q-1)\big] \\ &- p(\delta^{+})\left(2\xi\mathbb{1}_{q=0} - \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\big[h_{k+1}(\mathbf{j},q-1) + h_{k+1}(\mathbf{j},q+1) - 2h_{k+1}(\mathbf{j},q)\big]\right)\Big\}\end{aligned} \tag{3.106}$$

And similarly, the optimal buy posting depth is given by:

$$\begin{aligned}\delta^{+*} = \max\Big\{0\ ;\ &\frac{1}{\kappa} - \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\mathbb{1}_{q\leq -1} + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\big[h_{k+1}(\mathbf{j},q) - h_{k+1}(\mathbf{j},q+1)\big] \\ &- p(\delta^{-})\left(2\xi\mathbb{1}_{q=0} - \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\big[h_{k+1}(\mathbf{j},q-1) + h_{k+1}(\mathbf{j},q+1) - 2h_{k+1}(\mathbf{j},q)\big]\right)\Big\}\end{aligned} \tag{3.107}$$

For ease of notation we'll write $\aleph(q) = \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\left[h_{k+1}(\mathbf{j},q-1) + h_{k+1}(\mathbf{j},q+1) - 2h_{k+1}(\mathbf{j},q)\right]$. Now, assuming we behave optimally on both the buy and sell sides simultaneously, we can substitute equation 3.107 into equation 3.106, while evaluating both at δ^{+*} and δ^{-*} to obtain the optimal posting depth in feedback form:

$$\begin{aligned}\delta^{-*} = &\frac{1}{\kappa} + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\mathbb{1}_{q\geq 1} + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[h_{k+1}(\mathbf{j},q) - h_{k+1}(\mathbf{j},q-1)\Big] \\ &- (1 - e^{\mu^{-}(\boldsymbol{z})\Delta t})e^{-\kappa \max\left\{0\ ;\ \frac{1}{\kappa} - \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\mathbb{1}_{q\leq -1} + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\left[h_{k+1}(\mathbf{j},q) - h_{k+1}(\mathbf{j},q+1)\right] - (1-e^{\mu^{+}(\boldsymbol{z})\Delta t})e^{-\kappa\delta^{-*}}(2\xi\mathbb{1}_{q=0} - \aleph(q))\right\}}\,(2\xi\mathbb{1}_{q=0} - \aleph(q))\end{aligned} \tag{3.108}$$

This equation will need to be solved numerically due to the difficulty in isolating δ^{-*} on one side of the equality. Once a solution has been obtained, the value can be substituted back into Equation (3.107) to solve for δ^{+*}.

3.2.3 Simplifying the DPE

We now turn to simplifying the DPE in Equation (3.63) by substituting in the optimal posting depths as written in recursive form: Equation (3.107) and Equation (3.106). In doing so we see a incredible amount of cancellation and simplification, and we obtain the rather elegant, and surprisingly simple form of the DPE:

$$
\begin{aligned}
h_k(\boldsymbol{z}, q) = \max\Bigg\{ & q\mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] + \frac{1}{\kappa}(p(\delta^{+*}) + p(\delta^{-*})) + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} h_{k+1}(\mathbf{j}, q) \\
& \quad + p(\delta^{+*})p(\delta^{-*}) \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \left[h_{k+1}(\mathbf{j}, q-1) + h_{k+1}(\mathbf{j}, q+1) - 2h_{k+1}(\mathbf{j}, q)\right] \; ; \\
& - 2\xi \cdot \mathbb{1}_{q\geq 0} + h_k(\boldsymbol{z}, q+1) \; ; \\
& - 2\xi \cdot \mathbb{1}_{q\leq 0} + h_k(\boldsymbol{z}, q-1) \Bigg\}
\end{aligned} \tag{3.109}
$$

As was the case in continuous time, Equation (3.109) yields that whilst in the continuation region, we have

$$h_k(\boldsymbol{z}, q) \leq h_k(\boldsymbol{z}, q+1) - 2\xi \cdot \mathbb{1}_{q\geq 0} \tag{3.110}$$

$$h_k(\boldsymbol{z}, q) \leq h_k(\boldsymbol{z}, q-1) - 2\xi \cdot \mathbb{1}_{q\leq 0} \tag{3.111}$$

And these inequalities again give us

$$-2\xi \cdot \mathbb{1}_{q\geq 0} \leq h_k(\boldsymbol{z}, q) - h_k(\boldsymbol{z}, q+1) \leq 2\xi \cdot \mathbb{1}_{q\leq -1} \tag{3.112}$$

$$-2\xi \cdot \mathbb{1}_{q\leq 0} \leq h_k(\boldsymbol{z}, q) - h_k(\boldsymbol{z}, q-1) \leq 2\xi \cdot \mathbb{1}_{q\geq 1} \tag{3.113}$$

$$h_k(\boldsymbol{z}, q) \overset{\substack{\text{sell if } =\\ \downarrow}}{\leq} h_k(\boldsymbol{z}, q+1) \overset{\substack{\text{buy if } =\\ \downarrow}}{\leq} h_k(\boldsymbol{z}, q) + 2\xi, \qquad q \geq 0 \tag{3.114}$$

$$h_k(\boldsymbol{z}, q) \underset{\substack{\uparrow\\ \text{buy if } =}}{\leq} h_k(\boldsymbol{z}, q-1) \underset{\substack{\uparrow\\ \text{sell if } =}}{\leq} h_k(\boldsymbol{z}, q) + 2\xi, \qquad q \leq 0 \tag{3.115}$$

Recalling the boundary condition $h_k(\boldsymbol{z}, 0) = 0$, Equation (3.114) and Equation (3.115) tell us that the function h is non-negative everywhere.

At terminal time T, we liquidate our position at a cost of $(s - xi\,\mathrm{sgn}(q) - \alpha q)$ per share, whereas at $T-1$, we can liquidate at the regular cost of $(s - \xi\,\mathrm{sgn}(q)$. It is thus never optimal to wait until maturity to liquidate the position, and instead we force liquidation one step earlier by setting $h(T-1, \boldsymbol{z}, q) = 0 \; \forall q$. This allows us to effectively ignore the terminal condition, and avoids a contradiction with the finding that $h \geq 0$.

We now have an explicit means of numerically solving for the optimal posting depths. Since we know the function h at the terminal timesteps T and $T-1$, we can take one step back to $T - 2$ and solve for both the optimal posting depths. With these values we are then able to calculate the value function h_{T-2} using Equation (3.109), and in doing so determine whether to execute market orders in addition to posting limit orders. This process then repeats for each step backward.

Chapter 4

Results

In this chapter we explore the dynamics of the continuous time and discrete time models, and perform in-sample and out-of-sample backtests to compare the performance of the continuous and discrete time solutions to the stochastic optimal control problem.

4.1 Calibration

All tests in this chapter were run using the following global set of parameters:

time window for computing price change	Δt_S	1000ms
time window for averaging order imbalance	Δt_I	1000ms
number of imbalance bins	$\#_{bins}$	5
fill probability constant	κ	100

For each daily calibration, we then computed the remaining parameters utilizing the following formulae:

infinitesimal generator matrix	$\boldsymbol{G}$	Equation 2.16
transition probability matrix	$\boldsymbol{P}$	Equation 2.9
market order arrival intensities	$\mu^{\pm}$	Equation 2.21

Additionally, ξ was computed as half of the simple average of the bid-ask spread observed during the trading day, rounded to the nearest half-cent; and the imbalance bins ρ were computed as the partitioning of the interval $[-1, 1]$ into percentile bins symmetric around zero, where the percentile interval was $100 \div \#_{bins}$.

As was mentioned in Chapter 2, the exploratory data analysis done on the data made use of an unorthodox Markov chain, where its state at time t was actually not determinable at time t because the price change $\Delta S(t)$ was computed over the *future* time interval Δt_S. (See Section 2.3.) In the optimal stochastic control formulations, the Markov chain was defined instead such the price change was computed over the *past* interval Δt_S. However, it was of interest to explore what results would be obtained if the calibration was still done using the non $\mathcal{F}$-predictable method. A justification for doing so is that in a given Markov state Z, there is a state-dependent arrival rate of price updates, and there is a state-dependent distribution of jumps when a price update occurs. So that although the price change is measured over the future when calibrating, this really is a way of getting at the state-dependence of those price changes. In the following tests, this calibration method is denoted 'w nFPC', standing for 'with non-$\mathcal{F}$-predictable calibration'.

4.2 Dynamics of $\delta^{\pm *}$

We have solved the same stochastic control problem using both continuous and discrete time, which have yielded markedly different resulting formulae for the optimal limit order posting depths. In this section we explore the calibrated results obtained on an in-sample backtest on `ORCL`, both calibrating and backtesting using data from 2013-05-15. The dynamics of $\delta^{\pm}$ were obtained using a time-to-maturity of 100sec to best depict the behavior as we approached the end of day trading; as the time-to-maturity horizon increases, the postings depths tend to stabilize.

The surge in market orders seen for the Cts w nFPC strategy can be explained by the difference in the $\delta^{\pm}$ plots. From the stochastic analysis chapter, we know that if $q < 0$ and $\delta^{+} = 0$, or if $q > 0$ and $\delta^{+} = 1/\kappa$, then we execute a buy MO. Likewise, if $q < 0$ and $\delta^{-} = 1/\kappa$, or if $q > 0$ and $\delta^{-} = 0$, then we execute a sell MO. In Figure 4.1 we see that we have $\delta^{+} = 0$ for almost all inventory values, and in Figure 4.3 we have $\delta^{+} = 1/\kappa$ for almost all inventory values. This tells us that when the Markov chain state is in one of the non-neutral states, the Cts w nFPC strategy will execute market orders when possible, as it expects prices to move in the corresponding direction.

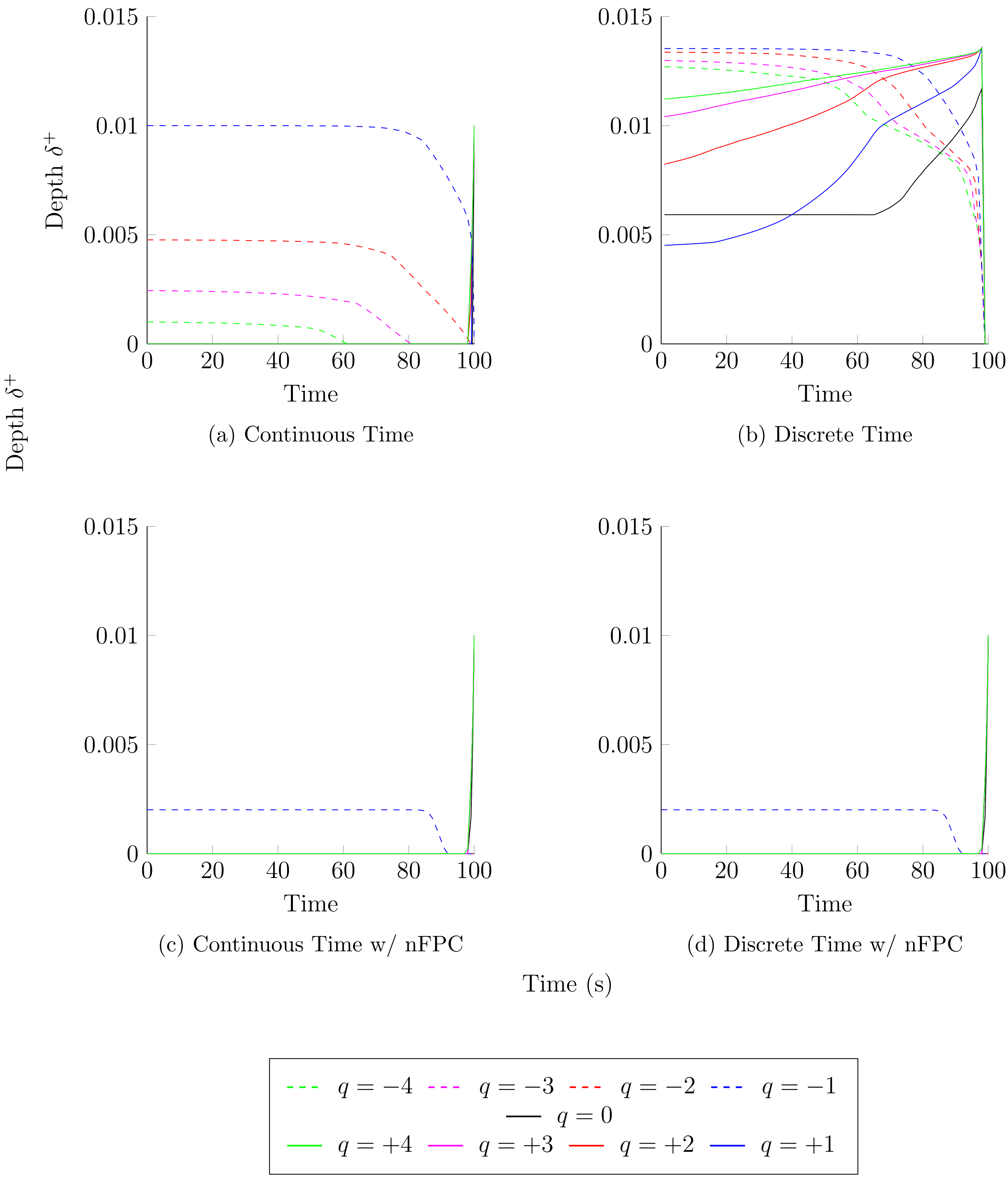


Figure 4.1: Optimal buy depths δ^+ for Markov state $Z = (\rho = -1, \Delta S = -1)$, implying heavy imbalance in favor of sell pressure, and having previously seen a downward price change. We expect the midprice to fall.

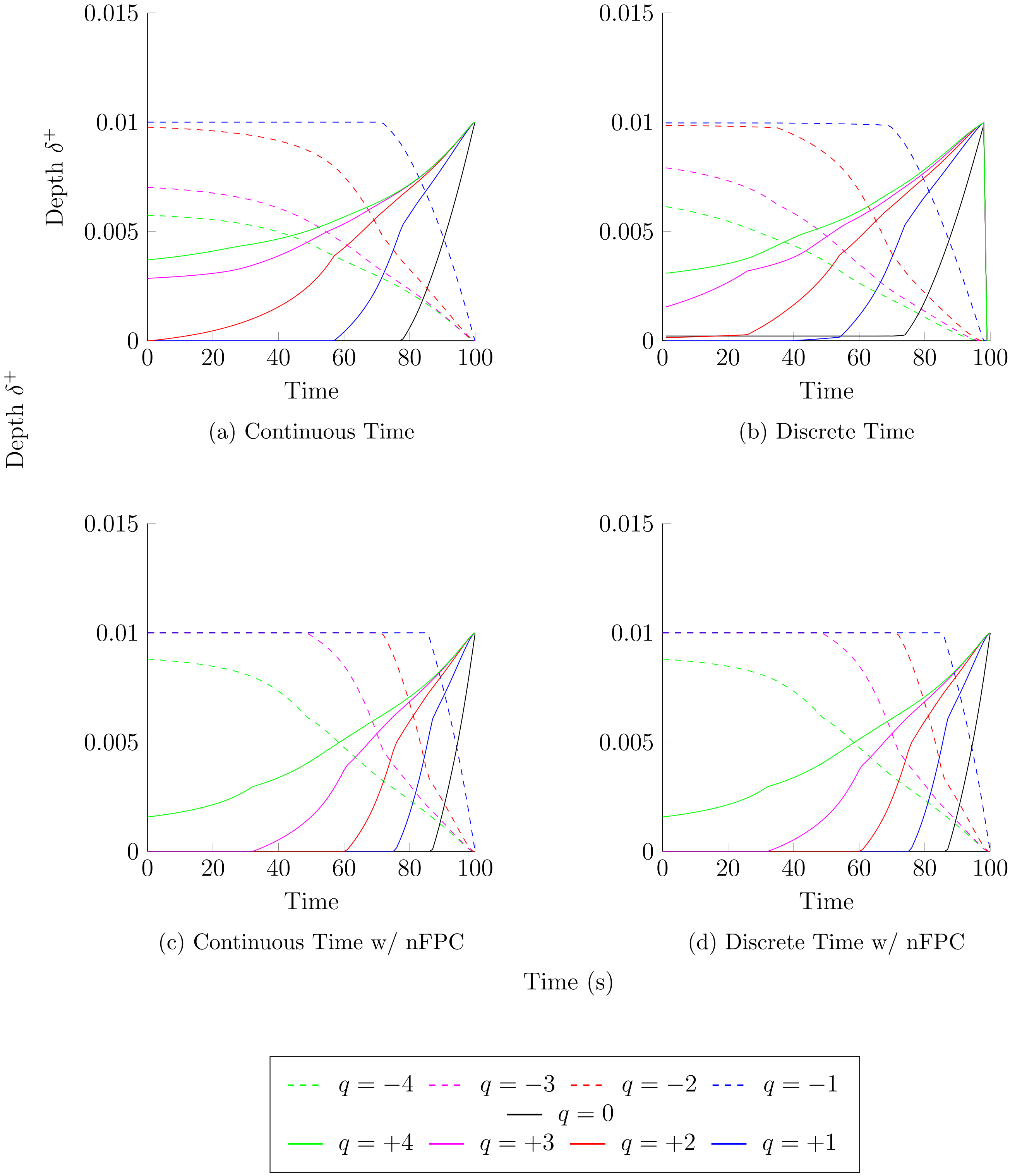
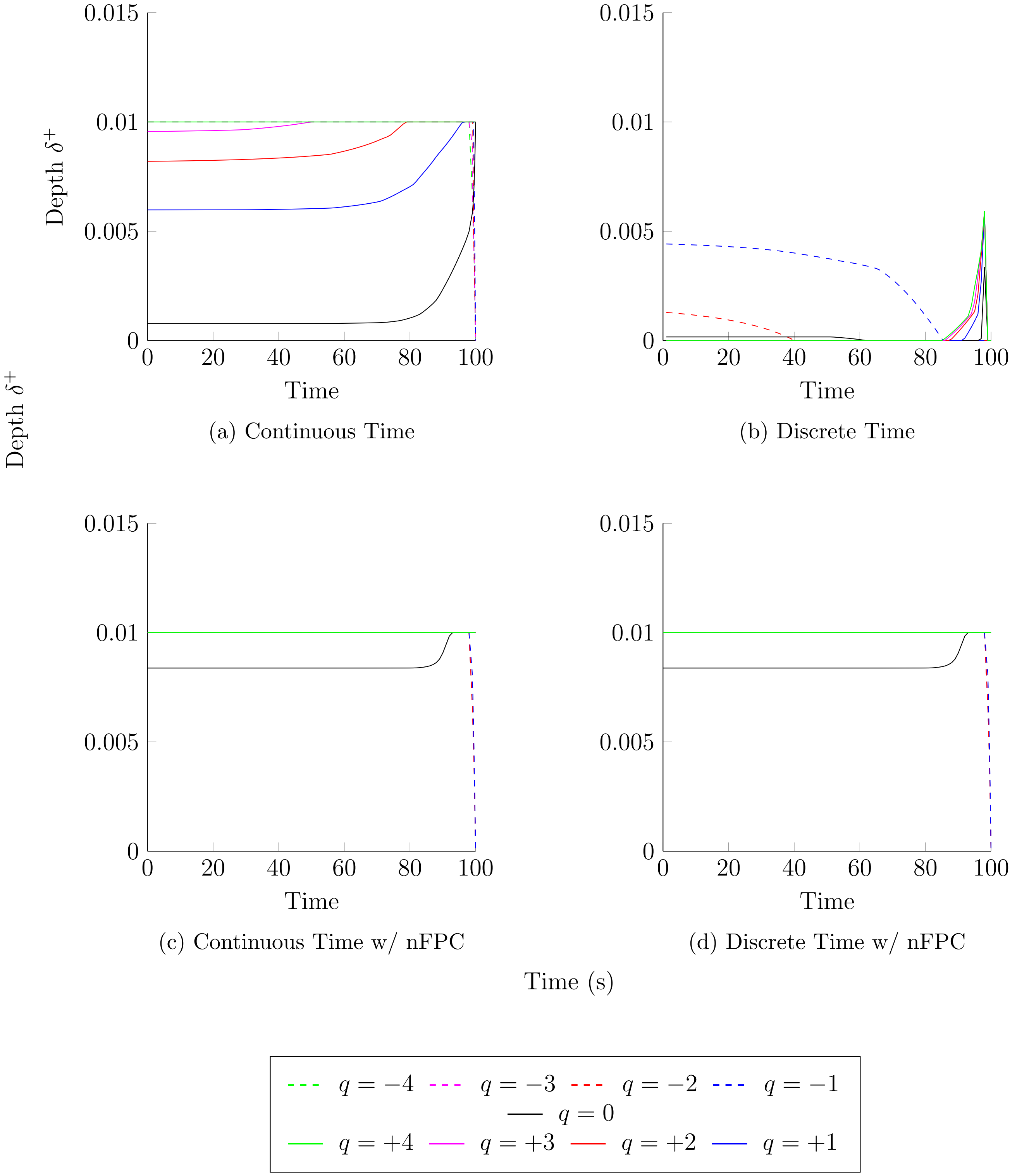


Figure 4.2: Optimal buy depths δ^+ for Markov state $Z = (\rho = 0, \Delta S = 0)$, implying neutral imbalance and no previous price change. We expect no change in midprice.



(a) Continuous Time

(b) Discrete Time

(c) Continuous Time w/ nFPC

(d) Discrete Time w/ nFPC

Figure 4.3: Optimal buy depths δ^+ for Markov state $Z = (\rho = +1, \Delta S = +1)$, implying heavy imbalance in favor of buy pressure, and having previously seen an upward price change. We expect the midprice to rise.

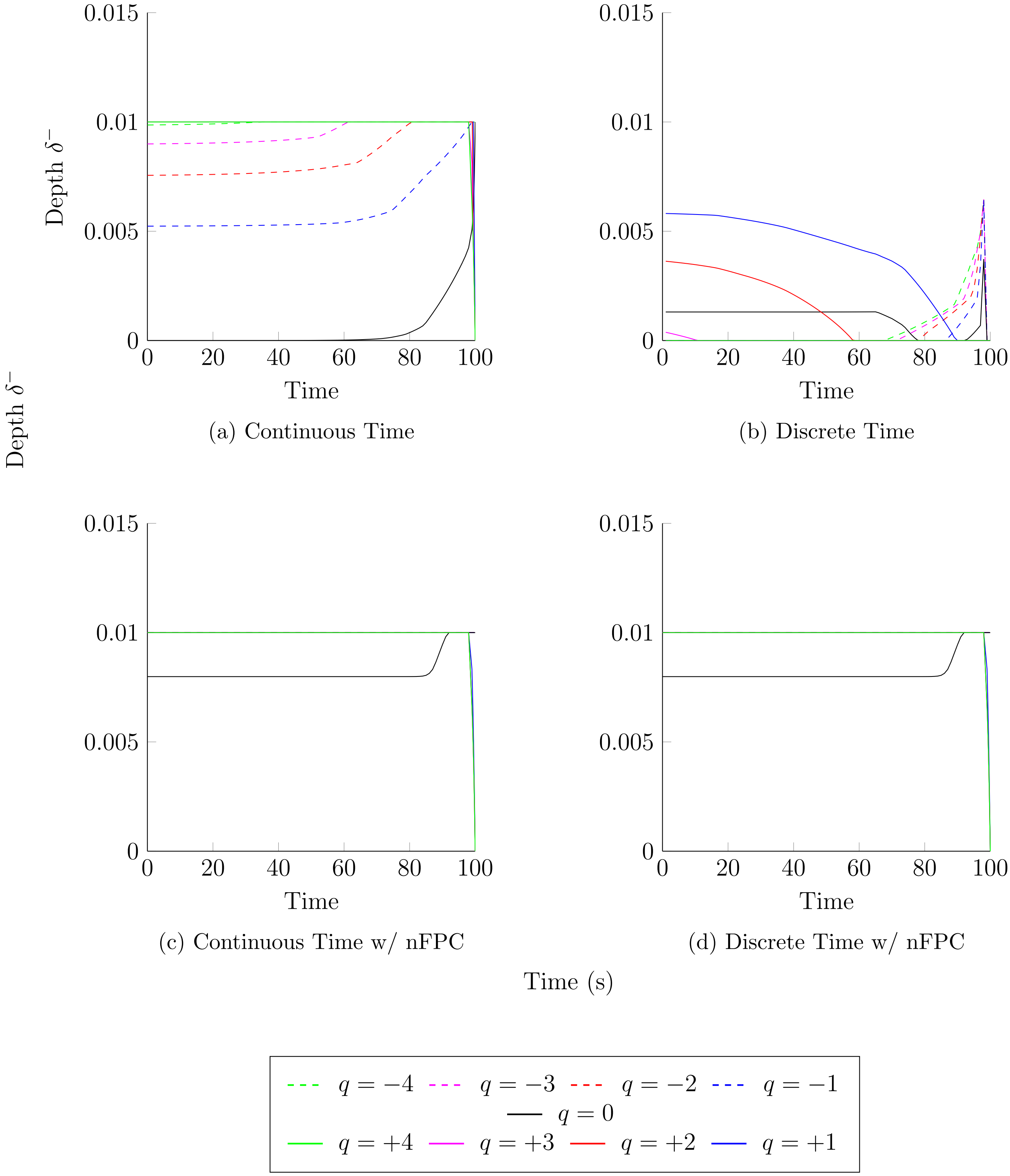


Figure 4.4: Optimal sell depths δ^- for Markov state $Z = (\rho = -1, \Delta S = -1)$, implying heavy imbalance in favor of sell pressure, and having previously seen a downward price change. We expect the midprice to fall.

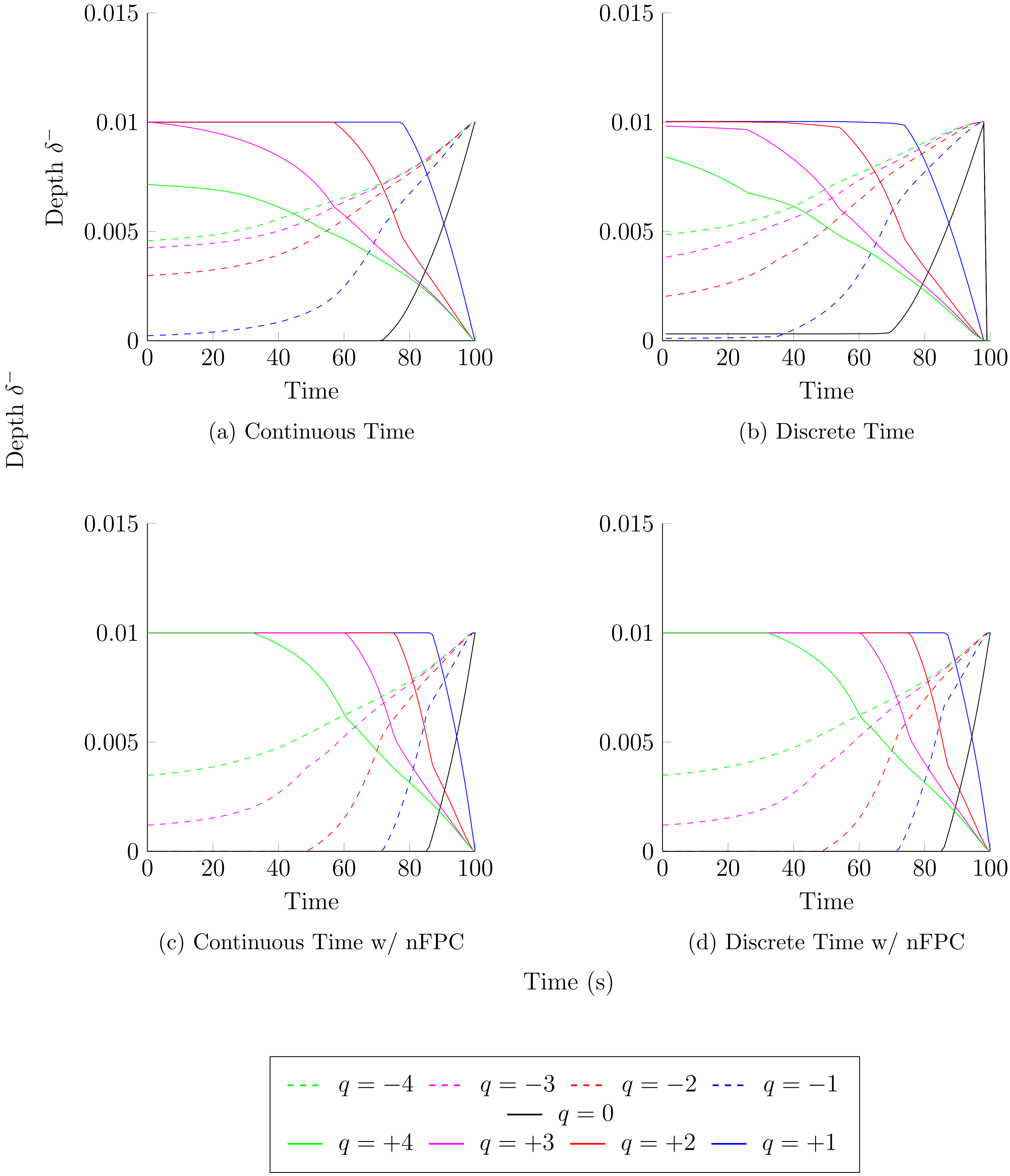
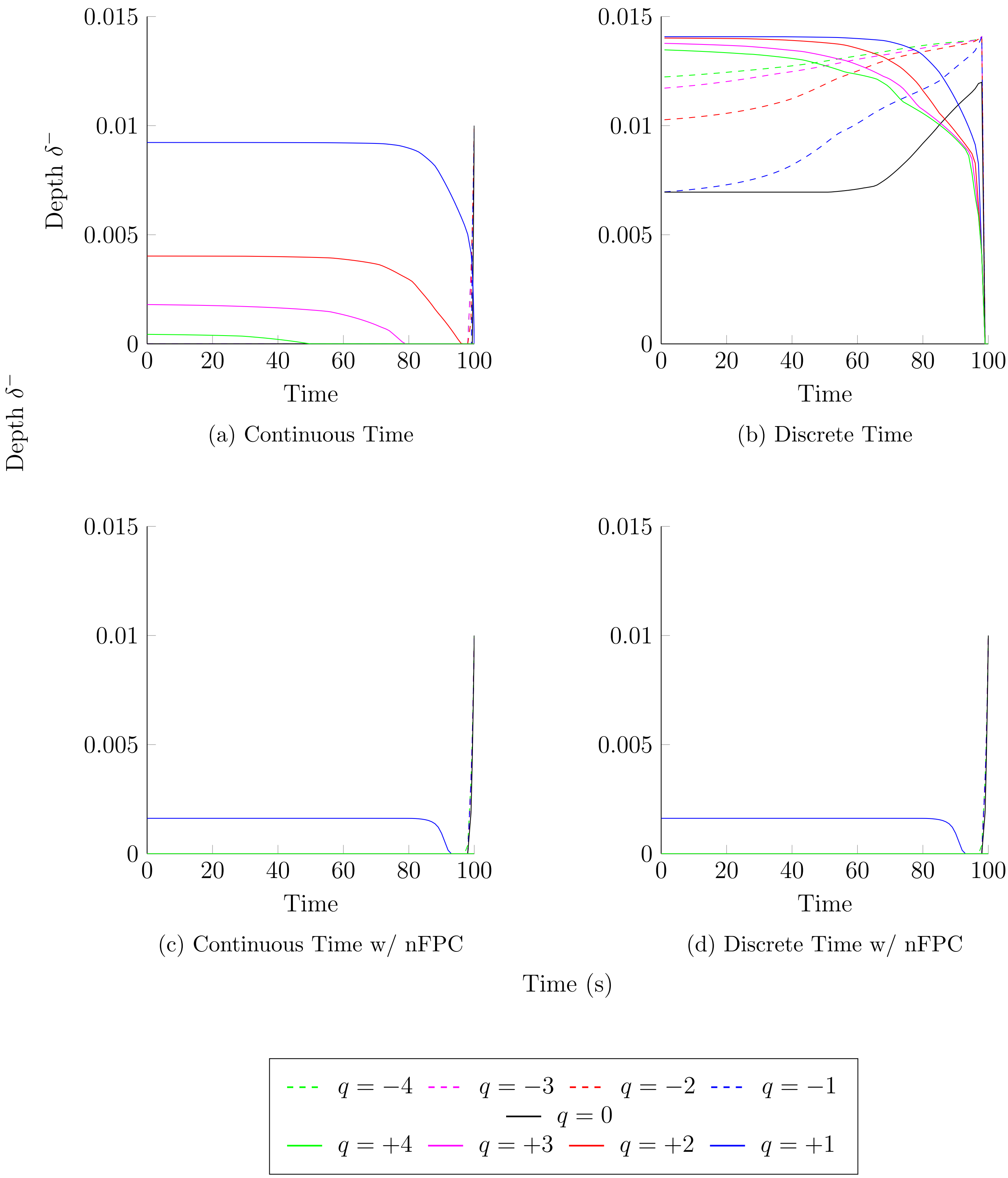


Figure 4.5: Optimal sell depths δ^- for Markov state $Z = (\rho = 0, \Delta S = 0)$, implying neutral imbalance and no previous price change. We expect no change in midprice.



(a) Continuous Time

(b) Discrete Time

(c) Continuous Time w/ nFPC

(d) Discrete Time w/ nFPC

Figure 4.6: Optimal sell depths δ^- for Markov state $Z = (\rho = +1, \Delta S = +1)$, implying heavy imbalance in favor of buy pressure, and having previously seen an upward price change. We expect the midprice to rise.

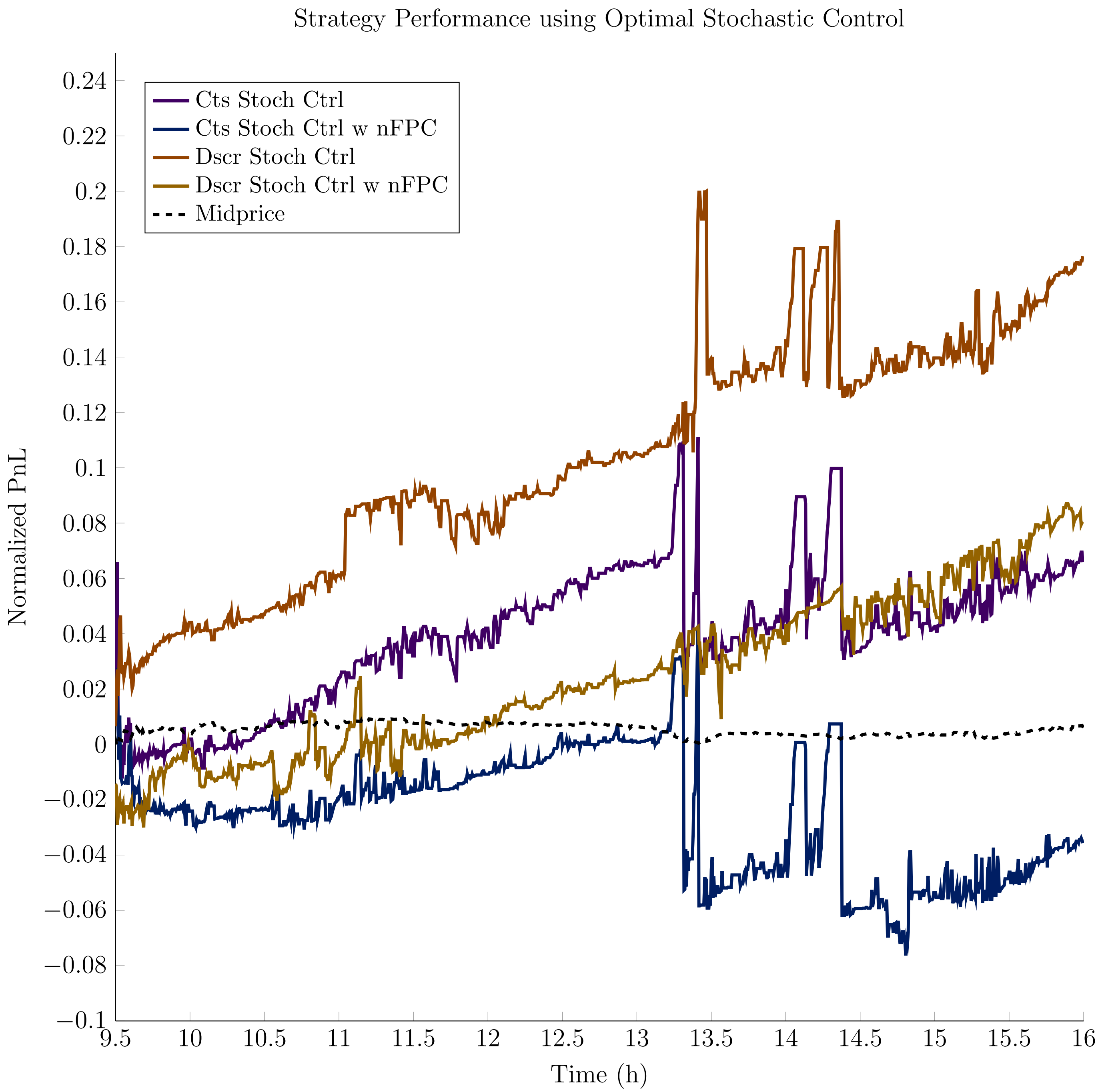
4.2.1 Comparing Performance

In Figure 4.7a we plot the normalized PnL for the four strategies, calibrated and back-tested using data for `ORCL` from 2013-05-15. At a glance the four plots show obvious similarities in trajectory, as well as the distinct spikes between 13h and 14.5h. Nevertheless the correlation of arithmetic returns, Table 4.1, shows that the strategies were uncorrelated. Indeed, while the overall paths are similar, the individual returns to show markedly different behavior.

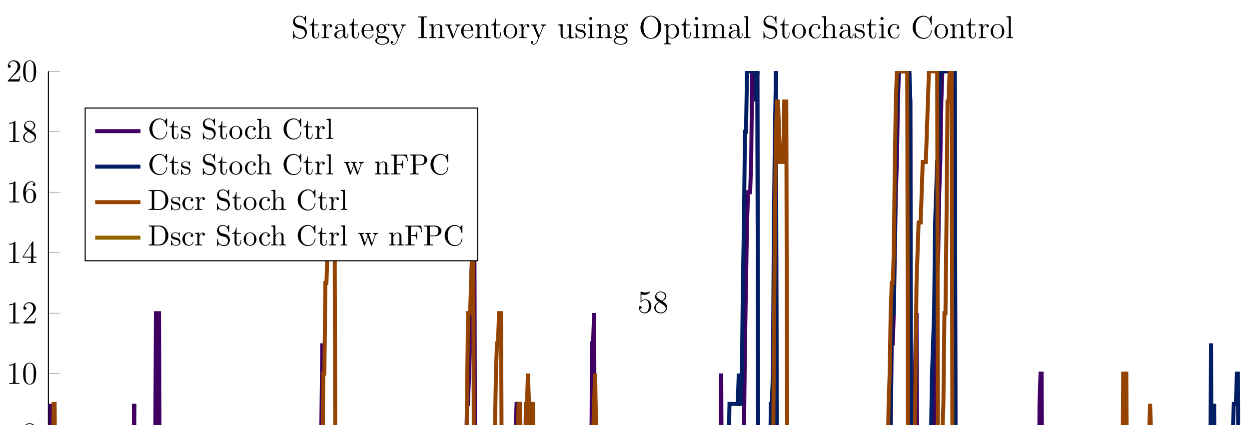
	Cts	Cts w nFPC	Dscr	Dscr w nFPC
Cts	1.0000			
Cts w nFPC	-0.0109	1.0000		
Dscr	-0.0120	0.0122	1.0000	
Dscr w nFPC	-0.0015	0.0034	-0.0165	1.0000

Table 4.1: Correlation of returns

However, we find instead that the returns are co-integrated. On running the Engle-Granger cointegration test with statistics computed using an augmented Dickey-Fuller test of residuals, the τ-test and z-test both returned p-values of 0.001, thus rejecting the null hypothesis of no co-integration. Indeed, the co-integration relation plotted in Figure 4.8 does display stationarity.



(a) Performance comparison of the four stochastic control methods.



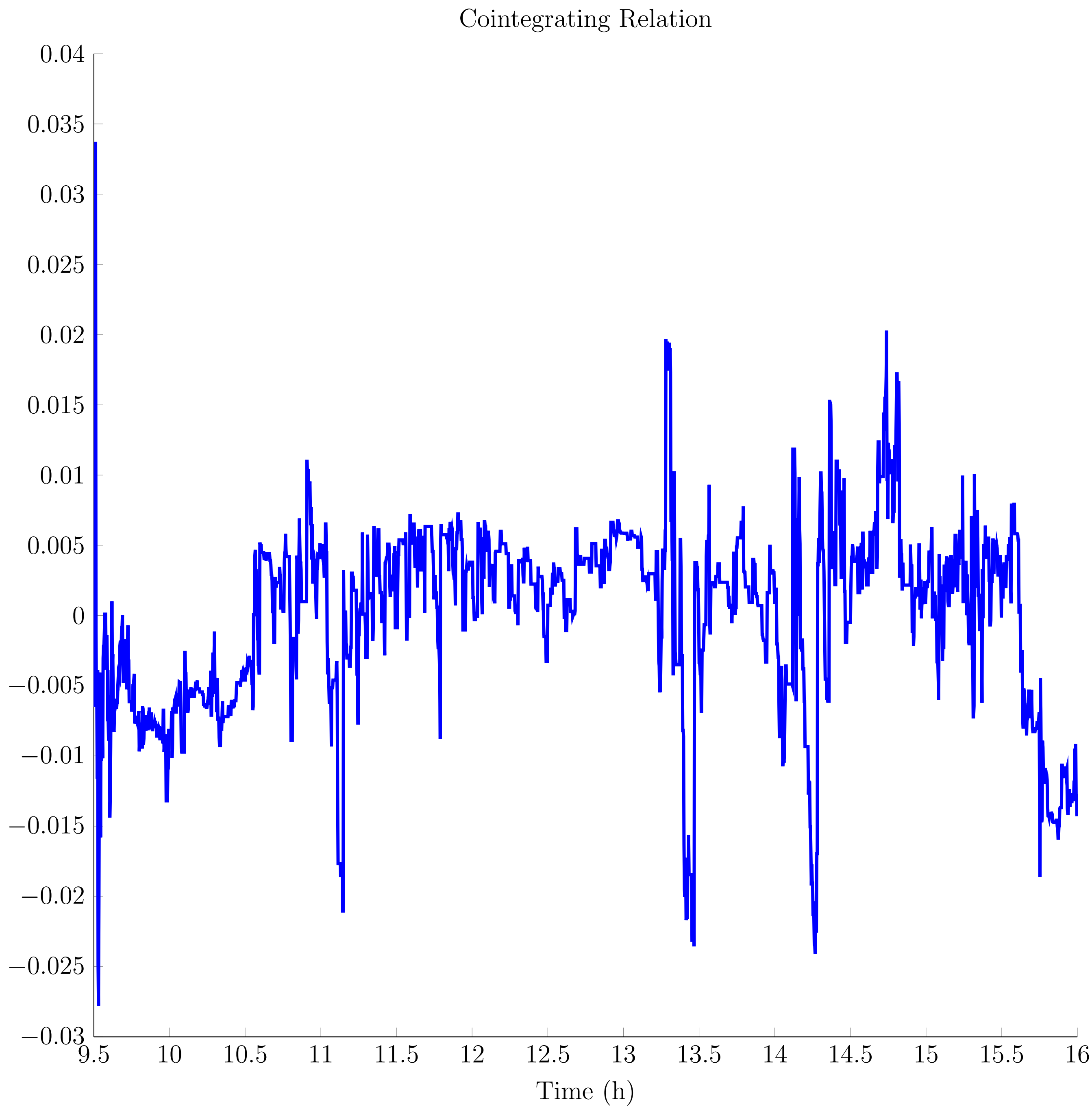


Figure 4.8: Co-integration relation of the four stochastic control methods.

From the inventory plot in Figure 4.7b we see that the strategies always avoid maintaining zero inventory. This is not surprising: in both the discrete and continuous cases we found that the value function ansatz $h(t, z, q)$ was non-negative, and was equal to zero at zero inventory. The interpretation is that there is no added value to having zero

	Market Orders	Limit Orders
Cts	536	1280
Cts w nFPC	1010	1306
Dscr	559	1285
Dscr w nFPC	523	1287

Table 4.2: Number of trades comparison of the four stochastic control methods.

inventory, whereas non-zero inventory can at worst have zero value. Thus it is always profitable, from a value-function standpoint, to have non-zero inventory. Further, we see that the strategies rarely cross the zero-inventory barrier. This is likely attributed to the backtesting algorithm itself, which gives priority to executing buy market orders above sell market orders - we suspect that once the strategy gets into either positive or negative inventory territory, the ansatz function h rarely produces the circumstances to cross the inventory sign barrier by virtue of the non-linear mark-to-market behavior on either side of zero inventory.

Concerning trade execution, the number of executed market orders and filled limit orders generated by each strategy are presented in Table 4.2. The surge in market orders seen for the Cts w nFPC strategy can be explained by the difference in the $\delta^{\pm}$ plots. As mentioned already, from the stochastic analysis chapter, we know that if $q < 0$ and $\delta^{+} = 0$, or if $q > 0$ and $\delta^{+} = 1/\kappa$, then we execute a buy MO. Likewise, if $q < 0$ and $\delta^{-} = 1/\kappa$, or if $q > 0$ and $\delta^{-} = 0$, then we execute a sell MO. In Figure 4.1 we see that we have $\delta^{+} = 0$ for almost all inventory values, and in Figure 4.3 we have $\delta^{+} = 1/\kappa$ for almost all inventory values. This tells us that when the Markov chain state is in one of the non-neutral states, the Cts w nFPC strategy will execute market orders when possible, as it expects prices to move in the corresponding direction.

4.3 In-Sample Backtesting

4.3.1 Same-Day Calibration

We begin our in-sample backtesting same-day calibration: calibration was run for each ticker and each trading day of 2013, and backtesting was then done for each strategy using the same day's calibration. Each backtest would yield the end of day PnL, average inventory held during the day, and the number of executed market orders and filled limit orders. In Table 4.3 we show performance values for several metrics of interest, while Figure 4.9 compares the day-over-day performance of the various strategies.

Since we are calibrating and backtesting using the same underlying data, the calibration should be best attuned to the price dynamics for that praticular day, and hence we expect the performance using same-day calibration to exceed that of the weekly offset calibration and the annual calibration (detailed in the sections that follow). Looking at the % Win column in Table 4.3 we see that trading on `FARO` very rarely produces positive PnL. This is not surprising and was mentioned at the conclusion of the exploratory data analysis chapter: `FARO` is highly illiquid, with daily volume hovering around 200k, and its bid-ask spread averages approximately 20 cents. This makes it never profitable to execute market orders (due to crossing the spread), and because our optimal strategies still force us to post limit orders at depths between 0 and $1/\kappa = 0.01 = 1$cent, the most probable occurrence is that our limit orders are lifted adversely. `NTAP` seems to be a borderline case for liquidity, with average volumes around 4m, and here the strategies exhibit weak regularity of profits. The most liquid stocks, `ORCL` and `NTAP`, with average volumes around 15m and 30m respectively, post extremely promising results: the stochastic control strategies produce positive EOD PnL more than 90% of the time. The discrete time controller outperforms its continuous time counterpart, and in particular we highlight that in the case of `INTC`, we attain a very good Sharpe ratio of 2.5.

Strategy	Return	Sharpe	Trades	Inv	% Win	Max Loss	Max Win
FARO							
Naive	-0.879	-0.808	413	0.47	0.07	-7.109	5.715
Naive+	0.101	0.107	213	2.45	0.74	-8.797	5.336
Naive++	0.002	0.021	7	0.17	0.50	-0.842	0.320
Cont	-0.059	-0.551	201	0.09	0.18	-0.912	0.071
Dscr	-0.064	-0.695	210	-0.02	0.08	-0.914	0.440
Cont w nFPC	-0.063	-0.571	204	0.08	0.14	-1.161	0.077
Dscr w nFPC	-0.060	-0.662	209	-0.03	0.09	-0.716	0.539
NTAP							
Naive	-0.188	-0.316	842	-9.81	0.23	-3.238	3.524
Naive+	0.388	0.169	3562	-9.73	0.74	-19.367	10.201
Naive++	-0.005	-0.012	157	-0.90	0.54	-2.888	2.558
Cont	-0.006	-0.062	2265	0.40	0.56	-0.441	0.215
Dscr	0.099	0.767	1872	4.74	0.86	-0.126	1.042
Cont w nFPC	-0.141	-0.951	2897	0.65	0.14	-0.935	0.244
Dscr w nFPC	0.121	0.881	1738	2.82	0.89	-0.139	0.962
ORCL							
Naive	-0.105	-0.253	484	1.40	0.28	-1.837	2.180
Naive+	-0.034	-0.011	4086	-55.18	0.61	-17.501	18.400
Naive++	0.002	0.006	132	0.61	0.52	-0.798	2.636
Cont	0.115	1.348	1874	1.94	0.92	-0.217	0.521
Dscr	0.135	1.620	1898	3.93	0.98	-0.063	0.515
Cont w nFPC	-0.010	-0.100	2455	1.32	0.48	-0.478	0.503
Dscr w nFPC	0.144	1.501	1759	2.85	0.97	-0.032	0.573
INTC							
Naive	-0.082	-0.228	258	-5.21	0.33	-1.465	1.425
Naive+	0.365	0.134	3962	-32.50	0.63	-11.202	11.669
Naive++	-0.001	-0.003	74	-0.84	0.48	-1.314	1.264
Cont	0.214	2.159	1577	5.17	0.97	-0.213	0.487
Dscr	0.232	2.528	1642	4.48	0.98	-0.217	0.611
Cont w nFPC	0.114	1.218	1894	2.01	0.90	-0.244	0.416
Dscr w nFPC	0.226	2.202	1569	4.28	0.98	-0.220	0.650

Table 4.3: Strategy performance results: in-sample backtesting using same-day calibration.

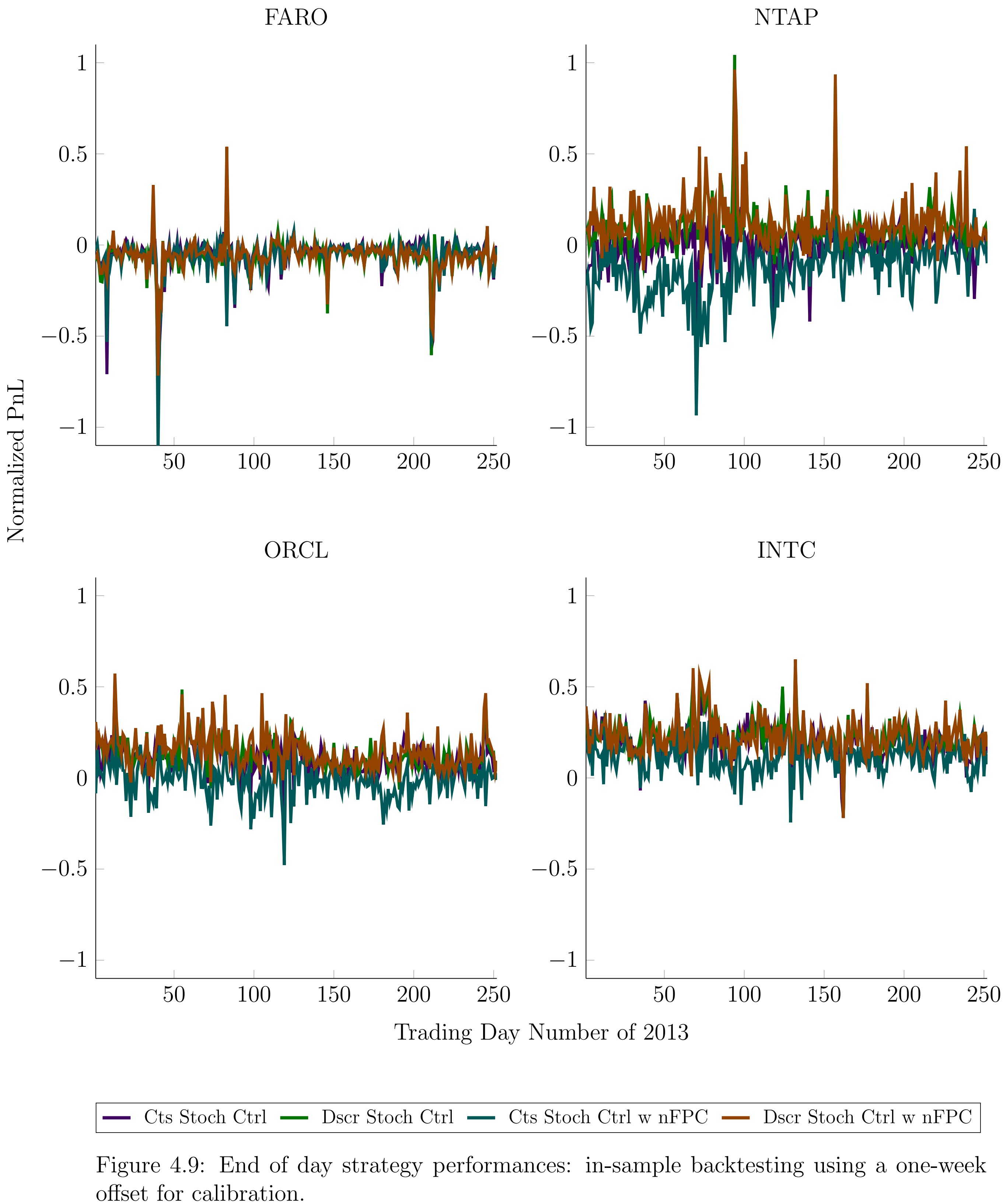


Figure 4.9: End of day strategy performances: in-sample backtesting using a one-week offset for calibration.

4.3.2 Week Offset Calibration

The next type of in-sample backtesting done was to calibrate for each ticker and each trading day of 2013, and to use the results to backtest on the date given by the calibration date shifted forward 7 days. Thus, the calibration obtained on Monday, January 2, 2013 would be used to backtest on Monday, January 9, 2013. Performance values are given in Table 4.4, and Figure 4.10 compares the day-over-day performance of the various strategies.

Most of the observations from the previous section apply here. Chiefly, the illiquid stock `FARO` produces negative PnL and the low-liquidity stock `NTAP` approximately breaks even. As expected, the week offset calibration underperforms same-day calibration, but remarkably the difference is very small: in the case of `INTC`, the discrete time controller still generates a Sharpe ratio of approximately 2.5, and in this case actually produced positive PnL for every single day of the trading year.

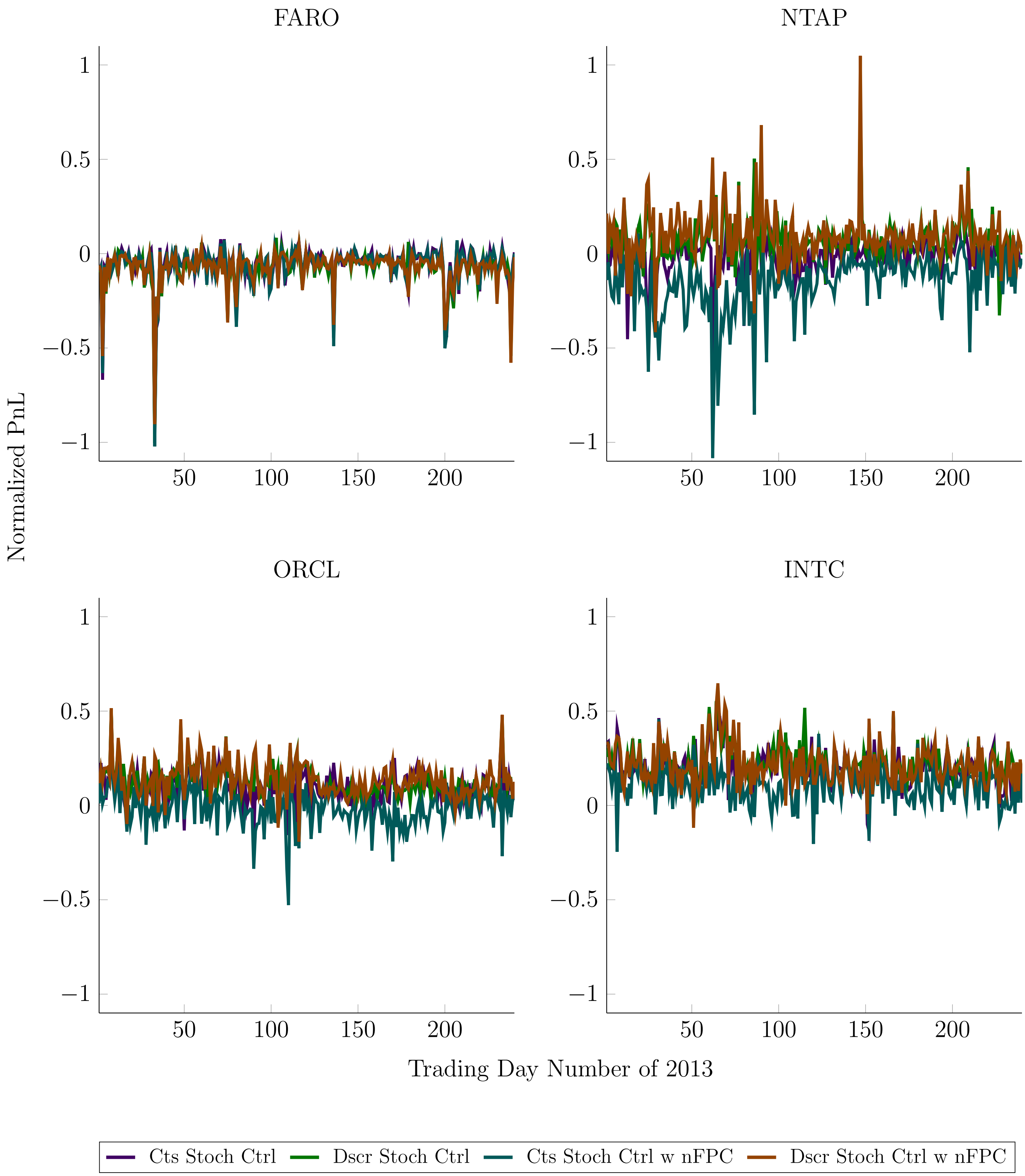


Figure 4.10: End of day strategy performances: in-sample backtesting using a one-week offset for calibration.

Strategy	Return	Sharpe	# MO	# LO	Inv	% Win	Max Loss	Max Win
`FARO`								
Naive	-1.072	-0.444	435	0	1.74	0.08	-34.276	2.984
Naive+	0.045	0.046	0	213	2.26	0.74	-8.764	5.363
Naive++	-0.003	-0.027	0	6	0.25	0.50	-1.138	0.444
Cts	-0.060	-0.565	53	149	0.08	0.20	-0.955	0.077
Dscr	-0.076	-0.764	80	133	-0.07	0.07	-1.004	0.084
Cts w nFPC	-0.065	-0.590	56	149	0.09	0.17	-1.022	0.072
Dscr w nFPC	-0.076	-0.778	78	133	0.07	0.07	-0.904	0.053
`NTAP`								
Naive	-0.303	-0.316	854	0	-10.96	0.20	-9.463	4.349
Naive+	0.290	0.122	0	3537	-13.83	0.73	-19.806	10.266
Naive++	-0.048	-0.084	0	156	-1.79	0.52	-6.310	2.670
Cts	-0.016	-0.165	830	1405	0.37	0.52	-0.612	0.158
Dscr	0.070	0.593	460	1388	5.06	0.79	-0.355	0.858
Cts w nFPC	-0.156	-0.987	1506	1425	0.70	0.09	-1.083	0.106
Dscr w nFPC	0.091	0.656	332	1401	3.16	0.85	-0.416	1.048
`ORCL`								
Naive	-0.112	-0.248	492	0	3.66	0.28	-3.197	2.452
Naive+	0.066	0.022	0	4049	-50.06	0.64	-17.396	18.873
Naive++	0.002	0.005	0	134	0.64	0.49	-1.691	2.537
Cts	0.098	1.181	545	1318	1.86	0.90	-0.203	0.310
Dscr	0.126	1.547	578	1310	4.01	0.97	-0.097	0.505
Cts w nFPC	-0.013	-0.130	1069	1365	1.39	0.47	-0.528	0.500
Dscr w nFPC	0.135	1.459	416	1338	3.16	0.96	-0.192	0.516
`INTC`								
Naive	-0.057	-0.179	274	0	-3.63	0.31	-0.954	1.766
Naive+	0.375	0.138	0	3925	-25.43	0.65	-11.060	11.465
Naive++	0.013	0.055	0	77	-0.47	0.53	-1.815	1.126
Cts	0.202	1.995	423	1139	4.76	0.98	-0.139	0.513
Dscr	0.226	2.494	501	1136	4.62	1.00	-0.029	0.560
Cts w nFPC	0.107	1.111	681	1187	1.64	0.87	-0.245	0.457
Dscr w nFPC	0.215	2.027	401	1156	3.78	0.99	-0.118	0.647

Table 4.4: Strategy performance results: in-sample backtesting using a one-week offset for calibration.

4.3.3 Annual Calibration

The second type of out-of-sample backtesting done was to calibrate using data amalgamated from the entire 2013 trading year. This was a very rich calibration source, as it effectively ensured that every possible state of the Markov chain would have had sufficient observations. Further, this caused us to fix the imbalance bins ρ for the entire year, rather than having bins (and hence what it means to be 'heavy buy imbalance' and 'neutral imbalance') vary each day. Performance values are given in tables **??** and **??**, and **??** compares the day-over-day performance of the various strategies.

4.4 Out-of-Sample Backtesting

Bibliography

[Almgren and Chriss, 2001] Almgren, R. and Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40.

[Bak et al., 1997] Bak, P., Paczuski, M., and Shubik, M. (1997). Price variations in a stock market with many agents. *Physica A: Statistical Mechanics and its Applications*, 246(3):430–453.

[Bensoussan, 2008] Bensoussan, A. (2008). Impulse control in discrete time. *Georgian Mathematical Journal*, 15(3):439–454.

[Booth, 2015] Booth, A. (2015). *Automated Algorithmic Trading: Machine Learning and Agent-Based Modelling in Complex Adaptive Financial Markets*. PhD thesis, University of Southampton.

[Cartea et al., 2015] Cartea, À., Jaimungal, S., and Penalva, J. (2015). *Algorithmic and High-Frequency Trading*. Cambridge University Press.

[Cont et al., 2010] Cont, R., Stoikov, S., and Talreja, R. (2010). A stochastic model for order book dynamics. *Operations Research*, 58(3):549–563.

[Gould et al., 2013] Gould, M. D., Porter, M. A., Williams, S., McDonald, M., Fenn, D. J., and Howison, S. D. (2013). Limit order books. *Quantitative Finance*, 13(11):1709–1742.

[Kurtz, 2004] Kurtz, T. (2004). Mat833 martingale problems and stochastic equations for markov processes. Lecture Notes, University of Wisconsin - Madison, Department of Mathematics.

[Kwong, 2015] Kwong, R. (2015). Ece1639 analysis and control of stochastic systems. Lecture Notes, University of Toronto, Department of Electrical & Computer Engineering.

[Kyle, 1989] Kyle, A. S. (1989). Informed speculation with imperfect competition. *The Review of Economic Studies*, 56(3):317–355.

[Laughlin et al., 2014] Laughlin, G., Aguirre, A., and Grundfest, J. (2014). Information transmission between financial markets in chicago and new york. *Financial Review*, 49(2):283–312.

[Takahara, 2014] Takahara, G. (2014). Stat455 stochastic processes. Lecture Notes, Queen's University, Department of Mathematics and Statistics.

[Tan and Ylmaz, 2002] Tan, B. and Ylmaz, K. (2002). Markov chain test for time dependence and homogeneity: An analytical and empirical evaluation. *European Journal of Operational Research*, 137(3):524–543.

[Weißbach and Walter, 2010] Weißbach, R. and Walter, R. (2010). A likelihood ratio test for stationarity of rating transitions. *Journal of Econometrics*, 155(2):188–194.