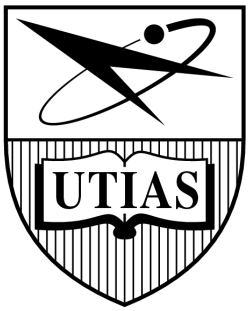
Statistical Arbitrage Using Limit Order Book Imbalance



Anton D. 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 Using Limit Order Book Imbalance

Anton D. Rubisov

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

2015

Abstract

This dissertation demonstrates that there is high revenue potential in using limit order book imbalance as a state variable in an algorithmic trading strategy. Beginning with the hypothesis that imbalance of bid/ask order volumes is an indicator for future price changes, exploratory data analysis suggests that modelling the joint distribution of imbalance and observed price changes as a continuous-time Markov chain presents a monetizable opportunity. The arbitrage problem is then formalized mathematically as a stochastic optimal control problem using limit orders and market orders with the aim of maximizing terminal wealth. The problem is solved in both continuous and discrete time using the dynamic programming principle, which produces both conditions for market order execution, as well as limit order posting depths, as functions of time, inventory, and imbalance. The optimal controls are calibrated and backtested on historical NASDAQ ITCH data, which produces consistent and substantial revenue.

Acknowledgements

I couldn't begin to enumerate the times and ways this dissertation could have failed to be. I am deeply grateful to several individuals for making it a reality, and I'd like to acknowledge them here.

Dr. Gabriele D'Eleuterio, my supervisor at UTIAS, is the epitomical supernatural mentor of the hero journey, without whom I would have abandoned my studies. Gabe was the inspirational upholder of academic values, who relentlessly stressed the subtleties of scientific presentation and communication, and guided me through a pretty tough time. Thank you for your support, guidance, and for every coffee.

Dr. Sebastian Jaimungal, professor of statistics at the university, took me on as a surrogate student half-way through my studies, and guided every step of this dissertation. Despite having absolutely no obligation to do so, he had faith in me and patience with me in navigating an entirely unfamiliar subject. Thank you for repeatedly making yourself available to help with all the math.

Dr. Jack Carlyle, then a PhD student in solar physics at UCL, inspired me to quit my job and go back to school for a graduate degree. It wasn't an easy decision, and the jury is out on whether it was a good one. But thanks for all the tea and crumpets.

Ted Herman, a friend that redefines friendship, has read more of my research than anyone ever will. At a crucial time, we struck the Fran's Accord that involved prewritten $100 cheques being shredded on a weekly basis upon receipt of a progress update. Legitimately, this is probably the number one reason that my dissertation got done.

Thank you **mom** for always loving, supporting, and inspiring me to my best. Thank you **dad** for pretty well being a supervisor too.

Contents

List of Figures

List of Tables

Nomenclature

$\#_{bins}$	Number of states into which the smoothed imbalance is discretized
$\mathcal{F}$	Filtration of a probability space
$\mathcal{L}$	Infinitesimal generator operator
$\Delta S(t)$	Sign of the observed price change at time t
Δt_I	Size of rolling window for imbalance smoothing
Δt_S	Size of window for price change calculation
$\delta^{\pm}$	The depth at which we post a buy (δ^+) and sell (δ^-) limit order
$\eta_{0,z}$	Magnitude of the midprice change (a random variable)
$\mathbb{1}$	Indicator function
κ	Limit order fill probability constant
$\lambda^{\pm}$	Arrival rates of buy (λ^+) and sell (λ^-) orders
$\boldsymbol{G}$	Continuous-time Markov chain generator matrix
$\boldsymbol{P}$	Continuous-time Markov chain transition probability matrix
$\mu^{\pm}$	Rate of arrival of other agents' buy (μ^+) and sell (μ^-) market orders
∂_x	Partial derivative with respect to x
$\rho(t)$	Imbalance ratio, smoothed by averaging over a rolling window
τ	A stopping time at which we execute a market order
ξ	Bid-Ask Half-Spread

$H(t, x, s, \boldsymbol{z}, q)$ Continuous-time value function

$H^{\boldsymbol{\tau},\delta^{\pm}}(t, x, s, \boldsymbol{z}, q)$ Continuous-time performance criterion of the controls $\boldsymbol{\tau}$ and $\delta^{\pm}$

$I(t)$ Imbalance ratio, raw

$K^{\pm}$ Other agents' buy (K^{+}) and sell (K^{-}) market orders

$L^{\pm}$ A counting process of how many of our buy (L^{+}) and sell (L^{-}) limit orders are filled

$M^{\pm}$ Execution of our buy (M^{+}) and sell (M^{-}) market orders

$Q_t^{\boldsymbol{\tau},\delta}$ Our inventory at time t having used controls $\boldsymbol{\tau}$ and δ

$S(t)$ Midprice

$V_k(x, s, \boldsymbol{z}, q)$ Discrete-time value function

$V_k^{\delta^{\pm}}(x, s, \boldsymbol{z}, q)$ Discrete-time performance criterion of the controls $\boldsymbol{\tau}$ and $\delta^{\pm}$

$W_t^{\boldsymbol{\tau},\delta}$ Our wealth (cash, value of assets, and liquidation penalty) at time t having used controls $\boldsymbol{\tau}$ and δ

$X_t^{\boldsymbol{\tau},\delta}$ Our cash at time t having used controls $\boldsymbol{\tau}$ and δ

$Z(t)$ Continuous-time Markov chain variable

`AAPL` Stock ticker for Apple Inc.

`FARO` Stock ticker for FARO Technologies Inc.

`INTC` Stock ticker for Intel Corporation

`MMM` Stock ticker for 3M Company

`NTAP` Stock ticker for NetApp, Inc.

`ORCL` Stock ticker for Oracle Corporation

CTMC Continuous-time Markov chain

DPE Dynamic programming equation

ITCH Market data feed provided by NASDAQ

LO Limit order

LOB Limit-order book

MO Market order

nFPC Non-$\mathcal{F}$-predictable calibration

P&L Profit and Loss

1

Introduction

With the introduction of mathematical models, participation in financial markets evolved from an art to a science. Beginning with Harry Markowitz's modern portfolio theory, moving through the capital asset pricing model and the Black-Scholes option pricing model, modern finance is now replete with models for pricing derivatives, credit scores, and costs of capital; what was once speculation is now calculation. Crucially, models lead to algorithms, which remove human error and add the ability to digest large sets of data.

The process of running computer algorithms to execute orders on an electronic exchange such as NASDAQ is known as *algorithmic trading*. Speed of execution is typically crucial, often requiring running the algorithms on servers directly wired to the exchange, known as *colocation*. Closely related is *high-frequency trading*, which refers simply to the timescale, generally milliseconds, on which the algorithms submit orders. In theory, high-frequency trading is encompassed by algorithmic trading, while not all algorithmic trading need be high frequency; in practice, the two terms are often used interchangeably.

The particular algorithms used in algorithmic trading vary greatly across the different types of strategies employed. Non-revenue-generating algorithmic trading is generally aimed at transaction cost reduction, with the primary theoretical papers on the subject being due to Bertsimas and Lo (1998) and 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 term 'large' is used 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 strategy is then to split 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.

Conversely, an example of algorithmic trading that capitalises on arbitrage opportunities to generate revenue is cross-exchange arbitrage, which uses simple, low-latency algorithms to profit from price discrepancies of a single stock dual-listed on two exchanges. The server running the algorithm 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 6 milliseconds via optical fibres that send information at about 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 4 milliseconds (Laughlin et al., 2014). Another class of strategies for generating revenue using algorithmic trading are statistical arbitrage strategies, which use complex algorithms to profit from observed statistical patterns of a single stock on a single exchange. In statistical arbitrage, the aim is to exploit predictable statistical patterns in the available data provided by the exchange, such as predicting stock price movements from prices observed thus far. This method too requires colocation, and operates on the scale of milliseconds. It is this type of high-frequency trading that is explored 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 high frequency trading activity. Nevertheless, as they are still required to provide liquidity to markets via *market-making* (simultaneously quoting both buy and sell prices on a range of financial instruments), banks use algorithmic trading to determine the bid/ask bands they will send to exchanges. Exploiting arbitrage opportunities using high frequency trading remains unrestricted for hedge funds, and notably has been used by Renaissance Technologies LLC's flagship Medallion fund to generate an average 71.8 percent annual return, before fees, from 1994 through mid-2014 (Rubin and Collins, 2015). However, as it remains exclusive to only Renaissance employees and family members, it serves instead as a reminder of the revenue potential of high frequency algorithmic trading.

1.1 The Limit-Order Book

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 key that uniquely identifies the agent. The exchange runs a trade-matching engine that uses the LOB to pair buy and sell requests that concur on price, even if only for partial volume. Orders remain in effect until they are modified, cancelled, 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 visualised 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* (alternatively referred to as *lifted*) according to a first-in-first-out priority queue for each price level. The price levels can also be referred to by their *depth*, where the best bid and ask prices are called *at-the-touch* and have a depth of zero, and depth increases in either direction according to the absolute price difference from the at-the-touch depths; the buy limit order at \$28.92 is at a depth of \$0.02. *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), a fee is charged for executing a market order.

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-based approach, with which we are more concerned here, assumes zero-intelligence, provides simplified conceptual 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 one-dimensional price lattice (Bak et al., 1997) to independent Poisson processes governing the arrival, modification, and cancellation of orders (Cont et al., 2010). An excellent literature survey on LOB modelling can be found in Gould et al. (2013).

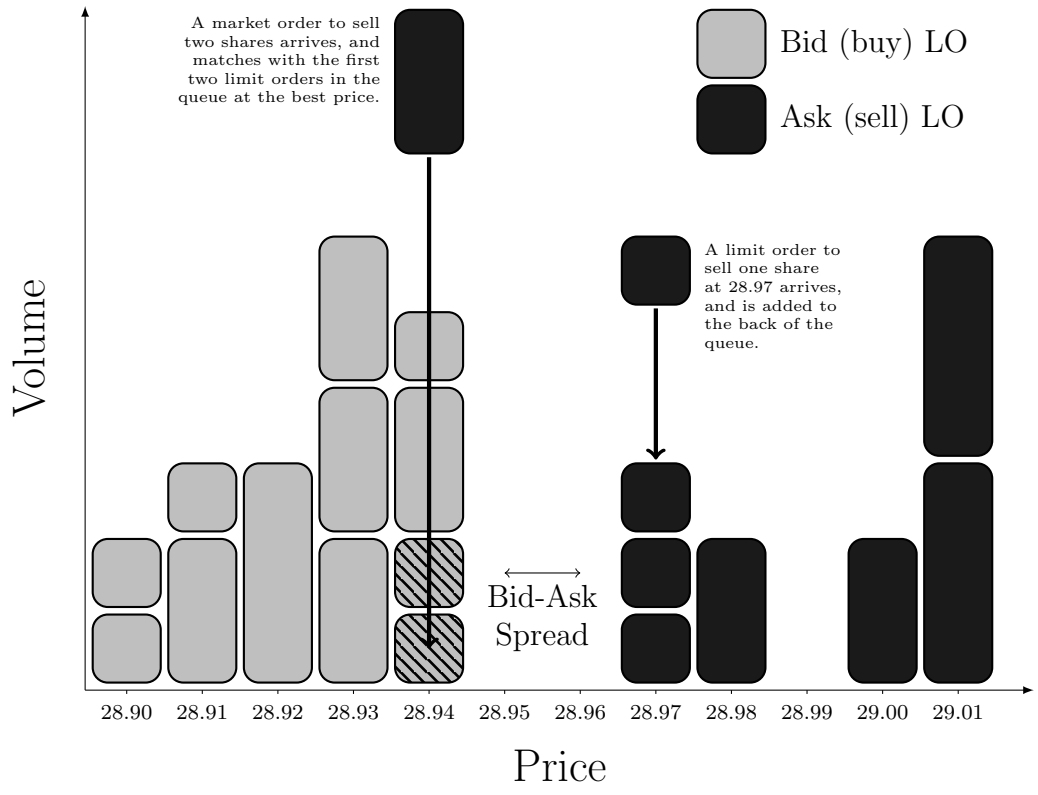


Figure 1.1: Structure and mechanics of the limit-order book, adapted from Booth (2015). Each block represents an order, of varying volumes, submitted by various agents participating in the market.

1.2 ITCH Data Set

The underlying data that will be used in this work to generate imbalance and price change timeseries comes from the NASDAQ Historical TotalView-ITCH. ITCH[1] is a direct data-feed protocol that makes it possible for those with a paid subscription to track the status of every order from arrival until cancellation or execution. The Historical TotalView is simply a historical record of the events in the live feed. In the data used in this work, timestamps are provided to 1 millisecond, though newer versions of the feed offer nanosecond precision. Our data has been converted to MATLAB format; below, the structure of the event feed is described in detail (retaining only relevant fields):

Time	Order ID	Event	Volume	Price
⋮	⋮	⋮	⋮	⋮
39960699	72408630	66	100	1107000
39960710	72408630	68	100	1107000
⋮	⋮	⋮	⋮	⋮

Time: Order arrival time in milliseconds from midnight.
Order ID: Unique order reference number.
Event: Event type:

66 – Add buy order
83 – Add sell order
69 – Execute outstanding order in part
67 – Cancel outstanding order in part
70 – Execute outstanding order in full
68 – Delete outstanding order in full
88 – Bulk volume for the cross event
84 – Execute non-displayed order

Volume: Number of shares.
Price: Dollar price times 10,000.

[1]Remarkably, according to a representative of the NASDAQ, 'ITCH' does not stand for anything.

Thus, in the above example, we have a buy order arriving at 11:06am for 100 shares at \$100.70, and being cancelled 11 milliseconds later. From this feed we are able to reconstruct the entire limit-order book at any point in time, which amounts to being able to generate a plot as in Figure 1.1 detailing the exact liquidity available at each order depth.

1.3 Order Imbalance

A core component of this dissertation is limit-order book imbalance. *Imbalance* is a ratio of limit order volumes between the bid and ask side, and in the work that follows is calculated as

$$I(t) = \frac{V_{bid}(t) - V_{ask}(t)}{V_{bid}(t) + V_{ask}(t)} \tag{1.1}$$

where $I(t) \in [-1, 1]$, and both V_{bid} and V_{ask} are computed as the weighted average volumes at the three lowest depths having non-zero volume, using exponentially decreasing weights. As a sample calculation, the imbalance of the sample LOB presented in Figure 1.1 (prior to order arrivals) would be

$$\begin{aligned}
V_{bid} &= \text{weight}(28.94) \cdot \text{volume}(28.94) \\
&\quad + \text{weight}(28.93) \cdot \text{volume}(28.93) \\
&\quad + \text{weight}(28.92) \cdot \text{volume}(28.92) \\
&= e^{-0.5(0)} \cdot 5 + e^{-0.5(1)} \cdot 6 + e^{-0.5(2)} \cdot 3 \\
&= 1.0000 \cdot 5 + 0.6065 \cdot 6 + 0.3679 \cdot 3 \\
&= 9.7428 \\
V_{ask} &= 4.9488 \\
I(t) &= \frac{9.7428 - 4.9488}{9.7428 + 4.9488} \\
&= 0.3263
\end{aligned}$$

In the figure we see that there are more limit orders on the bid side than on the ask, and the above value confirms that there is a medium imbalance in favour of the bid side.

1.4 Roadmap

Our objective in this dissertation is to use limit order book imbalance as a state variable in an algorithmic trading strategy, and to demonstrate that an effective statistical arbitrage strategy can be constructed around it. Our novel approach is to model imbalance and price change with a joint distribution, and to solve the resulting stochastic optimal control problem independently in both continuous time and discrete time.

The remainder of this dissertation is structured as follows:

Chapter 2 begins with the hypothesis that imbalance of bid/ask order volumes is an indicator for future price changes, and explores the possibility of constructing naive trading strategies using the statistical properties arising from modelling imbalance as a continuous time Markov chain; a basic understanding of probability theory is required.

Chapter 3 casts the same statistical arbitrage problem into a stochastic optimal control framework, and solves it in both continuous and discrete time; familiarity with stochastic calculus and dynamic programming is assumed.

Chapter 4 presents calibrations of the optimal controls and explores their dynamics. In-sample and out-of-sample backtests are conducted on historical ITCH data, which show a 877% return on investment over 2014.

Chapter 5 concludes the dissertation by restating the primary backtesting results in the context of operational costs, and summarizes the key assumptions and simplifications that have been made. The conclusions are intended as considerations for future work on this topic.

2

Exploratory Data Analysis

In this chapter we explore the possibility of constructing naive trading strategies from the statistical properties that emerge from our model of the limit-order book imbalance. The strategies will be tested on the stocks listed in Table 2.1. The chosen stocks are a good sample group of the population of stocks listed on the NASDAQ in that they span the spectrum of low liquidity to high liquidity stocks.

Table 2.1: List of stocks used in the exploratory data analysis, along with the daily average trading volume of each.

Ticker	Company	Average Daily Volume
FARO	FARO Technologies Inc.	200,000
MMM	3M Company	2,000,000
NTAP	NetApp, Inc.	4,000,000
ORCL	Oracle Corporation	15,000,000
INTC	Intel Corporation	30,000,000

2.1 Modelling Imbalance: Continuous Time Markov Chain

The aim of this research project is to use 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 used herein, is to discretise 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 finite or countable state space* K*, so that* $X(t) \in K$*, 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 nondecreasing sequence of* $n+1$ *times, and* $i_1, \ldots, i_{n-1}, i, j \in K$ *are any* $n+1$ *states.*

Definition 2. *A continuous-time Markov chain* $\{X(t) \mid t \geq 0\}$ *is* time homogeneous *if for any* $s \leq t$ *and any states* $i, j \in K$*,*

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

Properties of a CTMC. *Let* $\{X(t) \mid t \geq 0\}$ *be a time-homogeneous CTMC. Its key defining quantities are*

1. *the* transition rates q_{ij}*, which specify the rate at which* X *jumps from state* i *to* j*;*
2. *the* conditional transition probabilities p_{ij}*, which specify the probability with which* X *transitions to state* j *conditional on leaving state* i*;*
3. *the* holding times v_i*, where upon entering state* i*, the amount of time that* X *will spend in state* i *prior to transitioning is exponentially distributed with rate* v_i*.*

These quantities are related by

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

$$q_{ij} = v_i \cdot p_{ij} \quad (2.4)$$

$$p_{ij} = \frac{q_{ij}}{v_i} \quad (2.5)$$

Definition 3. *A continuous-time Markov chain* $\{X(t) \mid t \geq 0\}$ *has an* infinitesimal generator matrix $\boldsymbol{G}$*, whose entries are*

$$g_{ij} = \begin{cases} q_{ij}, & i \neq j \\ -v_i, & i = j \end{cases} \quad (2.6)$$

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

$$\dot{\boldsymbol{P}}(t) = \boldsymbol{G}\boldsymbol{P}(t) \tag{2.7}$$

$$\boldsymbol{P}(t) = e^{\boldsymbol{G}t} \tag{2.8}$$

Conditional on $X(t) = k$, 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 (respectively, 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 as in Figure 2.1.

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

2.2 Maximum Likelihood Estimate of a Markov-Modulated Poisson Process

2.2.1 Infinitesimal Generator Matrix

Let $\boldsymbol{G}$ be the generator matrix for a CTMC $X(t)$ with finite-dimensional state space K. From observations, e.g., the fictional events in the timeline given in Figure 2.1, we

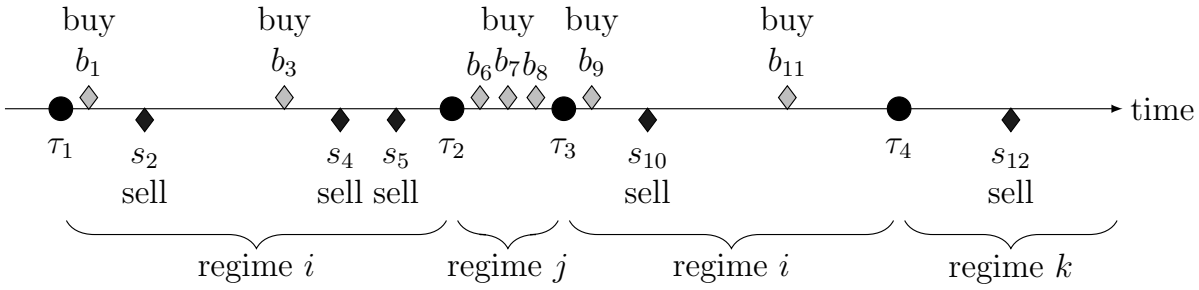


Figure 2.1: Hypothetical timeline of market orders arriving during changing order imbalance regimes. The τ nodes represent regime switch times, and appear in unequally spaced intervals. Regime i occurs twice, and market order arrivals behave similarly in both instances.

want to estimate the entries of $\boldsymbol{G}$. As the holding time in a given state i has probability density function $f(t; v_i) = v_i e^{-v_i t}$, the likelihood function (allowing for repetition of terms) is therefore

$$\mathcal{L}(\boldsymbol{G}) = (v_i e^{-v_i(\tau_2 - \tau_1)} p_{ij})(v_j e^{-v_j(\tau_3 - \tau_2)} p_{ji})(v_i e^{-v_i(\tau_4 - \tau_3)} p_{ik}) \dots \tag{2.9}$$

$$= \prod_{i=1}^{K} \prod_{i \neq j} (v_i p_{ij})^{N_{ij}(T)} e^{-v_i H_i(T)} \tag{2.10}$$

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

where $N_{ij}(T)$ is the number of transitions from regime i to j up to time T, and $H_i(T)$ is the holding time in regime i up to time T. Therefore, the log-likelihood becomes

$$\ln \mathcal{L}(\boldsymbol{G}) = \sum_{i=1}^{K} \sum_{i \neq j} \left[N_{ij}(T) \ln(q_{ij}) - v_i H_i(T) \right] \tag{2.12}$$

$$= \sum_{i=1}^{K} \sum_{i \neq j} \left[N_{ij}(T) \ln(q_{ij}) - \left(\sum_{i \neq k} q_{ik} H_i(T) \right) \right] \tag{2.13}$$

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:

$$\frac{\partial \ln \mathcal{L}(\boldsymbol{G})}{\partial \hat{q}_{ij}} = \frac{N_{ij}(T)}{q_{ij}} - H_i(T) = 0 \tag{2.14}$$

Solving gives

$$\hat{q}_{ij} = \frac{N_{ij}(T)}{H_i(T)} \tag{2.15}$$

as the maximum likelihood estimate. This has the simple and intuitive interpretation that the estimated rate of transition from state i to j is equal to the number of transitions from i to j divided by the total time spent in i.

2.2.2 Arrival Rates

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 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 interarrival time of events conditional on being in state k. The maximum likelihood derivation then follows just as it did for the generator matrix.

$$\mathcal{L}(\lambda_k^+; b_1, \ldots, b_N) = \prod_{i=2}^{N} \lambda_k^+ e^{-\lambda_k^+ (b_i - b_{i-1})} \tag{2.16}$$

$$= (\lambda_k^+)^{N_k^+(T)} e^{-\lambda_k^+ H_k(T)} \tag{2.17}$$

where $N_k^+(T)$ is the number of market order arrivals in regime k up to time T, and $H_k(T)$ is the holding time in regime k up to time T. Therefore, the log-likelihood becomes:

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

By setting the partial derivative of $\ln \mathcal{L}$ with respect to λ_k^+ equal to zero, we get that the maximum likelihood estimate for $\hat{\lambda}_k^+$ is:

$$\frac{\partial \ln \mathcal{L}}{\partial \lambda_k^+} = \frac{N_k^+(T)}{\hat{\lambda}_k^+} - H_k(T) = 0 \tag{2.19}$$

Solving gives

$$\hat{\lambda}_k^+ = \frac{N_k^+(T)}{H_k(T)} \tag{2.20}$$

as the maximum likelihood estimate. Again this has the intuitive interpretation that the estimated rate of arrivals while in state k is equal to the number of observed arrivals while in k divided by the total time spent in k.

2.3 Two-Dimensional CTMC

Next we consider a CTMC $Z(t)$ that jointly models the imbalance bin $\rho(t)$ and the price change $\Delta S(t)$. The raw imbalance timeseries is very erratic, so to smooth it we take the time-weighted average of imbalance over the past time inteval Δt_I. We compute price change as the *sign* of the change in midprice of the *future* time interval Δt_S. These time intervals are illustrated in Figure 2.2.

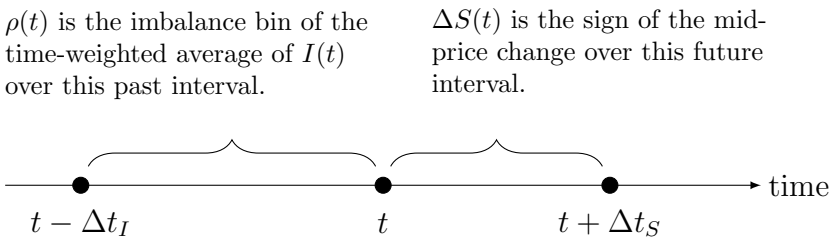


Figure 2.2: Time intervals for time-weighted averaging of imbalance and for price change.

Thus, the CTMC models the joint distribution $(\rho(t), \Delta S(t))$ where

$$\rho(t) \in \{1, 2, \ldots, \#_{bins}\}$$

is the bin corresponding to imbalance averaged over the interval $[t - \Delta t_I, t]$, and

$$\Delta S(t) = \text{sgn}(S(t + \Delta t_S) - S(t)) \in \{-1, 0, 1\}$$

For simplicity of computation, the pair $(\rho(t), \Delta S(t))$ is then reduced into one dimension with a simple encoding function φ. An example of this reduction using three bins is presented in Table 2.2.

Table 2.2: $\varphi(\rho(t), S(t))$: One-dimensional encoding of two-dimensional CTMC.

$Z(t)$	Bin $\rho(t)$	$\Delta S(t)$	$Z(t)$	Bin $\rho(t)$	$\Delta S(t)$	$Z(t)$	Bin $\rho(t)$	$\Delta S(t)$
1	Bin 1	-1	4	Bin 1	0	7	Bin 1	$+1$
2	Bin 2	-1	5	Bin 2	0	8	Bin 2	$+1$
3	Bin 3	-1	6	Bin 3	0	9	Bin 3	$+1$

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 Yılmaz (2002). The homogeneity hypothesis is given by (Weißbach and Walter, 2010)

$$H_0 = \forall i, j \in S \ : \ \exists q_{ij} \in \mathbb{R}^+ \ : \ q_{ij}(t) \equiv q_{ij} \ \forall t \in [0, T] \tag{2.21}$$

whereas the alternative hypothesis states that transition rates/probabilities are time-dependent. To test the hypothesis, we fix $\Delta t_I = \Delta t_S$ at some value, choose a number of imbalance bins, and calculate the maximum likelihood estimate of the infinitesimal generator matrix $\boldsymbol{G}$ on the full timeseries using (2.15). For a chosen error threshold ϵ, we use the relationship in (2.8) to calculate the number of timesteps n_{conv} of size Δt_I such that

$$||\boldsymbol{P}\left((n_{conv}+1)\Delta t_I\right) - \boldsymbol{P}\left(n_{conv}\Delta t_I\right)|| < \epsilon \tag{2.22}$$

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_{conv}. 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 $\boldsymbol{P}_k\left(\Delta t_I\right)$ are statistically different from those of the full period. For comparison, we also partitioned the timeseries into 8, 4, and 2 equal intervals. The asymptotically equivalent test statistic to the likelihood ratio test statistic is

$$D = -2\ln(\mathcal{L}) = 2\sum_k \sum_{i,j} n_{i,j}^k \left[\ln(p_{ij}^k) - \ln(p_{ij})\right] \tag{2.23}$$

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 $(K-1)(3\cdot\#_{bins})(3\cdot\#_{bins}-1)$ degrees of freedom. The tests were run for each ticker for each trading day of 2013, and averaged over the year. Table 2.3 shows the p-value scores for the tests. Considering the standard cut-off p-value of 0.05, the cross-validation results show a strong case for the rejection of the homogeneity hypothesis. However, using a nonhomogeneous 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. The severity of proceeding with the homogeneity hypothesis is not known *a priori*, and may instead emerge with the backtesting results done later in this chapter and in Chapter 4.

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.

Table 2.3: 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 $\epsilon = 1 \times 10^{-10}$.

		subintervals					subintervals			
Δt_I	n_{conv}	K	8	4	2	n_{conv}	K	8	4	2
			FARO					ORCL		
$\#_{bins} = 3$										
100ms	4933	0.000	0.000	0.000	0.003	1803	0.000	0.000	0.000	0.000
1000ms	727	0.000	0.002	0.001	0.005	303	0.000	0.000	0.000	0.001
10000ms	149	0.000	0.005	0.010	0.017	84	0.000	0.007	0.005	0.010
$\#_{bins} = 5$										
100ms	6450	0.000	0.001	0.002	0.004	2503	0.000	0.000	0.000	0.000
1000ms	941	0.000	0.001	0.003	0.006	404	0.000	0.001	0.002	0.003
10000ms	187	0.000	0.000	0.000	0.005	103	0.000	0.000	0.001	0.009
			NTAP					INTC		
$\#_{bins} = 3$										
100ms	1320	0.000	0.000	0.000	0.000	2545	0.000	0.000	0.000	0.001
1000ms	237	0.000	0.000	0.000	0.000	408	0.000	0.001	0.001	0.002
10000ms	72	0.000	0.006	0.003	0.007	105	0.000	0.004	0.006	0.009
$\#_{bins} = 5$										
100ms	1777	0.000	0.000	0.000	0.000	3498	0.000	0.001	0.001	0.001
1000ms	308	0.000	0.001	0.000	0.001	771	0.000	0.001	0.002	0.002
10000ms	87	0.000	0.000	0.002	0.010	133	0.000	0.000	0.000	0.007

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

$$P_{ij} = \mathbb{P}\left[\varphi(\rho_{[t-\Delta t_I,t]}, \Delta S_{[t,t+\Delta t_S]}) = j \mid \varphi(\rho_{[t-2\Delta t_I,t-\Delta t_I]}, \Delta S_{[t-\Delta t_I,t]}) = i\right] \tag{2.24}$$

which can be expressed semantically as

$$P_{ij} = \mathbb{P}\left[\varphi(\rho_{\text{curr}}, \Delta S_{\text{future}}) = j \mid \varphi(\rho_{\text{prev}}, \Delta S_{\text{curr}}) = i\right] \tag{2.25}$$

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

$$= \mathbb{P}\left[\rho_{\text{curr}} = i, \Delta S_{\text{future}} = j \mid \rho_{\text{prev}} = k, \Delta S_{\text{curr}} = m\right] \tag{2.26}$$

$$= \mathbb{P}\left[\rho_{\text{curr}} = i, \Delta S_{\text{future}} = j \mid B\right] \tag{2.27}$$

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

$$\mathbb{P}\left[\Delta S_{\text{future}} = j \mid B, \rho_{\text{curr}} = i\right] = \frac{\mathbb{P}\left[\rho_{\text{curr}} = i, \Delta S_{\text{future}} = j \mid B\right]}{\mathbb{P}\left[\rho_{\text{curr}} = i \mid B\right]} \tag{2.28}$$

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

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

The left-hand-side value in (2.28) 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 (2.28).

The $\boldsymbol{Q}$ matrix in Table 2.4 was obtained using data for `MMM` from 2013-05-15, averaging imbalance and price change timewindows $\Delta t_I = \Delta t_S = 1000$ms, and $K = 3$ imbalance bins.

The three middle rows of Table 2.4 contain the majority of values > 0.5, showing that in most cases we are expecting no price change. The only other cases in which the probability of a price change is greater than 0.5 show evidence of *momentum*; for example, the

value in row 1, column 1 can be interpreted as: if $\rho_{\text{prev}} = \rho_{\text{curr}} = 1$ and previously we saw a downward price change, then we expect to again see a downward price change. The boldface diagonal values in the table lend themselves to the empirical conclusion

$$\mathbb{P}\left[\Delta S_{\text{future}} = \Delta S_{\text{curr}} \mid \rho_{\text{curr}} = \rho_{\text{prev}}\right] > 0.5 \tag{2.30}$$

2.5 Naive Trading Strategies

Using the key insight drawn from (2.30), we implemented several naive trading strategies, descriptions of which follow:

Naive Trading Strategy. This strategy seeks to profit from expected price changes by using market orders. Using the conditional probabilities obtained from $\boldsymbol{Q}$, we will execute a buy (respectively sell) market order if the probability of an upward (respectively downward) price change is > 0.5. Pseudocode for this strategy is given in Algorithm 1. In lines 4-6 we are forecasting a downward price move, and therefore sell one share with a market order at the best bid price. In lines 7-9, we are forecasting an upward price move, and buy one share with a market order at the best ask price. In lines 12-16 we have reached the end of the trading day, and liquidate our position at the at-the-touch price.

Algorithm 1 Naive Trading Strategy

$cash = 0$
$asset = 0$
for $t = 2$: `length`$(timeseries)$ **do**
  if $\mathbb{P}\left[\Delta S_{future} < 0 \mid \rho_{curr}, \rho_{prev}, \Delta S_{curr}\right] > 0.5$ **then**
    $cash \mathrel{+}= BidPrice(t)$
    $asset \mathrel{-}= 1$
  else if $\mathbb{P}\left[\Delta S_{future} > 0 \mid \rho_{curr}, \rho_{prev}, \Delta S_{curr}\right] > 0.5$ **then**
    $cash \mathrel{-}= AskPrice(t)$
    $asset \mathrel{+}= 1$
  end if
end for
if $asset > 0$ **then**
  $cash \mathrel{+}= asset \times BidPrice(EndOfDay)$
else if $asset < 0$ **then**
  $cash \mathrel{+}= asset \times AskPrice(EndOfDay)$
end if

Table 2.4: The $\boldsymbol{Q}$ matrix: conditional probabilities of future price changes, conditioned on current imbalance, current price change, and previous imbalance.

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

Naive+ Trading Strategy. This strategy seeks to profit from no expected changes in price by using limit orders. If no change in midprice is expected then we'll post buy and sell limit orders at the touch. For every market order that arrives, we'll assume that our limit order is always filled, following which we will immediately repost whichever limit order was filled. Pseudocode for this strategy is given in Algorithm 2. In lines 11-13, since a sell market order is arriving, our buy limit order is being filled and we are paying the bid price.

Algorithm 2 Naive+ Trading Strategy

1: $cash = 0$
2: $asset = 0$
3: **for** $t = 2$: `length`$(timeseries)$ **do**
4: **if** $\mathbb{P}\left[\Delta S_{curr} = 0 \mid \rho_{curr}, \rho_{prev}, \Delta S_{prev}\right] > 0.5$ **then**
5: $LO_{posted} =$ `True`
6: **else**
7: $LO_{posted} =$ `False`
8: **end if**
9: **if** LO_{posted} **then**
10: **for** $MO \in ArrivedMarketOrders(t, t+1)$ **do**
11: **if** $MO == Sell$ **then**
12: $cash \mathrel{-}= BidPrice(t)$
13: $asset \mathrel{+}= 1$
14: **else if** $MO == Buy$ **then**
15: $cash \mathrel{+}= AskPrice(t)$
16: $asset \mathrel{-}= 1$
17: **end if**
18: **end for**
19: **end if**
20: **end for**
21: **if** $asset > 0$ **then**
22: $cash \mathrel{+}= asset \times BidPrice(EndOfDay)$
23: **else if** $asset < 0$ **then**
24: $cash \mathrel{+}= asset \times AskPrice(EndOfDay)$
25: **end if**

Naive++ Trading Strategy. Like the Naive strategy, this strategy seeks to profit from expected price changes, but using limit orders (and therefore buying/selling at better prices). If we expect a downward (respectively upward) price change then we'll post an at-the-touch sell (respectively 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 (2.30) could be used. We found that $\#_{bins} = 4$ provided the highest expected number of trades for most tickers. However, as we were using 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 intraday risk-adjusted return (average of returns divided by standard deviation of returns), then using the calibrated parameters to backtest the entire year.

2.7 Conclusions from the Naive Trading Strategies

Figure 2.3 plots the average daily performance of the naive strategies for four stocks listed on the NASDAQ. The performance is normalized via division by the initial stock price on each trading day. The Naive strategy lost revenue on average, the Naive+ strategy (at-the-touch limit orders when no change was expected) generated revenue on average, and the Naive++ strategy (using limit orders to adversely select agents that traded against the price change momentum) neither profited nor lost on average.

Why is the Naive strategy producing, on average, normalized losses? On calibration, we see that our intraday risk-adjusted return is about 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 to the calibration, meaning the parameters are not specifically tuned to that data, and thus the parameters are (likely) not optimal. This adds evidence to rejecting the time homogeneity assumption, and in particular suggests that not every day can be modelled by the same Markov chain. The problem may be exaggerated by the fact that we are calibrating on the first trading day of the calendar year, when we might expect reduced, or

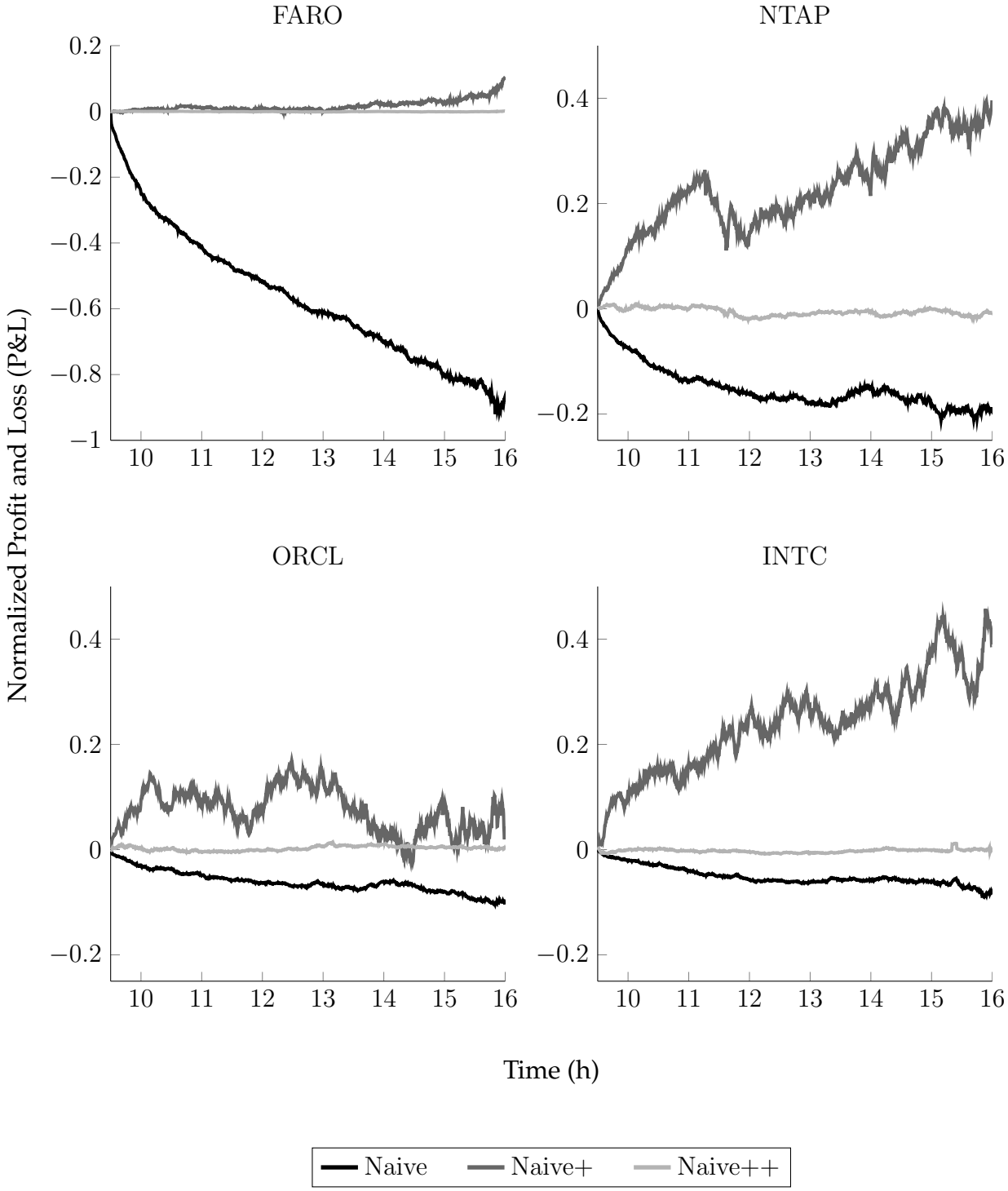


Figure 2.3: Comparison of Naive (black), Naive+ (dark gray), and Naive++ (light gray) trading strategies. Plotted are normalized book values, averaged across the trading year, between 0930h (market open) and 1600h (market close). Book value is the sum of cash and market value of assets. Each day begins with zero book value, and all book values are normalized by dividing by the stock price at the start of each trading day.

at least nonrepresentative, 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 monetisable 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.

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 (a) uses market orders, and (b) executes when we predict nonzero changes.

In (a), the difference between limit orders and market orders 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 the 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.

In (b), executing when predicting nonzero changes 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 2.5 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.

Table 2.5: Hypothetical timeline of adverse selection with market orders.

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

3

Maximizing Wealth via Continuous-Time Stochastic Control

Leveraging the insights gained from the exploratory data analysis in the previous chapter, we turn our attention now to casting the statistical arbitrage problem into a mathematical framework. With our underlying process of interest being a Markov chain, the problem lends itself naturally to being considered in the context of stochastic optimal control. Stochastic control problems are common in finance, where many of the underlying processes have a random nature; the goal is to maximize the expectation of some target function, representing profit, by converging on a set of optimal controls that drive the dynamics of the stochastic system to whichever state attains that maximum expectation. [Of course, the problem can conversely be aimed at minimizing expected cost.]

The principal tool in stochastic optimal control is the dynamic programming principle. Under the requisite conditions, the principle allows the optimal controls to be solved from the terminal timestep backward, one step at a time, rather than attempting to simultaneously solve for the controls over the entire time horizon. In most cases, where an analytic solution does not exist, this produces a lookup table for the optimal control conditional on all state variables.

Because a continuous-time Markov chain has a so-called embedded discrete-time Markov chain, we are able to consider the stochastic control problem in both continuous time and discrete time, and an interesting byproduct of the analysis will be in considering how the emerging dynamics differ.

For all the analysis that follows, we fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$ satisfying the usual conditions of being complete ($\mathcal{F}_0$ contains all ω such that $\mathbb{P}(\omega) = 0$) and right-continuous ($\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$).

3.1 System Description

In order to set up the optimization problem we require an aforementioned 'profit' function, and a description of the stochastic system which we are attempting to control. Below, we identify and describe the variables that will be used in the analysis that follows.

Imbalance Averaging Time Δt_I

A constant, specifying the time window over which the imbalance ratio $I(t)$ (1.1) will be averaged.

Price Change Time Δt_S

A constant, specifying the time window over which price changes will be computed.

Number of Imbalance Bins $\#_{bins}$

A constant, specifying the number of bins (spaced by percentiles, symmetric around zero) into which $I(t)$ will be sorted.

Imbalance ρ_t

The finite, discrete stochastic process that results from sorting $I(t)$ into the imbalance bins $\{1, \ldots, \#_{bins}\}$, and which evolves in accordance with the CTMC $\boldsymbol{Z}$.

Midprice S_t

A stochastic process that evolves in accordance with the CTMC $\boldsymbol{Z}$.

Midprice Change $\Delta S_t = \operatorname{sgn}(S_t - S_{t-\Delta t_S})$

Imbalance & Midprice Change $\boldsymbol{Z}_t = (\rho_t, \Delta S_t)$

A continuous-time Markov chain with generator $\boldsymbol{G}$.

Bid-Ask Half-Spread ξ

As a modelling assumption, this is taken to be a constant. 2ξ is equal to the best ask price minus the best bid price.

Midprice Change when $\Delta S \neq 0$ $\{\eta_{0,z}, \eta_{1,z}, \ldots\} \sim F_z$

Independent, identically-distributed random variables, where the distribution is dependent on the Markov chain state. These variables represent the sign and magnitude of the midprice change when such a change occurs.

Other Agent Market Orders $K_t^{\pm}$

Poisson processes with rate $\mu^{\pm}(\boldsymbol{Z}_t)$. K^+ represents the arrival of another agent's buy market order.

Our Limit Order Posting Depth $\delta_t^{\pm}$

One of the $\mathcal{F}$-predictable processes that we will control. The value of δ^+ dictates how deep on the buy side we will post our buy limit order; if $\delta^+ = 0$ then we are posting at-the-touch.

Our Limit Order Fill Count $L_t^{\pm}$

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

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

Note that L^-, our sell limit order being filled, depends on K^+, an external buy order arriving. To determine the value of the constant κ, we consider the average volume available at the first few depths in relation to the distribution of volumes of incoming market orders; κ can then be fitted to satisfy the relation.

Our Market Orders $M_t^{\pm}$

Our other controlled process. M^+ represents our executing a buy market order. In executing market orders, we assume that the volume of the order is small enough to achieve the best bid/ask price.

Our Market Order Execution Times $\boldsymbol{\tau}^{\pm} = \{\tau_k^{\pm} : k = 1, \dots\}$

An increasing sequence of $\mathcal{F}$-stopping times, representing the time at which we execute market orders.

Cash $X_t^{\boldsymbol{\tau},\delta}$

A stochastic variable representing our cash, initially zero, that evolves according to

$$\begin{aligned}\mathrm{d}X_t^{\boldsymbol{\tau},\delta} = &\underbrace{(S_t + \xi + \delta_t^-)\,\mathrm{d}L_t^-}_{\text{sell limit order}} - \underbrace{(S_t - \xi - \delta_t^+)\,\mathrm{d}L_t^+}_{\text{buy limit order}} \\ &+ \underbrace{(S_t - \xi)\,\mathrm{d}M_t^-}_{\text{sell market order}} - \underbrace{(S_t + \xi)\,\mathrm{d}M_t^+}_{\text{buy market order}}\end{aligned} \tag{3.2}$$

Inventory $Q_t^{\boldsymbol{\tau},\delta}$

A stochastic process representing our assets, initially zero, that satisfies

$$Q_0^{\tau,\delta} = 0, \qquad Q_t^{\tau,\delta} = L_t^+ + M_t^+ - L_t^- - M_t^- \tag{3.3}$$

We define a new variable for our net present value (NPV) at time t, call it $W_t^{\tau,\delta}$, and hence $W_T^{\tau,\delta}$ at terminal time T is our 'terminal wealth.' It is the norm in high-frequency algorithmic trading to finish each trading day with zero inventory in order to avoid overnight positional risk, and thus we 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 $S - \xi \operatorname{sgn} Q - \alpha Q$, where $\xi \operatorname{sgn} Q$ represents crossing the spread in the direction of trading, and α is a penalty constant so that αQ represents receiving a worse price linearly in Q due to walking the book. Hence, $W_t^{\tau,\delta}$ satisfies:

$$W_t^{\tau,\delta} = \underbrace{X_t^{\tau,\delta}}_{\text{cash}} + \underbrace{Q_t^{\tau,\delta}\left(S_t - \xi \operatorname{sgn}(Q_t^{\tau,\delta})\right)}_{\text{book value of assets}} - \underbrace{\alpha\left(Q_t^{\tau,\delta}\right)^2}_{\text{liquidation penalty}} \tag{3.4}$$

The set of admissible trading strategies is the product of the set $\mathcal{T}$ of all $\mathcal{F}$-stopping times, with the set $\mathcal{A}$ of all $\mathcal{F}$-predictable, bounded-from-below depths δ. Informally, this has the intuitive interpretation that for any time $\tau \in \mathcal{T}$, one can recognize when one is at τ without any knowledge of the future. Compare 'the time at which I have won the lottery' with 'the time at which one week later I will have won the lottery'; the former is a stopping time while the latter is not. Likewise, δ being $\mathcal{F}$-predictable implies that it cannot 'see into the future' and is measurable with respect to the information at an earlier time; by being bounded from below, the infimum of the set exists. We will only consider δ with the lower bound $\delta^{\pm} \geq 0$, 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.

To derive an optimal trading strategy via dynamic programming, we will measure the performance of our controls $\boldsymbol{\tau}, \delta$ according to a *performance criterion function* $H^{\tau,\delta}$ that maximizes terminal wealth. As a shorthand, we will write our conditional expectation as $\mathbb{E}_{t,x,s,z,q}[\cdot] = \mathbb{E}\left[\cdot \mid X_{t^-}^{\delta} = x, S_{t^-}^{\delta} = s, Z_{t^-} = z, Q_{t^-}^{\delta} = q\right]$. With the above notation, the

performance criteria function can be written

$$H^{\boldsymbol{\tau},\delta}(t,x,s,\boldsymbol{z},q) = \mathbb{E}_{t,x,s,z,q}\left[W_T^{\boldsymbol{\tau},\delta}\right] \tag{3.5}$$

The *value function*, in turn, is given by

$$H(t,x,s,\boldsymbol{z},q) = \sup_{\boldsymbol{\tau}\in\mathcal{T}_{[t,T]}} \sup_{\delta\in\mathcal{A}_{[t,T]}} H^{\boldsymbol{\tau},\delta}(t,x,s,\boldsymbol{z},q) \tag{3.6}$$

where the subscript $[t,T]$ on the sets of admissible strategies means that we are considering strategies only from the time t, at which the function is being evaluated, to maturity T.

3.2 Dynamic Programming

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

Theorem 4. Dynamic Programming Principle for Optimal Stopping and Control. *(Cartea et al., 2015) If an agent's performance criterion function 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

$$H(t,\boldsymbol{x}) = \sup_{\tau\in\mathcal{T}_{[t,T]}} \sup_{\boldsymbol{u}\in\mathcal{A}_{[t,T]}} \mathbb{E}_{t,\boldsymbol{x}}\left[G(X_\tau^{\boldsymbol{u}})\mathbb{1}_{\tau<\theta} + H(\theta, X_\theta^{\boldsymbol{u}})\mathbb{1}_{\tau\geq\theta}\right] \tag{3.7}$$

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

The function $\mathbb{1}$ in the above definition is the *indicator function*, defined by

$$\mathbb{1}_A = \begin{cases} 1, & \text{if } A \text{ holds;} \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

Theorem 5. Dynamic Programming Equation for Optimal Stopping and Control. *(Cartea et al., 2015) 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*

$$\max \left\{ \partial_t H + \sup_{\boldsymbol{u} \in \mathcal{A}_t} \mathcal{L}_t^{\boldsymbol{u}} H \; ; \; G - H \right\} = 0 \tag{3.9}$$

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

The quasi-variational inequality in (3.9) 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.

3.3 Maximizing Terminal Wealth

In this section we solve the dynamic programming equation (DPE) that results from using the maximal terminal wealth performance criterion. We'll solve for $\mathrm{d}H$ inside the continuation region, and hence $\mathrm{d}M^{\pm} = 0$, in order to then extract out the infinitesimal generator. We'll make use of the shorthand notations $H(\cdot) = H(t, x, s, \boldsymbol{z}, q)$ and $\partial_x = \frac{\partial}{\partial x}$.

$$\mathrm{d}H(t, x, s, \boldsymbol{z}, q) = \sum_i \partial_{x_i} H \, \mathrm{d}x_i \tag{3.10}$$

$$
\begin{aligned}
= \partial_t H \, \mathrm{d}t + & \underbrace{e^{-\kappa\delta^-}\left[H(t, x + (s + \xi + \delta^-), s, \boldsymbol{z}, q - 1) - H(\cdot)\right] \mathrm{d}K^+}_{\text{probability of our sell limit order being filled, times the change in value}} \\
& + \underbrace{e^{-\kappa\delta^+}\left[H(t, x - (s - \xi - \delta^+), s, \boldsymbol{z}, q + 1) - H(\cdot)\right] \mathrm{d}K^-}_{\text{probability of our buy limit order being filled, times the change in value}} \\
& + \sum_{\mathbf{j}} \underbrace{\mathbb{E}\left[H(t, x, s + \eta_{0,\mathbf{j}}, \mathbf{j}, q) - H(\cdot)\right] \, \mathrm{d}Z_{\boldsymbol{z},\mathbf{j}}}_{\text{change in value resulting from a CTMC state switch}}
\end{aligned}
\tag{3.11}
$$

where the first equality is the total differential of H, and the second equality is from applying Itô's formula for Poisson processes and Markov chains to write the differential of H as the sum of increments in H whenever jumps in the respective processes arrive. The summation index $\mathbf{j}$ is over the possible two-dimensional Markov states. The three processes K^+, K^-, and $\boldsymbol{Z}$ account for all the changes in H due to changes in any of the state variables. $\boldsymbol{Z}$ accounts for changes in $\boldsymbol{z}$ and s, and because we are in the continuation region and therefore not executing market orders, $K^\pm$ accounts for all changes in x and q.

We substitute in the identities relating each process to its corresponding compensated process, each of which is a continuous-time martingale (informally, the expectation of a future value is equal to the present value). For Poisson processes we have (Cartea et al., 2015)

$$
\mathrm{d}K^\pm = \mathrm{d}\tilde{K}^\pm + \mu^\pm(\boldsymbol{z}) \, \mathrm{d}t \tag{3.12}
$$

while for the Markov chain, this is (Kurtz, 2004)

$$
\mathrm{d}Z_{\boldsymbol{z},\mathbf{j}} = \mathrm{d}\tilde{Z}_{\boldsymbol{z},\mathbf{j}} + G_{\boldsymbol{z},\mathbf{j}} \, \mathrm{d}t \tag{3.13}
$$

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

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

$$
\begin{aligned}
\mathcal{L}_t^\delta H = & \mu^+(\boldsymbol{z}) e^{-\kappa\delta^-} \left[H(t, x + (s + \xi + \delta^-), s, \boldsymbol{z}, q - 1) - H(\cdot) \right] \\
& + \mu^-(\boldsymbol{z}) e^{-\kappa\delta^+} \left[H(t, x - (s - \xi - \delta^+), s, \boldsymbol{z}, q + 1) - H(\cdot) \right] \\
& + \sum_{\mathbf{j}} G_{\boldsymbol{z},\mathbf{j}} \mathbb{E} \left[H(t, x, s + \eta_{0,\mathbf{j}}, \mathbf{j}, q) - H(\cdot) \right]
\end{aligned}
\tag{3.15}
$$

Now, our DPE has the form

$$
\begin{aligned}
0 = \max \Big\{ & \partial_t H + \sup_{\boldsymbol{u} \in \mathcal{A}_t} \mathcal{L}_t^{\boldsymbol{u}} H \,;\, H(t, x - (s + \xi), s, \boldsymbol{z}, q + 1) - H(\cdot) \,; \\
& H(t, x + (s - \xi), s, \boldsymbol{z}, q - 1) - H(\cdot) \Big\}
\end{aligned}
\tag{3.16}
$$

with boundary conditions

$$
H(T, x, s, \boldsymbol{z}, q) = x + q(s - \xi \operatorname{sgn}(q)) - \alpha q^2 \tag{3.17}
$$

$$
H(t, x, s, \boldsymbol{z}, 0) = x \tag{3.18}
$$

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.

We introduce the ansatz $H(\cdot) = x + q(s - \xi \operatorname{sgn}(q)) + h(t, \boldsymbol{z}, q)$. The first two terms are the wealth plus book value of assets, hence a mark-to-market of the current po-

sition, whereas the $h(t, \boldsymbol{z}, q)$ captures value due to the optimal trading strategy. The corresponding boundary conditions on h are

$$h(T, \boldsymbol{z}, q) = -\alpha q^2 \tag{3.19}$$

$$h(t, \boldsymbol{z}, 0) = 0 \tag{3.20}$$

Substituting this ansatz into (3.15), we get:

$$\begin{aligned}\mathcal{L}_t^\delta H = \mu^+(\boldsymbol{z}) e^{-\kappa\delta^-} \big[\delta^- + \xi[1 + \operatorname{sgn}(q-1) + q(\operatorname{sgn}(q) - \operatorname{sgn}(q-1))] \\ + h(t, \boldsymbol{z}, q-1) - h(t, \boldsymbol{z}, q)\big] \\ + \mu^-(\boldsymbol{z}) e^{-\kappa\delta^+} \big[\delta^+ + \xi[1 - \operatorname{sgn}(q+1) + q(\operatorname{sgn}(q) - \operatorname{sgn}(q+1))] \\ + h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q)\big] \\ + \sum_{\mathbf{j}} G_{\boldsymbol{z},\mathbf{j}} \left[q\mathbb{E}[\eta_{0,\mathbf{j}}] + h(t, \mathbf{j}, q) - h(t, \boldsymbol{z}, q)\right]\end{aligned} \tag{3.21}$$

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

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

This gives us the simplified infinitesimal generator term

$$\begin{aligned}\mathcal{L}_t^\delta H = \mu^+(\boldsymbol{z}) e^{-\kappa\delta^-} \big[\delta^- + 2\xi \cdot \mathbb{1}_{q\ge 1} + h(t, \boldsymbol{z}, q-1) - h(t, \boldsymbol{z}, q)\big] \\ + \mu^-(\boldsymbol{z}) e^{-\kappa\delta^+} \big[\delta^+ + 2\xi \cdot \mathbb{1}_{q\le -1} + h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q)\big] \\ + \sum_{\mathbf{j}} G_{\boldsymbol{z},\mathbf{j}} \left[q\mathbb{E}[\eta_{0,\mathbf{j}}] + h(t, \mathbf{j}, q) - h(t, \boldsymbol{z}, q)\right]\end{aligned} \tag{3.22}$$

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:

$$0 = \partial_{\delta^+} \left[e^{-\kappa\delta^{+*}} \big[\delta^{+*} + 2\xi \cdot \mathbb{1}_{q\le -1} + h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q)\big] \right] \tag{3.23}$$

$$= -\kappa e^{-\kappa\delta^{+*}} \big[\delta^{+*} + 2\xi \cdot \mathbb{1}_{q\le -1} + h(t, \boldsymbol{z}, q+1) - h(t, \boldsymbol{z}, q)\big] + e^{-\kappa\delta^{+*}} \tag{3.24}$$

$$= e^{-\kappa\delta^+}\left[-\kappa\left(\delta^{+*} + 2\xi\cdot\mathbb{1}_{q\leq -1} + h(t,\boldsymbol{z},q+1) - h(t,\boldsymbol{z},q)\right) + 1\right] \tag{3.25}$$

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

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

$$\delta^{+*} = \frac{1}{\kappa} - 2\xi\cdot\mathbb{1}_{q\leq -1} - h(t,\boldsymbol{z},q+1) + h(t,\boldsymbol{z},q) \tag{3.27}$$

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

$$\delta^{+*} = \max\left\{0\,;\ \frac{1}{\kappa} - 2\xi\cdot\mathbb{1}_{q\leq -1} - h(t,\boldsymbol{z},q+1) + h(t,\boldsymbol{z},q)\right\} \tag{3.28}$$

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

$$\delta^{-*} = \max\left\{0\,;\ \frac{1}{\kappa} - 2\xi\cdot\mathbb{1}_{q\geq 1} - h(t,\boldsymbol{z},q-1) + h(t,\boldsymbol{z},q)\right\} \tag{3.29}$$

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

$$\begin{aligned}
&H(t,x-(s+\xi),s,\boldsymbol{z},q+1) - H(\cdot)\\
&\quad= x - s - \xi + (q+1)(s-\xi\,\mathrm{sgn}(q+1)) + h(t,\boldsymbol{z},q+1) &(3.30)\\
&\qquad - \left[x + q(s-\xi\,\mathrm{sgn}(q)) + h(t,\boldsymbol{z},q)\right]\\
&\quad= -\xi\left[(q+1)\,\mathrm{sgn}(q+1) - q\,\mathrm{sgn}(q) + 1\right] + h(t,\boldsymbol{z},q+1) - h(t,\boldsymbol{z},q) &(3.31)\\
&\quad= -2\xi\cdot\mathbb{1}_{q\geq 0} + h(t,\boldsymbol{z},q+1) - h(t,\boldsymbol{z},q) &(3.32)
\end{aligned}$$

and similarly,

$$H(t,x+(s-\xi),s,\boldsymbol{z},q-1) - H(\cdot) = -2\xi\cdot\mathbb{1}_{q\leq 0} + h(t,\boldsymbol{z},q-1) - h(t,\boldsymbol{z},q) \tag{3.33}$$

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

$$\begin{aligned} 0 = \max\Big\{ & \partial_t h + \mu^+(z) e^{-\kappa\delta^{-*}} \left(\delta^{-*} + 2\xi \cdot \mathbb{1}_{q\geq 1} + h(t,z,q-1) - h(t,z,q)\right) \\ & + \mu^-(z) e^{-\kappa\delta^{+*}} \left(\delta^{+*} + 2\xi \cdot \mathbb{1}_{q\leq -1} + h(t,z,q+1) - h(t,z,q)\right) \\ & + \sum_{\mathbf{j}} G_{z,\mathbf{j}} \left[ql\mathbb{E}\left[\eta_{0,\mathbf{j}}\right] + h(t,\mathbf{j},q) - h(t,z,q)\right] \ ; \\ & - 2\xi \cdot \mathbb{1}_{q\geq 0} + h(t,z,q+1) - h(t,z,q) \ ; \\ & - 2\xi \cdot \mathbb{1}_{q\leq 0} + h(t,z,q-1) - h(t,z,q) \Big\} \end{aligned} \tag{3.34}$$

3.4 Interpreting the DPE

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

$$h(\tau_q^+, z, q+1) - h(\tau_q^+, z, q) = 2\xi \cdot \mathbb{1}_{q\geq 0} \tag{3.35}$$

and a sell market order whenever

$$h(\tau_q^+, z, q-1) - h(\tau_q^+, z, q) = 2\xi \cdot \mathbb{1}_{q\leq 0} \tag{3.36}$$

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 (3.34) yields that while in the continuation region, we instead have

$$h(\tau_q^+, z, q+1) - h(\tau_q^+, z, q) \leq 2\xi \cdot \mathbb{1}_{q \geq 0} \tag{3.37}$$

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

Taken together, these inequalities yield

$$-2\xi \cdot \mathbb{1}_{q \geq 0} \leq h(t, z, q) - h(t, z, q+1) \leq 2\xi \cdot \mathbb{1}_{q \leq -1} \tag{3.39}$$

$$-2\xi \cdot \mathbb{1}_{q \leq 0} \leq h(t, z, q) - h(t, z, q-1) \leq 2\xi \cdot \mathbb{1}_{q \geq 1} \tag{3.40}$$

or alternatively,

$$h(t, z, q) \underset{\text{sell if =}}{\overset{\downarrow}{\leq}} h(t, z, q+1) \underset{\text{buy if =}}{\overset{\downarrow}{\leq}} h(t, z, q) + 2\xi, \qquad q \geq 0 \tag{3.41}$$

$$h(t, z, q) \underset{\text{buy if =}}{\overset{\uparrow}{\leq}} h(t, z, q-1) \underset{\text{sell if =}}{\overset{\uparrow}{\leq}} h(t, z, q) + 2\xi, \qquad q \leq 0 \tag{3.42}$$

Recalling the boundary condition $h(t, z, 0) = 0$, (3.41) and (3.42) tell us that the function h is non-negative everywhere. Furthermore, noting the feedback form of our optimal buy limit order depth given in (3.28), together with the inequalities in (3.39) and (3.39), we obtain bounds on our posting depths given by

$$\delta^{+*} = \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q \leq -1} - h(t, z, q+1) + h(t, z, q) \tag{3.43}$$

$$\geq \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q \leq -1} - 2\xi \cdot \mathbb{1}_{q \geq 0} \tag{3.44}$$

$$= \frac{1}{\kappa} - 2\xi \tag{3.45}$$

and

$$\delta^{+*} \leq \frac{1}{\kappa} - 2\xi \cdot \mathbb{1}_{q \leq -1} + 2\xi \cdot \mathbb{1}_{q \leq -1} \tag{3.46}$$

$$= \frac{1}{\kappa} \tag{3.47}$$

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

$$\boxed{\frac{1}{\kappa} - 2\xi \leq \delta^{\pm *} \leq \frac{1}{\kappa}} \tag{3.48}$$

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

Combining (3.28) and (3.35), we know that if δ^+ is determined by its feedback form rather than being floored at zero, then a buy market order is executed if and only if

$$-2\xi \cdot \mathbb{1}_{q \geq 0} = h(t, \boldsymbol{z}, q) - h(t, \boldsymbol{z}, q+1) \tag{3.49}$$

$$= \delta^{+*} - \frac{1}{\kappa} + 2\xi \cdot \mathbb{1}_{q \leq -1} \tag{3.50}$$

$$\left.\begin{matrix} -2\xi, & q \geq 0 \\ 0, & q < 0 \end{matrix}\right\} = \begin{cases} \delta^{+*} - \dfrac{1}{\kappa}, & q \geq 0 \\ \delta^{+*} - \dfrac{1}{\kappa} + 2\xi, & q < 0 \end{cases} \tag{3.51}$$

$$\delta^{+*} = \begin{cases} \dfrac{1}{\kappa} - 2\xi, & q \geq 0 \\ \dfrac{1}{\kappa} - 2\xi, & q < 0 \end{cases} \tag{3.52}$$

$$\delta^{+*} = \frac{1}{\kappa} - 2\xi \tag{3.53}$$

An identical derivation holds for sell market orders. In the next chapter on results, this equality will allow us to gauge the market order behaviour by viewing only the limit order posting depths.

To solve for our ansatz h, we can use any finite differences method (Coleman and Jarrow, 1998) to numerically solve the quasi-variational inequality in (3.34). Since we know that at the terminal time we have $h(T, \boldsymbol{z}, q) = -\alpha q^2$, we can work backward in small time increments of ε, and use a forward approximation for the time derivative, given by $\partial_t h(s) = \frac{h(s+\varepsilon) - h(s)}{\varepsilon}$. This will also require bounding the inventory levels at arbitrary 'large enough' levels at which the behaviour of the function is seen to stabilize. Empirically, we found that $|q| \leq 20$ was sufficient.

4

Maximizing Wealth via Discrete-Time Stochastic Control

4.1 System Description

We now consider the same optimization problem but in discrete time, and we will attempt to reuse the same variable definitions and notation where it makes sense; namely, the constants $\Delta t_I, \Delta t_S, \#_{bins}, \xi$ are defined as before. We will be analysing the embedded discrete time Markov chain, which for any time interval of size Δt can be obtained from the CTMC by considering the transition probability matrix obtained by $\boldsymbol{P} = e^{\boldsymbol{G}\Delta t}$. We have derived the below results with the consideration that $\Delta t = \Delta t_I = \Delta t_S = 1000ms$, though this is not strictly necessary. For convenience, we re-list in discrete-time form the processes we will consider for this control problem:

Imbalance ρ_k

The finite, discrete stochastic process that results from sorting the imbalance ratio $I(t)$ into $\{1, \dots, \#_{bins}\}$, and which evolves in accordance with the Markov chain $\boldsymbol{z}$.

Midprice S_k

A stochastic process that evolves in accordance with the Markov chain $\boldsymbol{z}$.

Midprice Change $\Delta S_k = \text{sgn}(S_k - S_{k-1})$

Imbalance & Midprice Change $\boldsymbol{z}_k = (\rho_k, \Delta S_k)$

A discrete-time 2-dimensional time-homogeneous Markov chain with transition probabilities $\boldsymbol{P}_{ij}$.

Midprice Change when $\Delta S \neq 0$ $\{\eta_{0,\boldsymbol{z}}, \eta_{1,\boldsymbol{z}}, \dots\} \sim F_{\boldsymbol{z}}$

Independent, identically-distributed random variables, where the distribution is dependent on the Markov chain state. This is the price change that accompanies a change in Markov chain state.

Other Agent Market Orders $K_k^{\pm}$

The sum of the Poisson process with rate $\mu^{\pm}(z_k)$ over the past time interval Δt. This allows us to consider a continuous time process in discrete time by looking at how many arrivals there were since the previous timestep. K^+ represents the arrival of other agents' buy market orders.

Our Limit Order Posting Depth $\delta_k^{\pm}$

One of the $\mathcal{F}$-predictable processes that we will control. The value of δ^+ dictates how deep on the buy side we will post our buy limit order - if $\delta^+ = 0$ then we are posting at-the-touch.

Our Limit Order fill count $L_k^{\pm}$

A binary random processes (not Poisson) identifying whether our buy (L^+) or sell (L^-) limit orders were filled. This process is considered in greater detail later in this section.

Our Market Orders $M_k^{\pm}$

Our other controlled process. M^+ represents our executing a buy market order. In executing market orders, we assume that the volume of the order is small enough to achieve the best bid/ask price, and not walk the book.

Cash $X_k^{\tau,\delta}$

A stochastic variable representing our cash, initially zero.

Inventory $Q_k^{\tau,\delta}$

A stochastic process representing our assets, initially zero.

As per a typical discrete-time stochastic control problem, we will consider the following state, control, and random vectors:

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

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

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

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:

$$\boldsymbol{z}_{k+1} = T(\boldsymbol{z}_k, \omega_k) = \sum_{i=0}^{|\Gamma|} i \cdot \mathbb{1}_{\omega_k \in \left[\sum_{j=0}^{i-1} \boldsymbol{P}_{\boldsymbol{z}_k, j}, \sum_{j=0}^{i} \boldsymbol{P}_{\boldsymbol{z}_k, j}\right]} \tag{4.1}$$

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 $\vec{x}_{k+1} = f\left(\vec{x}_k, \vec{u}_k, \vec{w}_k\right)$, can be written explicitly as

$$\begin{pmatrix} x_{k+1} \\ s_{k+1} \\ \boldsymbol{z}_{k+1} \\ q_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ s_k + \eta_{k+1, T(\boldsymbol{z}_k, \omega_k)} \\ T(\boldsymbol{z}_k, \omega_k) \\ q_k \end{pmatrix} + \begin{pmatrix} s_k + \xi + \delta_k^- \\ 0 \\ 0 \\ -1 \end{pmatrix} L_k^- + \begin{pmatrix} -(s_k - \xi - \delta_k^+) \\ 0 \\ 0 \\ 1 \end{pmatrix} L_k^+ \tag{4.2}$$

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

$$\begin{pmatrix} x_k \\ s_k \\ z_k \\ q_k \end{pmatrix} = \begin{pmatrix} x_k \\ s_k \\ z_k \\ q_k \end{pmatrix} + \begin{pmatrix} s_k - \xi \\ 0 \\ 0 \\ -1 \end{pmatrix} M_k^- + \begin{pmatrix} -(s_k + \xi) \\ 0 \\ 0 \\ 1 \end{pmatrix} M_k^+ \tag{4.3}$$

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.

4.2 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 6. Dynamic Programming with Impulse Control in Discrete Time. *(Bensoussan, 2008) Consider a controlled Markov Chain with state space $X = \mathbb{R}^d$, transition 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

$$J_x(V, W) = \mathbb{E}^{V,W,x}\left[\sum_{n=1}^{\infty} \alpha^{n-1} l(y_n, v_n, w_n)\right] \tag{4.4}$$

and the corresponding Bellman equation

$$u(x) = \inf_{\substack{v\in U\\ w\geq 0}} [l(x+w,v) + c(w) + \alpha\Phi^v u(x+w)] \tag{4.5}$$

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 (4.5) belonging to the space of uniformly continuous and bounded functions. Further, this solution is identical to that of

$$u(x) = \min\left\{K + \inf_{w\geq 0}[c_0(w) + u(x+w)]\ ;\ \inf_{v\in U}[l(x,v) + \alpha\Phi^v u(x)]\right\} \tag{4.6}$$

4.3 Maximizing Terminal Wealth

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

$$V_k^{\delta^\pm}(x,s,\boldsymbol{z},q) = \mathbb{E}_{k,x,s,\boldsymbol{z},q}\left[W_T^{\delta^\pm}\right] \tag{4.7}$$

$$= \mathbb{E}_{k,x,s,\boldsymbol{z},q}\left[X_T^{\delta^\pm} + Q_T^{\delta^\pm}(S_T - \xi\,\mathrm{sgn}(Q_T^{\delta^\pm})) - \alpha(Q_T^{\delta^\pm})^2\right] \tag{4.8}$$

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

$$V_T(x,s,\boldsymbol{z},q) = x + q(s - \xi\,\mathrm{sgn}(q)) - \alpha q^2 \tag{4.9}$$

$$V_k(x,s,\boldsymbol{z},q) = \max\Bigg\{ \sup_{\delta^\pm}\big\{\mathbb{E}_{\mathbf{w}}\left[V_{k+1}(f((x,s,\boldsymbol{z},q),\boldsymbol{u},\mathbf{w}_k)\right]\big\}\ ; \\ V_k(x+s_k-\xi,s_k,\boldsymbol{z}_k,q_k-1)\ ; \\ V_k(x-s_k-\xi,s_k,\boldsymbol{z}_k,q_k+1)\Bigg\} \tag{4.10}$$

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:

$$V_k(x,s,\boldsymbol{z},q) = x + q(s - \xi\,\mathrm{sgn}(q)) + h_k(\boldsymbol{z},q) \tag{4.11}$$

with boundary condition $h_k(\boldsymbol{z},0) = 0$ and terminal condition $h_T(\boldsymbol{z},q) = -\alpha q^2$. Substituting this ansatz into (4.10), we obtain

$$0 = \max\Bigg\{ \sup_{\delta^\pm}\big\{\mathbb{E}_{\mathbf{w}}\left[V_{k+1}(f((x,s,\boldsymbol{z},q),\boldsymbol{u},\mathbf{w}_k)\right] - V_k(x,s,\boldsymbol{z},q)\big\}\ ; \\ V_k(x+s_k-\xi,s_k,\boldsymbol{z}_k,q_k-1) - V_k(x,s,\boldsymbol{z},q)\ ; \\ V_k(x-s_k-\xi,s_k,\boldsymbol{z}_k,q_k+1) - V_k(x,s,\boldsymbol{z},q)\Bigg\} \tag{4.12}$$

$$0 = \max\Bigg\{ \sup_{\delta^\pm}\big\{\mathbb{E}_{\mathbf{w}}\big[(s+\xi+\delta^-)L_k^- - (s-\xi-\delta^+)L_k^+ \\ + (L_k^+ - L_k^-)\big(s + \eta_{0,T(\boldsymbol{z},\omega)} - \xi\,\mathrm{sgn}(q + L_k^+ - L_k^-)\big) \\ + q\left(\eta_{0,T(\boldsymbol{z},\omega)} - \xi\left(\mathrm{sgn}(q+L_k^+-L_k^-) - \mathrm{sgn}(q)\right)\right) \\ + h_{k+1}(T(\boldsymbol{z},\omega), q + L_k^+ - L_k^-) - h_k(\boldsymbol{z},q)\big]\big\}\ ; \\ -2\xi\cdot\mathbb{1}_{q\geq 0} + h_k(\boldsymbol{z},q+1)\ ; \\ -2\xi\cdot\mathbb{1}_{q\leq 0} + h_k(\boldsymbol{z},q-1)\Bigg\} \tag{4.13}$$

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

$$\begin{aligned}\sup_{\delta^{\pm}}\Big\{\mathbb{E}_{\mathbf{w}}\Big[&(s+\xi+\delta^{-})L_k^{-}-(s-\xi-\delta^{+})L_k^{+}\\&+(L_k^{+}-L_k^{-})\big(s+\eta_{0,T(\boldsymbol{z},\omega)}-\xi\,\mathrm{sgn}(q+L_k^{+}-L_k^{-})\big)\\&+q\left(\eta_{0,T(\boldsymbol{z},\omega)}-\xi\left(\mathrm{sgn}(q+L_k^{+}-L_k^{-})-\mathrm{sgn}(q)\right)\right)\\&+h_{k+1}(T(\boldsymbol{z},\omega),q+L_k^{+}-L_k^{-})-h_k(\boldsymbol{z},q)\Big]\Big\}\end{aligned}\tag{4.14}$$

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

$$\mathbb{P}[K_k^{+}=0]=\frac{e^{-\mu^{+}(\boldsymbol{z})\Delta t}(\mu^{+}(\boldsymbol{z})\Delta t)^{0}}{0!}=e^{-\mu^{+}(\boldsymbol{z})\Delta t}\tag{4.15}$$

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

$$\mathbb{P}[K_k^{+}>0]=1-e^{-\mu^{+}(\boldsymbol{z})\Delta t}\tag{4.16}$$

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:

$$\mathbb{P}[L_k^{-}=1|K_k^{+}>0]=e^{-\kappa\delta^{-}}\tag{4.17}$$

$$\mathbb{P}[L_k^{-}=0|K_k^{+}>0]=1-e^{-\kappa\delta^{-}}\tag{4.18}$$

$$\mathbb{E}[L_k^{-}]=\mathbb{P}[L_k^{-}=1|K_k^{+}>0]\cdot\mathbb{P}[K_k^{+}>0]\tag{4.19}$$

$$=(1-e^{-\mu^{+}(\boldsymbol{z})\Delta t})e^{-\kappa\delta^{-}}\tag{4.20}$$

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

$$\mathbb{P}[L_k^{-}=1]=p(\delta^{-})=\mathbb{E}[L_k^{-}]\tag{4.21}$$

$$\mathbb{P}[L_k^{-}=0]=1-p(\delta^{-})\tag{4.22}$$

$$\partial_{\delta^{-}}\mathbb{P}[L_k^{-}=1]=-\kappa p(\delta^{-})\tag{4.23}$$

$$\partial_{\delta^{-}}\mathbb{P}[L_k^{-}=0]=\kappa p(\delta^{-})\tag{4.24}$$

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 upper-case Greek letter as shorthand, as will be evident from the analysis.

$$\begin{aligned}\mathbb{E}[\operatorname{sgn}(q + L_k^+ - L_k^-)] &= \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 1] \cdot \operatorname{sgn}(q) \\ &\quad + \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 0] \cdot \operatorname{sgn}(q-1) \\ &\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 1] \cdot \operatorname{sgn}(q+1) \\ &\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 0] \cdot \operatorname{sgn}(q) \end{aligned} \tag{4.25}$$

$$\begin{aligned}&= p(\delta^-)p(\delta^+)\operatorname{sgn}(q) \\ &\quad + p(\delta^-)(1-p(\delta^+))\operatorname{sgn}(q-1) \\ &\quad + (1-p(\delta^-))p(\delta^+)\operatorname{sgn}(q+1) \\ &\quad + (1-p(\delta^-))(1-p(\delta^+))\operatorname{sgn}(q)\end{aligned} \tag{4.26}$$

$$\begin{aligned}&= \operatorname{sgn}(q)\big[1 - p(\delta^+) - p(\delta^-) + 2p(\delta^+)p(\delta^-)\big] \\ &\quad + \operatorname{sgn}(q-1)\big[p(\delta^-) - p(\delta^+)p(\delta^-)\big] \\ &\quad + \operatorname{sgn}(q+1)\big[p(\delta^+) - p(\delta^+)p(\delta^-)\big]\end{aligned} \tag{4.27}$$

$$= \begin{cases} 1 & q \geq 2 \\ 1 - p(\delta^-)(1-p(\delta^+)) & q = 1 \\ p(\delta^+) - p(\delta^-) & q = 0 \\ -\big[1 - p(\delta^+)(1-p(\delta^-))\big] & q = -1 \\ -1 & q \leq -2 \end{cases} \tag{4.28}$$

$$= \Phi(q, \delta^+, \delta^-) \tag{4.29}$$

Similarly:

$$\begin{aligned}\mathbb{E}[L_k^+ \operatorname{sgn}(q + L_k^+ - L_k^-)] &= \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 1] \cdot \operatorname{sgn}(q) \\ &\quad + \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 0] \cdot 0 \operatorname{sgn}(q-1) \\ &\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 1] \cdot \operatorname{sgn}(q+1) \\ &\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 0] \cdot 0 \operatorname{sgn}(q)\end{aligned} \tag{4.30}$$

$$= p(\delta^+)\big[p(\delta^-)\operatorname{sgn}(q) + (1 - p(\delta^-)\operatorname{sgn}(q+1)\big] \tag{4.31}$$

$$= p(\delta^+) \begin{cases} 1 & q \geq 2 \\ 1 & q = 1 \\ (1 - p(\delta^-)) & q = 0 \\ -p(\delta^-) & q = -1 \\ -1 & q \leq -2 \end{cases} \tag{4.32}$$

$$= p(\delta^+)\Psi(q, \delta^-) \tag{4.33}$$

and

$$\mathbb{E}[L_k^- \operatorname{sgn}(q + L_k^+ - L_k^-)] = p(\delta^-)\left[p(\delta^+)\operatorname{sgn}(q) + (1 - p(\delta^+))\operatorname{sgn}(q-1)\right] \tag{4.34}$$

$$= p(\delta^-) \begin{cases} 1 & q \geq 2 \\ p(\delta^+) & q = 1 \\ -(1 - p(\delta^+)) & q = 0 \\ -1 & q = -1 \\ -1 & q \leq -2 \end{cases} \tag{4.35}$$

$$= p(\delta^-)\Upsilon(q, \delta^+) \tag{4.36}$$

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^+.

$$\partial_{\delta^-}\mathbb{E}[\operatorname{sgn}(q + L_k^+ - L_k^-)] = \partial_{\delta^-}\Phi(q, \delta^+, \delta^-) = \kappa p(\delta^-) \begin{cases} 0 & q \geq 2 \\ (1 - p(\delta^+)) & q = 1 \\ 1 & q = 0 \\ p(\delta^+) & q = -1 \\ 0 & q \leq -2 \end{cases} \tag{4.37}$$

$$= \kappa p(\delta^-)\Phi_-(q, \delta^+) \tag{4.38}$$

$$\partial_{\delta^+}\mathbb{E}[\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^+}\Phi(q,\delta^+,\delta^-) = \kappa p(\delta^+)\begin{cases}0 & q\geq 2\\ -p(\delta^-) & q=1\\ -1 & q=0\\ -(1-p(\delta^-)) & q=-1\\ 0 & q\leq -2\end{cases} \tag{4.39}$$

$$= \kappa p(\delta^+)\Phi_+(q,\delta^-) \tag{4.40}$$

$$\partial_{\delta^-}\mathbb{E}[L_k^+\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^-}p(\delta^+)\Psi(q,\delta^-) = \kappa p(\delta^+)p(\delta^-)\begin{cases}0 & q\geq 2\\ 0 & q=1\\ 1 & q=0\\ 1 & q=-1\\ 0 & q\leq -2\end{cases} \tag{4.41}$$

$$= \kappa p(\delta^+)p(\delta^-)\Psi_-(q) \tag{4.42}$$

$$\partial_{\delta^+}\mathbb{E}[L_k^+\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^+}p(\delta^+)\Psi(q,\delta^-) = -\kappa p(\delta^+)\Psi(q,\delta^-) \tag{4.43}$$

$$\partial_{\delta^-}\mathbb{E}[L_k^-\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^-}p(\delta^-)\Upsilon(q,\delta^+) = -\kappa p(\delta^-)\Upsilon(q,\delta^+) \tag{4.44}$$

$$\partial_{\delta^+}\mathbb{E}[L_k^-\operatorname{sgn}(q+L_k^+-L_k^-)] = \partial_{\delta^+}p(\delta^-)\Upsilon(q,\delta^+) = \kappa p(\delta^+)p(\delta^-)\begin{cases}0 & q\geq 2\\ -1 & q=1\\ -1 & q=0\\ 0 & q=-1\\ 0 & q\leq -2\end{cases} \tag{4.45}$$

$$= \kappa p(\delta^+)p(\delta^-)\Upsilon_+(q) \tag{4.46}$$

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:

$$\begin{aligned}\mathbb{E}[h_{k+1}(T(\boldsymbol{z},\omega),q+L_k^+-L_k^-)] = \sum_{\mathbf{j}}\boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[&h_{k+1}(\mathbf{j},q)\left[1-p(\delta^+)-p(\delta^-)+2p(\delta^+)p(\delta^-)\right]\\ &+h_{k+1}(\mathbf{j},q-1)\left[p(\delta^-)-p(\delta^+)p(\delta^-)\right]\\ &+h_{k+1}(\mathbf{j},q+1)\left[p(\delta^+)-p(\delta^+)p(\delta^-)\right]\Big]\end{aligned} \tag{4.47}$$

and its partial derivatives as

$$\begin{aligned}\partial_{\delta^-}\mathbb{E}[h_{k+1}(T(\boldsymbol{z},\omega),q+L_k^+-L_k^-)] = \sum_{\mathbf{j}}\boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[&h_{k+1}(\mathbf{j},q)\left[\kappa p(\delta^-)-2\kappa p(\delta^+)p(\delta^-)\right]\\&+h_{k+1}(\mathbf{j},q-1)\left[-\kappa p(\delta^-)+\kappa p(\delta^+)p(\delta^-)\right]\\&+h_{k+1}(\mathbf{j},q+1)\left[\kappa p(\delta^+)p(\delta^-)\right]\Big]\end{aligned} \tag{4.48}$$

$$\begin{aligned}= \kappa p(\delta^-)\sum_{\mathbf{j}}\boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[&h_{k+1}(\mathbf{j},q)\left[1-2p(\delta^+)\right]\\&+h_{k+1}(\mathbf{j},q-1)\left[-1+p(\delta^+)\right]\\&+h_{k+1}(\mathbf{j},q+1)\left[p(\delta^+)\right]\Big]\end{aligned} \tag{4.49}$$

$$\begin{aligned}\partial_{\delta^+}\mathbb{E}[h_{k+1}(T(\boldsymbol{z},\omega),q+L_k^+-L_k^-)] = \kappa p(\delta^+)\sum_{\mathbf{j}}\boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[&h_{k+1}(\mathbf{j},q)\left[1-2p(\delta^-)\right]\\&+h_{k+1}(\mathbf{j},q-1)\left[p(\delta^-)\right]\\&+h_{k+1}(\mathbf{j},q+1)\left[-1+p(\delta^-)\right]\Big]\end{aligned} \tag{4.50}$$

Now we tackle solving the supremum in (4.14) 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.

$$\begin{aligned}0 = \partial_{\delta^-}\Big\{&(s+\xi+{\delta^-}^*)\mathbb{E}[L_k^-]-(s-\xi-\delta^+)\mathbb{E}[L_k^+]\\&+\mathbb{E}[L_k^+]\left(s+\mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]\right)-\xi\mathbb{E}\left[L_k^+\operatorname{sgn}(q+L_k^+-L_k^-)\right]\\&-\mathbb{E}[L_k^-]\left(s+\mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]\right)+\xi\mathbb{E}\left[L_k^-\operatorname{sgn}(q+L_k^+-L_k^-)\right]\\&+q\mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]-q\xi\mathbb{E}[\operatorname{sgn}(q+L_k^+-L_k^-)]+q\xi\operatorname{sgn}(q)\\&+\mathbb{E}\left[h_{k+1}(T(\boldsymbol{z},\omega),q+L_k^+-L_k^-)\right]-h_k(\boldsymbol{z},q)\Big\}\end{aligned} \tag{4.51}$$

$$
\begin{aligned}
&= \partial_{\delta^-}\Big\{(s+\xi+\delta^{-*})\mathbb{E}[L_k^-] - \xi\mathbb{E}\left[L_k^+ \operatorname{sgn}(q+L_k^+-L_k^-)\right] \\
&\qquad - \mathbb{E}[L_k^-]\left(s+\mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]\right) + \xi\mathbb{E}\left[L_k^- \operatorname{sgn}(q+L_k^+-L_k^-)\right] \\
&\qquad - q\xi\mathbb{E}[\operatorname{sgn}(q+L_k^+-L_k^-)] + \mathbb{E}\left[h_{k+1}(T(\boldsymbol{z},\omega),q+L_k^+-L_k^-)\right]\Big\} \\
&= p(\delta^{-*}) - \kappa p(\delta^{-*})(s+\xi+\delta^{-*}) - \xi\kappa p(\delta^+)p(\delta^{-*})\Psi_-(q) \\
&\quad + \kappa p(\delta^{-*})\left(s+\mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}]\right) - \xi\kappa p(\delta^{-*})\Upsilon(q,\delta^+) - q\xi\kappa p(\delta^{-*})\Phi_-(q,\delta^+) \\
&\quad + \kappa p(\delta^{-*})\sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[h_{k+1}(\mathbf{j},q)\left[1-2p(\delta^+)\right] + h_{k+1}(\mathbf{j},q-1)\left[-1+p(\delta^+)\right] \\
&\qquad\qquad + h_{k+1}(\mathbf{j},q+1)\left[p(\delta^+)\right]\Big]
\end{aligned} \tag{4.52, 4.53}
$$

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

$$
\begin{aligned}
\delta^{-*} &= \frac{1}{\kappa} + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - \xi\left(1 + p(\delta^+)\Psi_-(q) + \Upsilon(q,\delta^+) + q\Phi_-(q,\delta^+)\right) \\
&\quad + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[h_{k+1}(\mathbf{j},q)\left[1-2p(\delta^+)\right] + h_{k+1}(\mathbf{j},q-1)\left[-1+p(\delta^+)\right] \\
&\qquad\qquad + h_{k+1}(\mathbf{j},q+1)\left[p(\delta^+)\right]\Big] \qquad (4.54) \\
&= \frac{1}{\kappa} + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\left(\mathbb{1}_{q\geq 1} + p(\delta^+)\mathbb{1}_{q=0}\right) \\
&\quad + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[h_{k+1}(\mathbf{j},q)\left[1-2p(\delta^+)\right] + h_{k+1}(\mathbf{j},q-1)\left[-1+p(\delta^+)\right] \\
&\qquad\qquad + h_{k+1}(\mathbf{j},q+1)\left[p(\delta^+)\right]\Big] \qquad (4.55)
\end{aligned}
$$

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

$$
\begin{aligned}
\delta^{-*} = \max\Big\{0\,;\; &\frac{1}{\kappa} + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\mathbb{1}_{q\geq 1} + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\left[h_{k+1}(\mathbf{j},q) - h_{k+1}(\mathbf{j},q-1)\right] \\
&- p(\delta^+)\left(2\xi\mathbb{1}_{q=0} - \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\left[h_{k+1}(\mathbf{j},q-1) + h_{k+1}(\mathbf{j},q+1) - 2h_{k+1}(\mathbf{j},q)\right]\right)\Big\}
\end{aligned} \tag{4.56}
$$

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

$$\delta^{+*} = \max\Bigg\{0\ ;\ \frac{1}{\kappa} - \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\mathbb{1}_{q\leq -1} + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\left[h_{k+1}(\mathbf{j},q) - h_{k+1}(\mathbf{j},q+1)\right] \\ - p(\delta^-)\left(2\xi\mathbb{1}_{q=0} - \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\left[h_{k+1}(\mathbf{j},q-1) + h_{k+1}(\mathbf{j},q+1) - 2h_{k+1}(\mathbf{j},q)\right]\right)\Bigg\} \tag{4.57}$$

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 (4.57) into (4.56), while evaluating both at δ^{+*} and δ^{-*} to obtain the optimal posting depth in feedback form:

$$\delta^{-*} = \max\Bigg\{0\ ;\ \frac{1}{\kappa} + \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\mathbb{1}_{q\geq 1} + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\Big[h_{k+1}(\mathbf{j},q) - h_{k+1}(\mathbf{j},q-1)\Big] \\ - (1-e^{\mu^-(\boldsymbol{z})\Delta t})e^{-\kappa \max\left\{0\ ;\ \frac{1}{\kappa} - \mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] - 2\xi\mathbb{1}_{q\leq -1} + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}}\left[h_{k+1}(\mathbf{j},q) - h_{k+1}(\mathbf{j},q+1)\right]\right.} \\ \left. -(1-e^{\mu^+(\boldsymbol{z})\Delta t})e^{-\kappa\delta^-}\left(2\xi\mathbb{1}_{q=0} - \aleph(q)\right)\right\}\ \left(2\xi\mathbb{1}_{q=0} - \aleph(q)\right)\Bigg\} \tag{4.58}$$

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 (4.57) to solve for δ^{+*}.

4.4 Simplifying and Interpreting the DPE

We now turn to simplifying the DPE in (4.13) by substituting in the optimal posting depths as written in recursive form: (4.57) and (4.56). 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:

$$h_k(\boldsymbol{z},q) = \max\Bigg\{ q\mathbb{E}[\eta_{0,T(\boldsymbol{z},\omega)}] + \frac{1}{\kappa}(p(\delta^{+*}) + p(\delta^{-*})) + \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} h_{k+1}(\mathbf{j},q) \\ + p(\delta^{+*})p(\delta^{-*}) \sum_{\mathbf{j}} \boldsymbol{P}_{\boldsymbol{z},\mathbf{j}} \left[h_{k+1}(\mathbf{j},q-1) + h_{k+1}(\mathbf{j},q+1) - 2h_{k+1}(\mathbf{j},q) \right] \; ; \\ - 2\xi \cdot \mathbb{1}_{q\geq 0} + h_k(\boldsymbol{z},q+1) \; ; \\ - 2\xi \cdot \mathbb{1}_{q\leq 0} + h_k(\boldsymbol{z},q-1) \Bigg\} \tag{4.59}$$

As was the case in continuous time, (4.59) yields that whilst in the continuation region, we have

$$h_k(\boldsymbol{z},q) \leq h_k(\boldsymbol{z},q+1) - 2\xi \cdot \mathbb{1}_{q\geq 0} \tag{4.60}$$
$$h_k(\boldsymbol{z},q) \leq h_k(\boldsymbol{z},q-1) - 2\xi \cdot \mathbb{1}_{q\leq 0} \tag{4.61}$$

And these inequalities again give us

$$-2\xi \cdot \mathbb{1}_{q\geq 0} \leq h_k(\boldsymbol{z},q) - h_k(\boldsymbol{z},q+1) \leq 2\xi \cdot \mathbb{1}_{q\leq -1} \tag{4.62}$$
$$-2\xi \cdot \mathbb{1}_{q\leq 0} \leq h_k(\boldsymbol{z},q) - h_k(\boldsymbol{z},q-1) \leq 2\xi \cdot \mathbb{1}_{q\geq 1} \tag{4.63}$$

$$h_k(\boldsymbol{z},q) \overset{\text{sell if =}}{\underset{\downarrow}{\leq}} h_k(\boldsymbol{z},q+1) \overset{\text{buy if =}}{\underset{\downarrow}{\leq}} h_k(\boldsymbol{z},q) + 2\xi, \qquad q \geq 0 \tag{4.64}$$
$$h_k(\boldsymbol{z},q) \underset{\uparrow}{\overset{\text{buy if =}}{\leq}} h_k(\boldsymbol{z},q-1) \underset{\uparrow}{\overset{\text{sell if =}}{\leq}} h_k(\boldsymbol{z},q) + 2\xi, \qquad q \leq 0 \tag{4.65}$$

Recalling the boundary condition $h_k(\boldsymbol{z},0) = 0$, (4.64) and (4.65) 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 (4.59), and in doing so determine whether to execute market orders in addition to posting limit orders. This process then repeats for each step backward.

5

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. As a reminder, the dataset over which we calibrate and backtest is the NASDAQ Historical TotalView-ITCH, timestamped to the millisecond, and the backtests are conducted on the stock tickers listed in Table 5.1.

Table 5.1: List of stocks used in this chapter for stochastic optimal control backtesting, along with the daily average trading volume of each. The chosen stocks are a good sample group of the population of stocks listed on the NASDAQ in that they span the spectrum of low liquidity to high liquidity stocks.

Ticker	Company	Average Daily Volume
FARO	FARO Technologies Inc.	200,000
NTAP	NetApp, Inc.	4,000,000
ORCL	Oracle Corporation	15,000,000
INTC	Intel Corporation	30,000,000
AAPL	Apple Inc.	50,000,000

5.1 Calibration

All tests in this chapter were run using the global set of parameters found in Table 5.2. For each calibration, we then computed the remaining parameters using the formulae in Table 5.3.

In addition, 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 calculated as to partition the interval $[-1, 1]$ into percentile bins symmetric

Table 5.2: List of global parameters used in this chapter for backtesting.

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

Table 5.3: Reference list of calculated parameter formulae for backtesting.

Parameter	Equation	Description
$\boldsymbol{G}$	(2.15)	infinitesimal generator matrix
$\boldsymbol{P}$	(2.8)	transition probability matrix
$\mu^{\pm}$	(2.20)	market order arrival intensities

around zero, where the percentile interval was $100/\#_{bins}$.

As 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 is of interest to explore what results would be obtained if the calibration were 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 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 'with nFPC', standing for 'with non-$\mathcal{F}$-predictable calibration'.

5.2 Dynamics of the Optimal Posting Depths

In Chapters 3 and 4 we derived optimal stochastic controls for the same terminal wealth performance criterion in both continuous and discrete time, which yielded dif-

ferent formulae for the optimal limit order posting depths. In this section we explore the divergence in the resulting dynamics as obtained in an in-sample backtest on `INTC`, calibrating on amalgamated data for the entire 2013 trading year. The dynamics of $\delta^{\pm}$ were obtained using a time-to-maturity of 600 seconds to best depict the behaviour as we approached the end of day trading; as the time-to-maturity horizon increases further, the postings depths tend to stabilize.

From the depth plots we can see that a symmetry that has emerged between δ^+ and δ^- in 'opposite' Markov states, where $Z = (-1, -1)$ is opposite $Z = (+1, +1)$, and $Z = (0, 0)$ is opposite itself. Therefore, compare Figure 5.1 with Figure 5.6, Figure 5.2 with Figure 5.5, and Figure 5.3 with Figure 5.4. Thus, we focus the discussion here on the behaviour of δ^+.

As expected, the plots all have a lower bound of $\delta^+ = 0$. In this calibration we have the fixed parameter $\kappa = 100$ and the calibrated parameter $\xi = 0.005$. From (3.53), we thus know that a necessary but not sufficient condition for a buy market order to be executed is $\delta^+ = \frac{1}{\kappa} - 2\xi = 0$. With respect to the upper bound, both continuous time calibrations respect the $\frac{1}{\kappa} = 0.1$ bound derived in (3.48). However, no such bound was found analytically in the discrete time case; indeed, we see in Figure 5.1 that the discrete time upper bound does exceed this threshold. Nevertheless, it is interesting that the discrete depths still *are* bounded, and that in Figure 5.2 and Figure 5.3 the same $\frac{1}{\kappa}$ upper bound is respected even in the absence of such an analytic result.

Each of the plots display consistent behaviour with respect to changes in the inventory level q. Posting depths increase and converge as q increases from -20 to 0 (light green to purple), and again increase and converge as inventory increases from 0 to 20. However, there is a discontinuity at zero inventory; for example, in Figure 5.2, posting depths jump down from 0.01 at $q = -1$ to 0 at $q = 1$.

Comparing the two calibration methods, wee see that the regular calibration remains stable for longer as time approaches maturity, showing considerable variation only at 400 seconds, whereas the nFPC calibration does so at about 200 seconds. Further, regular calibration results in posting depths that are concentrated slightly closer around the convergence levels. Otherwise, within both continuous and discrete time, the nFPC calibration produces near-identical results compared with the regular calibration method.

Comparing continuous time with discrete time, we see a very large divergence in be-

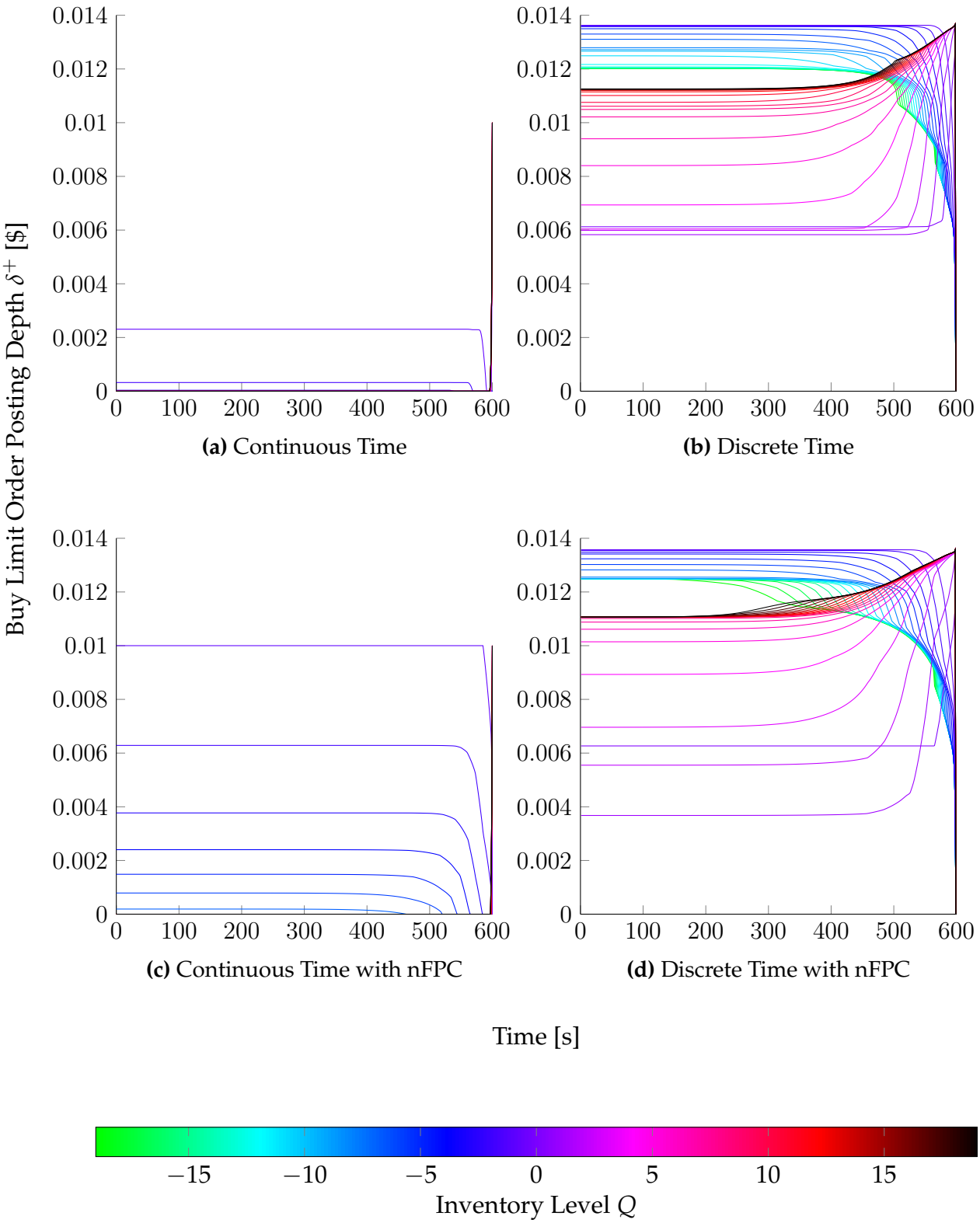


Figure 5.1: Optimal buy depths δ^+ for Markov state $Z = (\rho = -1, \Delta S = -1)$, implying heavy imbalance in favor of sell pressure, and having previously seen a downward price change. We expect the midprice to fall.

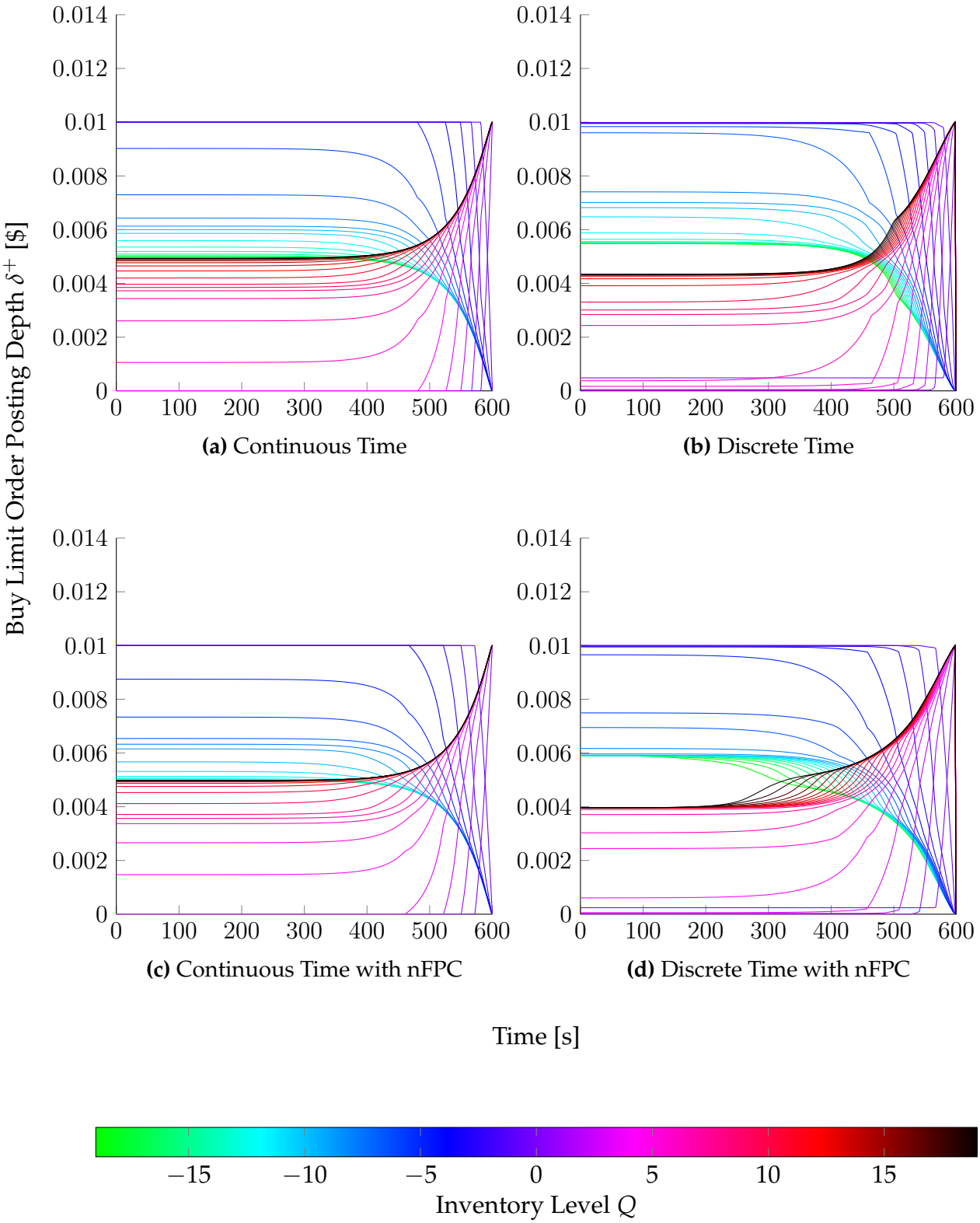


Figure 5.2: Optimal buy depths δ^+ for Markov state $Z = (\rho = 0, \Delta S = 0)$, implying neutral imbalance and no previous price change. We expect no change in midprice.

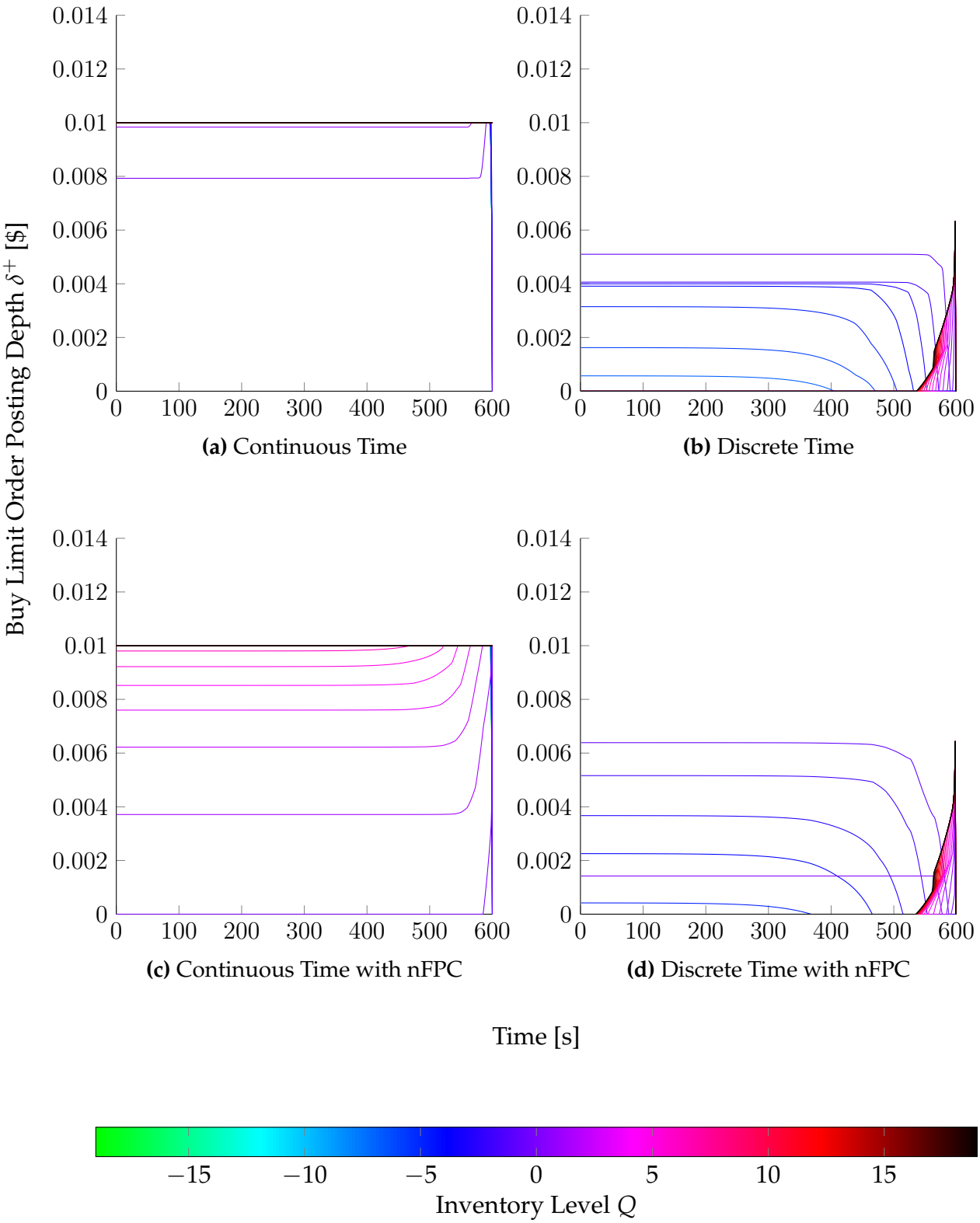


Figure 5.3: Optimal buy depths δ^+ for Markov state $Z = (\rho = +1, \Delta S = +1)$, implying heavy imbalance in favor of buy pressure, and having previously seen an upward price change. We expect the midprice to rise.

haviour. In Figure 5.1, the continuous control produces posting depths concentrated around zero, which suggests that acquiring stock is optimal, whereas the discrete control is posting substantially deeper into the limit order book to avoid purchasing. In Figure 5.2 we see near-identical model behaviour. Since the majority of the day is spent in a Markov state for which $\Delta S = 0$, as seen here, any similarities or correlations in backtesting performance can likely be attributed to the similar behaviour observed in this Markov state. (Recall that ρ, by contrast, is computed via evenly spaced percentiles symmetric around zero, so that time spent in each imbalance state is evenly distributed.) In Figure 5.3 the two models display near opposite behaviours to one another, mirroring what was seen in Figure 5.1. Here the discrete model is posting at depths near zero, mimicking the behaviour of the Naive++ strategy by suggesting that purchasing is optimal, while the continuous model is posting deeper into the book to avoid purchasing.

The continuous-time parameters in Figure 5.1, Figure 5.3, Figure 5.4, and Figure 5.6 display spikes at the terminal time. A possible reason for this occurrence is a so-called *face-lift*, wherein the left-limit of the solution differs from the terminal condition. In this case, we have the terminal condition that h is equal to $-\alpha q^2$, whereas at all other times h is strictly non-negative. It is precisely this difference that may result in the observed spike. As it is actually optimal to liquidate the position prior to reaching the terminal time (so as to avoid the terminal penalty), we would expect that any implementation would avoid posting orders at the instantaneous moment of the terminal time, if not earlier.

Finally, we note that we see relative stability in the posting depths at a time horizon of 600 seconds. This is consistent with the findings in Table 2.3, where we saw that the transition probability matrix $\boldsymbol{P}(t)$ converged for `INTC` to an error threshold of 10^{-10} within 771 timesteps of 1 second each.

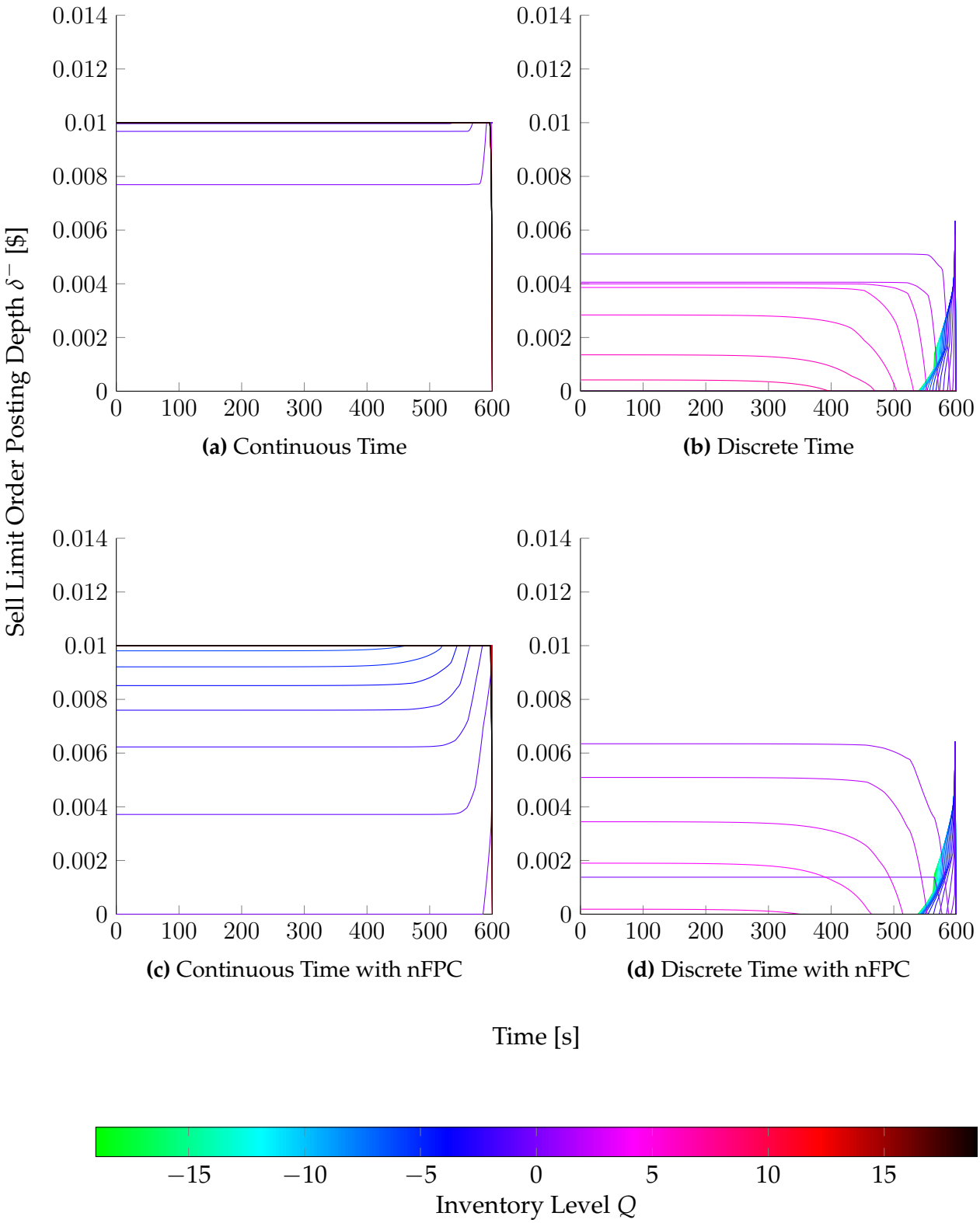


Figure 5.4: Optimal sell depths δ^- for Markov state $Z = (\rho = -1, \Delta S = -1)$, implying heavy imbalance in favor of sell pressure, and having previously seen a downward price change. We expect the midprice to fall.

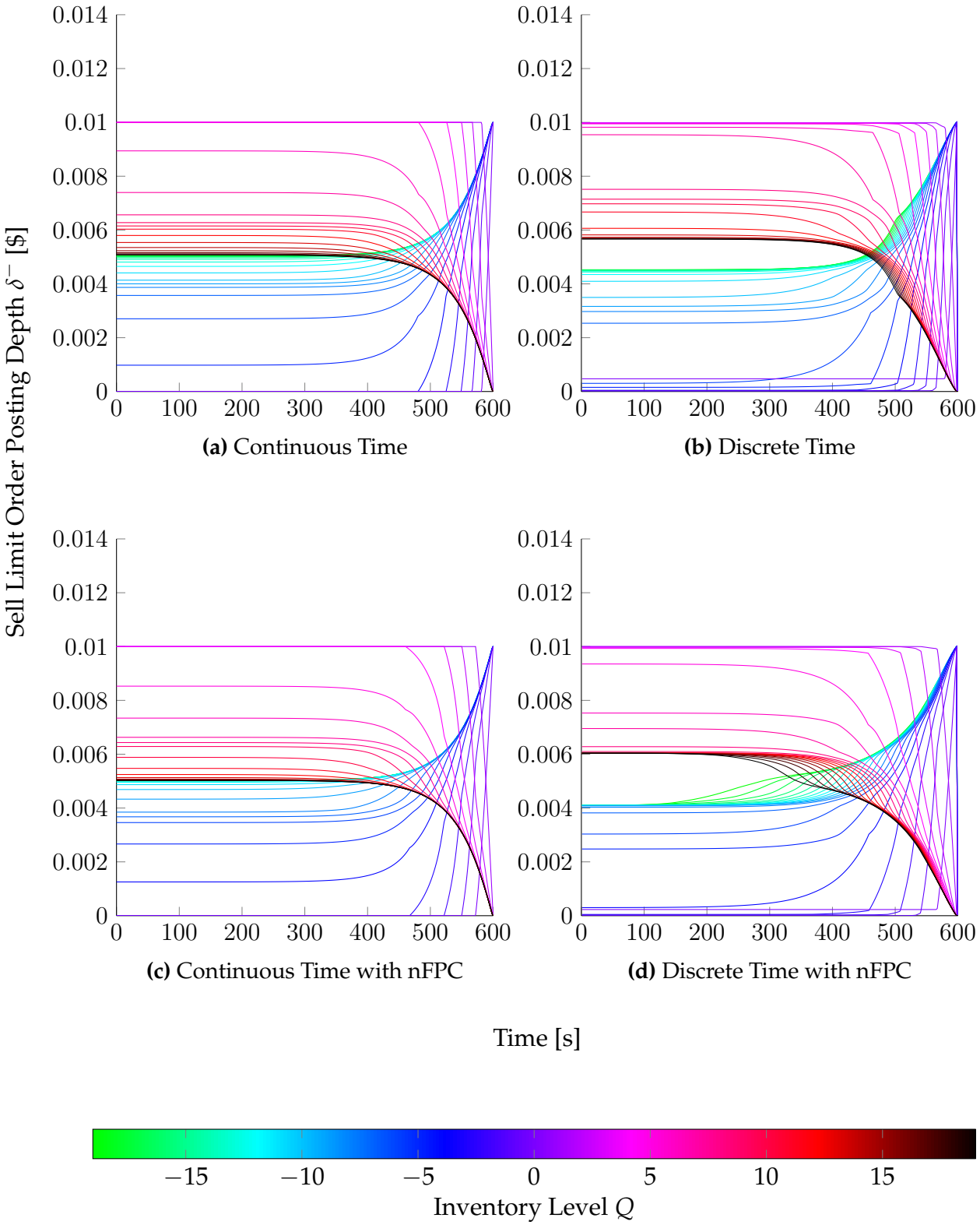


Figure 5.5: Optimal sell depths δ^- for Markov state $Z = (\rho = 0, \Delta S = 0)$, implying neutral imbalance and no previous price change. We expect no change in midprice.

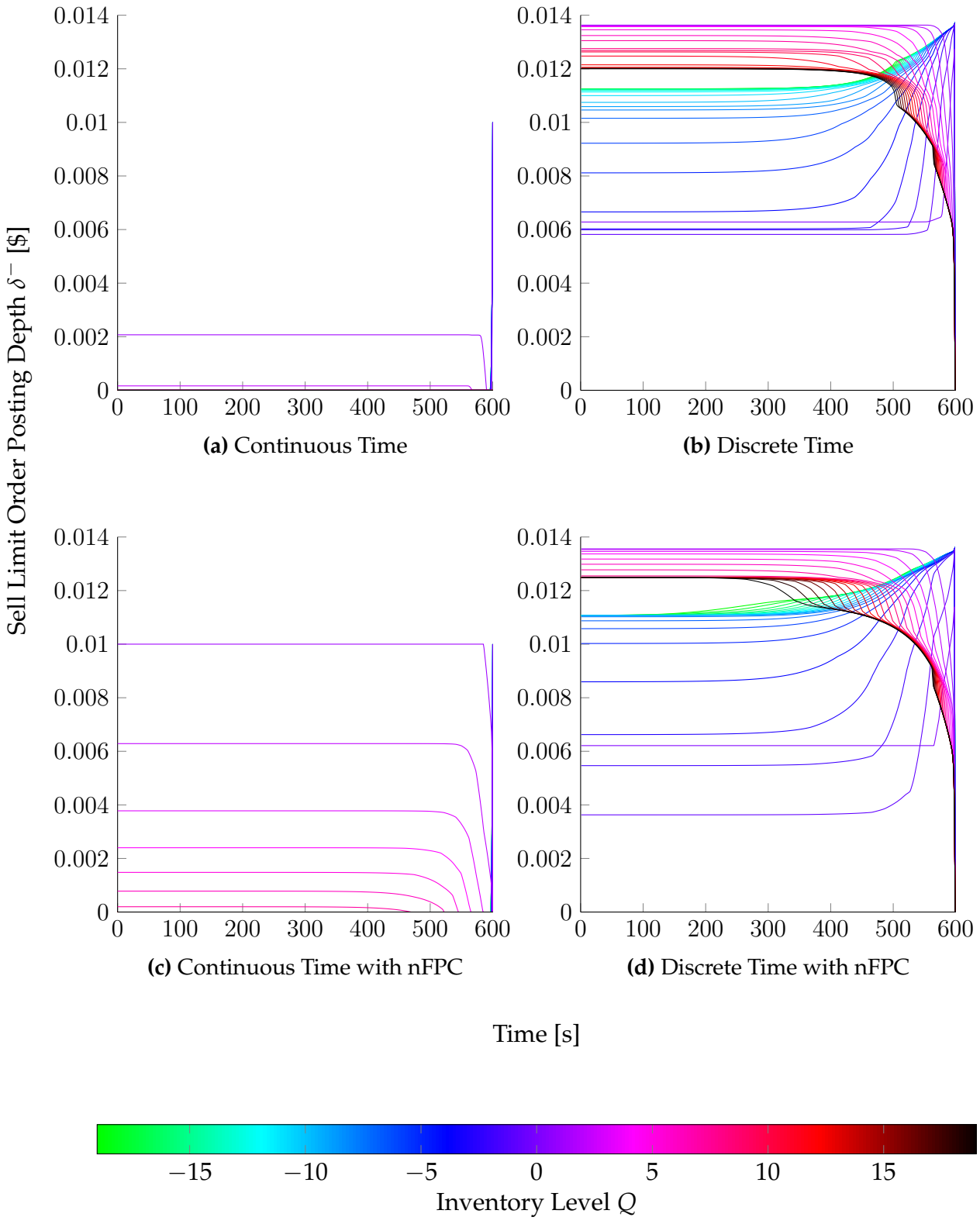


Figure 5.6: Optimal sell depths δ^- for Markov state $Z = (\rho = +1, \Delta S = +1)$, implying heavy imbalance in favor of buy pressure, and having previously seen an upward price change. We expect the midprice to rise.

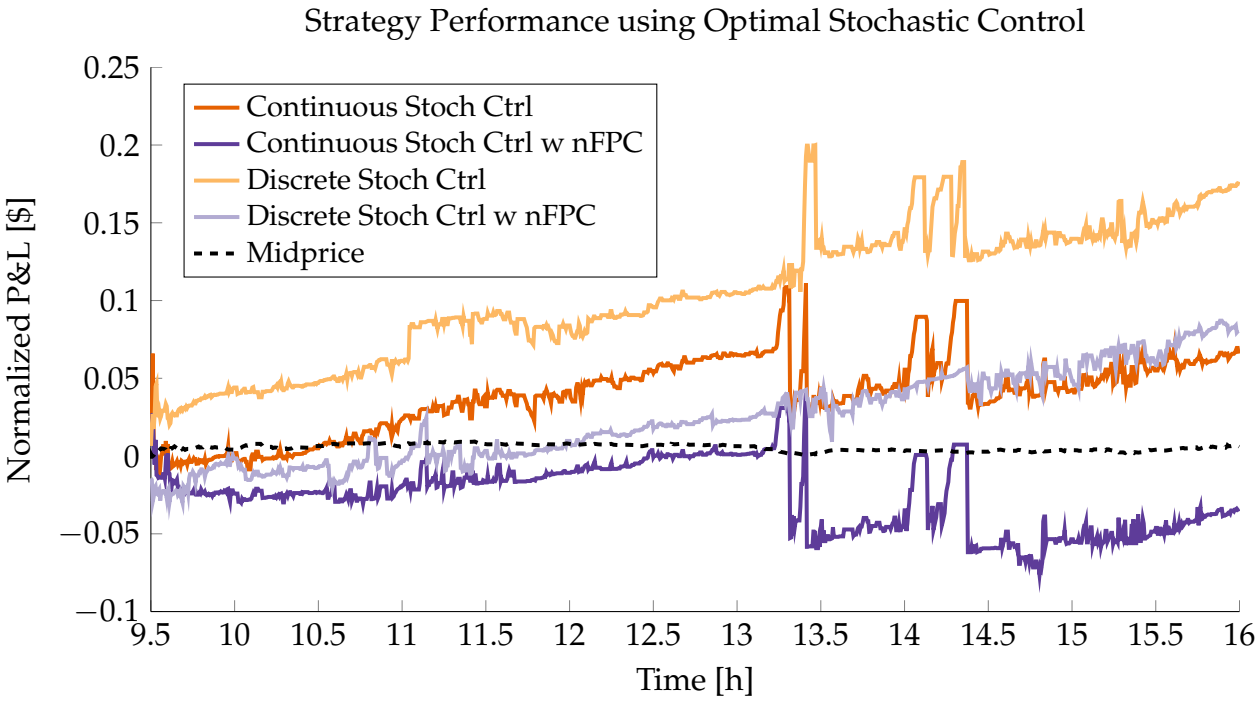
5.2.1 Comparing Optimal Control Performance

In Figure 5.7a we plot the normalized profit and loss (P&L) for the four strategies, calibrated and backtested using data for `ORCL` on an arbitrary single day (2013-05-15). The performance plots display similarities in trajectory, as well as in the distinct spikes between 13 h and 14.5 h. Nevertheless the correlation of arithmetic returns, Table 5.4, shows that the strategies' returns were uncorrelated. Indeed, while the overall paths are similar, on close inspection the individual returns do show differing behaviour.

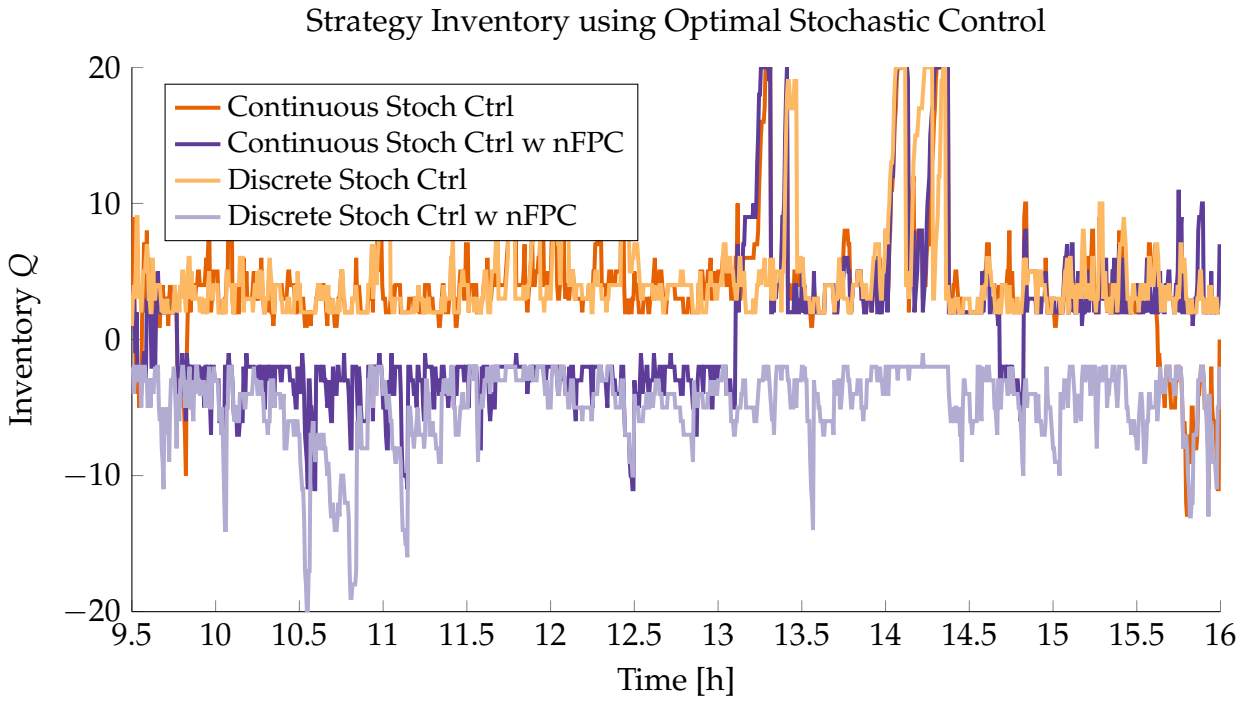
Table 5.4: Correlation matrix of returns for the four stochastic optimal control methods, showing they are uncorrelated.

	Continuous	Continuous with nFPC	Discrete	Discrete with nFPC
Continuous	1.0000			
Continuous with nFPC	-0.0109	1.0000		
Discrete	-0.0120	0.0122	1.0000	
Discrete with nFPC	-0.0015	0.0034	-0.0165	1.0000

We find instead that the returns are cointegrated. On running the Engle-Granger cointegration test (Engle and Granger, 1987) with statistics computed using an augmented Dickey-Fuller test of residuals (Dickey and Fuller, 1979), the τ-test and z-test both returned p-values of 0.001, thus rejecting the null hypothesis of no cointegration. The cointegration relation plotted in Figure 5.8 displays stationarity, thus confirming the existence of a cointegration relation. This result allows us to conclude that although the optimal controls formulae and calibrations are different, there is a a fundamental similarity in the behaviour of the backtested strategies.



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



(b) Inventory comparison of the four stochastic control methods.

Figure 5.7: Comparison of the four stochastic optimal control methods.

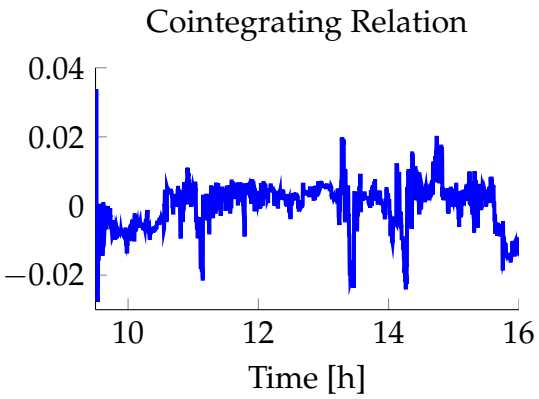


Figure 5.8: Cointegration relation of the four stochastic control methods.

From the inventory plot in Figure 5.7b we see that the strategies 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 nonnegative, and was equal to zero at zero inventory. The interpretation is that there is no added value to having zero inventory, whereas nonzero 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. It is likely that once the strategy enters either positive or negative inventory, the ansatz function h rarely produces the circumstances to cross the inventory sign barrier, perhaps by virtue of the non-linear mark-to-market behaviour on either side of zero inventory.

Table 5.5: Number of trades comparison of the four stochastic control methods.

	Market Orders	Limit Orders
Continuous	536	1280
Continuous with nFPC	1010	1306
Discrete	559	1285
Discrete with nFPC	523	1287

Concerning trade execution, the number of executed market orders and filled limit orders generated by each strategy are presented in Table 5.5. The surge in market orders seen for the Continuous with 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 5.1 we see that we have $\delta^+ = 0$ for almost all inventory values, and in Figure 5.3 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 Continuous with nFPC strategy will execute market orders when possible, as it expects prices to move in the corresponding direction. Regarding overall number of trades, recall that the cost of market order execution was not included in the stochastic control problem. Thus, actual performance would have been negatively affected in proportion to the number of market orders listed.

To better see how the strategies differ in behaviour, in the figures that follow we show a short sample path on a fine timescale spanning about 2 minutes. In Figure 5.9 we plot the midprice path (black line), the optimal posting depths on either side of the real bid/ask prices (gray lines), our execution of market orders (dark blue and dark green), and track incoming external market orders (light blue and light green) that either fill our limit orders (solid lines) or do not (dashed lines). Figure 5.10 plots just the optimal depths as they react to the changing Markov state, allowing a better comparison of the behaviours, as well as highlighting the almost-symmetric behaviour between δ^+ and δ^-. The left and right columns of Figure 5.11 show the effect on P&L and inventory, respectively.

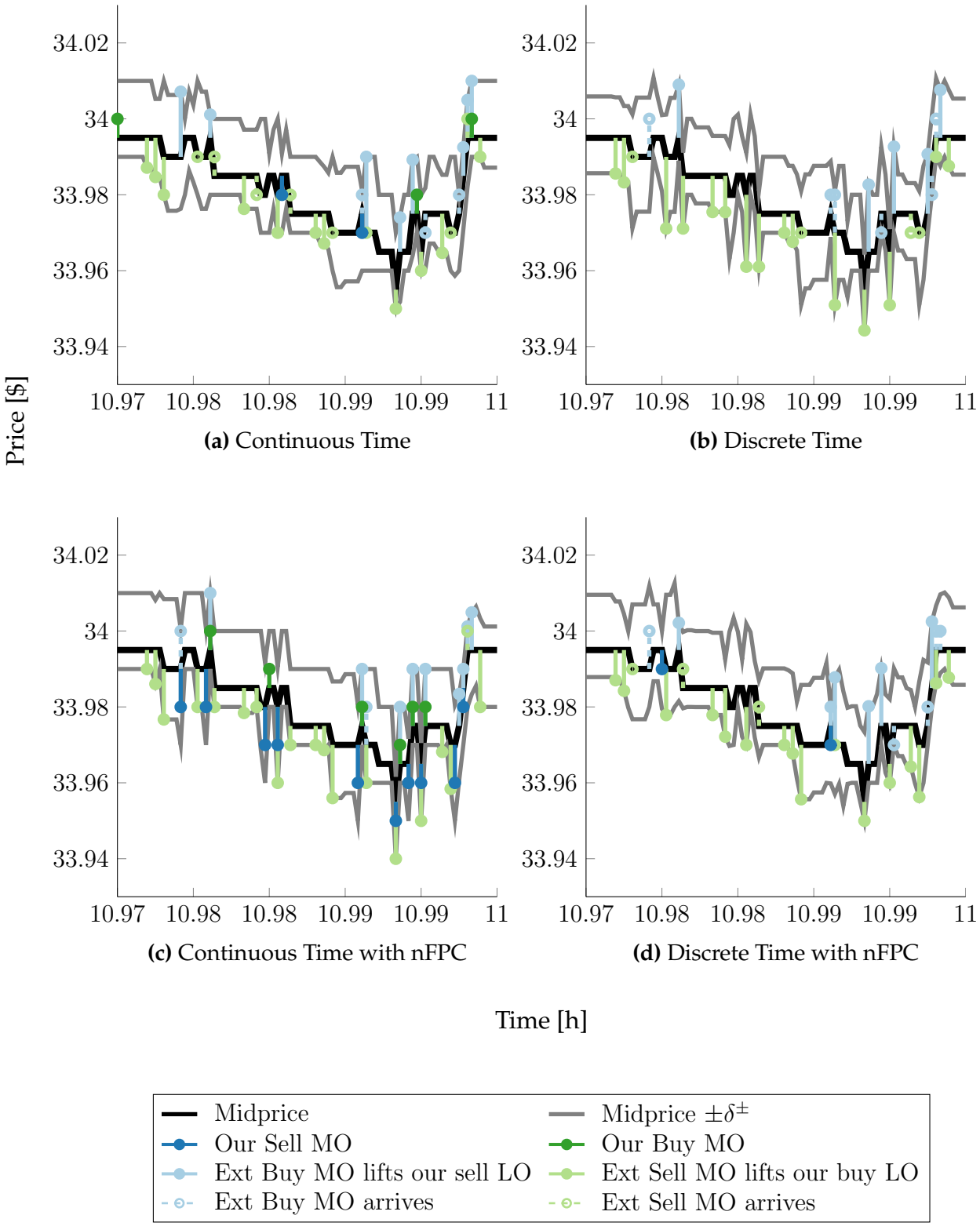


Figure 5.9: Sample paths of the stochastic optimal control strategies, showing price, limit order posting depths, executed market orders, and filled limit orders.

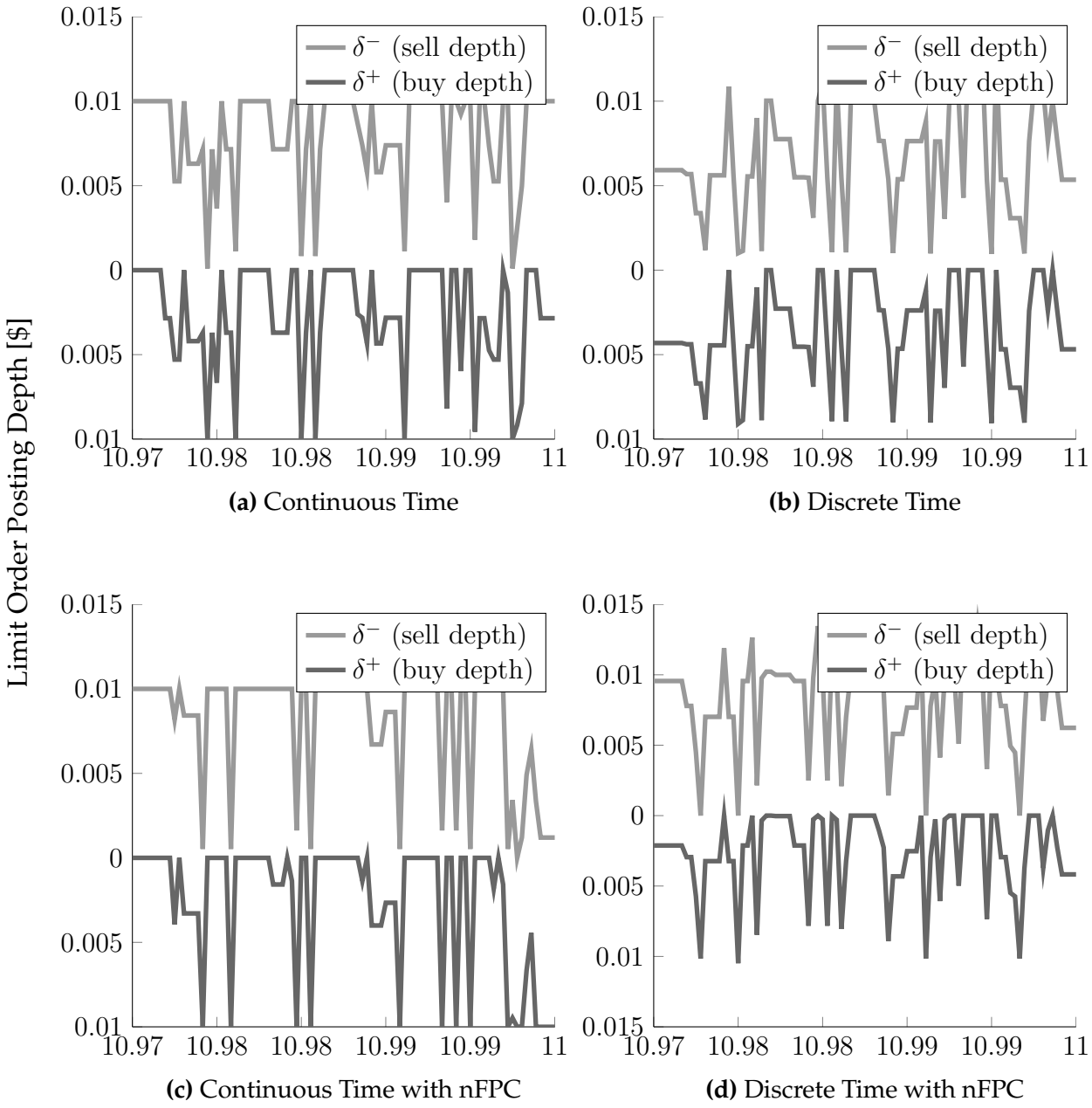
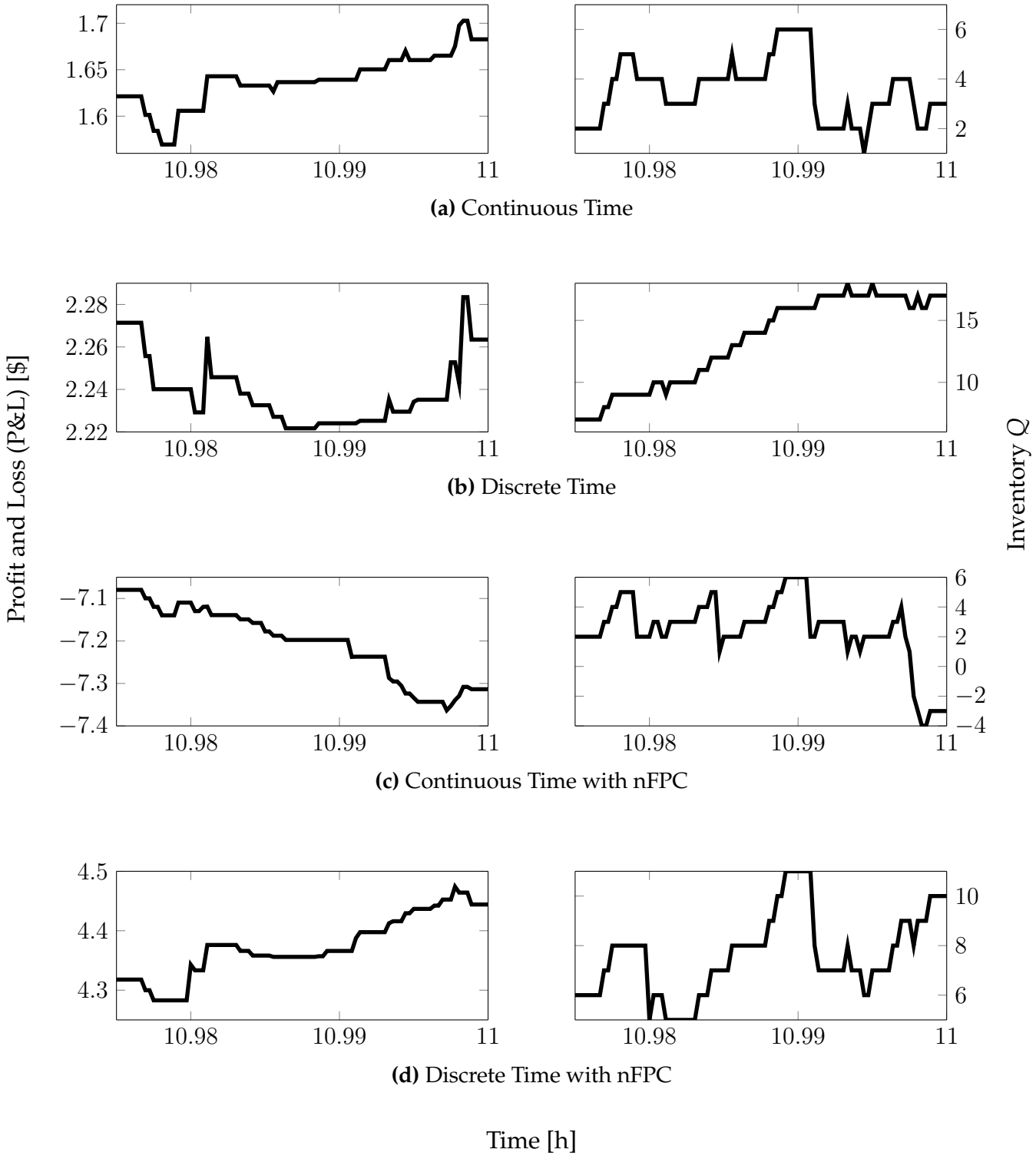


Figure 5.10: The optimal posting depths for each of the stochastic optimal control strategies, isolated from the sample paths in Figure 5.9. Note the scale on the vertical axis: the values are positive and symmetric around zero, intended to show the tandem movements of the optimal posting depths.



(a) Continuous Time

(b) Discrete Time

(c) Continuous Time with nFPC

(d) Discrete Time with nFPC

Figure 5.11: Profit and loss and inventory levels resulting from each of the stochastic optimal control strategies, corresponding to the sample paths in Figure 5.9.

5.3 In-Sample Backtesting

5.3.1 Same-Day Calibration

The first in-sample backtest was done using same-day calibration: calibration was run for each ticker and each trading day of 2013, and each strategy was backtested using the same day's calibration. Each backtest would yield the end of day P&L, average inventory held during the day, and the number of executed market orders and filled limit orders. Figure 5.12 compares the day-over-day performance of the four strategies, while in Table 5.6 we show performance values for several metrics of interest; in particular, risk-adjusted return is calculated by dividing the average end-of-day return by the standard deviation of end-of-day returns.

Since we are calibrating and backtesting using the same underlying data, the calibration should be best attuned to the price dynamics for that particular 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 5.6 we see that trading on `FARO` seldom produces positive P&L. This is not surprising and was mentioned at the conclusion of the exploratory data analysis chapter: `FARO` is illiquid, with daily average volume of about 200,000, and its bid-ask spread is about \$0.20 on average. As executing market orders results in 'crossing the spread', the wide bid-ask spread makes it not profitable to execute market orders. Because our optimal strategies still force us to post limit orders at depths between 0 and $1/\kappa = \$0.01$, the most probable occurrence is that our limit orders are lifted adversely. The strategies exhibit weak regularity of profits when run on `NTAP`, suggesting that its average daily volume of about 4,000,000 is at the threshold required to generate revenue. The strategies performed best when run on the stocks with highest liquidity, `ORCL` and `NTAP`, which have average volumes around 15,000,000 and 30,000,000 respectively. Positive end of day P&L was produced more than 90% of the time, with an average daily return between 0.1 and 0.2 times the initial stock price. The discrete time controller on averaged outperformed its continuous time counterpart; in particular, in the case of `INTC`, a risk-adjusted return of 2.5 was attained.

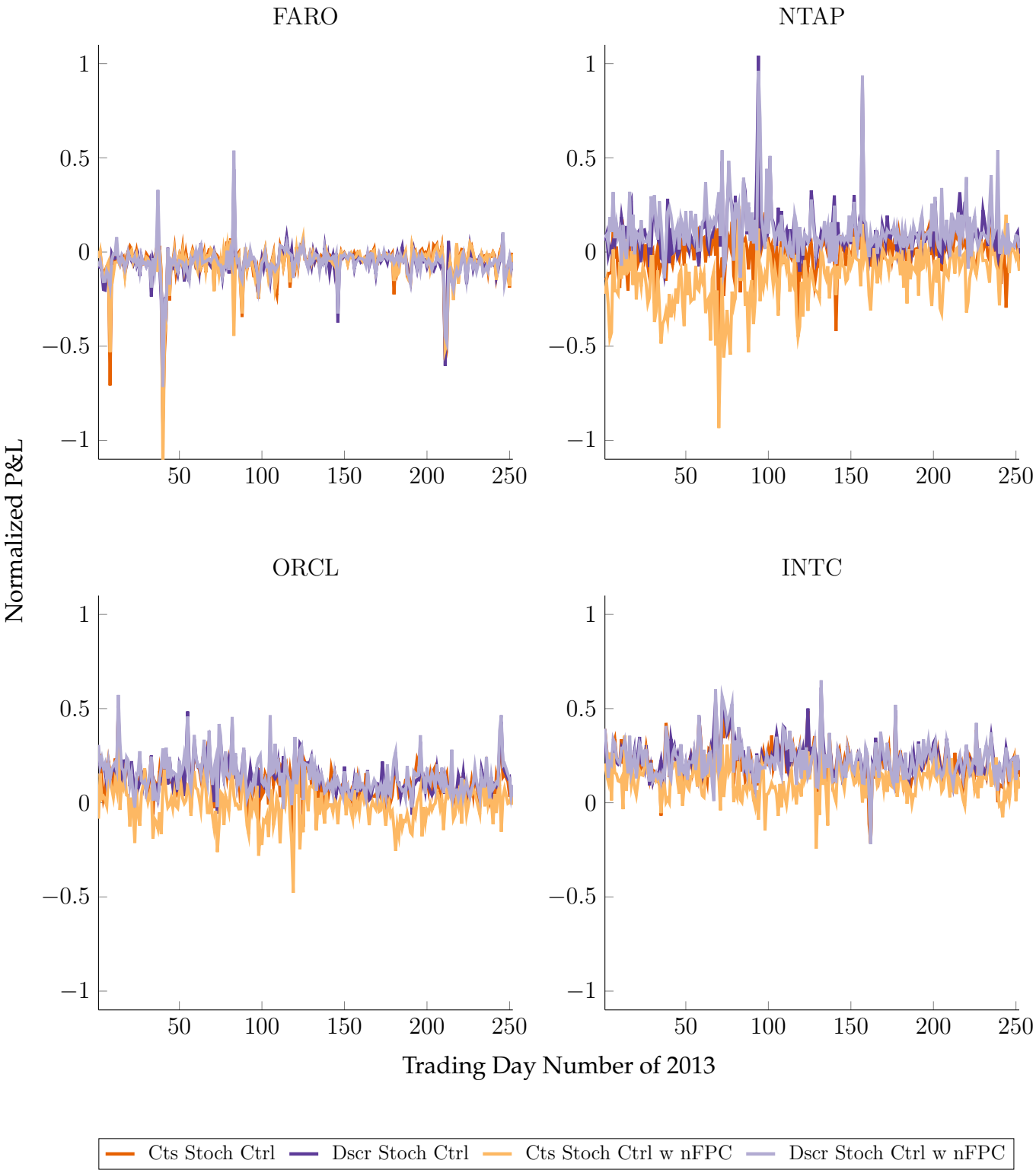


Figure 5.12: End of day strategy performances: in-sample backtesting using same-day calibration.

Table 5.6: Averaged strategy performance results: in-sample backtesting using same-day calibration.

Strategy	Average Return	Risk Adj Return	# Trades	Average Inventory	% Win
FARO					
Naive	-0.879	-0.808	413	0.47	7%
Naive+	0.101	0.107	213	2.45	74%
Naive++	0.002	0.021	7	0.17	50%
Continuous	-0.059	-0.551	201	0.09	18%
Discrete	-0.064	-0.695	210	-0.02	8%
Continuous with nFPC	-0.063	-0.571	204	0.08	14%
Discrete with nFPC	-0.060	-0.662	209	-0.03	9%
NTAP					
Naive	-0.188	-0.316	842	-9.81	23%
Naive+	0.388	0.169	3562	-9.73	74%
Naive++	-0.005	-0.012	157	-0.90	54%
Continuous	-0.006	-0.062	2265	0.40	56%
Discrete	0.099	0.767	1872	4.74	86%
Continuous with nFPC	-0.141	-0.951	2897	0.65	14%
Discrete with nFPC	0.121	0.881	1738	2.82	89%
ORCL					
Naive	-0.105	-0.253	484	1.40	28%
Naive+	-0.034	-0.011	4086	-55.18	61%
Naive++	0.002	0.006	132	0.61	52%
Continuous	0.115	1.348	1874	1.94	92%
Discrete	0.135	1.620	1898	3.93	98%
Continuous with nFPC	-0.010	-0.100	2455	1.32	48%
Discrete with nFPC	0.144	1.501	1759	2.85	97%
INTC					
Naive	-0.082	-0.228	258	-5.21	33%
Naive+	0.365	0.134	3962	-32.50	63%
Naive++	-0.001	-0.003	74	-0.84	48%
Continuous	0.214	2.159	1577	5.17	97%
Discrete	0.232	2.528	1642	4.48	98%
Continuous with nFPC	0.114	1.218	1894	2.01	90%
Discrete with nFPC	0.226	2.202	1569	4.28	98%

5.3.2 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. For example, the calibration obtained on Monday, January 2, 2013 would be used to backtest on Monday, January 9, 2013. Performance values are given in Table 5.7, and Figure 5.13 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 P&L and the low-liquidity stock `NTAP` approximately breaks even. As expected, the week offset calibration underperformed same-day calibration, but remarkably the difference is very small: in the case of `INTC`, the discrete time controller still generates a risk-adjusted return of approximately 2.5, and in this case only returned negative P&L once during the trading year.

The similarity of the results can be interpreted in several ways. First, it is possible that trading behaviour is stable across days of the week, such that substituting one Monday for another yields a similar calibration. This is readily testable by calibrating on a given day and backtesting on the subsequent trading day, instead of a one-week offset. On the other hand, even with dissimilar data, it's possible that the calculation of $\delta^{\pm}$ is stable with respect to day-over-day fluctuations of data.

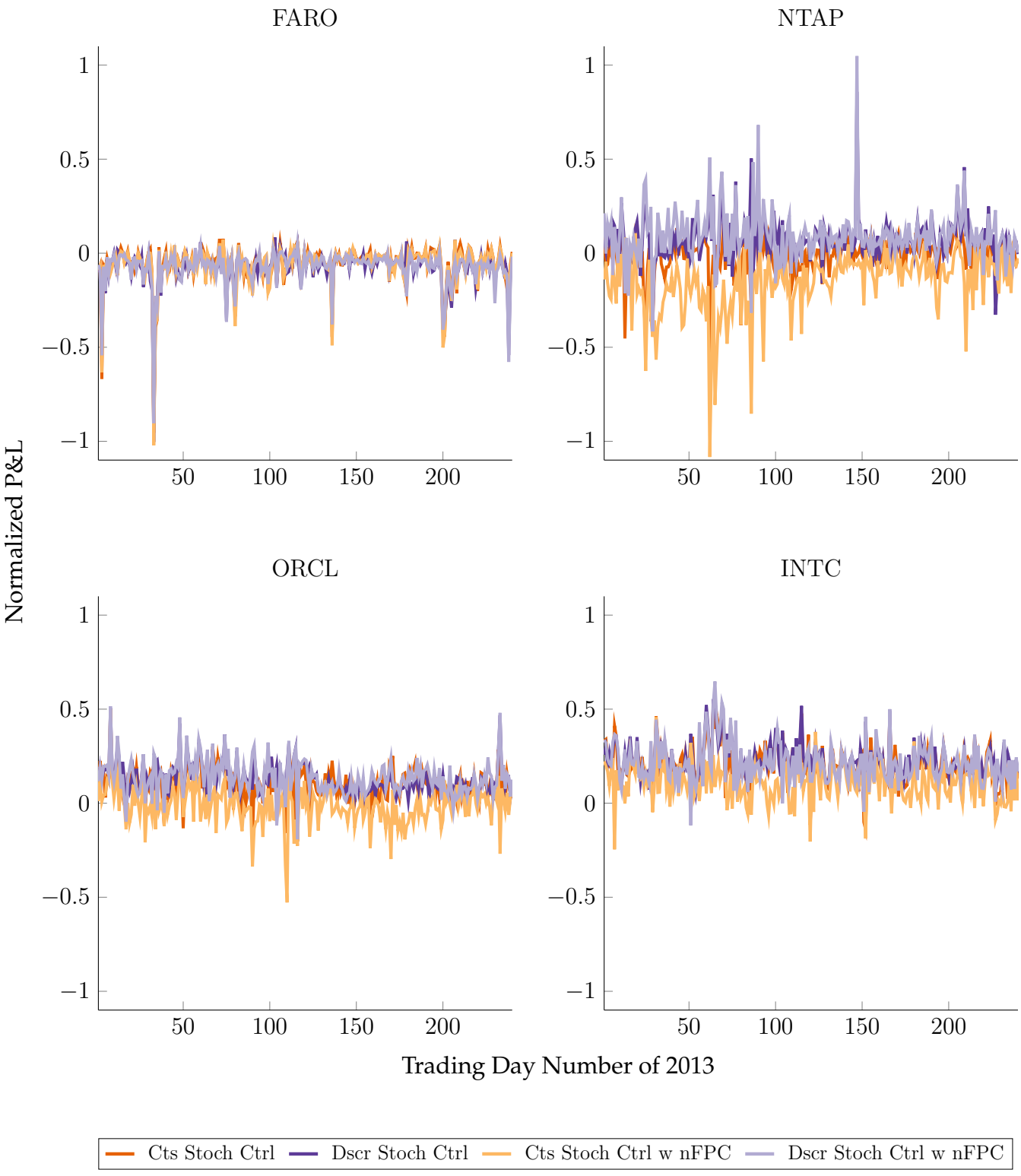


Figure 5.13: End of day strategy performances: in-sample backtesting using a one-week offset for calibration.

Table 5.7: Averaged strategy performance results: in-sample backtesting using a one-week offset for calibration.

Strategy	Average Return	Risk Adj Return	# MO	# LO	Average Invntry	% Win
FARO						
Naive	-1.072	-0.444	435	0	1.74	8%
Naive+	0.045	0.046	0	213	2.26	74%
Naive++	-0.003	-0.027	0	6	0.25	50%
Continuous	-0.060	-0.565	53	149	0.08	20%
Discrete	-0.076	-0.764	80	133	-0.07	7%
Continuous with nFPC	-0.065	-0.590	56	149	0.09	17%
Discrete with nFPC	-0.076	-0.778	78	133	0.07	7%
NTAP						
Naive	-0.303	-0.316	854	0	-10.96	20%
Naive+	0.290	0.122	0	3537	-13.83	73%
Naive++	-0.048	-0.084	0	156	-1.79	52%
Continuous	-0.016	-0.165	830	1405	0.37	52%
Discrete	0.070	0.593	460	1388	5.06	79%
Continuous with nFPC	-0.156	-0.987	1506	1425	0.70	9%
Discrete with nFPC	0.091	0.656	332	1401	3.16	85%
ORCL						
Naive	-0.112	-0.248	492	0	3.66	28%
Naive+	0.066	0.022	0	4049	-50.06	64%
Naive++	0.002	0.005	0	134	0.64	49%
Continuous	0.098	1.181	545	1318	1.86	90%
Discrete	0.126	1.547	578	1310	4.01	97%
Continuous with nFPC	-0.013	-0.130	1069	1365	1.39	47%
Discrete with nFPC	0.135	1.459	416	1338	3.16	96%
INTC						
Naive	-0.057	-0.179	274	0	-3.63	31%
Naive+	0.375	0.138	0	3925	-25.43	65%
Naive++	0.013	0.055	0	77	-0.47	53%
Continuous	0.202	1.995	423	1139	4.76	98%
Discrete	0.226	2.494	501	1136	4.62	99%
Continuous with nFPC	0.107	1.111	681	1187	1.64	87%
Discrete with nFPC	0.215	2.027	401	1156	3.78	99%

5.3.3 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 Table 5.8, and Figure 5.14 compares the day-over-day performance of the various strategies.

Here we backtest only the more liquid of the stocks, `ORCL` and `INTC`. In comparing Table 5.8 with Table 5.7, we note some interesting observations. Again we see the most liquid stock, `INTC`, posting on average the better results using the strategies, suggesting that using a liquid stock is key. (`INTC` started the year at \$21.38 and gained 21.42% over the year, while `ORCL` started at \$34.69 and climbed 10.29%. However, `NTAP` started at \$34.30 and gained 19.94%, similar in performance to `INTC`, and yet performed substantially worse.) Whereas we have seen thus far that the nFPC strategies underperform the regular calibration, here the roles were reversed in terms of performance, number of market orders used, and average inventory held. Across the strategies we see stability in the number of limit orders used, which suggests that this isn't so much strategy dependent as it is externally dependent on outside agents submitting their market orders. In the case of `ORCL` we see that the continuous stochastic control strategy was particularly susceptible to the sharp downward spikes on days 55, 100, and 119, which corresponded to large sell-offs in the market.

From the in-sample backtesting we draw several conclusions regarding the performance of the optimal control strategies:

- average return increases as the underlying stock liquidity increases;
- average return increases as the underlying stock bid-ask spread decreases;
- average return is stable and risk-adjusted return is improved when calibrating over a larger period of time, and is therefore preferred;
- there is no clear victor between regular calibration and the nFPC method.

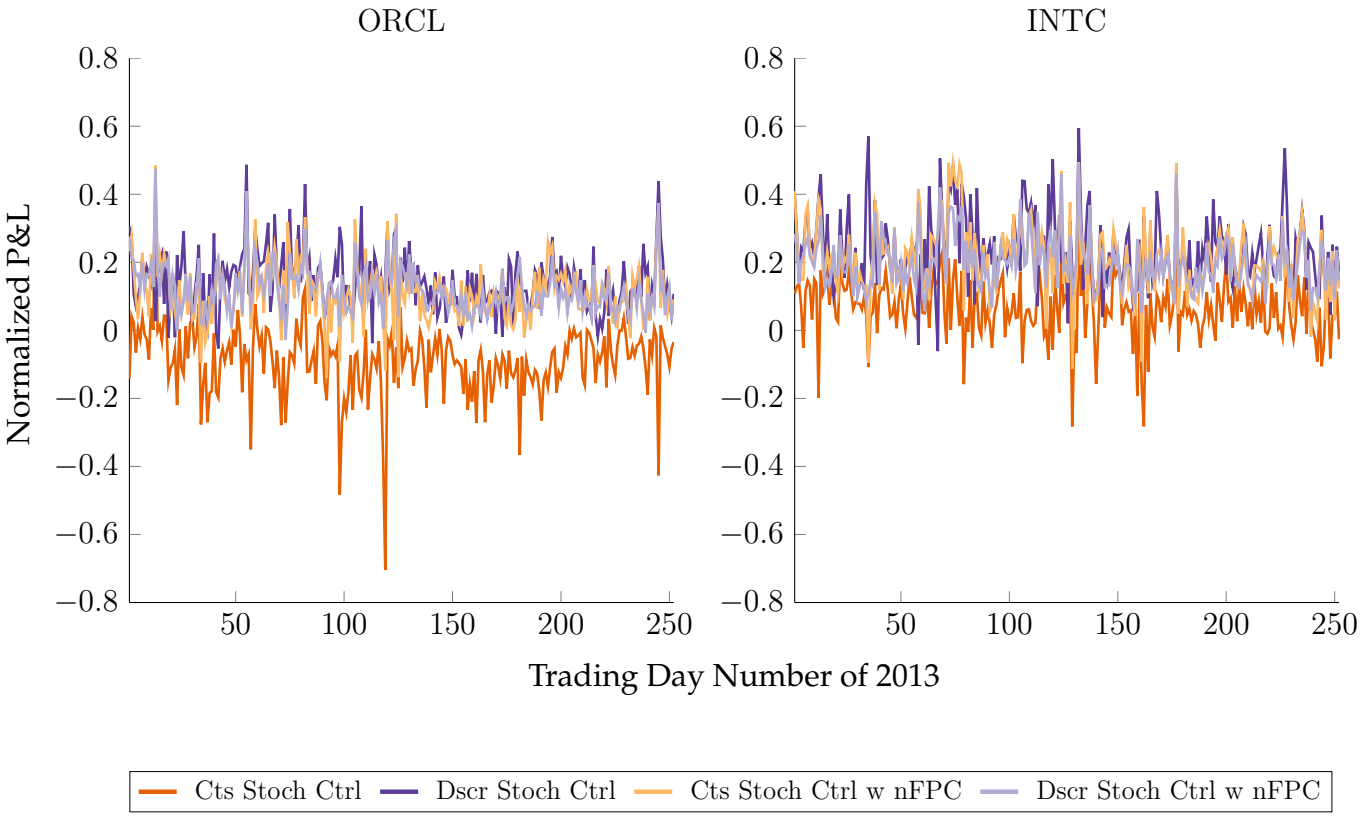


Figure 5.14: End of day strategy performances: in-sample backtesting on 2013 data, using amalgamated annual 2013 data for calibration.

Table 5.8: Averaged strategy performance results: in-sample backtesting on 2013 data, using amalgamated annual 2013 data for calibration.

Strategy	Average Return	Risk Adj Return	# MO	# LO	Average Invntry	% Win
ORCL						
Continuous	-0.089	-0.875	1540	1383	1.19	14%
Discrete	0.140	1.596	368	1344	0.46	96%
Continuous with nFPC	0.113	1.327	476	1338	2.67	94%
Discrete with nFPC	0.118	1.735	590	1337	3.43	99%
INTC						
Continuous	0.065	0.743	888	1207	1.22	84%
Discrete	0.235	2.189	380	1170	1.19	99%
Continuous with nFPC	0.209	2.030	396	1160	5.58	98%
Discrete with nFPC	0.197	2.588	576	1164	3.78	100%

5.4 Out-of-Sample Backtesting

For out-of-sample backtesting we move to using 2014 data that has hitherto remained untouched. We elect to test all four strategies on two stocks, `INTC` and `AAPL`, that have average daily trading volumes of 30m and 45m respectively, and each have a typical bid-ask spread of the minimum 1 cent. `AAPL` underwent a 7-for-1 stock split on 2014-06-09, and prices are adjusted correspondingly in the underlying data; although prior to the split the bid-ask spread was an order of magnitude greater than 1 cent, we retain the 1 cent spread assumption on the adjusted pre-split prices to stay consistent with what was observed after the split. Additionally, we use a sliding calibration window of 1 month (21 trading days) and thus begin trading on the 22nd trading day of the year.

Table 5.9 shows the results of out-of-sample backtesting. The strategies post strong returns and risk-adjusted returns, and provide positive returns on almost every day that was tested. The strategies were run assuming our trade volume was 1 stock for every market order or limit order executed. Thus, to quantify these results in dollar terms, we can do a back of the envelope calculation. For example, for the Discrete with nFPC strategy, we assume that we trade 100 shares at a time, and multiply the average return (0.515 and 0.764 for `INTC` and `AAPL`, respectively) by the average share price during the year ($30 and $95, respectively) by 249 (the total number of trading days in 2014). Thus, we conclude:

Trading `INTC` would have generated revenue of $384,705.

Trading `AAPL` would have generated revenue of $1,807,200.

The capital requirements would have been the full price of the shares, $30 and $95 each, multiplied by the maximum long/short inventory of 20×100 shares - thus $250,000. This represents a return on investment (ROI) of 877%.

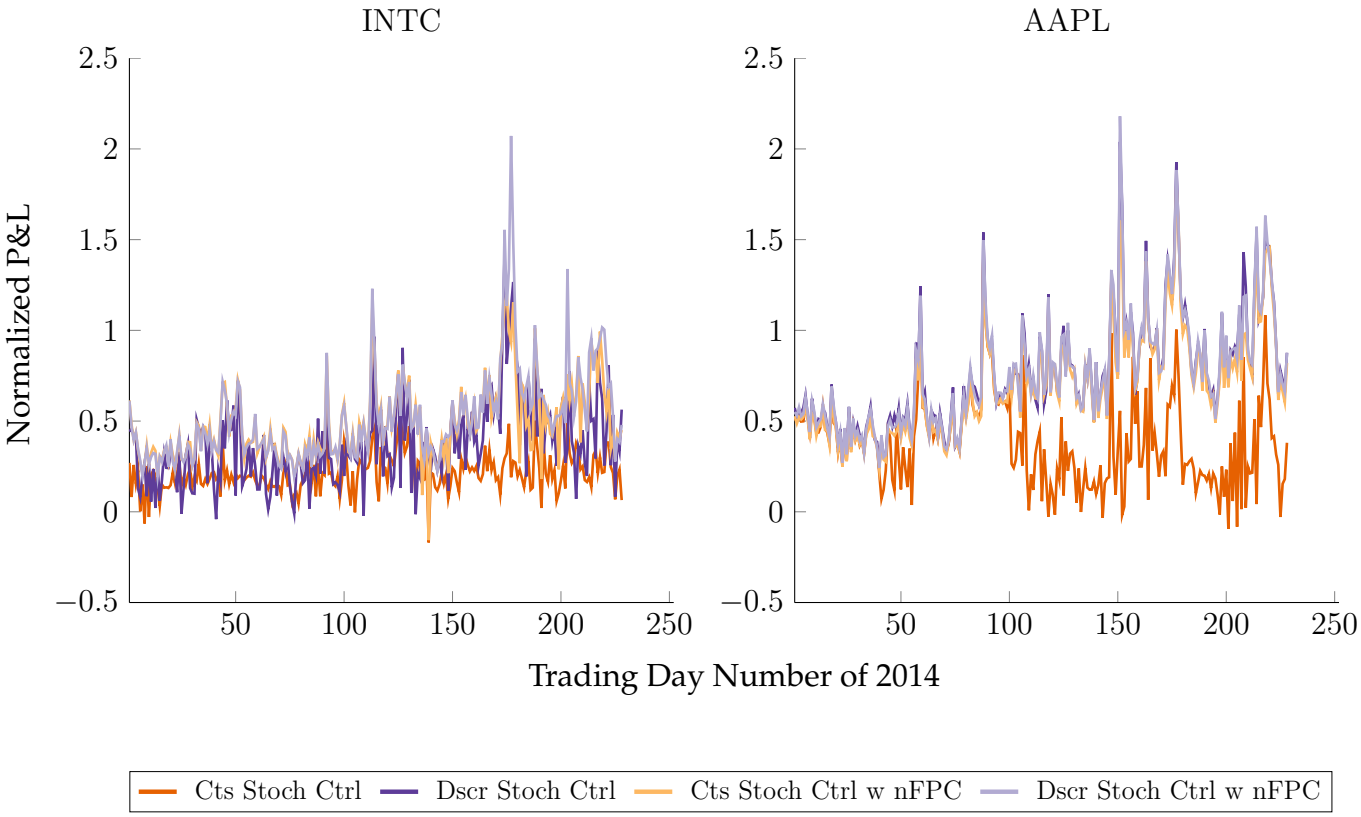


Figure 5.15: End of day strategy performances: out-of-sample backtesting on 2014 data, using amalgamated annual 2013 data for calibration.

Table 5.9: Averaged strategy performance results: out-of-sample backtesting on 2014 data, using amalgamated annual 2013 data for calibration.

Strategy	Average Return	Risk Adj Return	# MO	# LO	Average Invntry	% Win
INTC						
Continuous	0.209	2.112	2118	1758	0.44	98%
Discrete	0.372	1.591	949	1770	-5.89	98%
Continuous with nFPC	0.483	2.364	704	1693	1.46	100%
Discrete with nFPC	0.515	2.033	490	1629	2.81	100%
AAPL						
Continuous	0.378	1.571	3853	6297	-5.80	96%
Discrete	0.761	2.457	830	5566	4.05	100%
Continuous with nFPC	0.710	2.479	1276	5689	2.93	100%
Discrete with nFPC	0.764	2.442	796	5559	3.85	100%

6

Conclusion

The backtesting results indicate that implementing the stochastic strategies on just two stocks would have generated a 877% return on investment in 2014. In addition, however, recurring fees would have included the co-location fees of storage rental and a cross-link to the exchange, approximately $15,000 per month, and a subscription to the ITCH feed as well as access to NASDAQ's order submission protocol, another $15,000 per month. These fee approximations reduce the return on investment down to 359%. However, considering that there are other highly liquid shares with small bid-ask spreads listed on the NASDAQ, such as `DELL` and `MSFT`, of which some may produce similar backtesting results, the results strongly indicate that the stochastic control strategies are monetisable.

Of course, this is a major oversimplification of reality. There are a number of assumptions and modelling choices that were made over the course of this research that must first be revisited prior to attempting forward performance testing (paper trading), let alone taking the system live. In what follows, I address what I consider to be the critical items among the list, each of which indicate a potential future research direction.

Market order costs. Perhaps the easiest change of all, the dynamic programming equations need to account for the fact that market order executions come with a cost from the exchange, as they are essentially taking away liquidity. (Posting, modifying, and cancelling limit orders are all free transactions.) Specifically, the market order cost c would have to appear in the stopping regions/impulse controls of the DPEs. The presumed effect would be widening the upper bound on $\delta^{\pm}$ from $1/\kappa$ to $1/\kappa + c$. This would also presumably decrease the overall incidence of market order executions by the strategies.

Discrete posting depths in increments of 1 tick. Presently, the posting depths $\delta^{\pm}$ are continuous variables, and for example in Figure 5.2 were seen to be between 0 and \$0.01. In reality, one can only post orders in depth increments of 1 tick, which for the majority of stocks is equal to \$0.01. Thus, the current results would have us post at illegal depths. This can partially be solved by rounding the results to the nearest cent, but those rounded depths would no longer have the support of the mathematical derivations, and would likely result in a marked decrease in revenue.

Short term price impact. In order to realize the \$2 million profit, we would be executing orders in sizes of 100 shares at a time. In the case of `AAPL` this would have us trading 600k-700k shares per day. `AAPL` has an average daily trading volume of 50m shares, thus we would be contributing more than 1% of daily volume. It is very likely that due to the large quantity of shares we would be transacting each day, our trading would have an impact on the stock price, which generally has an adverse effect on P&L (Almgren, 2003). Modelling this price impact would be an essential component of either generating more realistic returns, or reducing our daily number of trades.

Accounting for non-homogeneity. Currently the calibration method uses entire trading days to determine the value function H and the posting depths $\delta^{\pm}$, which themselves are applied over the whole trading day. However, we saw in the cross-validation section that there are strong grounds for rejecting the time-homogeneity assumption in the underlying data. A feasible extension would be to have independent calibrations for the first hour of trading, last hour of trading, and mid-day trading, as the three are known to have very distinct behaviours.

Early cut-off with optimal liquidation. As was noted in the stochastic optimal control sections, it is never optimal to wait until maturity and pay the liquidation penalty ϕ per share - thus market orders were executed in bulk immediately prior with no penalty. Thus in addition to independently calibrating the last hour of trading, we could determine an early cut-off time, on the order of minutes, at which the wealth maximization strategy is ended, and proceed to the end of the day with an optimal liquidation/acquisition strategy.

Backtesting engine: tracking LOB queue position. The backtesting engine fills our limit orders with probability $e^{-\kappa\delta}$, an assumption that was made in the stochastic optimal control chapter. Largely, this is because we do not presently track our position in the limit order book queue, and this provides a workaround with some empirical credence. Thus, at present, random numbers effectively determine whether our limit orders get filled, and in particular, a depth of $\delta = 0$ implies guaranteed execution. However, as we have the ability to reconstruct the entire limit order book from the ITCH data, we thus have all the data we need to actually track our position in the queue, and know with certain whether our order would have been partially or fully executed. Additionally, it's currently implicit that at every timestep we cancel our existing order and repost, even if at the same depth; queue-tracking would force us to make this explicit, adding the option of modifying or keeping our existing orders from the previous timestep.

Backtesting engine: information latency. The backtesting engine allows for immediate execution of market orders and posting of limit orders, which ignores the time that a signal takes to be sent from the trading system to the exchange server. As was mentioned in the introduction, minimizing latency is a critical consideration in high-frequency algorithmic trading, and is the justification for the large co-location expenses. Thus it is also paramount to account for its existence in the backtesting engine. A simple 2-4ms lag in execution would realistically simulate the time taken to learn of an event, generate a response, and have the exchange act on the response (Hasbrouck and Saar, 2013).

Backtesting engine: algorithm latency. The historical ITCH data that was fed to the backtesting engine had already been transformed into a reconstructed limit order book. This process will need to be done in real-time as data comes in from the exchange in order to be able to calculate the present Markov chain state. The latency of this algorithm itself must also be minimized and accounted for when backtesting.

Bibliography

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

Almgren, R. F. (2003). Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied mathematical finance*, 10(1):1–18.

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

Bertsimas, D. and Lo, A. W. (1998). Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50.

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

Cartea, Á. and Jaimungal, S. (2013). Modelling asset prices for algorithmic and high-frequency trading. *Applied Mathematical Finance*, 20(6):512–547.

Cartea, Á. and Jaimungal, S. (2015). Incorporating order-flow into optimal execution. *Available at SSRN 2557457*.

Cartea, Á., Jaimungal, S., and Penalva, J. (2015). *Algorithmic and High-Frequency Trading*. Cambridge University Press.

Cartea, Á., Jaimungal, S., and Ricci, J. (2014). Buy low, sell high: A high frequency trading perspective. *SIAM Journal on Financial Mathematics*, 5(1):415–444.

Coleman, T. and Jarrow, R. A. (1998). Cs522 computational tools and methods for finance. Lecture Notes, Cornell University, Department of Computer Science.

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

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276.

Gosavi, A. (2009). Reinforcement learning: A tutorial survey and recent advances. *INFORMS Journal on Computing*, 21(2):178–192.

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.

Guilbaud, F. and Pham, H. (2013). Optimal high-frequency trading with limit and market orders. *Quantitative Finance*, 13(1):79–94.

Hasbrouck, J. and Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, 16(4):646–679.

Inamura, Y. (2006). Estimating continuous time transition matrices from discretely observed data. Technical report, Bank of Japan.

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

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

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

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

Lorenz, J. M. (2008). *Optimal trading algorithms: Portfolio transactions, multiperiod portfolio selection, and competitive online search*. PhD thesis, Technische Universität München.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. Technical report, McKinsey Global Institute.

Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286.

Rubin, R. and Collins, M. (2015). How an exclusive hedge fund turbocharged its retirement plan. *Bloomberg Business*.

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

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