Nash Equilibria in Market Impact Models

Differential Game, Transient Price Impact and Transaction Costs

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

vorgelegt von

Dipl.-Math. oec. Tao Zhang
aus Fuzhou, China

Mannheim, 2014

Dekan:	Professor Dr. Heinz Jürgen Müller, Universität Mannheim
Referent:	Professor Dr. Alexander Schied, Universität Mannheim
Korreferent:	Professor Dr. Martin Schlather, Universität Mannheim

Tag der mündlichen Prüfung: 3. Juni 2014

Abstract

Market impact is the effect caused by transactions that can move asset prices. Nash equilibria describe an optimal state for the players in a non-cooperative game. In this thesis, we combine these two concepts to analyze the competing behavior of two or more large traders in a financial market.

We first consider n risk-averse agents who compete for liquidity in an Almgren–Chriss market impact model. Mathematically, this situation can be described by a Nash equilibrium for a certain linear-quadratic differential game with state constraints. The state constraints enter the problem as terminal boundary conditions for finite and infinite time horizons. We prove existence and uniqueness of Nash equilibria and give closed-form solutions in some special cases. We also analyze qualitative properties of the equilibrium strategies and provide corresponding financial interpretations.

Motivated by the observation that high-frequency traders may use oscillatory trading strategies, which result in a "hot-potato game", we propose quadratic transaction costs in order to make the market more stable and efficient. We identify a critical value for the size of the transaction costs, above which all oscillations disappear and strategies become buy-only or sell-only. Numerical simulations show that for both traders the expected costs can be lower with transaction costs than without. The liquidation costs can increase with trading frequency when there are no transaction costs, but decrease with trading frequency when transaction costs are sufficiently high.

Moreover, we extend this model in several aspects including incorporation of permanent impact, unequal splitting of combined liquidation costs, optimal closed-loop strategies, and we introduce a continuous-time version of the model. In particular, we prove that a Nash equilibrium for continuous-time strategies exists only if the transaction costs are exactly equal to a critical value. For the nonexistence of optimal strategies we give intuitive and mathematical explanations.

Zusammenfassung

Markteinfluss wird durch Transaktionen verursacht, die Vermögenspreise verändern können. Nash-Gleichgewichte beschreiben einen optimalen Zustand in einem nichtkooperativen Spiel für mehrere Spieler. In dieser Arbeit kombinieren wir diese zwei Überlegungen, um das Wettbewerbsverhalten von zwei oder mehreren großen Händlern in einem Finanzmarkt zu analysieren.

Zunächst betrachten wir n risikoaverse Agenten, die in einem Almgren–Chriss Markteinflussmodell um Liquidität konkurrieren. Mathematisch kann diese Situation durch ein Nash-Gleichgewicht für ein linear-quadratisches Differenzialspiel mit Zustandsbeschränkungen beschrieben werden. Die Zustandsbeschränkungen liefern Randbedingungen für endliche und unendliche Zeithorizonte an. Wir beweisen die Existenz und Eindeutigkeit des Nash-Gleichgewichts und geben die Lösung in geschlossener Form in einigen Sonderfällen an. Wir analysieren auch qualitative Eigenschaften der Gleichgewichtsstrategien und geben entsprechende finanzielle Interpretationen an.

Anschließend beobachten wir, dass Hochfrequenzhändler oszillierende Handlungsstrategien verwenden, die ein "hot-potato game" verursachen können. Motiviert durch diese Beobachtung schlagen wir quadratische Transaktionskosten vor, damit der Markt an Stabilität und Effizienz gewinnen kann. Wir bestimmen einen kritischen Wert für die Größe der Transaktionskosten, oberhalb derer alle Oszillationen verschwinden und alle Handelsstrategien nur aus Käufen oder Verkäufen bestehen. Nummerische Simulationen zeigen, dass die erwarteten Liquidationskosten mit Transaktionskosten für beide Händler niedriger als ohne Transaktionskosten sein können. Außerdem erhöhen sich die Liquidationskosten mit der Handelsfrequenz, wenn keine Transaktionskosten existieren und umgekehrt genauso, wenn die Transaktionskosten ausreichend hoch sind.

Darüber hinaus erweitern wir dieses Modell um mehrere Aspekte: den Einbau von permanentem Preiseinfluss, eine ungleiche Aufteilung von kombinierten Liquidationskosten und optimale closed-loop Strategien. Wir leiten auch eine zeitstetige Version des Modells ab. Insbesondere zeigen wir, dass ein Nash-Gleichgewicht für zeitstetige Strategien nur dann existiert, wenn die Transaktionskosten genau gleich dem kritischen Wert entsprechen. Zur Nichtexistenz von optimalen Strategien führen wir sowohl intuitive als auch mathematische Erklärungen an.

Acknowledgment

Foremost, I want to express my deepest gratitude to my advisor Professor Alexander Schied. He not only gave me the opportunity to explore in such a stimulating area of market impact models and Nash equilibria, but also taught me how to do research in an efficient way. He was always friendly and willing to discuss with me about various problems concerning my work, from simple question to mathematical trouble, from technical detail to the whole picture of my research. By following his "learning by doing" style, I improved my mathematical understanding and strengthened the ability of solving problem greatly. Without his guidance and persistent help, this thesis would definitely not have been possible.

Furthermore, I am grateful to Professor Martin Schlather for accepting to be referee of this thesis. Research funding by Deutsche Forschungsgemeinschaft for the project "Mathematical Aspects of Market Impact Modeling" is also appreciated.

Many thanks go to my colleagues Dr. Florian Klöck and Mourad Lazgham for the discussions. They also provided me much useful information when I just came to Mannheim and gave me great support during my work. The wonderful time with them is highly appreciated.

Finally, I am fully grateful to my parents. Thank you so much for supporting my study in Germany for a long time. It can not be counted how much I owe you. To my little daughter Sarina, I am so grateful for your birth and the incomparable happiness that you bring. Much special appreciation I want to express to my wife Su, for your love, patient and sacrifice for our family.

Contents

Chapter 1

Introduction

1.1 Overview

In financial markets, *market impact* is the effect on asset prices that a market participant causes when buying or selling an asset. The transaction moves the price against the buyer or seller, i.e., upward when buying and downward when selling. Market impact is closely related to market liquidity. In many cases "liquidity" and "market impact" are used synonymously. In classical financial market models, markets are usually assumed to be frictionless with infinite liquidity. This assumption works well if trading volumes are small and trading actions do not affect asset prices. However, if trading volumes are sufficiently large, they do impact asset prices. A common observation is that for large execution it is more efficient to split orders into several smaller orders and spread them over a certain time horizon. The problem of how to split and spread orders in an optimal way to minimize liquidation costs is known as *optimal execution problem*.

A first model for optimal execution was introduced by Bertsimas and Lo [1998]. Dynamic discrete-time optimal trading strategies which minimize the expected cost of trading a large block of equity over a fixed time horizon is studied. Almgren and Chriss [1999, 2000] determine optimal trading strategies for liquidation of a large single-asset portfolio to minimize a combination of volatility risk and market impact costs. Its continuous-time variant is studied by Almgren [2003]. In all these models, market impact is described by *temporary* and *permanent impact*. Temporary impact affects only the individual trade that has triggered it, which affects only the current trade and does not last, while permanent impact affects all current and future trades equally and does not change until new large trading comes or new information is revealed. However, only considering temporary and permanent impact may dismiss some features of market impact, especially when the trading time scale is finer. One observes that price impact can be neither temporary nor permanent, but transient. That is, each order triggers immediate price impact which then gradually decays over time. A first quantitative model for *transient impact* is proposed by Obizhaeva and Wang [2013]. In that work, orders are assumed to be block-shaped and the transient impact is identified by exponential resilience, which is linear in the size of the orders. The optimal strategies which minimize the expected execution costs for buying a target position of shares are given explicitly. This model is further extended by Alfonsi et al. [2008], Alfonsi et al. [2010], Alfonsi and Schied [2010], Gatheral [2010], Gatheral et al. [2012] and Alfonsi et al. [2012]. We refer to Gatheral and Schied [2013] and Lehalle [2013] for recent surveys on the price impact literature.

On the other hand, there are usually more than one agent generating market impact in a financial market. In a competing situation each of these agents will try to maximize their own liquidation payoff or utility. Therefore, each agent must optimize their trading strategy by taking into account the market impact caused by the other agents. A nature question is whether there exists a stable point at which *each agent* maximizes their payoff or utility given the strategies of the other agents. This leads to consider a *Nash equilibrium*, see Nash et al. [1950], Nash [1951], the most commonly used concept in a non-cooperative or competing game to analyze the outcome of the strategic interaction of multiple players. In a Nash equilibrium no player can benefit by changing strategies while the other players keep theirs unchanged.

There are several previous works in this area. Brunnermeier and Pedersen [2005] discuss predatory trading by analyzing optimal strategies in Nash equilibria. They find that if a trader has to liquidate a large position of an asset, other strategic traders may take advantage by knowing this. For example, one agent will short sell an asset and then buy back to gain an arbitrage profit if this agent knows that a large position of this asset is being sold, since large liquidation usually pulls down asset price. Carlin et al. [2007] also find this phenomenon by applying an Almgren–Chriss market impact model for risk-neutral agents who have the same liquidation time horizon. In this modeling framework equilibrium strategies are given explicitly by linear combinations of exponential functions. This model is further extended by Schöneborn and Schied [2009]. There competitors are allowed to have a longer time horizon than the seller who has a large liquidation. It is found that depending on market conditions, there is either predatory trading or liquidity provision in Nash equilibria. By extending the model of Obizhaeva and Wang [2013] for two competing agents, Schöneborn [2008] observes that the equilibrium strategies can exhibit strong oscillations in both of open-loop and closed-loop models. An intuitive reason for the oscillatory strategies is to protect against possible predatory trading by the other agent.

Furthermore, Moallemi et al. [2012] consider a dynamic game for one seller and one predator with asymmetric information. An algorithm for computing perfect Bayesian equilibrium is presented. Carmona and Yang [2011] use numerical methods to analyze a system of coupled HJB equations arising from a closed-loop Nash equilibrium for two utility-maximizing agents. Using a mean field game approach, Carmona and Webster [2012] analyze Nash equilibria for an infinite player problem and Lachapelle et al. [2013] consider a stochastic market impact model for institutional investors and high frequency traders.

1.2 Statement of results

In this thesis we consider Nash equilibria in two main types of market impact models. The main goal is to determine the existence and uniqueness of Nash equilibria and to analyze the properties of optimal strategies. In Chapter 2 we study n risk-averse agents who compete for liquidity in an Almgren-Chriss market impact model. In Chapter 3 we analyze a Nash equilibrium between two high-frequency traders in a simple market impact model with transient price impact and additional quadratic transaction costs. In the following we state the main results of this thesis.

Chapter 2

We consider a standard continuous-time Almgren and Chriss [2000] framework for n investors who are active over a fixed time period or an infinite time horizon. We prove the existence and uniqueness of Nash equilibria and give closed-form solutions in some special cases. We also analyze qualitative properties of the equilibrium strategies and provide corresponding financial interpretations. This chapter is based on Schied and Zhang [2013b].

In Section 2.1, we assume that n investors are active over a fixed time period $[0,T]$. The trading strategy employed by the i-th investor is denoted by $X_i = (X_i(t))_{t\in[0,T]} \in \mathcal{X}(x_i,T)$, where $\mathcal{X}(x_i,T)$ denotes the class of all admissible strategies with an initial value x_i. When the n investors use the respective strategies $X_1,\dots,X_n$, the price process is given by

$$S^{X_1,\dots,X_n}(t) := S^0(t) + \gamma\sum_{j=1}^n (X_j(t)-X_j(0)) + \lambda\sum_{j=1}^n \dot X_j(t), \qquad t\in[0,T],$$

where S^0 follows a Bachelier model with an extra drift b,

$$S^0(t) = S_0 + \sigma W(t) + \int_0^t b(s)\,ds.$$

We analyze open-loop Nash equilibria for mean-variance optimization and constant absolute risk aversion (CARA) utility maximization.

Definition. *Suppose that $n\in\mathbb{N}$, $x_1,\dots,x_n\in\mathbb{R}$ are initial asset positions, and that $\alpha_1,\dots,\alpha_n$ are nonnegative coefficients of risk aversion.*

- *A Nash equilibrium for mean-variance optimization consists of a collection $X_1^*,\dots,X_n^*$ of deterministic strategies such that for each i and $\boldsymbol{X}_{-i}^* = \{X_1^*,\dots,X_{i-1}^*,X_{i+1}^*,\dots,X_n^*\}$ the strategy $X_i^*\in\mathcal{X}_{det}(x_i,T)$ maximizes the mean-variance functional*

$$\mathbb{E}[\,\mathcal{R}(X|\boldsymbol{X}_{-i}^*)\,] - \frac{\alpha_i}{2}\operatorname{var}(\mathcal{R}(X|\boldsymbol{X}_{-i}^*)).$$

- *A Nash equilibrium for CARA utility maximization consists of a collection $X_1^*,\dots,X_n^*$ of admissible strategies such that for each i the strategy $X_i^*\in\mathcal{X}(x_i,T)$ maximizes the expected utility*

$$\mathbb{E}[\,u_{\alpha_i}(\mathcal{R}(X|\boldsymbol{X}_{-i}^*))\,]$$

  over all $X\in\mathcal{X}(x_i,T)$.

We then derive the existence and uniqueness of Nash equilibria for mean-variance optimization.

Result (Theorem 2.1.2)**.** *For given $n\in\mathbb{N}$, $\alpha_1,\dots,\alpha_n\ge 0$, and $x_1,\dots,x_n$ there exists a unique Nash equilibrium $X_1^*,\dots,X_n^*$ for mean-variance optimization. It is given as the unique solution of the following second-order system of differential equations*

$$\alpha_i\sigma^2 X_i(t) - 2\lambda\ddot X_i(t) = b(t) + \gamma\sum_{j\neq i}\dot X_j(t) + \lambda\sum_{j\neq i}\ddot X_j(t),$$

with two-point boundary conditions

$$X_i(0) = x_i \text{ and } X_i(T) = 0$$

for $i = 1, 2, \ldots, n$.

This result is one of the main contributions of this thesis. We see that this situation can be described by a linear-quadratic differential game with state constraints. The state constraints enter the problem as terminal boundary conditions. To prove this result, we first show by way of contradiction that there exists at most one Nash equilibrium for mean-variance optimization. Then we show that in a Nash equilibrium, each strategy must be a solution of a second order differential equation with boundary conditions if they are sufficiently smooth. The proof is completed by showing the existence of a solution to be an n-dimensional two-point boundary value problem. As a consequence, we also show that the Nash equilibrium for mean-variance optimization is also a Nash equilibrium for CARA utility maximization. It follows as in Schied et al. [2010] that each Nash equilibrium in the class of deterministic strategies is also a Nash equilibrium in the class of adaptive strategies.

If the risk aversions of the n investors are identical and there is no price drift, we obtain a closed form of the optimal strategies in a Nash equilibrium.

Result (Theorem 2.1.8)**.** *Assume that* $\alpha_1 = \cdots = \alpha_n = \alpha > 0$ *and* $b = 0$. *We define*

$$\theta_\pm = \frac{\gamma \pm \sqrt{\gamma^2 + 4\alpha\sigma^2\lambda}}{2\lambda} \quad \text{and} \quad \rho_\pm = -\frac{(n-1)\gamma}{2(n+1)\lambda} \pm \widehat{\rho}$$

for

$$\widehat{\rho} = \frac{\sqrt{(n-1)^2\gamma^2 + 4(n+1)\alpha\sigma^2\lambda}}{2(n+1)\lambda}.$$

Then the i^{th} *equilibrium strategy* X_i^* *is of the form*

$$X_i^*(t) = c_i(\theta_+)e^{\theta_+ t} + c_i(\theta_-)e^{\theta_- t} + c(\rho_+)e^{\rho_+ t} + c(\rho_-)e^{\rho_- t},$$

where, for $\overline{x}_n := \frac{1}{n}\sum_{j=1}^n x_j$,

$$c_i(\theta_+) = \frac{\overline{x}_n - x_i}{e^{2\widehat{\theta}T} - 1}, \quad c_i(\theta_-) = \frac{-(\overline{x}_n - x_i)}{1 - e^{-2\widehat{\theta}T}}, \quad c(\rho_+) = \frac{-\overline{x}_n}{e^{2\widehat{\rho}T} - 1}, \quad c(\rho_-) = \frac{\overline{x}_n}{1 - e^{-2\widehat{\rho}T}}.$$

Moreover, $\Sigma(t) = \sum_{i=1}^n X_i^*(t)$, *which solves the two-point boundary value problem* (2.20), *is given by*

$$\Sigma(t) = \frac{n\overline{x}_n}{2\sinh(\widehat{\rho}T)}\Big(e^{\widehat{\rho}T}e^{\rho_- t} - e^{-\widehat{\rho}T}e^{\rho_+ t}\Big).$$

This result can be further simplified if $n = 2$. For this case, we give numerical simulations and qualitative discussion.

In Section 2.2, we consider the infinite time case. To simplify the discussion, we assume that the drift b vanishes. The following result is on the existence and uniqueness of Nash equilibria for mean-variance optimization.

Result (Theorem 2.2.2)**.** *Suppose that one of the following two conditions holds:*

1. $n \in \mathbb{N}$ is arbitrary and $\alpha_1 = \cdots = \alpha_n = \alpha > 0$;

2. $n = 2$ and α_1, α_2 are distinct and strictly positive.

Then for all $x_1, \ldots, x_n \in \mathbb{R}$ there exists a unique Nash equilibrium for mean-variance optimization with infinite time horizon, which is also a Nash equilibrium for CARA utility maximization.

In case 1 the optimal strategies are given by

$$X_i^*(t) = (x_i - \overline{x}_n)e^{\theta_- t} + \overline{x}_n e^{\rho_- t},$$

where $\overline{x}_n = \frac{1}{n}\sum_{j=1}^n x_j$; ρ_- and θ_- are as in (2.22).

In case 2 the fourth-order equation

$$\tau^4 - \frac{2\gamma}{3\lambda}\tau^3 - \frac{\gamma^2 + 2\lambda\sigma^2(\alpha_1 + \alpha_2)}{3\lambda^2}\tau^2 + \frac{\sigma^4\alpha_1\alpha_2}{3\lambda^2} = 0$$

has precisely two distinct strictly negative roots, τ_1, τ_2, and the equilibrium strategies $X_1^*(t)$ and $X_2^*(t)$ are linear combinations of the exponential functions $e^{\tau_1 t}$ and $e^{\tau_2 t}$.

We also prove that the optimal strategies $X_1^{(T)}, \ldots, X_n^{(T)}$ in a Nash equilibrium of finite time period $[0, T]$ are convergent to the optimal strategies for infinite time horizon as $T \uparrow \infty$. Finally we discuss some qualitative properties of Nash equilibria on predatory trading.

Chapter 3

Following the observation by the report CFTC-SEC [2010] and Schöneborn [2008], we first consider a straightforward two-agent extension of the market impact model of Obizhaeva and Wang [2013] and we call it *primary model*. Then we extend the primary model in three aspects: incorporation of permanent impact, splitting of combined liquidation costs and closed-loop strategies. Finally we analyze a continuous-time version of the primary model.

Section 3.1 is based on Schied and Zhang [2013a]. In this section we assume that when the two agents X and Y apply respective strategies $\boldsymbol{\xi} \in \mathcal{X}(X_0, \mathbb{T})$ and $\boldsymbol{\eta} \in \mathcal{X}(Y_0, \mathbb{T})$, the asset price is given by

$$S_t^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_t^0 - \lambda \sum_{t_k < t} e^{-\rho(t - t_k)}(\xi_k + \eta_k),$$

where S^0 is a martingale and $\mathcal{X}(Z_0, \mathbb{T})$ denotes the class of all admissible strategies in $\mathbb{T}$ with an initial value Z_0. Under this assumption, we observe that if one agent executes its order, say η_k, immediately after the order, say ξ_k, of the other agent, this will result in an additional cost term $\lambda\xi_k\eta_k$ for the slower agent. Assuming that none of the two agents has an advantage in latency over the other, we define the liquidation costs and the corresponding Nash equilibrium.

Definition. *Suppose that $\mathbb{T} = \{t_0, t_1, \ldots, t_N\}$, X_0 and Y_0 are given. Let furthermore $(\varepsilon_i)_{i=0,1,\ldots}$ be an i.i.d. sequence of Bernoulli$(\frac{1}{2})$-distributed random variables that are independent of $\sigma(\bigcup_{t\geq 0}\mathcal{F}_t)$. Then the liquidation costs of $\boldsymbol{\xi} \in \mathcal{X}(X_0, \mathbb{T})$ given $\boldsymbol{\eta} \in \mathcal{X}(Y_0, \mathbb{T})$*

are defined as

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}) = X_0 S_0^0 + \sum_{k=0}^{N} \Big(\frac{\lambda}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k + \varepsilon_k \lambda \xi_k \eta_k + \theta \xi_k^2 \Big),$$

and the liquidation costs of $\boldsymbol{\eta}$ given $\boldsymbol{\xi}$ are

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}) = Y_0 S_0^0 + \sum_{k=0}^{N} \Big(\frac{\lambda}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + (1-\varepsilon_k) \lambda \xi_k \eta_k + \theta \eta_k^2 \Big).$$

A Nash equilibrium is a pair $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ of strategies in $\mathcal{X}(X_0,\mathbb{T}) \times \mathcal{X}(Y_0,\mathbb{T})$ such that

$$\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,] = \inf_{\boldsymbol{\xi}\in\mathcal{X}(X_0,\mathbb{T})} \mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,] \qquad \text{and} \qquad \mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)\,] = \inf_{\boldsymbol{\eta}\in\mathcal{X}(Y_0,\mathbb{T})} \mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,].$$

Note that each trade ζ_k incurs quadratic transaction costs of the form $\theta\zeta_k^2$, where θ is a nonnegative parameter.

Following this model setup, we derive the existence and uniqueness of Nash equilibria in the class of adapted strategies.

Result (Theorem 3.1.6)**.** *Let $\rho > 0$, $\lambda > 0$, and $\theta \geq 0$ be given. For any time grid $\mathbb{T}$ and initial values X_0, $Y_0 \in \mathbb{R}$, there exists a unique Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*) \in \mathcal{X}(X_0,\mathbb{T}) \times \mathcal{X}(Y_0,\mathbb{T})$. The optimal strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ are deterministic and given by*

$$\begin{aligned} \boldsymbol{\xi}^* &= \frac{1}{2}(X_0+Y_0)\boldsymbol{v} + \frac{1}{2}(X_0-Y_0)\boldsymbol{w}, \\ \boldsymbol{\eta}^* &= \frac{1}{2}(X_0+Y_0)\boldsymbol{v} - \frac{1}{2}(X_0-Y_0)\boldsymbol{w}, \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{v} &= \frac{1}{\mathbf{1}^\top(\lambda G + \lambda \widetilde{G} + 2\theta\,\mathrm{Id})^{-1}\mathbf{1}} (\lambda G + \lambda \widetilde{G} + 2\theta\,\mathrm{Id})^{-1}\mathbf{1}, \\ \boldsymbol{w} &= \frac{1}{\mathbf{1}^\top(\lambda G - \lambda \widetilde{G} + 2\theta\,\mathrm{Id})^{-1}\mathbf{1}} (\lambda G - \lambda \widetilde{G} + 2\theta\,\mathrm{Id})^{-1}\mathbf{1}. \end{aligned}$$

To prove this result, first we show by way of contradiction that there exists at most one Nash equilibrium in the class of adapted strategies. Then we show that a Nash equilibrium in the class of deterministic strategies is also a Nash equilibrium in the class of adapted strategies. Finally we find a unique Nash equilibrium of deterministic strategies.

In numerical simulations we find that optimal strategies in a Nash equilibrium exhibit strong oscillations when there are no transaction costs. This phenomenon is also observed by Schöneborn [2008, Section 9.3]. These oscillations can be interpreted as a of hedging against predatory trading by the other agent.

Since oscillatory strategies are neither efficient nor stable, we suggest a critical value of transaction costs in the following result such that all oscillations disappear completely when transaction costs are above this critical value.

Result (Theorem 3.1.14)**.** *Suppose that λ, $T > 0$ and $\mathbb{T}_N := \Big\{ \frac{kT}{N} \,\Big|\, k = 0, 1, \ldots, N \Big\}$. Then the following conditions are equivalent.*

(a) For every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{v}$ are nonnegative.

(b) For every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{w}$ are nonnegative.

(c) $\theta \geq \theta^* = \lambda/4$.

The proof of this result is complex and relies on *M-matrix* and the corresponding theorems in Berman and Plemmons [1994]. An astonishing fact is that, if both agents have a same initial value, the expected liquidation costs for both agents are decreasing as the trading frequency N increases if $\theta = \theta^*$, while the expected liquidation costs increase if $\theta = 0$. This observation indicates that the transaction costs lead not only smooth but also efficient strategies.

In Section 3.2, we assume that when the two financial agents X and Y apply respective strategies $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, the asset prices are given by

$$S_t^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_t^0 - \lambda \sum_{t_k<t} e^{-\rho(t-t_k)}(\xi_k + \eta_k) - \gamma \sum_{t_k<t}(\xi_k + \eta_k).$$

The first two extensions we consider are incorporation of permanent impact and splitting of combined liquidation costs. We derive the existence and uniqueness of Nash equilibria for these two extensions in the class of adapted strategies.

Result (Theorems 3.2.3 and 3.2.12)**.** *Given $\rho > 0$, $\lambda > 0$ and $\gamma > 0$. For any time grid $\mathbb{T}$ and initial values X_0, $Y_0 \in \mathbb{R}$, there exists a unique Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*) \in \mathcal{X}(X_0, \mathbb{T}) \times \mathcal{X}(Y_0, \mathbb{T})$ in the class of adapted strategies. The optimal strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ are deterministic.*

In particular we also derive the critical value of transaction costs such that all oscillations of optimal strategies disappear completely if permanent impact is present. We see in the following result that permanent impact plays a similar role as transient impact does in the interaction with transaction costs.

Result (Theorem 3.2.5)**.** *Given λ, $T > 0$ and $\mathbb{T}_N := \left\{ \frac{kT}{N} \,\middle|\, k = 0, 1, \ldots, N \right\}$. The following conditions are equivalent:*

(a) for every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{v}$ are nonnegative,

(b) for every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{w}$ are nonnegative,

(c) $\theta \geq \theta^* = (\lambda + \gamma)/4$,

where

$$\boldsymbol{v} = \frac{\left(\lambda G + 2\theta\,\mathrm{Id} + \lambda \widetilde{G} - \gamma \widetilde{H}^\top\right)^{-1}\mathbf{1}}{\mathbf{1}^\top\left(\lambda G + 2\theta\,\mathrm{Id} + \lambda \widetilde{G} - \gamma \widetilde{H}^\top\right)^{-1}\mathbf{1}}, \quad \boldsymbol{w} = \frac{\left(\lambda \widetilde{G}^\top + 2\theta\,\mathrm{Id} + \gamma \widetilde{H}^\top\right)^{-1}\mathbf{1}}{\mathbf{1}^\top\left(\lambda \widetilde{G}^\top + 2\theta\,\mathrm{Id} + \gamma \widetilde{H}^\top\right)^{-1}\mathbf{1}}.$$

The last extension is the closed-loop model, in which for each execution time all trades of the other agent at previous execution time will be considered. We use dynamic programming to derive the existence and uniqueness of Nash equilibria.

Result (Theorem 3.2.14)**.** *There exists a unique Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ in the class of closed-loop strategies. At each execution time point t_n, the optimal orders are linear in remained asset positions X_n, Y_n, and transient impact I_n, i.e.,*

$$\xi_n^* = A_n^X X_n + B_n^X Y_n + C_n^X I_n, \quad \eta_n^* = A_n^Y X_n + B_n^Y Y_n + C_n^Y I_n.$$

The cost functionals of the Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ for $n \in 0, 1, \dots, N$ are

$$\begin{aligned} J_n^X(X_n, Y_n, I_n) =& a_n^X X_n^2 + b_n^X Y_n^2 + c_n^X I_n^2 + u_n^X X_n Y_n + v_n^X X_n I_n + w_n^X Y_n I_n + \gamma(X_0 + Y_0) X_n, \\ J_n^Y(X_n, Y_n, I_n) =& a_n^Y X_n^2 + b_n^Y Y_n^2 + c_n^Y I_n^2 + u_n^Y X_n Y_n + v_n^Y X_n I_n + w_n^Y Y_n I_n + \gamma(X_0 + Y_0) Y_n. \end{aligned}$$

All of the coefficients

$$A_n^X, B_n^X, C_n^X, a_n^X, b_n^X, c_n^X, u_n^X, v_n^X, w_n^X, A_n^Y, B_n^Y, C_n^Y, a_n^Y, b_n^Y, c_n^Y, u_n^Y, v_n^Y, w_n^Y \in \mathbb{R}$$

can be recursively computed for each time t_n, $n \in \{0, 1, \dots, N-1\}$.

In numerical simulations we observe that there exists certain cooperation between the two agents in this closed-loop model.

In Section 3.3, we consider a continuous-time version of the primary model. When the two agents X and Y are active, the price process is given by

$$S_t^{X,Y} = S_t^0 + \lambda \int_{[0,t)} e^{-\rho(t-s)}\, dX_s + \lambda \int_{[0,t)} e^{-\rho(t-s)}\, dY_s.$$

This model is an extended version of the continuous-time model introduced in Gatheral et al. [2012] with the exponential decay $G(t) = e^{-\rho t}$ for two agents. This model can also be regarded as a continuous-time version of the model introduced in Section 3.1 that comes from Obizhaeva and Wang [2013] originally.

Through two different approaches, we obtain the following expression of liquidation costs, which is consistent with the primary model.

Definition. *Given initial asset positions $x, y \in \mathbb{R}$ and $T > 0$. The liquidation costs of $X \in \mathcal{X}(x, [0,T])$ given $Y \in \mathcal{X}(y, [0,T])$ are defined as*

$$\begin{aligned} \mathcal{C}(X|Y) = \frac{1}{2} \int_{[0,T]} \int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t + \int_{[0,T]} \int_{[0,t)} \lambda e^{-\rho(t-s)}\, dY_s\, dX_t \\ + \frac{\lambda}{2} \sum_{t \in [0,T]} \Delta X_t \Delta Y_t + \theta \sum_{t \in [0,T]} (\Delta X_t)^2, \end{aligned}$$

and the liquidation costs of $Y \in \mathcal{X}(y, [0,T])$ given $X \in \mathcal{X}(x, [0,T])$ are defined as

$$\begin{aligned} \mathcal{C}(Y|X) = \frac{1}{2} \int_{[0,T]} \int_{[0,T]} \lambda e^{-\rho|t-s|}\, dY_s\, dY_t + \int_{[0,T]} \int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s\, dY_t \\ + \frac{\lambda}{2} \sum_{t \in [0,T]} \Delta X_t \Delta Y_t + \theta \sum_{t \in [0,T]} (\Delta Y_t)^2, \end{aligned}$$

where $\mathcal{X}(z, [0,T])$ denotes the class of all admissible strategies in the time horizon $[0,T]$ with an initial value z.

We assume that each agent seeks to minimize their expected liquidation costs accordingly. In the following result, we prove that for any nontrivial case, a unique Nash equilibrium exits if and only if the transaction costs are exactly equal to a critical value.

Result (Theorem 3.3.6)**.** *Given $\rho > 0$, $\lambda > 0$, $T > 0$ and $x, y \in \mathbb{R}$ with $x \neq 0$ or $y \neq 0$. There exists a unique Nash equilibrium (X^*, Y^*) in the class $\mathcal{X}(x, [0,T]) \times \mathcal{X}(y, [0,T])$ of adapted strategies if and only if $\theta = \lambda/4$. The optimal strategies X^* and Y^* are deterministic and given by*

$$X_t^* = \frac{1}{2}(x+y)V_t + \frac{1}{2}(x-y)W_t,$$
$$Y_t^* = \frac{1}{2}(x+y)V_t - \frac{1}{2}(x-y)W_t,$$

where

$$V_t = \frac{e^{3\rho T}\big(6\rho(T-t)+4\big) - 4e^{3\rho t}}{2e^{3\rho T}(3\rho T+5)-1} \quad \text{if } t \in [0,T], \text{ and } V_{0-} = 1,$$
$$W_t = \frac{\rho(T-t)+1}{\rho T+1} \quad \text{if } t \in [0,T), \text{ and } W_{0-} = 1, \quad W_T = 0.$$

There are several steps to prove this result. First we derive an equivalent condition for the optimality of a deterministic strategy X^* given another deterministic strategy Y. This condition helps us to find Nash equilibria in the class of deterministic strategies. Then we find a Nash equilibrium in the class of deterministic strategies if the transaction costs θ are exactly equal to $\lambda/4$. Then we show that this deterministic Nash equilibrium is also a Nash equilibrium in the class of adapted strategies and it is unique. At last we show by way of contradiction that there exists no Nash equilibrium in the class of adapted strategies if $\theta \neq \lambda/4$.

To understand this result intuitively, numerical simulations of the primary model suggest that: when $\theta < \theta^*$, oscillatory strategies are not convergent to a continuous-time strategy; when $\theta > \theta^*$, the components $\boldsymbol{v}$ and $\boldsymbol{w}$ are convergent to the functions V and W. However, strategies consisting of V and W form a Nash equilibrium if and only if $\theta = \theta^*$.

Following this observation, we take a review of the single-agent model, where the price process is given by

$$S_t = S_t^0 - \lambda \sum_{t_k < t} e^{-\rho(t-t_k)}\xi_k.$$

We prove that there is no optimal strategy minimizing the expected liquidation costs

$$\mathbb{E}[\mathcal{C}(X)] := \mathbb{E}\bigg[\frac{1}{2}\Big(\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\,dX_s\,dX_t + 2\theta \sum_{t\in[0,T]} \big(\Delta X_t\big)^2\Big)\bigg], \qquad \theta > 0.$$

Result (Proposition 3.3.15, Corollary 3.3.16)**.** *Let $h : [0,\infty) \to [0,\infty)$ be a positive definite decay kernel. If $h(t)$ is continuous in $(0,\infty)$ but $h(0) \neq \lim_{t\downarrow 0} h(t)$, then for all strategies $X \in \mathcal{X}_{det}(x,T)$ there is no constant $\eta \in \mathbb{R}$, such that X solves the generalized Fredholm integral equation,*

$$\int_{[0,T]} h(|t-s|)\,dX_s = \eta.$$

Particularly, given $\lambda > 0$, $\rho > 0$, $T > 0$, $\theta > 0$ and an initial asset position $x \neq 0$, there is no optimal strategy in the class $\mathcal{X}_{det}(x, [0, T])$ of deterministic strategies.

The key point for the nonexistence of optimal strategies is that the decay kernel is not continuous at $t = 0$ if $\theta > 0$. In this case one would try to use continuous strategies to approximate jumps infinitesimally to avoid the quadratic transaction costs incurred by jumps. However it has been shown in Gatheral et al. [2012] that an optimal strategy X^* with $x \neq 0$ must have jumps at $t = 0$ and $t = T$ if $\theta = 0$. These two conflicting situations lead to the nonexistence of optimal strategies.

Chapter 2

A state-constrained differential game in the Almgren-Chriss framework

We analyze a state-constrained differential game that arises for risk-averse economic agents aiming to liquidate a given asset position by a given time $T > 0$. Agents face both price impact and volatility risk. For a single agent there is hence a trade-off between slow trading so as to reduce transaction costs from price impact and fast liquidation in view of volatility risk. Beginning with Bertsimas and Lo [1998] and Almgren and Chriss [2000], a large numbers of papers has recently been studying the corresponding single-agent optimization problems in various settings; we refer to Lehalle [2013] and Gatheral and Schied [2013] for recent overviews and more complete lists of references. The problem becomes even more interesting when considering not just one but n agents who are aware of each others initial positions, a situation that is not unlikely to occur in reality; see Carlin et al. [2007] and Schöneborn and Schied [2009]. Together with Brunnermeier and Pedersen [2005], these two papers were among the first to consider a corresponding game theoretic approach, but only consider open-loop Nash equilibria for risk-neutral agents applying deterministic strategies. Moallemi et al. [2012] give an extension to a model with asymmetric information. Carmona and Yang [2011] use numerical simulations to study a system of coupled HJB equations arising from a closed-loop Nash equilibrium for two utility-maximizing agents. Lachapelle et al. [2013] apply mean-field games to modeling the price formation process in the presence of high-frequency traders.

Here we consider agents maximizing a mean-variance functional in a continuous-time Almgren and Chriss [2000] framework, which is a very common setup for portfolio liquidation in practice. It leads to a linear-quadratic differential game, which has the interesting additional feature of a terminal state constraint arising from the liquidation constraint imposed in portfolio liquidation. This state constraint leads to two-point boundary problems in place of the usual initial value problems connected with unconstrained differential games. Aside from the financial interpretation of our results, this chapter thus also contributes a natural case study for a class of state-constrained differential games.

Our main results provide existence and uniqueness statements for the corresponding Nash equilibria with both finite and infinite time horizon. In several cases we can also give closed-form solutions of the equilibrium strategies. These formulas allow us to discuss some qualitative properties of the Nash equilibrium. Some of these properties are surprising as they show that certain monotonicity properties that are discussed in the finance literature may break down under certain market conditions. See Lebedeva et al. [2012]

for discussions and for an empirical analysis of a large data set of portfolio liquidations by large investors.

This chapter is based on Schied and Zhang [2013b] and is organized as follows. In Subsection 2.1.1 we recall some background material on portfolio liquidation in the Almgren-Chriss framework. Existence, uniqueness, and representation results for Nash equilibria with finite time horizon are stated in Subsection 2.1.2. Subsection 2.1.3 contains a discussion of the qualitative properties of the corresponding two-player Nash equilibrium. Nash equilibria with infinite time horizon are discussed in Section 2.2.

2.1 Nash equilibrium with finite time horizon

2.1.1 Background

We consider a standard continuous-time Almgren and Chriss [2000] framework for investors who are active over a fixed time period $[0,T]$. An investor may hold an initial position of x shares and is required to close this position by time T. The information flow available to an investor is modeled by a filtration $(\mathcal{F}_t)_{t\geq 0}$ on a given probability space $(\Omega,\mathcal{F},\mathbb{P})$. The trading strategy employed by the investor is denoted by $X=(X(t))_{t\in[0,T]}$. It needs to satisfy the following conditions of admissibility:

- X satisfies the liquidation constraint $X(T)=0$;
- X is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$;
- X is absolutely continuous in the sense that there exists a progressively measurable process $(\dot{X}(t))_{t\in[0,T]}$ such that for all $\omega\in\Omega$, $\int_0^T(\dot{X}(t,\omega))^2\,dt<\infty$ and
$$X(t,\omega)=X(0,\omega)+\int_0^t\dot{X}(s,\omega)\,ds,\qquad t\in[0,T];$$
- there exists a constant $c\geq 0$ such that $|X(t,\omega)|\leq c$ for all t and ω.

The class of all strategies that are admissible in this sense and satisfy $X(0)=x$ for given $x\in\mathbb{R}$ will be denoted by $\mathscr{X}(x,T)$. Let us also introduce the subclass $\mathscr{X}_{\mathrm{det}}(x,T)$ of all strategies in $\mathscr{X}(x,T)$ that are deterministic in the sense that they do not depend on ω.

The "unaffected price process" S^0 describes the fluctuations of asset prices perceived by an investor who has no inside information on large trades carried out by other market participants during the time interval $[0,T]$. In the Almgren–Chriss model it is usually assumed that S^0 follows a Bachelier model. Here we are sometimes also going to allow for an extra drift to describe current price trends. Thus,

$$S^0(t)=S_0+\sigma W(t)+\int_0^t b(s)\,ds,$$

where S_0 is a constant, W is a standard Brownian motion, $\sigma\geq 0$, and b is deterministic and continuous.

When an investor is using the strategy $X \in \mathcal{X}(x,T)$, the strategy X will influence the prices at which assets are traded. In the linear Almgren–Chriss framework, one assumes that the resulting price is given by

$$S^X(t) := S^0(t) + \gamma(X(t) - X(0)) + \lambda\dot{X}(t), \qquad t \in [0,T], \tag{2.1}$$

where the constants $\gamma \geq 0$ and $\lambda > 0$ describe the respective permanent and temporary price impact components. At each time $t \in [0,T]$, the infinitesimal amount of $-\dot{X}(t)\,dt$ shares are sold at price $S^X(t)$. The total revenues generated by the strategy $X \in \mathcal{X}(x,T)$ are therefore given by

$$\mathcal{R}(X) := -\int_0^T \dot{X}(t)S^X(t)\,dt.$$

The optimal trade execution problem consists in maximizing a cost-risk functional of the revenues over all admissible strategies in $\mathcal{X}(x,T)$. One possibility is the maximization of expected revenues,

$$\text{maximize} \quad \mathbb{E}[\,\mathcal{R}(X)\,] \tag{2.2}$$

as considered in many paper on optimal execution and, with the notable exception of Carmona and Yang [2011], all other papers dealing with corresponding Nash equilibria. Bertsimas and Lo [1998] were among the first to propose this problem. In practice, it is common to take into account the volatility risk arising from late execution by maximizing a mean-variance criterion:

$$\text{maximize} \quad \mathbb{E}[\,\mathcal{R}(X)\,] - \frac{\alpha}{2}\operatorname{var}(\mathcal{R}(X)); \tag{2.3}$$

here α is a nonnegative risk-aversion parameter. When dealing with the problem (2.3), admissible strategies are usually restricted to the class $\mathcal{X}_{\text{det}}(x,T)$ of deterministic strategies. Except for the results in Lorenz and Almgren [2011], little is known when general adapted strategies are used in (2.3); to the knowledge of the authors, not even the existence of maximizers has been established to date. The main reason for this is the lack of time consistency of the variance functional, which does not fit well into a context of dynamic optimization. On the other hand, Schied et al. [2010] show that the maximization of (2.3) over deterministic strategies $X \in \mathcal{X}_{\text{det}}(x,T)$ is equivalent to the maximization of the expected utility of revenues,

$$\text{maximize} \quad \mathbb{E}[\,u_\alpha(\mathcal{R}(X))\,], \tag{2.4}$$

over all strategies in $\mathcal{X}(x,T)$ when

$$u_\alpha(x) := \begin{cases} \frac{1}{\alpha}(1 - e^{-\alpha x}) & \text{if } \alpha > 0, \\ x & \text{if } \alpha = 0; \end{cases} \tag{2.5}$$

is a CARA utility function with absolute risk aversion $\alpha \geq 0$. We refer to Lehalle [2013] and Gatheral and Schied [2013] for recent overview on portfolio liquidation and related issues of market microstructure.

2.1.2 Mean-variance and CARA utility optimization

Now suppose that n investors are active in the market, using the respective strategies $X_1, \ldots, X_n$. As in (2.1), each strategy X_i will impact the price process S^0, thus leading to the following price with aggregated price impact:

$$S^{X_1,\ldots,X_n}(t) := S^0(t) + \gamma \sum_{j=1}^{n}(X_j(t) - X_j(0)) + \lambda \sum_{j=1}^{n} \dot{X}_j(t), \qquad t \in [0,T]. \tag{2.6}$$

Let us denote by $\boldsymbol{X}_{-i}$ the collection $\boldsymbol{X}_{-i} := \{X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n\}$ of strategies of all competitors of player i. Then player i will obtain the following revenues,

$$\mathcal{R}(X_i|\boldsymbol{X}_{-i}) = -\int_0^T \dot{X}_i(t) S^{X_1,\ldots,X_n}(t)\, dt,$$

and seek to maximize one of the objective functionals (2.2), (2.3), or (2.4). A natural question is whether there exists a Nash equilibrium in which each player maximizes their objective functional given the strategies of their competitors. For the maximization of the expected revenues and vanishing drift, this problem is solved in Carlin et al. [2007] within the class of deterministic strategies. It is later extended in Schöneborn and Schied [2009] to the case in which players have different time horizons and by Moallemi et al. [2012] to a situation with asymmetric information. A system of coupled HJB equations arising from a closed-loop Nash equilibrium for two utility-maximizing agents is studied through numerical simulations by Carmona and Yang [2011]. Here we will now conduct a mathematical analysis of n-player open-loop Nash equilibria for mean-variance optimization (2.3) and CARA utility maximization (2.4).

Definition 2.1.1. Suppose that $n \in \mathbb{N}$, $x_1, \ldots, x_n \in \mathbb{R}$ are initial asset positions, and that $\alpha_1, \ldots, \alpha_n$ are nonnegative coefficients of risk aversion.

- A *Nash equilibrium for mean-variance optimization* consists of a collection $X_1^*, \ldots, X_n^*$ of deterministic strategies such that for each i and $\boldsymbol{X}_{-i}^* = \{X_1^*, \ldots, X_{i-1}^*, X_{i+1}^*, \ldots, X_n^*\}$ the strategy $X_i^* \in \mathcal{X}_{\mathrm{det}}(x_i, T)$ maximizes the mean-variance functional
$$\mathbb{E}[\,\mathcal{R}(X|\boldsymbol{X}_{-i}^*)\,] - \frac{\alpha_i}{2}\operatorname{var}(\mathcal{R}(X|\boldsymbol{X}_{-i}^*)).$$
- A *Nash equilibrium for CARA utility maximization* consists of a collection $X_1^*, \ldots, X_n^*$ of admissible strategies such that for each i the strategy $X_i^* \in \mathcal{X}(x_i, T)$ maximizes the expected utility
$$\mathbb{E}[\,u_{\alpha_i}(\mathcal{R}(X|\boldsymbol{X}_{-i}^*))\,]$$
over all $X \in \mathcal{X}(x_i, T)$.

Let $X_i \in \mathcal{X}(x_i, T)$ be given and write $\boldsymbol{X}_{-i} := \{X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n\}$ for $i = 1, \ldots, n$. For $Y \in \mathcal{X}(y, T)$ we note first that, after integrating by parts,

$$\begin{aligned}\mathcal{R}(Y|\boldsymbol{X}_{-i}) = yS_0 - \frac{\gamma}{2}y^2 + \int_0^T Y(t)\Big(b(t) + \gamma \sum_{j\neq i} \dot{X}_j(t)\Big)\, dt \\ - \lambda \sum_{j\neq i}^{n} \int_0^T \dot{Y}(t)\dot{X}_j(t)\, dt - \lambda \int_0^T \dot{Y}(t)^2\, dt + \sigma \int_0^T Y(t)\, dW(t).\end{aligned}$$

When all X_i and Y are deterministic, it follows that

$$\mathbb{E}[\,\mathcal{R}(Y|\boldsymbol{X}_{-i})\,] - \frac{\alpha_i}{2}\operatorname{var}(\mathcal{R}(Y|\boldsymbol{X}_{-i})) = c + \int_0^T \mathcal{L}^i(t, Y(t), \dot{Y}(t)|\boldsymbol{X}_{-i})\,dt, \tag{2.7}$$

where $c = yS_0 - \frac{\gamma}{2}y^2$ and the Lagrangian $\mathcal{L}^i$ is given by

$$\mathcal{L}^i(t,q,p|\boldsymbol{X}_{-i}) = q\Big(b(t) + \gamma\sum_{j\neq i}\dot{X}_j(t)\Big) - \frac{\alpha_i\sigma^2}{2}q^2 - \lambda p\Big(\sum_{j\neq i}\dot{X}_j(t) + p\Big). \tag{2.8}$$

Note that the equilibrium strategies X_i^* for CARA utility maximization are allowed to be adapted and maximize the expected utility within the entire class $\mathcal{X}(x_i,T)$, whereas, for reasons explained above, only deterministic strategies are allowed in mean-variance optimization. We start by formulating a general existence and uniqueness result for the Nash equilibrium for mean-variance optimization.

Theorem 2.1.2. *For given $n \in \mathbb{N}$, $\alpha_1,\dots,\alpha_n \geq 0$, and $x_1,\dots,x_n$ there exists a unique Nash equilibrium $X_1^*,\dots,X_n^*$ for mean-variance optimization. It is given as the unique solution of the following second-order system of differential equations*

$$\alpha_i\sigma^2 X_i(t) - 2\lambda\ddot{X}_i(t) = b(t) + \gamma\sum_{j\neq i}\dot{X}_j(t) + \lambda\sum_{j\neq i}\ddot{X}_j(t), \tag{2.9}$$

with two-point boundary conditions

$$X_i(0) = x_i \text{ and } X_i(T) = 0 \tag{2.10}$$

for $i = 1,2,\dots,n$.

To prove Theorem 2.1.2, we need the following auxiliary lemmas.

Lemma 2.1.3. *In the context of Theorem 2.1.2 there exists at most one Nash equilibrium for mean-variance optimization.*

Proof. We assume by way of contradiction that $(X_1^0,\dots,X_n^0)$ and $(X_0^1,\dots,X_n^1)$ are two distinct Nash equilibria with $X_i^k \in \mathcal{X}(x_i,T)$ for $i = 1,\dots,n$ and $k = 0,1$. For $\beta \in [0,1]$ let $X_i^\beta := \beta X_i^1 + (1-\beta)X_i^0$ and define

$$f(\beta) := \sum_{i=1}^n \int_0^T \Big(\mathcal{L}^i(t, X_i^\beta(t), \dot{X}_i^\beta(t)|\boldsymbol{X}_{-i}^0) + \mathcal{L}^i(t, X_i^{1-\beta}(t), \dot{X}_i^{1-\beta}(t)|\boldsymbol{X}_{-i}^1)\Big)\,dt.$$

By assumption, the strategy X_i^k maximizes the functional $Y \mapsto \int_0^T \mathcal{L}^i(t,Y(t),\dot{Y}(t)|\boldsymbol{X}_{-i}^k)\,dt$ within the class $\mathcal{X}_{\det}(x_i,T)$ for $k = 0,1$. We therefore must have $f(\beta) \leq f(0)$ for $\beta > 0$, which implies that

$$\frac{d}{d\beta}\Big|_{\beta=0} f(\beta) \leq 0. \tag{2.11}$$

On the other hand, by interchanging differentiation and integration, which is permitted due to our assumptions on admissible strategies and due to the linear-quadratic form of the Lagrangian, a short computation shows that

$$\begin{aligned}&\frac{d}{d\beta}\Big|_{\beta=0} f(\beta)\\ &= \sum_{i=1}^n \int_0^T \Big[\gamma(X_i^1(t)-X_i^0(t))\sum_{j=1}^n(\dot X_j^0(t)-\dot X_j^1(t)) - \gamma(X_i^1(t)-X_i^0(t))(\dot X_i^0(t)-\dot X_i^1(t))\\ &\quad + \alpha_i\sigma^2(X_i^1(t)-X_i^0(t))^2 + \lambda(\dot X_i^1(t)-\dot X_i^0(t))\sum_{j=1}^n(\dot X_j^1(t)-\dot X_j^0(t)) + \lambda(\dot X_i^0(t)-\dot X_i^1(t))^2\Big]\,dt.\end{aligned}$$

We note next that

$$\int_0^T (X_i^1(t)-X_i^0(t))(\dot X_i^0(t)-\dot X_i^1(t))\,dt = \frac12(X_i^1(T)-X_i^0(T))^2 - \frac12(X_i^1(0)-X_i^0(0))^2 = 0.$$

Moreover, by the same argument,

$$\int_0^T (X_i^1(t)-X_i^0(t))(\dot X_j^0(t)-\dot X_j^1(t))\,dt = -\int_0^T (X_j^1(t)-X_j^0(t))(\dot X_i^0(t)-\dot X_i^1(t))\,dt,$$

and hence

$$\sum_{i=1}^n\sum_{j=1}^n \int_0^T (X_i^1(t)-X_i^0(t))(\dot X_j^0(t)-\dot X_j^1(t))\,dt = 0.$$

It follows that

$$\begin{aligned}&\frac{d}{d\beta}\Big|_{\beta=0} f(\beta)\\ &= \int_0^T \Big[\alpha_i\sigma^2\sum_{i=1}^n(X_i^1(t)-X_i^0(t))^2 + \lambda\sum_{i=1}^n(\dot X_i^0(t)-\dot X_i^1(t))^2 + \lambda\Big(\sum_{i=1}^n(\dot X_i^0(t)-\dot X_i^1(t))\Big)^2\Big]\,dt,\end{aligned}$$

which is strictly positive since the two Nash equilibria $(X_1^0,\dots,X_n^0)$ and $(X_0^1,\dots,X_n^1)$ are distinct. But this contradicts (2.11). □

Lemma 2.1.4. *For $i=1,\dots,n$ there exists at most one maximizer in $\mathcal{X}_{det}(y,T)$ of the functional $Y \mapsto \int_0^T \mathcal{L}^i(t,Y(t),\dot Y(t)|\boldsymbol{X}_{-i})\,dt$. If, moreover, $X_1,\dots,X_n \in C^2[0,T]$ then there exists a unique maximizer $Y^* \in \mathcal{X}_{det}(y,T)\cap C^2[0,T]$, which is given as the unique solution of the two-point boundary value problem*

$$\begin{cases} \alpha_i\sigma^2 Y(t) - 2\lambda\ddot Y(t) = b(t) + \gamma\sum_{j\neq i}\dot X_j(t) + \lambda\sum_{j\neq i}\ddot X_j(t),\\ Y(0)=y,\ \ Y(T)=0.\end{cases}$$

Proof. It follows from the strict concavity of the Lagrangian $\mathcal{L}^i$ and the convexity of the set $\mathcal{X}_{\rm det}(y,T)$ that there can be at most one maximizer in $\mathcal{X}_{\rm det}(y,T)$.

Now we show the existence of a maximizer under the additional assumption $X_1,\dots,X_n \in C^2[0,T]$. The Euler–Lagrange equation,

$$\mathcal{L}^i_q(t,Y(t),\dot Y(t)|\boldsymbol{X}_{-i}) = \frac{d}{dt}\mathcal{L}^i_p(t,Y(t),\dot Y(t)|\boldsymbol{X}_{-i})$$

is read as follows for our problem:

$$\alpha_i \sigma^2 Y(t) - 2\lambda \ddot{Y}(t) = b(t) + \gamma \sum_{j \neq i} \dot{X}_j(t) + \lambda \sum_{j \neq i} \ddot{X}_j(t) \tag{2.12}$$

with boundary condition $Y(0) = y$ and $Y(T) = 0$. Denoting the right-hand side of (2.12) by $u(t)$, the general solution of this second-order ODE is of the form

$$Y(t) = c_1 e^{-\kappa_i t} + c_2 e^{\kappa_i t} - \frac{1}{4\lambda\kappa_i} \int_0^t e^{\kappa_i(t-s)} u(s)\, ds + \frac{1}{4\lambda\kappa_i} \int_0^t e^{-\kappa_i(t-s)} u(s)\, ds, \tag{2.13}$$

where c_1 and c_2 are constants and $\kappa_i = \sqrt{\alpha_i \sigma^2 / 2\lambda}$. It is clear that the two constants c_1 and c_2 can be uniquely determined by imposing the boundary conditions $Y(0) = y$ and $Y(T) = 0$. From now on, let $Y^* \in \mathcal{X}_{\text{det}}(y, T) \cap C^2[0, T]$ denote the corresponding solution. We will now verify that Y^* is indeed a maximizer of our problem. To this end, let $Y \in \mathcal{X}_{\text{det}}(y, T)$ be arbitrary. Using first the concavity of $(q, p) \mapsto \mathcal{L}^i(t, q, p|\boldsymbol{X}_{-i})$ and then the fact that Y^* solves the Euler–Lagrange equation (2.12), we get

$$\begin{aligned}
&\mathcal{L}^i(t, Y^*(t), \dot{Y}^*(t)|\boldsymbol{X}_{-i})) - \mathcal{L}^i(t, Y(t), \dot{Y}(t)|\boldsymbol{X}_{-i}) \\
&\geq \mathcal{L}^i_q(t, Y^*(t), \dot{Y}^*(t)|\boldsymbol{X}_{-i}))(Y^*(t) - Y(t)) + \mathcal{L}^i_p(t, Y^*(t), \dot{Y}^*(t)|\boldsymbol{X}_{-i}))(\dot{Y}^*(t) - \dot{Y}(t)) \\
&= \Big(\frac{d}{dt}\mathcal{L}^i_p(t, Y^*(t), \dot{Y}^*(t)|\boldsymbol{X}_{-i})\Big)(Y^*(t) - Y(t)) + \mathcal{L}^i_p(t, Y^*(t), \dot{Y}^*(t)|\boldsymbol{X}_{-i}))(\dot{Y}^*(t) - \dot{Y}(t)) \\
&= \frac{d}{dt}\Big(\mathcal{L}^i_p(t, Y^*(t), \dot{Y}^*(t)|\boldsymbol{X}_{-i})(Y^*(t) - Y(t))\Big).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\int_0^T \mathcal{L}^i(t, Y^*(t), \dot{Y}^*(t)|\boldsymbol{X}_{-i})\, dt - \int_0^T \mathcal{L}^i(t, Y(t), \dot{Y}(t)|\boldsymbol{X}_{-i})\, dt \\
&\geq \int_0^T \frac{d}{dt}\Big(\mathcal{L}^i_p(t, Y^*(t), \dot{Y}^*(t)|\boldsymbol{X}_{-i})(Y^*(t) - Y(t))\Big)\, dt = 0,
\end{aligned}$$

where in the final step we have used that $Y^*(0) = Y(0)$ and $Y^*(T) = Y(T)$. This proves the lemma. □

Proof of Theorem 2.1.2. According to Lemma 2.1.3 there exists at most one Nash equilibrium. We will now show that there exists a Nash equilibrium $(X_1^*, \dots, X_n^*)$ such that each strategy X_i^* belongs to $\mathcal{X}_{\text{det}}(x_i, T) \cap C^2[0, T]$. By Lemma 2.1.4, each strategy X_i^* must then be a solution of the second-order differential equation

$$\alpha_i \sigma^2 X_i(t) - 2\lambda \ddot{X}_i(t) = b(t) + \gamma \sum_{j \neq i} \dot{X}_j(t) + \lambda \sum_{j \neq i} \ddot{X}_j(t), \tag{2.14}$$

with boundary conditions

$$X_i(0) = x_i \text{ and } X_i(T) = 0. \tag{2.15}$$

We can clearly combine the n differential equations (2.14) into a system of n coupled second-order linear ordinary differential equations for the vector $\boldsymbol{X}^* := (X_1^*, \dots, X_n^*)^\top$.

It follows again from Lemma 2.1.4 that every solution of the system (2.14), (2.15) is a Nash equilibrium. Therefore it will be sufficient to show the existence of a solution to the n-dimensional two-point boundary value problem (2.14), (2.15).

By introducing the auxiliary function $\boldsymbol{Y}(t)$ for the derivative $\dot{\boldsymbol{X}}(t)$, by letting $\boldsymbol{b}(t)$ be the vector with all components equal to $b(t)$, and by defining the $n \times n$ matrices $A := \sigma^2 \mathrm{diag}(\alpha_1, \dots, \alpha_n)$, the identity matrix $I = \mathrm{diag}(1, \dots, 1)$, and the matrix J with all entries equal to one, the system (2.14) can be re-written as follows:

$$\begin{pmatrix} A & -\gamma(J-I) \\ 0 & I \end{pmatrix} \begin{pmatrix} \boldsymbol{X}(t) \\ \boldsymbol{Y}(t) \end{pmatrix} - \begin{pmatrix} 0 & \lambda(J+I) \\ I & 0 \end{pmatrix} \begin{pmatrix} \dot{\boldsymbol{X}}(t) \\ \dot{\boldsymbol{Y}}(t) \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}(t) \\ \boldsymbol{0} \end{pmatrix}. \tag{2.16}$$

Clearly, $\boldsymbol{X}$ solves (2.14) if and only if $\binom{\boldsymbol{X}}{\dot{\boldsymbol{X}}}$ solves (2.16). In particular every solution of (2.16) with boundary conditions (2.15) yields a Nash equilibrium.

Now consider the homogeneous system (2.16), (2.15) with $\boldsymbol{b}(t) = 0$ and initial values $x_1 = \cdots = x_n = 0$. The corresponding boundary condition can be written as

$$(\boldsymbol{X}(0), \boldsymbol{Y}(0), \boldsymbol{X}(T), \boldsymbol{Y}(T))^\top \in V, \tag{2.17}$$

where $V \subset \mathbb{R}^{4n}$ is the $2n$-dimensional linear space

$$V = \big\{ (\boldsymbol{x}_0, \boldsymbol{y}_0, \boldsymbol{x}_1, \boldsymbol{y}_1)^\top \in \mathbb{R}^{4n} \,|\, \boldsymbol{x}_0 = \boldsymbol{x}_1 = \boldsymbol{0} \big\}.$$

It is clear that $\binom{\boldsymbol{X}}{\boldsymbol{Y}} = \binom{\boldsymbol{0}}{\boldsymbol{0}}$ is a solution. In fact this trivial solution is the only solution since every solution must be a Nash equilibrium, and Nash equilibria are unique by Lemma 2.1.3. It therefore follows from the general theory of linear boundary value problems for systems of ordinary differential equations that the two-point boundary value problem (2.16), (2.17) has a unique solution for every continuous $\boldsymbol{b} : [0,T] \to \mathbb{R}^n$ (and in fact for every continuous $\mathbb{R}^{2n}$-valued function substituting $\binom{\boldsymbol{b}(t)}{\boldsymbol{0}}$ on the right-hand side of (2.16)); see Kurzweil [1986, (9.22), p. 189]. Using this fact, we let $\binom{\boldsymbol{X}^0}{\boldsymbol{Y}^0}$ be the solution of (2.16), (2.17) when $\boldsymbol{b}(t)$ in (2.16) is replaced by $\boldsymbol{b}^0(t) = (b_1^0(t), \dots, b_n^0(t))$ for

$$b_i^0(t) = b(t) + \frac{T-t}{T} \alpha_i \sigma^2 x_i - \frac{\gamma}{T} \sum_{j \neq i} x_j.$$

One then checks that

$$X_i^*(t) := X_i^0(t) + \frac{T-t}{T} x_i, \qquad i = 1, \dots, n,$$

solves (2.14), (2.15) and is thus the desired Nash equilibrium. □

Remark 2.1.5. With $\boldsymbol{Z}(t) := (\boldsymbol{X}(t), \boldsymbol{Y}(t))^\top$ the system (2.16) can be written as

$$\dot{\boldsymbol{Z}}(t) = M \boldsymbol{Z}(t) + \boldsymbol{f}(t), \tag{2.18}$$

where

$$M := \begin{pmatrix} 0 & \lambda(J+I) \\ I & 0 \end{pmatrix}^{-1} \begin{pmatrix} A & -\gamma(J-I) \\ 0 & I \end{pmatrix},$$

and

$$\boldsymbol{f}(t) := -\begin{pmatrix} 0 & \lambda(J+I) \\ I & 0 \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{b}(t) \\ \mathbf{0} \end{pmatrix}.$$

Note that $J^2 = nJ$ and that hence

$$(J+I)\Big(I - \frac{1}{n+1}J\Big) = I = \Big(I - \frac{1}{n+1}J\Big)(J+I).$$

It follows that

$$\begin{pmatrix} 0 & \lambda(J+I) \\ I & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & I \\ \frac{1}{\lambda}(I - \frac{1}{n+1}J) & 0 \end{pmatrix}$$

and hence

$$M = \begin{pmatrix} 0 & I \\ \frac{1}{\lambda}(A - \frac{1}{n+1}JA) & -\frac{\gamma}{\lambda}(\frac{2}{n+1}J - I) \end{pmatrix}. \tag{2.19}$$

□

It will become clear from (2.7), (2.8) and below that, from a mathematical point of view, the Nash equilibrium constructed above is an open-loop linear-quadratic differential game with state constraints. The state constraints are provided by the liquidation constraints $X_i(T) = 0$, $i = 1, \ldots, n$. They are responsible for the fact that we cannot apply standard results on the existence and uniqueness of open-loop linear-quadratic differential games, and significantly complicate the proof for the existence of Nash equilibria, especially in the case of an infinite time horizon as studied in Section 2.2. It is also interesting to point out that the proof of the existence of solutions to (2.9), (2.10) rests on the uniqueness of Nash equilibria, which is established in Lemma 2.1.3.

Our next result states that the unique Nash equilibrium for mean-variance optimization is also a Nash equilibrium for CARA utility maximization. It is an open question, however, whether there may be more than one Nash equilibrium for CARA utility maximization.

Corollary 2.1.6. *For given $n \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_n \geq 0$, and $x_1, \ldots, x_n$ the Nash equilibrium for mean-variance optimization constructed in Theorem 2.1.2 is also a Nash equilibrium for CARA utility maximization.*

Proof. Let $X_1^*, \ldots, X_n^*$ be the unique Nash equilibrium for mean-variance optimization as constructed in Theorem 2.1.2. When $\boldsymbol{X}_{-i}^* := \{X_1^*, \ldots, X_{i-1}^*, X_{i+1}^*, \ldots, X_n^*\}$ is fixed, the i^{th} agent perceives

$$S^{\boldsymbol{X}_{-i}^*}(t) := S^0(t) + \gamma \sum_{j \neq i} (X_j(t) - X_j(0)) + \lambda \sum_{j \neq i} \dot{X}_j(t), \qquad t \in [0,T],$$

as "unaffected" price process. It is of the form

$$S^{\boldsymbol{X}_{-i}^*}(t) = S_0 + \sigma W(t) + \int_0^t b^i(s)\, ds$$

for a deterministic and continuous function $b^i : [0,T] \to \mathbb{R}$. Since the process $S^{\boldsymbol{X}_{-i}^*}$ has independent increments and $S_T^{\boldsymbol{X}_{-i}^*}$ has all exponential moments, i.e., $\mathbb{E}\big[\, e^{\beta S_T^{\boldsymbol{X}_{-i}^*}}\,\big] < \infty$ for all $\beta \in \mathbb{R}$, it follows as in Schied et al. [2010, Theorem 2.1] that for $\alpha_i > 0$

$$\sup_{X \in \mathcal{X}(x_i,T)} \mathbb{E}[\, u_{\alpha_i}(\mathcal{R}(X|\boldsymbol{X}_{-i}))\,] = \sup_{X \in \mathcal{X}_{\text{det}}(x_i,T)} \mathbb{E}[\, u_{\alpha_i}(\mathcal{R}(X|\boldsymbol{X}_{-i}))\,].$$

But for $\alpha_i > 0$ and $X \in \mathcal{X}_{\mathrm{det}}(x_i, T)$ we have

$$\mathbb{E}[\, u_{\alpha_i}(\mathcal{R}(X|\boldsymbol{X}_{-i}))\,] = \frac{1}{\alpha_i}\Big(1 - e^{-\alpha_i \mathbb{E}[\,\mathcal{R}(X|\boldsymbol{X}_{-i})\,] + \frac{\alpha_i^2}{2}\,\mathrm{var}\,(\mathcal{R}(X|\boldsymbol{X}_{-i})}\Big),$$

which shows that CARA utility maximization is equivalent to the maximization of the corresponding mean-variance functional. The corresponding result for $\alpha_i = 0$ is obvious. □

Let us now have a closer look at the system (2.9). It simplifies when all agents have the same risk aversion:

Corollary 2.1.7. *In the setting of Theorem 2.1.2 suppose that $\alpha_1 = \cdots = \alpha_n = \alpha \geq 0$. Then*

$$\Sigma(t) := \sum_{i=1}^n X_i^*(t)$$

is the unique solution of the following one-dimensional two-point boundary value problem

$$\alpha\sigma^2\Sigma(t) - (n-1)\gamma\dot{\Sigma}(t) - (n+1)\lambda\ddot{\Sigma}(t) = nb(t), \qquad \Sigma(0) = \sum_{i=1}^n x_i,\ \Sigma(T) = 0. \quad (2.20)$$

Given Σ, each equilibrium strategy X_i^* is equal to the unique solution of the following one-dimensional two-point boundary value problem,

$$\alpha\sigma^2 X_i(t) + \gamma\dot{X}_i(t) - \lambda\ddot{X}_i(t) = b(t) + \gamma\dot{\Sigma}(t) + \lambda\ddot{\Sigma}(t), \qquad X_i(0) = x_i,\ X_i(T) = 0. \quad (2.21)$$

Proof. Letting $\Sigma(t) := \sum_{j=1}^n X_j(t)$ and re-writing (2.9) yields

$$\alpha\sigma^2 X_i(t) + \gamma\dot{X}_i(t) - \lambda\ddot{X}_i(t) = b(t) + \gamma\dot{\Sigma}(t) + \lambda\ddot{\Sigma}(t),$$

and hence (2.21). Summing over i then implies (2.20). □

It is possible to obtain closed-form solutions of (2.20) and (2.21), but the corresponding expressions are quite involved. The situation simplifies when the drift b vanishes identically:

Theorem 2.1.8. *In the setting of Corollary 2.1.7 assume in addition that $b = 0$ and $\alpha > 0$. We define*

$$\theta_\pm = \frac{\gamma \pm \sqrt{\gamma^2 + 4\alpha\sigma^2\lambda}}{2\lambda} \quad \textit{and} \quad \rho_\pm = -\frac{(n-1)\gamma}{2(n+1)\lambda} \pm \widehat{\rho} \quad (2.22)$$

for

$$\widehat{\rho} = \frac{\sqrt{(n-1)^2\gamma^2 + 4(n+1)\alpha\sigma^2\lambda}}{2(n+1)\lambda}. \quad (2.23)$$

Then the i^{th} equilibrium strategy X_i^* is of the form

$$X_i^*(t) = c_i(\theta_+)e^{\theta_+ t} + c_i(\theta_-)e^{\theta_- t} + c(\rho_+)e^{\rho_+ t} + c(\rho_-)e^{\rho_- t}, \quad (2.24)$$

where, for $\overline{x}_n := \frac{1}{n}\sum_{j=1}^n x_j$,

$$c_i(\theta_+) = \frac{\overline{x}_n - x_i}{e^{2\widehat{\theta}T} - 1}, \quad c_i(\theta_-) = \frac{-(\overline{x}_n - x_i)}{1 - e^{-2\widehat{\theta}T}}, \quad c(\rho_+) = \frac{-\overline{x}_n}{e^{2\widehat{\rho}T} - 1}, \quad c(\rho_-) = \frac{\overline{x}_n}{1 - e^{-2\widehat{\rho}T}}. \tag{2.25}$$

Moreover, $\Sigma(t) = \sum_{i=1}^n X_i^*(t)$*, which solves the two-point boundary value problem* (2.20)*, is given by*

$$\Sigma(t) = \frac{n\overline{x}_n}{2\sinh(\widehat{\rho}T)}\Big(e^{\widehat{\rho}T}e^{\rho_- t} - e^{-\widehat{\rho}T}e^{\rho_+ t}\Big). \tag{2.26}$$

To prove Theorem 2.1.8, we need the following auxiliary lemma.

Lemma 2.1.9. *For* $\alpha_1 = \cdots = \alpha_n = \alpha > 0$ *the matrix* M *from* (2.19) *has four real eigenvalues* $\theta_+, \theta_-, \rho_+, \rho_-$ *given by* (2.22)*. Moreover, with* $\mathbf{1} \in \mathbb{R}^n$ *denoting the vector with all entries equal to* 1*, the corresponding eigenspaces are given by*

$$E(\rho_\pm) = \operatorname{span}\begin{pmatrix}\mathbf{1}\\ \rho_\pm\mathbf{1}\end{pmatrix} \qquad \text{and} \qquad E(\theta_\pm) = \left\{\begin{pmatrix}\boldsymbol{v}\\ \theta_\pm\boldsymbol{v}\end{pmatrix} \,\middle|\, \boldsymbol{v} \in \mathbb{R}^n,\ \boldsymbol{v} \perp \mathbf{1}\right\}.$$

Proof. Let us write an arbitrary vector in $\mathbb{R}^{2n}$ as $\binom{\boldsymbol{v}_1}{\boldsymbol{v}_2}$ for $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^n$. By applying M to $\binom{\boldsymbol{v}_1}{\boldsymbol{v}_2}$ we see that we must have $\boldsymbol{v}_2 = \tau\boldsymbol{v}_1$ for $\binom{\boldsymbol{v}_1}{\boldsymbol{v}_2}$ to be an eigenvector with eigenvalue τ. So let us consider vectors in $\mathbb{R}^{2n}$ of the form $\binom{\boldsymbol{v}}{\tau\boldsymbol{v}}$ for $\boldsymbol{v} \in \mathbb{R}^n$ and $\tau \in \mathbb{R}$. The equation $M\binom{\boldsymbol{v}}{\tau\boldsymbol{v}} = \tau\binom{\boldsymbol{v}}{\tau\boldsymbol{v}}$ is equivalent to

$$\Big(\frac{\alpha\sigma^2}{\lambda} + \frac{\tau\gamma}{\lambda}\Big)\boldsymbol{v} - \frac{\alpha\sigma^2 + \tau\gamma}{n+1}J\boldsymbol{v} = \tau^2\boldsymbol{v}. \tag{2.27}$$

When $\boldsymbol{v} = \mathbf{1}$ then $J\boldsymbol{v} = n\boldsymbol{v}$ and (2.27) becomes the quadratic equation

$$\alpha\sigma^2 + \gamma\tau - \frac{n(\alpha\sigma^2 + 2\gamma\tau)}{n+1} - \lambda\tau^2 = 0, \tag{2.28}$$

which is solved for $\tau = \rho_+$ and $\tau = \rho_-$. When $\boldsymbol{v} \perp \mathbf{1}$ then $J\boldsymbol{v} = 0$ and (2.27) becomes the quadratic equation

$$\alpha\sigma^2 + \gamma\tau - \lambda\tau^2 = 0, \tag{2.29}$$

which is solved for $\tau = \theta_+$ and $\tau = \theta_-$. Since the eigenvectors are found and thus span the entire space $\mathbb{R}^{2n}$, the proof is completed. □

Proof of Theorem 2.1.8. It follows from Theorem 2.1.2 and its proof that $X_1^*, \ldots, X_n^*$ are obtained from the solutions of (2.18) for $\boldsymbol{f}(t) = \mathbf{0}$. The general solution of this system is of the form $\boldsymbol{Z}(t) = e^{tM}\boldsymbol{Z}(0)$. By Lemma 2.1.9, M is diagonalizable and so every solution $\boldsymbol{Z}(t)$ must be a linear combination of exponential functions $e^{\tau t}$, where τ is an eigenvalue of M. Another application of Lemma 2.1.9 thus implies that each X_i^* can be represented as in (2.24). One finally checks that for $c_i(\theta_+), c_i(\theta_-), c(\rho_+), c(\rho_-)$ as in (2.25) the boundary conditions $X_i^*(0) = x_i$ and $X_i^*(T) = 0$ are satisfied. □

The formulas in Theorem 2.1.8 can be further simplified in a two-player setting:

Corollary 2.1.10. *In the setting of Theorem 2.1.8 assume in addition that $n = 2$. Then*

$$X_1^*(t) = \frac{1}{2}\big(\Sigma(t) + \Delta(t)\big) \qquad \text{and} \qquad X_2^*(t) = \frac{1}{2}\big(\Sigma(t) - \Delta(t)\big), \tag{2.30}$$

where

$$\Sigma(t) = (x_1 + x_2)e^{-\frac{\gamma t}{6\lambda}} \frac{\sinh\left(\frac{(T-t)\sqrt{\gamma^2+12\alpha\lambda\sigma^2}}{6\lambda}\right)}{\sinh\left(\frac{T\sqrt{\gamma^2+12\alpha\lambda\sigma^2}}{6\lambda}\right)}, \tag{2.31}$$

$$\Delta(t) = (x_1 - x_2)e^{\frac{\gamma t}{2\lambda}} \frac{\sinh\left(\frac{(T-t)\sqrt{\gamma^2+4\alpha\lambda\sigma^2}}{2\lambda}\right)}{\sinh\left(\frac{T\sqrt{\gamma^2+4\alpha\lambda\sigma^2}}{2\lambda}\right)}. \tag{2.32}$$

Proof. From (2.26) we have that $\Sigma(t) = X_1^*(t) + X_2^*(t)$ is given by (2.31). When letting $\Delta(t) := X_1^*(t) - X_2^*(t)$, we get from (2.21) that Δ solves the two-point boundary value problem

$$\alpha\sigma^2\Delta(t) + \gamma\dot{\Delta}(t) - \lambda\ddot{\Delta}(t) = 0, \qquad \Delta(0) = x_1 - x_2, \ \Delta(T) = 0.$$

This boundary value problem is solved by (2.32). □

The following mean-field limit is obtained in a straightforward manner by sending n to infinity in Theorem 2.1.8.

Corollary 2.1.11. *In the setting of Theorem 2.1.8 suppose that $\lim_{n\uparrow\infty} \frac{1}{n}\sum_{j=1}^n x_j = \overline{x} \in \mathbb{R}$. Then, as $n \uparrow \infty$, the equilibrium strategy of agent i converges to*

$$\frac{\overline{x} - x_i}{e^{2\widehat{\theta}T} - 1}e^{\theta_+ t} - \frac{\overline{x} - x_i}{1 - e^{-2\widehat{\theta}T}}e^{\theta_- t} + \frac{\overline{x}}{1 - e^{-\frac{\gamma T}{\lambda}}}e^{-\frac{\gamma t}{\lambda}} - \frac{\overline{x}}{e^{\frac{\gamma T}{\lambda}} - 1}, \tag{2.33}$$

where θ_+, θ_- and $\widehat{\theta}$ are as in (2.22) and (2.23).

Proof. Note that as $n \uparrow \infty$, we have $\widehat{\rho} = \frac{\gamma}{2\lambda}$, $\rho_+ = 0$ and $\rho_- = -\frac{\gamma}{\lambda}$. Putting this into (2.24) completes the proof. □

2.1.3 Qualitative discussion of the two-player Nash equilibrium

Throughout this subsection, (X_1^*, X_2^*) will denote the two-player Nash equilibrium constructed in Corollary 2.1.10. We first present numerical simulations of X_1^* and X_2^*. We assume that agent 1 uses deterministic strategies to sell the asset position of one unit in the time horizon $[0, 1]$, that is, $X_1^* \in \mathcal{X}_{det}(1, 1)$. In Figures 2.1, 2.2 and 2.3 we show equilibrium strategies for agent 1 under the following four different conditions of X_2^*:

- $X_2^* \in \mathcal{X}_{det}(1, 1)$: agent 2 wants to sell a same asset position;
- $X_2^* \in \mathcal{X}_{det}(0, 1)$: agent 2 uses an admissible round trip;

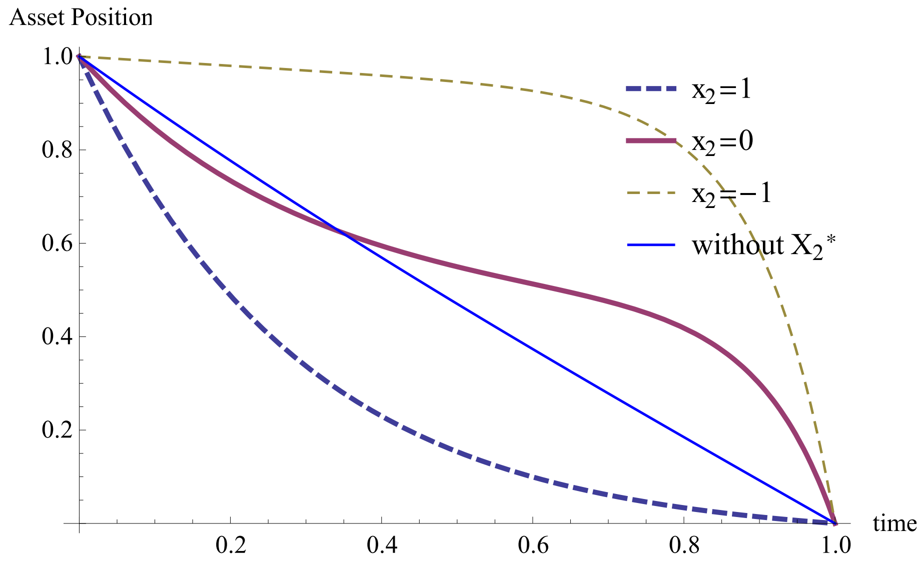


Figure 2.1: Optimal asset positions of X_1^* with initial position $x_1 = 1$ in Nash equilibrium in cases that $x_2 = 1, 0, -1$, and X_2^* is absent with $\lambda = \sigma = A = T = 1$, $\gamma = 10$.

- $X_2^* \in \mathcal{X}_{det}(-1, 1)$: agent 2 wants to buy the equal amount of asset position that agent 1 wants to sell;
- X_2^* is absent.

Note that the optimal strategy of a single agent without competitors is given by

$$X_0^*(t) = x_0 \frac{\sinh(\kappa(T-t))}{\sinh(\kappa T)}, \tag{2.34}$$

where $\kappa = \sqrt{\alpha\sigma^2/2\lambda}$; see Almgren [2003] or take $n = 1$ in (2.26).

By setting $A = \lambda = \sigma = 1, \gamma = 10$, Figure 2.1 displays the Nash equilibria with an enlargement of the permanent impact coefficient γ. A remarkable observation is that the optimal strategy in case that a round trip is present (the thick solid line) is not concave or convex with respect to time t any more. Liquidation of the asset position given an opposite investor (the thin dashed line) is significantly delayed while the normal asset position without X_2^* is nearly linear.

In Figure 2.2, we set $A = \lambda = \gamma = 1$, $\sigma = 10$. We observe that, an enlargement of the market fluctuation σ leads to an early liquidation. That means, agent 1 tends to liquidate earlier in an unstable market.

Furthermore, we see in Figure 2.3 that all the four lines will converge to a nearly linear curve if the coefficient of temporary impact λ increases.

To study the behavior of the strategies X_0^*, X_1^*, X_2^* we will need the following elementary fact.

Lemma 2.1.12. *For $0 < \nu < 1$ the function*

$$f(x) := \frac{\sinh(\nu x)}{\sinh(x)}$$

is strictly decreasing on $[0, \infty)$.

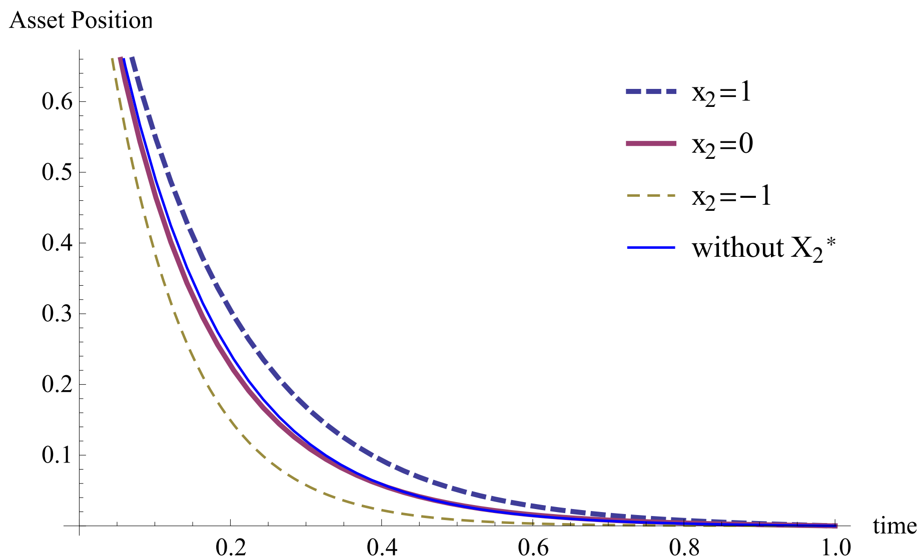


Figure 2.2: Optimal asset positions of X_1^* with initial position $x_1 = 1$ in Nash equilibrium in cases that $x_2 = 1, 0, -1$, and X_2^* is absent with $\lambda = \gamma = A = T = 1$, $\sigma = 10$.

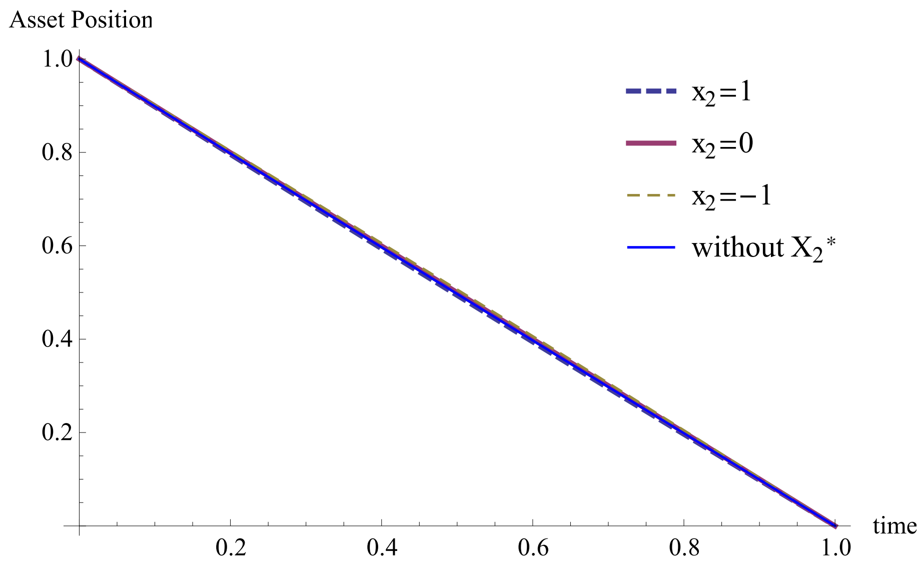
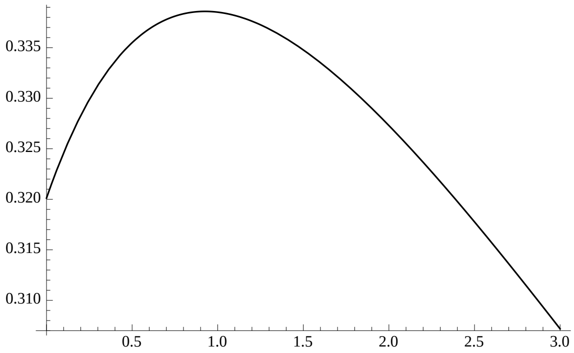


Figure 2.3: Optimal asset positions of X_1^* with initial position $x_1 = 1$ in Nash equilibrium in cases that $x_2 = 1, 0, -1$, and X_2^* is absent with $\sigma = \gamma = A = T = 1$, $\lambda = 10$.



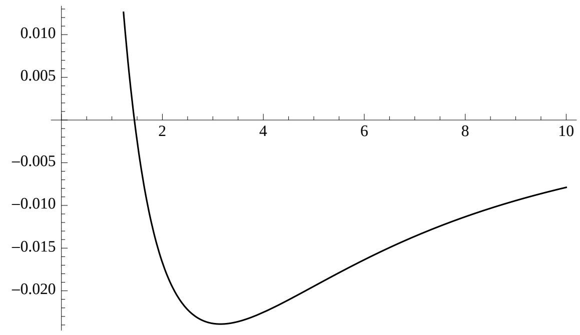


Figure 2.4: $X_1^*(1)$ as a function of $\alpha\sigma^2$. Left: $x_1 = 1.12$, $x_2 = 2.06$, $T = 2$, and $\lambda = \gamma = 1$; Right: $x_1 = 0.7$, $x_2 = -1.9$, $T = 2$, $\lambda = 0.2$, and $\gamma = 0.1$.

Proof. Note that

$$\begin{aligned} f'(x) &= \frac{\nu\cosh(\nu x)\sinh(x) - \cosh(x)\sinh(\nu x)}{(\sinh(x))^2} \\ &= \frac{(1+\nu)\sinh((1-\nu)x) - (1-\nu)\sinh((1+\nu)x)}{2(\sinh(x))^2}. \end{aligned}$$

With $\lambda := (1-\nu)/(1+\nu)$ the strict convexity of sinh implies that for $\nu \in (0,1)$ and $x > 0$

$$\sinh((1-\nu)x) < \lambda\sinh((1+\nu)x) + (1-\lambda)\sinh(0) = \frac{1-\nu}{1+\nu}\sinh((1+\nu)x).$$

Therefore $f'(x) < 0$. $\square$

It follows immediately from this fact that $X_0^*(t)$ is a strictly decreasing function of $\alpha\sigma^2$ when $x_0 > 0$ and $0 < t < T$. Economically, this means that the agent will liquidate the initial asset position faster when the perceived volatility risk increases, because $\operatorname{var}(\mathcal{R}(X_0^*))$ is proportional to $\alpha\sigma^2$ according to (2.7) and (2.8). So the first guess would be that also the equilibrium strategy X_1^* should be a decreasing function of $\alpha\sigma^2$ when $x_1 > 0$. This guess is also analyzed and tested empirically by Lebedeva et al. [2012] for a large data set of block executions by large insiders. Here, however, all we get from applying Lemma 2.1.12 to (2.30) is the following partial result.

Proposition 2.1.13. *If $x_1 \geq x_2 \geq 0$ then $X_1^*(t)$ is a strictly decreasing function of $\alpha\sigma^2$ for $0 < t < T$.*

As a matter of fact, the monotonicity in $\alpha\sigma^2$ may break down in the two-player Nash equilibrium when the conditions $x_1 \geq x_2$ and $x_2 \geq 0$ in Proposition 2.1.13 are not both satisfied; see Figure 2.4. An intuitive explanation for this failure of monotonicity can be understood from Figure 2.5. Here, agent 2 has a larger initial position than agent 1. When $\alpha\sigma^2$ increases from 0.1 to 0.8, agent 2 receives a relatively high increase in volatility risk and therefore increases the liquidation speed throughout the first part of $[0, T]$ while slowing down in the second part. The volatility risk of agent 1 also increases, but it does so less than for agent 2. On the other hand, the increased price pressure

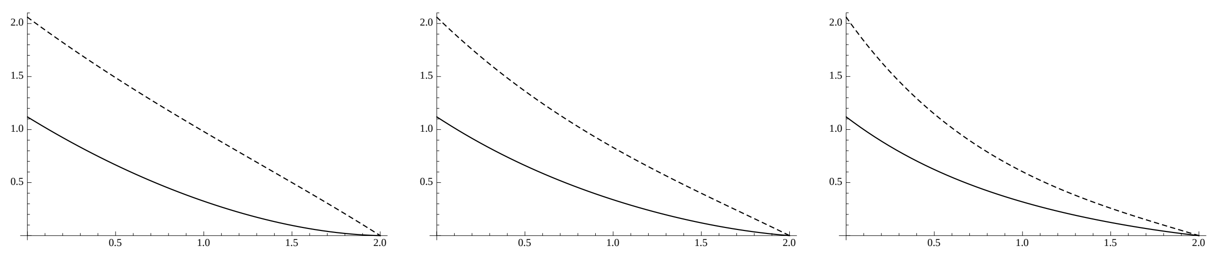


Figure 2.5: $X_1^*(t)$ (solid) and $X_2^*(t)$ (dashed) as functions of $t \in [0, T]$ for $x_1 = 1.12$, $x_2 = 2.06$, $T = 2$, $\lambda = \gamma = 1$, and $\alpha\sigma^2 = 0.1$ (left), $\alpha\sigma^2 = 0.8$ (center) and $\alpha\sigma^2 = 2.5$ (right).

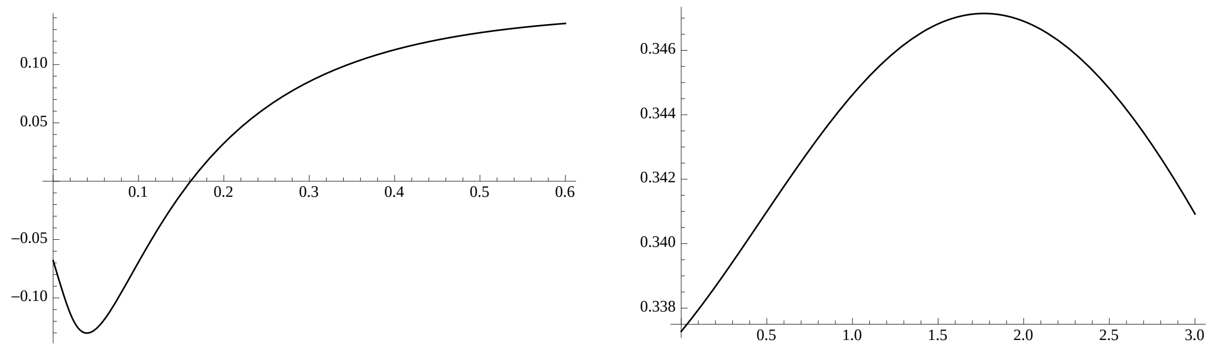


Figure 2.6: Left: $X_1^*(1)$ as a function of λ for $x_1 = 0.2$, $x_2 = 4$, $T = 2$, $\alpha\sigma^2 = 1$, and $\gamma = 0.3$; Right: $X_1^*(1)$ as a function of γ for $x_1 = 0.86$, $x_2 = 0.28$, $T = 2$, $\alpha\sigma^2 = 1$, and $\lambda = 1$.

from agent 2 leads to unfavorable asset prices for agent 1, and this latter effect outweighs the increased volatility risk. Therefore it is beneficial for agent 1 to delay selling in the first and accelerate the strategy in the second part of the time interval. This leads to the observed increase of the intermediate asset position $X_1^*(1)$. When $\alpha\sigma^2$ increases even further, the increase in volatility risk becomes dominant, and so $X_1^*(1)$ starts to decrease.

Next, $X_0^*(t)$ is independent of γ, whereas both two-player equilibrium strategies are nontrivial functions of γ. The intuitive reason for this dependence is the fact that the permanent price impact created by the liquidation strategy of one agent is perceived as an additional price trend by the other agent.

Moreover, $X_0^*(t)$ is an increasing function of λ by Lemma 2.1.12. The monotonicity in λ has the clear economic intuition that increasing the transaction costs from temporary price impact reduces the benefits from an early liquidation and thus drives the optimal strategy toward the linear liquidation strategy that is optimal in the risk-neutral case $\alpha = 0$. It is also tested and analyzed empirically by Lebedeva et al. [2012]. By applying Lemma 2.1.12 to (2.30) it is only possible to obtain the monotonic dependence of $X_1^*(t)$ on γ and λ when $x_1 = x_2$:

Proposition 2.1.14. *If $x_1 = x_2 \geq 0$ then $X_1^*(t) = X_2^*(t)$ is a strictly decreasing function of γ and a strictly increasing function of λ for $0 < t < T$.*

As shown in Figure 2.6, the monotonic dependence on γ or λ may break down when the condition $x_1 = x_2$ from Proposition 2.1.14 is not satisfied. The intuitive explanation for these effects are similar to the one for the breakdown of monotonicity for $\alpha\sigma^2$. For instance, when λ increases in a Nash equilibrium with $0 < x_1 \ll x_2$, both agents receive an incentive to reduce the curvature of their strategies, that is, to sell slower in the first part of the trading interval and to sell faster during the second part. Agent 2 will therefore create less price impact during the first part of $[0, T]$ and more price impact in the second part. In equilibrium, this change in price impact generated by one trader creates another incentive for the other trader with just the opposite effect, namely to increase trading speed during the first part of $[0, T]$ and to reduce it during the second part when the unfavorable price impact generated by the competitor is increased. When the position of agent 1 is smaller than the one of agent 2, this second effect can dominate the increase of transaction costs in the strategy of agent 1 so that we observe the decrease of $X_1^*(1)$ on the leftmost side of Figure 2.6.

2.2 Nash equilibrium with infinite time horizon

Now we consider mean-variance optimization and CARA utility maximization for an infinite time horizon $[0, \infty)$. Financially, this problem corresponds to a situation in which none of the agents faces a material time constraint. To simplify the discussion, we assume from the beginning that the drift $b(\cdot)$ vanishes identically. Then the unaffected price process is given by $S^0(t) = S_0 + \sigma W(t)$ for $t \geq 0$. Here we need to assume that $\sigma \neq 0$. If only one agent is active, we are in the situation of Schied and Schöneborn [2009], where the problem of maximizing the expected utility of revenues is discussed for an infinite time horizon. As discussed there, a strategy $(X(t))_{t\geq 0}$ should satisfy the following conditions of admissibility so that the utility-maximization problem is well-defined for a single agent:

- X is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$;
- X is absolutely continuous in the sense that $X(t) = X(0) + \int_0^t \dot{X}(s)\,ds$ for some progressively measurable process $\dot{X}(t))_{t\geq 0}$ for which
$$\int_0^\infty (\dot{X}(t))^2\,dt < \infty \qquad \mathbb{P}\text{-a.s.}; \tag{2.35}$$
- X is bounded and satisfies
$$\mathbb{E}\Big[\int_0^\infty X(t)^2\,dt\Big] < \infty \qquad \text{and} \qquad \lim_{t\uparrow\infty}(X(t))^2 t \log\log t = 0 \ \mathbb{P}\text{-a.s.} \tag{2.36}$$

The class of all strategies that are admissible in this sense and satisfy $X(0) = x$ for given $x \in \mathbb{R}$ will be denoted by $\mathcal{X}(x, \infty)$. As before we denote by $\mathcal{X}_{\det}(x, \infty)$ the subclass of all deterministic strategies in $\mathcal{X}(x, \infty)$. When the admissible strategy X is used, the affected price process is
$$S^X(t) = S^0(t) + \gamma(X(t) - X(0)) + \lambda\dot{X}(t).$$
It is shown in Schied and Schöneborn [2009, Section 3.1] that the total revenues of $X \in \mathcal{X}(x, \infty)$ are $\mathbb{P}$-a.s. well-defined as the limit
$$\mathcal{R}(X) := -\lim_{T\uparrow\infty}\int_0^T \dot{X}(t)S^X(t)\,dt = xS_0 - \frac{\gamma}{2}x^2 + \sigma\int_0^\infty X(t)\,dW(t) - \lambda\int_0^\infty (\dot{X}(t))^2\,dt$$

(see also Lemma 2.2.1 below). Moreover, for $\alpha > 0$, the unique strategy that maximizes the expected utility $\mathbb{E}[\,u_\alpha(\mathcal{R}(X))\,]$ over $X \in \mathcal{X}(x,\infty)$ is given by

$$X_0^*(t) = x \exp\Big(-t\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\Big), \qquad t \geq 0;$$

see Corollary 4.4 in Schied and Schöneborn [2009]. Since $\mathcal{R}(X)$ is a Gaussian random variable for $X \in \mathcal{X}_{\det}(x,\infty)$ one sees that

$$\mathbb{E}[\,u_\alpha(\mathcal{R}(X))\,] = \frac{1}{\alpha}\Big(1 - e^{-\alpha\mathbb{E}[\,\mathcal{R}(X)\,]+\frac{\alpha^2}{2}\operatorname{var}(\mathcal{R}(X))}\Big), \qquad X \in \mathcal{X}_{\det}(x,\infty),$$

and so X_0^* also maximizes the mean-variance functional $\mathbb{E}[\,\mathcal{R}(X)\,] - \frac{\alpha}{2}\operatorname{var}(\mathcal{R}(X))$ over $X \in \mathcal{X}_{\det}(x,\infty)$.

When n investors apply strategies $X_1, X_2, \dots, X_n$, the affected price $S^{X_1,\dots,X_n}(t)$ is again given by (2.6), as in the case of a finite time horizon. It will follow from Lemma 2.2.1 below that the admissibility of $X_1, X_2, \dots, X_n$ guarantees that the following limit exists $\mathbb{P}$-a.s.:

$$\mathcal{R}(X_i|\boldsymbol{X}_{-i}) := -\lim_{T\uparrow\infty}\int_0^T \dot X_i(t) S^{X_1,\dots,X_n}(t)\,dt.$$

The respective Nash equilibria for mean-variance optimization and CARA utility maximization can now be defined by taking $T = \infty$ in Definition 2.1.1.

Lemma 2.2.1. *For $X_i \in \mathcal{X}(x_i,\infty)$, $i = 1,\dots,n$, the limit*

$$\mathcal{R}(X_i|\boldsymbol{X}_{-i}) := -\lim_{T\uparrow\infty}\int_0^T \dot X_i(t) S^{X_1,\dots,X_n}(t)\,dt$$

exists and is given by

$$\begin{aligned}&\mathcal{R}(X_i|\boldsymbol{X}_{-i})\\ &\quad= x_iS_0 - \frac{\gamma}{2}x_i^2 + \int_0^\infty X_i(t)\,dW(t) + \gamma\sum_{j\neq i}\int_0^\infty X_i(t)\dot X_j(t)\,dt - \lambda\sum_{j=1}^n\int_0^\infty \dot X_i(t)\dot X_j(t)\,dt.\end{aligned}$$

Proof. Integrating by parts yields

$$\begin{aligned}&-\int_0^T \dot X_i(t) S^{X_1,\dots,X_n}(t)\,dt\\ &\quad= (x_i - X_i(T))S_0 - X(T)W(T) + \sigma\int_0^T X_i(t)\,dW(t) - \frac{\gamma}{2}(X_i(T)-X_i(0))^2\\ &\qquad - \gamma\sum_{j\neq i}X_i(T)(X_j(T)-X_j(0)) + \gamma\sum_{j\neq i}\int_0^T X_i(t)\dot X_j(t)\,dt - \lambda\sum_{j=1}^n\int_0^T \dot X_i(t)\dot X_j(t)\,dt.\end{aligned}$$

The assertion now follows by using the law of the iterated logarithm for W, (2.35), (2.36), and the Cauchy–Schwarz inequality. $\square$

Now let $X_i \in \mathcal{X}(x_i,\infty)$, $i = 1,\dots,n$, be given. As in (2.7), (2.8) we get that for $Y \in \mathcal{X}(y,\infty)$,

$$\mathbb{E}[\,\mathcal{R}(Y|\boldsymbol{X}_{-i})\,] - \frac{\alpha_i}{2}\operatorname{var}(\mathcal{R}(Y|\boldsymbol{X}_{-i})) = c + \int_0^\infty \mathcal{L}^i(t, Y(t), \dot Y(t)|\boldsymbol{X}_{-i})\,dt,$$

where $c = yS_0 - \frac{\gamma}{2}y^2$ and the Lagrangian $\mathcal{L}^i$ is given by (2.8). Note that we must have $\alpha_i\sigma^2 > 0$ to have a chance to obtain solutions of the mean-variance optimization problem for otherwise the Lagrangian is linear. Here is our result on the existence and uniqueness of a Nash equilibrium for mean-variance optimization with infinite time horizon.

Theorem 2.2.2. *Suppose that one of the following two conditions holds:*

1. *$n \in \mathbb{N}$ is arbitrary and $\alpha_1 = \cdots = \alpha_n = \alpha > 0$;*
2. *$n = 2$ and α_1, α_2 are distinct and strictly positive.*

Then for all $x_1, \dots, x_n \in \mathbb{R}$ there exists a unique Nash equilibrium for mean-variance optimization with infinite time horizon, which is also a Nash equilibrium for CARA utility maximization.

In case 1 the optimal strategies are given by

$$X_i^*(t) = (x_i - \overline{x}_n)e^{\theta_- t} + \overline{x}_n e^{\rho_- t}, \tag{2.37}$$

where $\overline{x}_n = \frac{1}{n}\sum_{j=1}^n x_j$; ρ_- and θ_- are as in (2.22).

In case 2 the fourth-order equation

$$\tau^4 - \frac{2\gamma}{3\lambda}\tau^3 - \frac{\gamma^2 + 2\lambda\sigma^2(\alpha_1 + \alpha_2)}{3\lambda^2}\tau^2 + \frac{\sigma^4\alpha_1\alpha_2}{3\lambda^2} = 0$$

has precisely two distinct strictly negative roots, τ_1, τ_2, and the equilibrium strategies $X_1^*(t)$ and $X_2^*(t)$ are linear combinations of the exponential functions $e^{\tau_1 t}$ and $e^{\tau_2 t}$.

To prove Theorem 2.2.2, we need the following auxiliary lemma.

Lemma 2.2.3. *For $i = 1, \dots, n$ and $\alpha_i > 0$ the functional $Y \mapsto \int_0^\infty \mathcal{L}^i(t, Y(t), \dot{Y}(t)|\boldsymbol{X}_{-i})\, dt$ has at most one maximizer in $\mathcal{X}_{det}(y, \infty)$. If, moreover, $X_1, \dots, X_n$ belong to $C^2[0,\infty)$ and are such that*

$$\int_0^T \Big|\gamma\sum_{j\neq i}\dot{X}_j(t) + \lambda\sum_{j\neq i}\ddot{X}_j(t)\Big|\, dt < \infty, \tag{2.38}$$

then there exists a unique maximizer $Y^* \in \mathcal{X}_{det}(y, T) \cap C^2[0,\infty)$, which is given as the unique solution of the boundary value problem

$$\alpha_i\sigma^2 Y(t) - 2\lambda\ddot{Y}(t) = \gamma\sum_{j\neq i}\dot{X}_j(t) + \lambda\sum_{j\neq i}\ddot{X}_j(t), \qquad Y(0) = y,\ \lim_{t\uparrow\infty} Y(t) = 0.$$

Moreover, Y satisfies $\int_0^\infty |\dot{Y}(t)| + |\ddot{Y}(t)|\, dt < \infty$.

Proof. It follows from the strict concavity of the Lagrangian $\mathcal{L}^i$ and the convexity of the set $\mathcal{X}_{\mathrm{det}}(y, \infty)$ that there can be at most one maximizer in $\mathcal{X}_{\mathrm{det}}(y, \infty)$.

Now we show the existence of a maximizer under the additional assumptions (2.38) and $X_1, \dots, X_n \in C^2[0,\infty)$. As noted in the proof of Lemma 2.1.4, the general solution of the Euler–Lagrange equation (2.12) is given by

$$Y(t) = c_1 e^{-\kappa_i t} + c_2 e^{\kappa_i t} - \frac{1}{4\lambda\kappa_i}\int_0^t e^{\kappa_i(t-s)}u(s)\, ds + \frac{1}{4\lambda\kappa_i}\int_0^t e^{-\kappa_i(t-s)}u(s)\, ds, \tag{2.39}$$

where $u(t) = \gamma \sum_{j \neq i} \dot{X}_j(t) + \lambda \sum_{j \neq i} \ddot{X}_j(t)$, c_1 and c_2 are constants, and $\kappa_i = \sqrt{\alpha_i \sigma^2 / 2\lambda} > 0$. One checks that (2.38) implies that $\int_0^t e^{-\kappa_i(t-s)} u(s)\, ds \to 0$ as $t \uparrow \infty$. Therefore, when letting

$$c_2 := \frac{1}{4\lambda\kappa_i} \int_0^\infty e^{-\kappa_i s} u(s)\, ds \tag{2.40}$$

and $c_1 := y - c_2$, one sees that Y solves (2.39).

To see that $\int_0^\infty |\dot{Y}(t)| + |\ddot{Y}(t)|\, dt < \infty$, note first that

$$\int_0^T \int_0^t e^{-\kappa_i(t-s)} |u(s)|\, ds\, dt = \frac{1}{\kappa_i} \int_0^T |u(s)|\, ds - \frac{1}{\kappa_i} \int_0^T e^{-\kappa_i(T-s)} |u(s)|\, ds$$

and

$$\int_0^T \int_t^\infty e^{\kappa_i(t-s)} |u(s)|\, ds\, dt = \frac{1}{\kappa_i} \int_0^\infty (e^{\kappa_i(s \wedge T - s)} - 1) |u(s)|\, ds,$$

both of which by (2.38) converge to finite limits for $T \uparrow \infty$. It thus follows from (2.39) and (2.40) that $\int_0^\infty |\dot{Y}(t)| + |\ddot{Y}(t)|\, dt < \infty$.

Finally, it is clear from the preceding arguments that $Y^* \in \mathscr{X}_{\mathrm{det}}(y, T) \cap C^2[0, \infty)$. The optimality of Y^* follows as in the second part of the proof of Lemma 2.1.4. □

Proof of Theorem 2.2.2. One first shows just as in Lemma 2.1.3 that there can be at most one Nash equilibrium for mean-variance optimization. Moreover, one shows as in the proof of Corollary 2.1.6 that a Nash equilibrium for mean-variance optimization is also a Nash equilibrium for CARA utility maximization.

Now we turn to the proof of existence of a Nash equilibrium for given initial values $x_1, \dots, x_n \in \mathbb{R}$. Let M be the $2n \times 2n$-matrix defined in Remark 2.1.5. As observed in the proof of Lemma 2.1.9, any eigenvector of M with eigenvalue τ must be of the form $\binom{\boldsymbol{v}}{\tau \boldsymbol{v}}$ for some $\boldsymbol{v} \in \mathbb{R}^n$. We will show below that in both cases, 1 and 2, there exists a basis $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ of $\mathbb{R}^n$ and numbers $\tau_1, \dots, \tau_n < 0$ (which are not necessarily distinct) such that $\binom{\boldsymbol{v}_1}{\tau_1 \boldsymbol{v}_1}, \dots, \binom{\boldsymbol{v}_n}{\tau_n \boldsymbol{v}_n}$ are eigenvectors of M. Taking this fact as given, let $c_1, \dots, c_n \in \mathbb{R}$ be such that $c_1 \boldsymbol{v}_1 + \dots + c_n \boldsymbol{v}_n = (x_1, \dots, x_n)^\top$ and define

$$\boldsymbol{Z}(0) := c_1 \begin{pmatrix} \boldsymbol{v}_1 \\ \tau_1 \boldsymbol{v}_1 \end{pmatrix} + \dots + c_n \begin{pmatrix} \boldsymbol{v}_n \\ \tau_n \boldsymbol{v}_n \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{Z}(t) := e^{tM} \boldsymbol{Z}(0).$$

We denote by $\boldsymbol{X}^*(t)$ the first n components of $\boldsymbol{Z}(t)$. As observed in the proof of Theorem 2.1.2 and Remark 2.1.5, $\boldsymbol{X}^*(t)$ will solve the system (2.14) of coupled Euler–Lagrange equations, which by Lemma 2.2.3 is sufficient for optimality in the infinite-horizon setting, provided that the components correspond to admissible strategies and satisfy the integrability conditions of Lemma 2.2.3. But each component of $\boldsymbol{X}^*(t)$ is by construction a linear combination of the decreasing exponential functions $e^{\tau_1 t}, \dots, e^{\tau_n t}$, and so these conditions are clearly satisfied.

Now we consider the situation in case 1. Then θ_- and ρ_- defined in (2.22) are strictly negative, and so the required existence of $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ follows from Lemma 2.1.9. It follows from the preceding part of the proof that each component of $\boldsymbol{X}^*(t)$ can be written as

$$X_i^*(t) = c_i(\theta_-) e^{\theta_- t} + c(\rho_-) e^{\rho_- t}.$$

Letting again $\Sigma(t) := \sum_{j=1}^n X_j^*(t)$ and arguing as in the proof of Theorem 2.1.8 yields first $\Sigma(t) = \sum_{i=1}^n x_i e^{\rho_- t}$ and then

$$c(\rho_-) = \frac{1}{n}\sum_{j=1}^n x_j \quad \text{and} \quad c_i(\theta_-) = x_i - c(\rho_-).$$

This establishes (2.37) and completes the proof of Theorem 2.2.2 under assumption 1.

Now we turn toward case 2. We may assume without loss of generality that $\sigma = 1$. The characteristic polynomial of the matrix M of the system (2.18) for $n = 2$ is

$$\chi(\tau) := \tau^4 - \frac{2\gamma}{3\lambda}\tau^3 - \frac{\gamma^2 + 2\lambda(\alpha_1+\alpha_2)}{3\lambda^2}\tau^2 + \frac{\alpha_1\alpha_2}{3\lambda^2}.$$

Its derivative, χ', has three distinct roots, t_0, t_+, t_-, which are given by

$$t_0 = 0, \quad t_\pm = \frac{3\gamma \pm \sqrt{33\gamma^2 + 48\,(\alpha_1+\alpha_2)\,\lambda}}{12\lambda}.$$

Note first that t_0 is a strictly positive local maximum of χ since

$$\chi(t_0) = \frac{\alpha_1\alpha_2}{3\lambda^2} > 0, \quad \chi''(t_0) = -\frac{2\,(2\alpha_1\lambda + 2\alpha_2\lambda + \gamma^2)}{3\lambda^2} < 0.$$

Next, $t_+ > 0$, $t_- < 0$, and

$$\begin{aligned}\chi(t_-) = \frac{1}{864\lambda^5}\Big(&-96\lambda^3\left(\alpha_1^2 - \alpha_1\alpha_2 + \alpha_2^2\right) - 168\gamma^2\lambda^2(\alpha_1+\alpha_2) - 69\gamma^4\lambda \\ &+ 16\gamma\lambda(\alpha_1+\alpha_2)\sqrt{48\lambda^3(\alpha_1+\alpha_2)+33\gamma^2\lambda^2} + 11\gamma^3\sqrt{48\lambda^3(\alpha_1+\alpha_2)+33\gamma^2\lambda^2}\Big).\end{aligned}$$

If we can show that $\chi(t_-) < 0$ then χ will have precisely two distinct strictly negative roots. It is, however, not easy to determine by direct inspection of our preceding formula whether $\chi(t_-) < 0$. But we already know that for $\alpha_1 = \alpha_2$ the matrix M has exactly two strictly negative (though not necessarily distinct) eigenvalues, ρ_- and θ_-. So in this case, both eigenvalues must be strictly negative roots of χ. We moreover know that $\chi(0) > 0$, $\lim_{\tau\downarrow-\infty}\chi(\tau) = +\infty$, and that t_- is the only strictly negative critical point of χ. It follows that we must have $\chi(t_-) \le 0$ when $\alpha_1 = \alpha_2$. Now suppose that $\alpha_1 \neq \alpha_2$ and let $\overline{\alpha} := \frac{1}{2}(\alpha_1+\alpha_2)$. Then $\alpha_1 + \alpha_2 = \overline{\alpha} + \overline{\alpha}$ and

$$\alpha_1^2 - \alpha_1\alpha_2 + \alpha_2^2 - \overline{\alpha}^2 = \frac{3}{4}(\alpha_1-\alpha_2)^2 > 0.$$

It therefore follows that $\chi(t_-) < \overline{\chi}(t_-)$, where $\overline{\chi}$ denotes the characteristic polynomial of M when both α_1 and α_2 have been substituted by $\overline{\alpha}$. Since the formula for t_- is left invariant by this substitution, we must have $\overline{\chi}(t_-) \le 0$ according to what has been said before, and so we arrive at $\chi(t_-) < 0$.

It follows from the preceding paragraph that M has two distinct strictly negative eigenvalues τ_1 and τ_2. Hence there exist corresponding eigenvectors of the form $\binom{\boldsymbol{v}_1}{\tau_1\boldsymbol{v}_1}$ and $\binom{\boldsymbol{v}_2}{\tau_2\boldsymbol{v}_2}$. But we still need to exclude the possibility that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are linearly dependent to complete the proof. To this end, note that it follows from (2.19) that we must have

$$\frac{1}{\lambda}\Big(A - \frac{1}{n+1}JA\Big)\boldsymbol{w} - \tau\frac{\gamma}{\lambda}\Big(\frac{2}{n+1}J - I\Big)\boldsymbol{w} = \tau^2\boldsymbol{w} \tag{2.41}$$

for $\binom{\boldsymbol{w}}{\tau\boldsymbol{w}}$ to be an eigenvector of M with eigenvalue τ.

Let us first suppose that the components w_1 and w_2 of $\boldsymbol{w}$ do not add up to zero: $w_1 + w_2 \neq 0$. Then taking the inner product of the vector equation (2.41) with the vector $\binom{1}{1}$ yields the equation

$$\alpha_1 w_1 + \alpha_2 w_2 - \tau\gamma(w_1 + w_2) = 3\tau^2\lambda(w_1 + w_2).$$

This quadratic equation in τ has the two possible roots

$$\tau_\pm = \frac{-\gamma \pm \sqrt{\gamma^2 + 12\lambda \frac{\alpha_1 w_1 + \alpha_2 w_2}{w_1 + w_2}}}{6\lambda},$$

one of which must be equal to τ. Since $\tau_- < 0 < \tau_+$ it follows that $\binom{\boldsymbol{w}}{\widetilde{\tau}\boldsymbol{w}}$ cannot be an eigenvector of M for any $\widetilde{\tau}$ that is different from τ and has the same sign as τ.

Let us now consider the case in which $w_1 = -w_2$. Taking the inner product of the equation (2.41) with the vector $\binom{-1}{1}$ and using the requirement $w_1, w_2 \neq 0$ yields the equation $\alpha_2 + \alpha_1 + 2\tau\gamma = 2\tau^2\lambda$, which is independent of w_1 and w_2. It has the roots

$$\frac{\gamma \pm \sqrt{\gamma^2 + 4\lambda(\alpha_1 + \alpha_2)}}{2\lambda},$$

which again have different signs. We thus conclude as in the case $w_1 + w_2 \neq 0$. □

On the one hand, the structure of equilibrium strategies for an infinite time horizon appears to be simpler than for the finite-time situation. On the other hand, the assumptions of Theorem 2.2.2 are more restrictive than those of Theorem 2.1.2. The reason is that all solutions $X_1(t), \dots, X_n(t)$ of the system (2.9) are linear combinations of exponential functions and thus can only take the limits $\pm\infty$ and 0 for $t \uparrow \infty$. We therefore cannot apply standard results on the existence of solutions for boundary value problems on non-compact intervals such as those in Cecchi et al. [1980], where it is required that the possible boundary values at $t = \infty$ include the full space $\mathbb{R}^n$. Instead, we show here that the eigenspaces associated with the negative eigenvalues of a certain non-symmetric matrix M are sufficiently rich. For $n > 2$ we are only able to understand these eigenspaces when $\alpha_1 = \cdots = \alpha_n$.

Remark 2.2.4. In the situation of part 1 of Theorem 2.2.2, consider the corresponding Nash equilibrium $X_1^{(T)}, \dots, X_n^{(T)}$ for the finite time interval $[0, T]$ as constructed in Theorem 2.1.8. Then we conclude from (2.24) and (2.25) that

$$\lim_{T\uparrow\infty} X_i^{(T)}(t) = X_i^*(t) \qquad \text{for } i = 1, \dots, n \text{ and } t \geq 0,$$

where X_i^* is as in (2.37). □

Let us finally discuss some qualitative properties of the Nash equilibrium in part 1 of Theorem 2.2.2. One issue that is discussed in Carlin et al. [2007] and Schöneborn and Schied [2009] is whether agents with zero initial capital, $x_i = 0$ for $i \neq 1$, engage in predatory trading or liquidity provision when another agent, say agent 1, is liquidating a large block of shares. Here, predatory trading refers to a strategy during which the asset

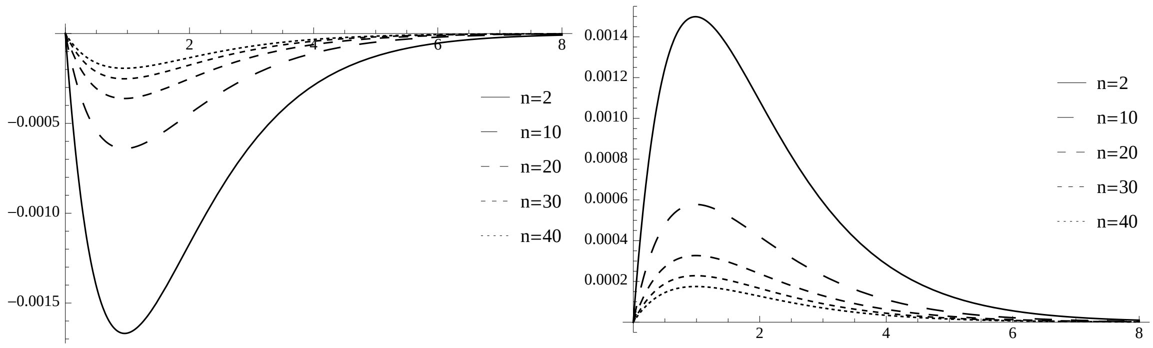


Figure 2.7: Strategies $X_i^*(t)$ for $\lambda = 0.15$ (left) and $\lambda = 0.16$ (right) for various choices of n and for $x_i = 0$, $\sum_{j=1}^n x_j = 1$, $\gamma = 0.16$, and $\alpha\sigma^2 = 0.33$.

is shortened at the initial high price and then bought back later when the sell strategy of agent 1 has depreciated the asset price. This strategy is "predatory" in the sense that the revenues it generates for agent i are made at the expense of agent 1. Liquidity provision refers to exactly the opposite strategy: agent i acquires a long position by first buying and later re-selling some of the shares agent 1 is liquidating. It can hence be seen as a cooperative behavior on behalf of agent i. Both Carlin et al. [2007] and Schöneborn and Schied [2009] consider risk-neutral agents who need to close their positions in finite time. In Carlin et al. [2007] all agents face the same time constraint, and in this case liquidity provision can only be observed if cooperation is enforced by repeating the game Schöneborn and Schied [2009] allow a longer time horizon for agent i than for agent 1 and in this case find that liquidity provision may be possible, depending on the market parameters and the number of competitors. Our result here is Corollary 2.2.5 below. It is illustrated in Figure 2.7. Note that by Remark 2.2.4 the two possibilities of predatory trading and liquidity provision must occur already for finite time horizons T, a fact that is markedly different from the risk-neutral case $\alpha = 0$ considered in Carlin et al. [2007] and Schöneborn and Schied [2009].

Corollary 2.2.5. *In the situation of part 1 of Theorem 2.2.2 suppose that $\sum_{i=1}^n x_i > 0$. Then an agent with $x_i = 0$ engages in liquidity provision in the sense that $X_i^*(t) > 0$ for all $t > 0$ if and only if $\alpha\sigma^2\lambda > 2\gamma^2$. When $\alpha\sigma^2\lambda < 2\gamma^2$ this agent engages in predatory trading, and for $\alpha\sigma^2\lambda = 2\gamma^2$ the agent does not trade at all.*

Proof of Corollary 2.2.5. It follows from (2.37) that $X_i^*(t)$ has the same sign as

$$\rho_- - \theta_- = \frac{\gamma}{2\lambda}\left(\frac{-2n}{n+1} + \sqrt{1+\xi} - \sqrt{\left(\frac{n-1}{n+1}\right)^2 + \frac{\xi}{n+1}}\right),$$

where $\xi = 4\alpha\sigma^2\lambda/\gamma^2$. The right-hand side is a strictly increasing function of ξ and vanishes for $\xi = 8$. □

Finally, we briefly discuss the behavior of equilibrium strategies as a function of the number n of agents active in the market. Lebedeva et al. [2012] discuss the following two hypothesis and analyze their validity for a large data set of block executions by large insiders:

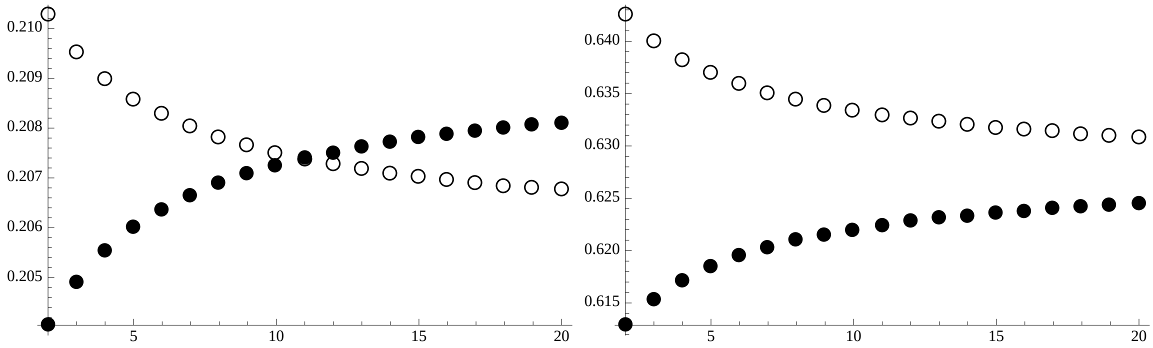


Figure 2.8: Left: $X_1^*(1.5)$ as a function of $n \in \{2, \ldots, 20\}$ for $\gamma = 0.155$ (circles) and $\gamma = 0.16$ (bullets) with $x_1 = 1$, $\sum_{i=1}^n x_i = 3.5$, $\lambda = 0.15$, and $\alpha\sigma^2 = 0.33$. Right: $X_1^*(1.5)$ as a function of $n \in \{2, \ldots, 20\}$ for $\sum_{i=1}^n x_i = -10$ (circles) and $\sum_{i=1}^n x_i = 10$ (bullets) with $x_1 = 3$, $\lambda = 0.15$, $\gamma = 1.6$, and $\alpha\sigma^2 = 0.33$.

- **Hypothesis 1:** "Trade duration decreases if several insiders compete for exploiting the same long-lived information."
- **Hypothesis 2:** "Trade duration increases if several insiders trade simultaneously in the same direction for liquidity reasons."

In the situation of part 1 of our Theorem 2.2.2 we typically do indeed find equilibrium strategies that are monotone in n, but the effective trade duration can be both increasing and decreasing in n; see Figure 2.8. Here, the effective trade duration can be defined as the time until a certain high percentage of the initial inventory has been liquidated. So both hypotheses from Lebedeva et al. [2012] are compatible with risk-averse agents in an Almgren–Chriss setting.

Chapter 3

A hot-potato game under transient price impact

According to the Report CFTC-SEC [2010] on the Flash Crash of May 6, 2010, the events that lead to the Flash Crash included a large sell order of E-Mini S&P 500 contracts:

> ... a large Fundamental Seller (...) initiated a program to sell a total of 75,000 E-Mini contracts (valued at approximately \$4.1 billion). ... [On another] occasion it took more than 5 hours for this large trader to execute the first 75,000 contracts of a large sell program. However, on May 6, when markets were already under stress, the Sell Algorithm chosen by the large Fundamental Seller to only target trading volume, and not price nor time, executed the sell program extremely rapidly in just 20 minutes.

The report CFTC-SEC [2010] furthermore suggests that a "hot-potato game" between high-frequency traders (HFTs) created artificial trading volume that contributed to the acceleration of the Fundamental Seller's trading algorithm:

> ... HFTs began to quickly buy and then resell contracts to each other—generating a "hot-potato" volume effect as the same positions were rapidly passed back and forth. Between 2:45:13 and 2:45:27, HFTs traded over 27,000 contracts, which accounted for about 49 percent of the total trading volume, while buying only about 200 additional contracts net.

See also Kirilenko et al. [2010] for additional background.

Schöneborn [2008] observes that the equilibrium strategies of two competing economic agents, who trade sufficiently fast in a simple market impact model with exponential decay of price impact, can exhibit strong oscillations. These oscillations have a striking similarity with the "hot-potato game" mentioned in CFTC-SEC [2010] and Kirilenko et al. [2010]. In each trading period, one agent sells a large asset position to the other agent and buys a similar position back in the next period. An intuitive reason for this hot-potato game is to protect against possible predatory trading by the other agent.

In this chapter, we pick up this observation by Schöneborn [2008]. Our first contribution is to extend the result by identifying a unique Nash equilibrium for two competing agents within a larger class of adaptive trading strategies and by giving an explicit formula for the equilibrium strategies. This formula allows us to establish the existence of oscillatory equilibrium strategies in some generality. We call this extended model the *primary model*.

Another new feature of the primary model is the addition of quadratic transaction costs, which can be thought of temporary price impact in the sense of Bertsimas and Lo [1998], Almgren and Chriss [2000] or as a transaction tax. The main goal of this chapter is to study the impact of the additional transaction costs on equilibrium strategies. Theorem 3.1.14, one of our main results, precisely identifies a critical threshold θ^* for the size θ of these transaction costs at which all oscillations disappear. That is, for transactions $\theta \geq \theta^*$ certain "fundamental" equilibrium strategies consist exclusively of all buy trades or of all sell trades. For $\theta < \theta^*$, the "fundamental" equilibrium strategies will contain both buy and sell trades when the resilience of price impact in between two trades is sufficiently small.

In addition, numerical simulations will exhibit some rather striking properties of equilibrium strategies. They reveal, for instance, that the expected costs of both agents can be a *decreasing* function of $\theta \in [0, \theta_0]$ when trading speed is sufficiently high. As a result, both agents can carry out their respective trades at a *lower cost* when there are transaction costs, compared to the situation without transaction costs. Even more interesting is the behavior of the costs as a function of the trading frequency. We will see that for $\theta = \theta^*$ the costs can *decrease* as the trading frequency goes up, whereas they can (essentially) *increase* for $\theta = 0$. In particular the latter effect is surprising, because at first glance a higher trading frequency suggests greater flexibility in the choice of a strategy and hence the possibility to use more efficient trading strategies. So why are the costs then increasing in the trading frequency? We will argue that the intuitive reason for this effect is a trade-off that occurs between the discouragement for predatory trading strategies by the other agent and the additional payment of transaction costs for one's own strategies. The discouragement of predatory trading strategies through increased transaction costs means that both agents can use more efficient strategies to carry out their trades, and the benefits of the effects outweigh the price to be paid in the higher transaction costs.

We also extend the primary model in three aspects. First we impose permanent impact and observe its interaction with transient impact and transaction costs. We will see that incorporating permanent impact increases the critical value of transaction costs, i.e., one needs to impose more transaction costs to prevent oscillatory strategies when there exists permanent impact. Another aspect we discuss is splitting of combined liquidation costs. We show the existence and uniqueness of Nash equilibria in the case that the combined liquidation costs are not equally split. Since the combined liquidation costs in our model will be taken by the one who trades slower than the other, the splitting of the combined liquidation costs can be regarded as a problem of latency impact. Then we simulate Nash equilibria in several cases. The last aspect we analyze is Nash equilibria in a closed-loop model. In comparison to open-loop strategies, for each execution time point t_n, closed-loop strategies are allowed to take into account the trades of the other agent at execution time points $t_0, \ldots, t_{n-1}$. We derive an explicit formula of optimal closed-loop strategies by dynamic programming. Through numerical simulations, we compare open-loop and closed-loop strategies. Furthermore, we also impose transaction costs, unequal splitting of combined liquidation costs and analyze their effects.

Another contribution of this chapter is to extend the primary model by considering its continuous-time version. By two different approaches we obtain the same definition of the liquidation costs, which is consistent with the primary model. A surprising result is that for any nontrivial case, a unique Nash equilibrium exists if and only if the transaction costs θ are exactly equal to a critical value θ^*. Mathematically, there exists no strategy that satisfies an equivalent condition for the existence of Nash equilibria if $\theta \neq \theta^*$. This equivalent condition is defined through an integral equation. Furthermore, we also show that the unique Nash equilibrium only consists of deterministic strategies. Intuitively, when $\theta \neq \theta^*$, we observe by our numerical simulations that both agents will try to use the strategies consisted of infinite tiny jumps to avoid the penalty of transaction costs. However, these strategies are convergent to strategies having large jumps at the beginning and the end of trading time horizon that lead to significant transaction costs. This causes the nonexistence of Nash equilibria.

This chapter builds on several research developments in the existing literature. First, there are several papers on predatory trading such as Brunnermeier and Pedersen [2005], Carlin et al. [2007] and Schöneborn and Schied [2009] dealing with Nash equilibria for several agents that are active in a market model with price impact. While predatory trading prevails in Brunnermeier and Pedersen [2005], Carlin et al. [2007], it is found in Schöneborn and Schied [2009] that, depending on market conditions, either predatory trading or liquidity provision is optimal. In contrast to these previous studies, the price impact model we use here goes back to Bouchaud et al. [2004] and Obizhaeva and Wang [2013]. It is further developed in Alfonsi et al. [2008, 2010], Gatheral [2010], Predoiu et al. [2011], Alfonsi et al. [2012], Lorenz and Schied [2012], to mention only a few related papers. We refer to Gatheral and Schied [2013], Lehalle [2013] for recent surveys on the price impact literature and extended bibliographies.

This chapter is organized as follows. In Section 3.1 we explain our modeling framework of the primary model. The existence and uniqueness theorem for the Nash equilibrium is stated and we analyze the oscillations of equilibrium strategies. Here we state one of the main results on the critical threshold for the disappearance of oscillations. Section 3.1 is based on Schied and Zhang [2013a]. In Section 3.2, three extensions of the primary model are stated. Here we consider incorporation of permanent impact, unequal splitting of combined liquidation costs and closed-loop strategies. For each extension we also present numerical simulations. In Section 3.3, we define the liquidation costs and analyze the continuous-time version of the primary model. Then we state another main result on the existence of Nash equilibria and analyze why Nash equilibria exist only for one critical value of transaction costs. At last we take a review on the effects of transaction costs in single-agent models.

3.1 The primary model

3.1.1 Modeling framework

We consider two financial agents, X and Y, who are active in a market impact model for one risky asset. When none of the two agents is active, asset prices are described by a right-continuous martingale[1] $S^0 = (S^0_t)_{t\geq 0}$ on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$, for which $\mathcal{F}_0$ is $\mathbb{P}$-trivial. The process S^0 is often called the unaffected price process.

[1]The martingale assumption is natural from an economic point of view, because we are interested here in high-frequency trading over short time intervals $[0, T]$. See also the discussion in Alfonsi et al. [2012] for additional arguments.

Trading takes place at the discrete trading times of a *time grid* $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$, where $0 = t_0 < t_1 < \cdots < t_N = T$. Both agents are assumed to use trading strategies that are admissible in the following sense.

Definition 3.1.1. Suppose that a time grid $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$ is given. An *admissible trading strategy* for $\mathbb{T}$ and $Z_0 \in \mathbb{R}$ is a vector $\boldsymbol{\zeta} = (\zeta_0, \dots, \zeta_N)$ of random variables such that

1. each ζ_i is $\mathcal{F}_{t_i}$-measurable and bounded,

2. $Z_0 = \zeta_0 + \cdots + \zeta_N$ $\mathbb{P}$-a.s.

The set of all admissible strategies for given $\mathbb{T}$ and Z_0 is denoted by $\mathcal{X}(Z_0, \mathbb{T})$.

For $\boldsymbol{\zeta} \in \mathcal{X}(Z_0, \mathbb{T})$, the value of ζ_i is taken as the number of shares traded at time t_i, with a positive sign indicating a sell order and a negative sign indicating a purchase. Thus, the requirement 2 in the preceding definition can be interpreted by saying that Z_0 is the initial asset position of the agent at time $t_0 = 0$ and that by time $t_N = T$ (e.g., the end of a trading day) the agent must have a zero inventory. The assumption that each ζ_i is bounded can be made without loss of generality from an economic point of view.

When the two agents X and Y apply respective strategies $\boldsymbol{\xi} \in \mathcal{X}(X_0, \mathbb{T})$ and $\boldsymbol{\eta} \in \mathcal{X}(Y_0, \mathbb{T})$, the asset price is given by

$$S_t^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_t^0 - \lambda \sum_{t_k < t} e^{-\rho(t-t_k)}(\xi_k + \eta_k). \tag{3.1}$$

That is, at each time $t_k \in \mathbb{T}$, the combined trading activities of the two agents move the current price by the amount $-\lambda(\xi_k + \eta_k)$, but this immediate price impact decays exponentially in time at rate $\rho \geq 0$. In this form, the model is a straightforward two-agent extension of the market impact model of Obizhaeva and Wang [2013]. One intuition behind this model is that $S^{\boldsymbol{\xi},\boldsymbol{\eta}}$ describes the price movements of the mid-price in a block-shaped limit order book of height $1/\lambda$ and without bid-ask spread; see, e.g., Alfonsi et al. [2008], Obizhaeva and Wang [2013].

Let us now discuss the definition of the liquidation costs incurred by each agent. When just one agent, say X, places a nonzero order at time t_k, then the price is moved from $S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}$ to $S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - \lambda\xi_k$ and the trade ξ_k incurs the following expenses:

$$\int_{S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}}^{S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - \lambda\xi_k} z\frac{1}{\lambda}\,dz = \frac{\lambda}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k.$$

Suppose now that the order η_k of agent Y is executed immediately after the order ξ_k. Then the price is moved from $S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - \lambda\xi_k$ to $S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - \lambda\xi_k - \lambda\eta_k$, and agent Y incurs the expenses

$$\int_{S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - \lambda\xi_k}^{S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - \lambda\xi_k - \lambda\eta_k} z\frac{1}{\lambda}\,dz = \frac{\lambda}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + \lambda\xi_k\eta_k.$$

So greater latency results in the additional cost term $\lambda\xi_k\eta_k$ for agent Y. Clearly, this term would appear in the expenses of agent X when the roles of X and Y are reversed. In the sequel, we are going to assume that none of the two agents has an advantage in latency over the other. Therefore, if both agents place nonzero orders at time t_k, execution priority is given to that agent who wins an independent coin toss.

In addition to the liquidation costs motivated above, we will also impose that each trade ζ_k incurs quadratic transaction costs of the form $\theta\zeta_k^2$, where θ is a nonnegative parameter.

Definition 3.1.2. Suppose that $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$, X_0 and Y_0 are given. Let furthermore $(\varepsilon_i)_{i=0,1,\dots}$ be an i.i.d. sequence of Bernoulli $(\frac{1}{2})$-distributed random variables that are independent of $\sigma(\bigcup_{t\geq 0}\mathcal{F}_t)$. Then the *liquidation costs of* $\boldsymbol{\xi} \in \mathcal{X}(X_0, \mathbb{T})$ *given* $\boldsymbol{\eta} \in \mathcal{X}(Y_0, \mathbb{T})$ are defined as

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}) = X_0 S_0^0 + \sum_{k=0}^{N}\Big(\frac{\lambda}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k + \varepsilon_k\lambda\xi_k\eta_k + \theta\xi_k^2\Big), \tag{3.2}$$

and the *liquidation costs of* $\boldsymbol{\eta}$ *given* $\boldsymbol{\xi}$ are

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}) = Y_0 S_0^0 + \sum_{k=0}^{N}\Big(\frac{\lambda}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + (1-\varepsilon_k)\lambda\xi_k\eta_k + \theta\eta_k^2\Big).$$

The term $X_0S_0^0$ corresponds to the face value of the position X_0 at time $t = 0$. When the position X_0 could be liquidated at face value, one would incur the expenses $-X_0S_0^0$. Therefore, the liquidation costs as defined in (3.2) are the difference of the actual accumulated expenses, as represented by the sum on the right-hand side of (3.2), and the expenses for liquidation at face value. In the following two remarks we comment on our model assumptions from an economic point of view.

Remark 3.1.3. The market impact model we are using here has often been linked to the placement of market orders in a block-shaped limit order book modeled in Obizhaeva and Wang [2013], Alfonsi et al. [2008]. In the model version we use here, this picture is simplified by neglecting bid-ask spread. As argued in Alfonsi et al. [2008], Alfonsi and Schied [2010], this simplification is irrelevant for strategies that consist exclusively of buy orders or exclusively of sell orders. For oscillatory strategies between buy and sell orders, however, neglecting the bid-ask spread will be unrealistic as long as these strategies consist exclusively of market orders. But in reality, strategies will involve a variety of different order types and one should think of the costs (3.2) as the costs *aggregated* over order types. For instance, while one may have to pay the spread when placing a market buy order, one essentially earns it back when a limit sell order is executed. Moreover, high-frequency traders often have access to a variety of more exotic order types that actually can pay rebates when executed, and they can use crossing networks or dark pools in which orders are executed at mid price. So for a setup of high-frequency trading, taking the bid-ask spread as zero is probably not unrealistic. The existence of hot-potato games in real-world markets, such as the one quoted from CFTC-SEC [2010], can be regarded as an empirical justification of the zero-spread assumption, because such a trading behavior could never be profitable when each trader had to pay the full spread upon each execution of an order. As a matter of fact, even though high-frequency traders engaged in a hot-potato game during the Flash Crash, Kirilenko et al. [2010, Figure 6] find that:

> High Frequency Traders are consistently profitable although they never accumulate a large net position. This does not change on May 6 as they appear to

have been even more successful despite the market volatility observed on that day.

Remark 3.1.4. We admit that we have chosen quadratic transaction costs because this choice makes our model mathematically tractable. Yet, there are several aspects why quadratic transaction costs may not be completely implausible from an economic point of view. For instance, these costs can be regarded as arising from temporary price impact in the spirit of Bertsimas and Lo [1998], Almgren and Chriss [2000], which is also quadratic in order size. Moreover, these costs can model a transaction tax that is subject to tax progression. With such a tax, small orders, such as those placed by small investors, are taxed at a lower rate than large orders, which may be placed with the intention of moving the market.

3.1.2 Nash equilibrium under transient price impact

We suppose now that each agent starts with a given initial position and minimizes the expected costs over admissible strategies. Obizhaeva and Wang [2013] considered this optimization problem when there is just one agent in the market; see Figure 3.2 for its solution. Suppose now that there are two agents, X and Y. In a partial equilibrium, only X has full knowledge of Y's strategy, whereas Y has no knowledge about X. The strategy of agent Y will then create dynamic price impact, which will be perceived as additional drift by agent X. So the optimization problem of agent X is equivalent to minimizing the expected costs for an unaffected price process with additional drift. This problem has been solved by Lorenz and Schied [2012] in a continuous-time version of the model considered here and under the assumption $\theta = 0$. They find that optimal strategies for agent X often may not exist, and if they do, they may strongly depend on the time derivative of the drift. It is shown in particular that agent X can make arbitrarily large expected profits when Y uses the strategy from Obizhaeva and Wang [2013] and X has slightly more time for closing the position than Y. This observation already indicates a certain instability of the transient price impact model without transaction costs.

Here we will assume that both agents X and Y have full knowledge of the other's trading strategy and maximize the expected liquidation costs of their strategies accordingly. In this situation, it is natural to define an optimality through the following notion of Nash equilibrium.

Definition 3.1.5. For given time grid $\mathbb{T}$ and initial values X_0, $Y_0 \in \mathbb{R}$, a *Nash equilibrium* is a pair $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ of strategies in $\mathcal{X}(X_0, \mathbb{T}) \times \mathcal{X}(Y_0, \mathbb{T})$ such that

$$\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,] = \inf_{\boldsymbol{\xi}\in\mathcal{X}(X_0,\mathbb{T})} \mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,] \qquad \text{and} \qquad \mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)\,] = \inf_{\boldsymbol{\eta}\in\mathcal{X}(Y_0,\mathbb{T})} \mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,].$$

The existence of a unique Nash equilibrium in the class of deterministic strategies is first shown in Schöneborn [2008, Theorem 9.1]. Here we extend this result first by including transaction costs and giving an explicit form of the deterministic Nash equilibrium and then by showing that this Nash equilibrium is also the unique Nash equilibrium in the class of adapted strategies.

To state out formula for this Nash equilibrium, we need to introduce the following notation. For a fixed time grid $\mathbb{T} = \{t_0, \dots, t_N\}$, we define the $(N+1) \times (N+1)$-matrix G by

$$G_{i+1,j+1} = e^{-\rho|t_i - t_j|}, \qquad i, j = 0, \dots, N. \tag{3.3}$$

We furthermore define the lower triangular matrix $\widetilde{G}$ by

$$\widetilde{G}_{ij} = \begin{cases} G_{ij}, & \text{if } i > j; \\ \frac{1}{2}, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

We will write $\mathbf{1}$ for the vector $(1, \dots, 1)^\top \in \mathbb{R}^{N+1}$. We furthermore define the two vectors

$$\begin{aligned} \boldsymbol{v} &= \frac{1}{\mathbf{1}^\top(\lambda G + \lambda\widetilde{G} + 2\theta\,\mathrm{Id})^{-1}\mathbf{1}}(\lambda G + \lambda\widetilde{G} + 2\theta\,\mathrm{Id})^{-1}\mathbf{1}, \\ \boldsymbol{w} &= \frac{1}{\mathbf{1}^\top(\lambda G - \lambda\widetilde{G} + 2\theta\,\mathrm{Id})^{-1}\mathbf{1}}(\lambda G - \lambda\widetilde{G} + 2\theta\,\mathrm{Id})^{-1}\mathbf{1}. \end{aligned} \tag{3.5}$$

It will be proved in Lemma 3.1.8 below that the occurring matrices are invertible and that the denominators in (3.5) are strictly positive so that $\boldsymbol{v}$ and $\boldsymbol{w}$ are well-defined.

A strategy $\boldsymbol{\zeta} = (\zeta_0, \dots, \zeta_N) \in \mathcal{X}(Z_0, \mathbb{T})$ will be identified with the $(N+1)$-dimensional random vector $(\zeta_0, \dots, \zeta_N)^\top$. Conversely, any vector $\boldsymbol{z} = (z_1, \dots, z_{N+1})^\top \in \mathbb{R}^{N+1}$ can be identified with the deterministic strategy $\boldsymbol{\zeta}$ with $\zeta_k = z_{k-1}$.

We can now state the following result on the existence and uniqueness of a Nash equilibrium. As explained above, this result extends Schöneborn [2008, Theorem 9.1].

Theorem 3.1.6. *Let $\rho > 0$, $\lambda > 0$, and $\theta \ge 0$ be given. For any time grid $\mathbb{T}$ and initial values X_0, $Y_0 \in \mathbb{R}$, there exists a unique Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*) \in \mathcal{X}(X_0, \mathbb{T}) \times \mathcal{X}(Y_0, \mathbb{T})$. The optimal strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ are deterministic and given by*

$$\begin{aligned} \boldsymbol{\xi}^* &= \frac{1}{2}(X_0 + Y_0)\boldsymbol{v} + \frac{1}{2}(X_0 - Y_0)\boldsymbol{w}, \\ \boldsymbol{\eta}^* &= \frac{1}{2}(X_0 + Y_0)\boldsymbol{v} - \frac{1}{2}(X_0 - Y_0)\boldsymbol{w}. \end{aligned} \tag{3.6}$$

In order to prove Theorem 3.1.6, we state the following series of lemmas.

Lemma 3.1.7. *The expected costs of an admissible strategy $\boldsymbol{\xi} \in \mathcal{X}(X_0, \mathbb{T})$ given another admissible strategy $\boldsymbol{\eta} \in \mathcal{X}(Y_0, \mathbb{T})$ are*

$$\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,] \;=\; \mathbb{E}\Big[\frac{1}{2}\boldsymbol{\xi}^\top(\lambda G + 2\theta\,\mathrm{Id})\boldsymbol{\xi} + \boldsymbol{\xi}^\top\lambda\widetilde{G}\boldsymbol{\eta}\Big]. \tag{3.7}$$

Proof. Since the sequence $(\varepsilon_i)_{i=0,1,\dots}$ is independent of $\sigma(\bigcup_{t\ge0}\mathcal{F}_t)$ and the two strategies $\boldsymbol{\xi}$

and $\boldsymbol{\eta}$ are measurable with respect to this σ-field, we get $\mathbb{E}[\,\varepsilon_k\lambda\xi_k\eta_k\,] = \frac{\lambda}{2}\mathbb{E}[\,\xi_k\eta_k\,]$. Hence,

$$\begin{aligned}
\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,] - X_0S_0^0 &= \mathbb{E}\Big[\sum_{k=0}^{N}\Big(\frac{\lambda}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k + \varepsilon_k\lambda\xi_k\eta_k + \theta\xi_k^2\Big)\Big] \\
&= \mathbb{E}\Big[\sum_{k=0}^{N}\Big(\frac{\lambda}{2}\xi_k^2 + \frac{\lambda}{2}\xi_k\eta_k - \xi_k\Big(S_{t_k}^0 - \lambda\sum_{m=0}^{k-1}(\xi_m+\eta_m)e^{-\rho(t_k-t_m)}\Big) + \theta\xi_k^2\Big)\Big] \\
&= \mathbb{E}\Big[-\sum_{k=0}^{N}\xi_kS_{t_k}^0 + \frac{\lambda}{2}\sum_{k=0}^{N}\xi_k^2 + \lambda\sum_{k=0}^{N}\xi_k\sum_{m=0}^{k-1}\xi_m e^{-\rho(t_k-t_m)} \\
&\qquad + \sum_{k=0}^{N}\Big(\xi_k\Big(\frac{\lambda}{2}\eta_k + \lambda\sum_{m=0}^{k-1}\eta_m e^{-\rho(t_k-t_m)}\Big) + \theta\xi_k^2\Big)\Big].
\end{aligned}$$

Since each ξ_k is $\mathcal{F}_{t_k}$-measurable and S^0 is a martingale, we get from condition 2 in Definition 3.1.1 that

$$\mathbb{E}\Big[\sum_{k=0}^{N}\xi_kS_{t_k}^0\Big] = \mathbb{E}\Big[\sum_{k=0}^{N}\xi_kS_T^0\Big] = X_0\mathbb{E}[\,S_T^0\,] = X_0S_0^0.$$

Moreover,

$$\frac{\lambda}{2}\sum_{k=0}^{N}\xi_k^2 + \lambda\sum_{k=0}^{N}\xi_k\sum_{m=0}^{k-1}\xi_m e^{-\rho(t_k-t_m)} = \frac{\lambda}{2}\sum_{k,m=0}^{N}\xi_k\xi_m e^{-\rho|t_k-t_m|} = \frac{1}{2}\boldsymbol{\xi}^\top\lambda G\boldsymbol{\xi},$$

and

$$\sum_{k=0}^{N}\xi_k\Big(\frac{\lambda}{2}\eta_k + \lambda\sum_{m=0}^{k-1}\eta_m e^{-\rho(t_k-t_m)}\Big) = \boldsymbol{\xi}^\top\lambda\widetilde{G}\boldsymbol{\eta}.$$

Putting everything together yields the assertion. □

We will use the convention of calling an $n\times n$-matrix A *positive definite* when $\boldsymbol{x}^\top A\boldsymbol{x} > 0$ for all nonzero $\boldsymbol{x}\in\mathbb{R}^n$, even when A is not necessarily symmetric. Clearly, for a positive definite matrix A there is no nonzero $\boldsymbol{x}\in\mathbb{R}^n$ for which $A\boldsymbol{x} = \mathbf{0}$, and so A is invertible. Moreover, writing a given nonzero $\boldsymbol{x}\in\mathbb{R}^n$ as $\boldsymbol{x} = A\boldsymbol{y}$ for $\boldsymbol{y} = A^{-1}\boldsymbol{x}\neq\mathbf{0}$, we see that $\boldsymbol{x}^\top A^{-1}\boldsymbol{x} = \boldsymbol{y}^\top A^\top\boldsymbol{y} = \boldsymbol{y}^\top A\boldsymbol{y} > 0$. So the inverse of a positive definite matrix is also positive definite.

Lemma 3.1.8. *For $\tau\geq 0$, the matrices G, $\widetilde{G}$, $G+\widetilde{G}+\tau\,\mathrm{Id}$, $G-\widetilde{G}+\tau\,\mathrm{Id}$ are positive definite. In particular, all terms in* (3.5) *are well-defined and the denominators in* (3.5) *are strictly positive.*

Proof. That G is positive definite is easy to see; two different proofs are given for instance in Theorem 3.3 of Alfonsi et al. [2008] or in Example 1 of Alfonsi et al. [2012]. Therefore, for nonzero $\boldsymbol{x}\in\mathbb{R}^{N+1}$,

$$0 < \boldsymbol{x}^\top G\boldsymbol{x} = \boldsymbol{x}^\top(\widetilde{G}+\widetilde{G}^\top)\boldsymbol{x} = \boldsymbol{x}^\top\widetilde{G}\boldsymbol{x} + \boldsymbol{x}^\top\widetilde{G}^\top\boldsymbol{x} = 2\boldsymbol{x}^\top\widetilde{G}\boldsymbol{x},$$

which shows that $\widetilde{G}$ is strictly positive definite. Next, $G-\widetilde{G} = \widetilde{G}^\top$ and so this matrix is also strictly positive definite. Clearly, the sum of two positive definite matrices is also positive definite, which shows that $G+\widetilde{G}+\tau\,\mathrm{Id}$ and $G-\widetilde{G}+\tau\,\mathrm{Id}$ are positive definite for $\tau\geq 0$. □

Lemma 3.1.9. *For given time grid $\mathbb{T}$ and initial values X_0 and Y_0 there exists at most one Nash equilibrium in the class $\mathcal{X}(X_0,\mathbb{T})\times\mathcal{X}(Y_0,\mathbb{T})$.*

Proof. We suppose by way of contradiction that there exist two distinct Nash equilibria $(\boldsymbol{\xi}^0,\boldsymbol{\eta}^0)$ and $(\boldsymbol{\xi}^1,\boldsymbol{\eta}^1)$ in $\mathcal{X}(X_0,\mathbb{T})\times\mathcal{X}(Y_0,\mathbb{T})$. Then we define for $\alpha\in[0,1]$

$$\boldsymbol{\xi}^\alpha := \alpha\boldsymbol{\xi}^1+(1-\alpha)\boldsymbol{\xi}^0 \qquad \text{and} \qquad \boldsymbol{\eta}^\alpha := \alpha\boldsymbol{\eta}^1+(1-\alpha)\boldsymbol{\eta}^0.$$

We furthermore let

$$f(\alpha):=\mathbb{E}\Big[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^\alpha|\boldsymbol{\eta}^0)+\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}^\alpha|\boldsymbol{\xi}^0)+\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^{1-\alpha}|\boldsymbol{\eta}^1)+\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}^{1-\alpha}|\boldsymbol{\xi}^1)\,\Big].$$

According to Lemma 3.1.8 the matrix $\lambda G+2\theta\,\mathrm{Id}$ is positive definite, the functional

$$\boldsymbol{\xi}\longmapsto\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,]=\mathbb{E}\Big[\,\frac{1}{2}\boldsymbol{\xi}^\top(\lambda G+2\theta\,\mathrm{Id})\boldsymbol{\xi}+\boldsymbol{\xi}^\top\lambda\widetilde{G}\boldsymbol{\eta}\,\Big]$$

is strictly convex with respect to $\boldsymbol{\xi}$. Since the two Nash equilibria $(\boldsymbol{\xi}^0,\boldsymbol{\eta}^0)$ and $(\boldsymbol{\xi}^1,\boldsymbol{\eta}^1)$ are distinct, $f(\alpha)$ must also be strictly convex in α and have its unique minimum in $\alpha=0$. That is,

$$f(\alpha)>f(0) \text{ for } \alpha>0. \tag{3.8}$$

It follows that

$$\lim_{h\downarrow 0}\frac{f(h)-f(0)}{h}=\frac{df(\alpha)}{d\alpha}\Big|_{\alpha=0+}\geq 0. \tag{3.9}$$

Let us introduce the shorthand notation $M:=\lambda G+2\theta\,\mathrm{Id}$. Then, by the symmetry of M,

$$\begin{aligned}\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^\alpha|\boldsymbol{\eta})\,] &= \mathbb{E}\Big[\,\frac{1}{2}(\boldsymbol{\xi}^\alpha)^\top M\boldsymbol{\xi}^\alpha+(\boldsymbol{\xi}^\alpha)^\top\lambda\widetilde{G}\boldsymbol{\eta}\,\Big]\\ &= \mathbb{E}\Big[\,\frac{1}{2}\alpha^2(\boldsymbol{\xi}^1)^\top M\boldsymbol{\xi}^1+\alpha(1-\alpha)(\boldsymbol{\xi}^1)^\top M\boldsymbol{\xi}^0+\frac{1}{2}(1-\alpha)^2(\boldsymbol{\xi}^0)^\top M\boldsymbol{\xi}^0\\ &\qquad +\alpha(\boldsymbol{\xi}^1)^\top\lambda\widetilde{G}\boldsymbol{\eta}+(1-\alpha)(\boldsymbol{\xi}^0)^\top\lambda\widetilde{G}\boldsymbol{\eta}\,\Big].\end{aligned}$$

Therefore,

$$\frac{d}{d\alpha}\Big|_{\alpha=0+}\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^\alpha|\boldsymbol{\eta})\,]=\mathbb{E}\Big[\,(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top M\boldsymbol{\xi}^0+(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\lambda\widetilde{G}\boldsymbol{\eta}\,\Big].$$

Hence, it follows that

$$\begin{aligned}&\frac{d}{d\alpha}\Big|_{\alpha=0+}f(\alpha)\\ &=\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top M\boldsymbol{\xi}^0+(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\lambda\widetilde{G}\boldsymbol{\eta}^0+(\boldsymbol{\xi}^0-\boldsymbol{\xi}^1)^\top M\boldsymbol{\xi}^1+(\boldsymbol{\xi}^0-\boldsymbol{\xi}^1)^\top\lambda\widetilde{G}\boldsymbol{\eta}^1\\ &\qquad+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top M\boldsymbol{\eta}^0+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top\lambda\widetilde{G}\boldsymbol{\xi}^0+(\boldsymbol{\eta}^0-\boldsymbol{\eta}^1)^\top M\boldsymbol{\eta}^1+(\boldsymbol{\eta}^0-\boldsymbol{\eta}^1)^\top\lambda\widetilde{G}\boldsymbol{\xi}^1\Big]\\ &=-\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top M(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top M(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big]\\ &\qquad+\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\lambda\widetilde{G}(\boldsymbol{\eta}^0-\boldsymbol{\eta}^1)+(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\lambda\widetilde{G}^\top(\boldsymbol{\eta}^0-\boldsymbol{\eta}^1)\Big]\\ &=-\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top M(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top M(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big]-\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\lambda G(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big].\end{aligned}$$

Now,

$$
\begin{aligned}
&(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top \lambda G(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)+\frac{1}{2}\Big((\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top M(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top M(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big)\\
&\quad\geq \frac{1}{2}\Big((\boldsymbol{\xi}^1-\boldsymbol{\xi}^0+\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top \lambda G(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0+\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big)\geq 0.
\end{aligned}
$$

Therefore, and because the two Nash equilibria $(\boldsymbol{\xi}^0,\boldsymbol{\eta}^0)$ and $(\boldsymbol{\xi}^1,\boldsymbol{\eta}^1)$ are distinct, we have

$$
\frac{d}{d\alpha}\Big|_{\alpha=0+} f(\alpha)\leq -\frac{1}{2}\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top \lambda G(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top \lambda G(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big]<0,
$$

which contradicts (3.9). Therefore, there can exist at most one Nash equilibrium in the class $\mathcal{X}(X_0,\mathbb{T})\times\mathcal{X}(Y_0,\mathbb{T})$. □

Now let us introduce the class

$$
\mathcal{X}_{\text{det}}(Z_0,\mathbb{T}):=\Big\{\boldsymbol{\zeta}\in\mathcal{X}(Z_0,\mathbb{T})\,\Big|\,\boldsymbol{\zeta}\text{ is deterministic}\Big\}
$$

of deterministic strategies in $\mathcal{X}(Z_0,\mathbb{T})$. A Nash equilibrium in the class $\mathcal{X}_{\text{det}}(X_0,\mathbb{T})\times\mathcal{X}_{\text{det}}(Y_0,\mathbb{T})$ is defined in the same way as in Definition 3.1.5.

Lemma 3.1.10. *A Nash equilibrium in the class $\mathcal{X}_{\text{det}}(X_0,\mathbb{T})\times\mathcal{X}_{\text{det}}(Y_0,\mathbb{T})$ of deterministic strategies is also a Nash equilibrium in the class $\mathcal{X}(X_0,\mathbb{T})\times\mathcal{X}(Y_0,\mathbb{T})$ of adapted strategies.*

Proof. Assume that $(\boldsymbol{\xi}^*,\boldsymbol{\eta}^*)$ is a Nash equilibrium in the class $\mathcal{X}_{\text{det}}(X_0,\mathbb{T})\times\mathcal{X}_{\text{det}}(Y_0,\mathbb{T})$ of deterministic strategies. We need to show that $\boldsymbol{\xi}^*$ minimizes $\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,]$ and $\boldsymbol{\eta}^*$ minimizes $\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,]$ in the respective classes $\mathcal{X}(X_0,\mathbb{T})$ and $\mathcal{X}(Y_0,\mathbb{T})$ of adapted strategies. To this end, let $\boldsymbol{\xi}\in\mathcal{X}(X_0,\mathbb{T})$ be given. We define $\overline{\boldsymbol{\xi}}\in\mathcal{X}_{\text{det}}(X_0,\mathbb{T})$ by $\overline{\xi}_k=\mathbb{E}[\,\xi_k\,]$ for $k=0,1,\dots,N$.

Using once again the shorthand notation $M:=\lambda G+2\theta\,\mathrm{Id}$ and applying Jensen's inequality to the convex function $\mathbb{R}^{N+1}\ni\boldsymbol{x}\mapsto\boldsymbol{x}^\top M\boldsymbol{x}$, we obtain

$$
\begin{aligned}
\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,]&=\mathbb{E}\Big[\frac{1}{2}\boldsymbol{\xi}^\top M\boldsymbol{\xi}+\boldsymbol{\xi}^\top\lambda\widetilde{G}\boldsymbol{\eta}^*\Big]=\mathbb{E}\Big[\frac{1}{2}\boldsymbol{\xi}^\top M\boldsymbol{\xi}\Big]+\overline{\boldsymbol{\xi}}^\top\lambda\widetilde{G}\boldsymbol{\eta}^*\\
&\geq\frac{1}{2}\overline{\boldsymbol{\xi}}^\top M\overline{\boldsymbol{\xi}}+\overline{\boldsymbol{\xi}}^\top\lambda\widetilde{G}\boldsymbol{\eta}^*=\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\overline{\boldsymbol{\xi}}|\boldsymbol{\eta}^*)\,]\\
&\geq\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,].
\end{aligned}
$$

This shows that $\boldsymbol{\xi}^*$ minimizes $\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,]$ over $\boldsymbol{\xi}\in\mathcal{X}(X_0,\mathbb{T})$. One can show analogously that $\boldsymbol{\eta}^*$ minimizes $\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,]$ over $\boldsymbol{\eta}\in\mathcal{X}(Y_0,\mathbb{T})$, which completes the proof. □

Remark 3.1.11. Before proving Theorem 3.1.6, we briefly explain how to derive the explicit form (3.6) of the equilibrium strategies. By Lemma 3.1.7 and the method of Lagrange multipliers, a necessary condition for $(\boldsymbol{\xi}^*,\boldsymbol{\eta}^*)$ to be a Nash equilibrium in $\mathcal{X}_{\text{det}}(X_0,\mathbb{T})\times\mathcal{X}_{\text{det}}(Y_0,\mathbb{T})$ is the existence of $\alpha,\beta\in\mathbb{R}$, such that

$$
\begin{cases}
M\boldsymbol{\xi}^*+\lambda\widetilde{G}\boldsymbol{\eta}^*=\alpha\mathbf{1};\\
M\boldsymbol{\eta}^*+\lambda\widetilde{G}\boldsymbol{\xi}^*=\beta\mathbf{1},
\end{cases}
\tag{3.10}
$$

where $M := \lambda G + 2\theta\,\mathrm{Id}$. By adding the equations in (3.10) we obtain

$$(M + \lambda\widetilde{G})(\boldsymbol{\xi}^* + \boldsymbol{\eta}^*) = (\alpha + \beta)\mathbf{1}. \tag{3.11}$$

By Lemma 3.1.8, the matrix $M + \lambda\widetilde{G}$ is positive definite and hence invertible, so that (3.11) can be solved for $\boldsymbol{\xi}^* + \boldsymbol{\eta}^*$. Since we must also have $\mathbf{1}^\top(\boldsymbol{\xi}^* + \boldsymbol{\eta}^*) = X_0 + Y_0$, we obtain

$$\boldsymbol{\xi}^* + \boldsymbol{\eta}^* = \frac{X_0 + Y_0}{\mathbf{1}^\top(M + \lambda\widetilde{G})^{-1}\mathbf{1}}(M + \lambda\widetilde{G})^{-1}\mathbf{1} = (X_0 + Y_0)\boldsymbol{v}.$$

Similarly, subtracting one equation from the other in (3.10) yields

$$(M - \lambda\widetilde{G})(\boldsymbol{\xi}^* - \boldsymbol{\eta}^*) = (\alpha - \beta)\mathbf{1}.$$

It follows again from Lemma 3.1.8 that $(M - \lambda\widetilde{G})$ is invertible, and so we have

$$\boldsymbol{\xi}^* - \boldsymbol{\eta}^* = \frac{X_0 - Y_0}{\mathbf{1}^T(M - \lambda\widetilde{G})^{-1}\mathbf{1}}(M - \lambda\widetilde{G})^{-1}\mathbf{1} = (X_0 - Y_0)\boldsymbol{w}.$$

Thus, $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ must necessarily be given by (3.6).

Proof of Theorem 3.1.6. By Lemmas 3.1.9 and 3.1.10 all we need to show is that (3.6) defines a Nash equilibrium in the class $\mathscr{X}_{\mathrm{det}}(X_0, \mathbb{T}) \times \mathscr{X}_{\mathrm{det}}(Y_0, \mathbb{T})$ of deterministic strategies. For $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathscr{X}_{\mathrm{det}}(X_0, \mathbb{T}) \times \mathscr{X}_{\mathrm{det}}(Y_0, \mathbb{T})$ we have, using the shorthand notation $M := \lambda G + 2\theta\,\mathrm{Id}$,

$$\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,] = \frac{1}{2}\boldsymbol{\xi}^\top M\boldsymbol{\xi} + \boldsymbol{\xi}^\top\lambda\widetilde{G}\boldsymbol{\eta}. \tag{3.12}$$

Therefore minimizing $\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,]$ over $\boldsymbol{\xi} \in \mathscr{X}_{\mathrm{det}}(X_0, \mathbb{T})$ is equivalent to the minimization of the quadratic form on the right-hand side of (3.12) over $\boldsymbol{\xi} \in \mathbb{R}^{N+1}$ under the constraint $\mathbf{1}^\top\boldsymbol{\xi} = X_0$.

Now we prove that the strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ given by (3.6) are indeed optimal. We have

$$M\boldsymbol{\xi}^* + \lambda\widetilde{G}\boldsymbol{\eta}^* = \frac{1}{2}(X_0 + Y_0)(M + \lambda\widetilde{G})\boldsymbol{v} + \frac{1}{2}(X_0 - Y_0)(M - \lambda\widetilde{G})\boldsymbol{w} = \mu\mathbf{1}, \tag{3.13}$$

where

$$\mu = \frac{X_0 + Y_0}{2\mathbf{1}^\top(M + \lambda\widetilde{G})\mathbf{1}} + \frac{X_0 - Y_0}{2\mathbf{1}^\top(M - \lambda\widetilde{G})\mathbf{1}}.$$

Now let $\boldsymbol{\xi} \in \mathscr{X}_{\mathrm{det}}(X_0, \mathbb{T})$ be arbitrary and define $\boldsymbol{\zeta} := \boldsymbol{\xi} - \boldsymbol{\xi}^*$. Then we have $\boldsymbol{\zeta}^\top\mathbf{1} = 0$. Hence, by the symmetry of M,

$$\begin{aligned}
\frac{1}{2}\boldsymbol{\xi}^\top M\boldsymbol{\xi} + \boldsymbol{\xi}^\top\lambda\widetilde{G}\boldsymbol{\eta}^* &= \frac{1}{2}(\boldsymbol{\xi}^*)^\top M\boldsymbol{\xi}^* + \frac{1}{2}\boldsymbol{\zeta}^\top M\boldsymbol{\zeta} + \boldsymbol{\zeta}^\top M\boldsymbol{\xi}^* + (\boldsymbol{\xi}^*)^\top\lambda\widetilde{G}\boldsymbol{\eta}^* + \boldsymbol{\zeta}^\top\lambda\widetilde{G}\boldsymbol{\eta}^* \\
&= \frac{1}{2}(\boldsymbol{\xi}^*)^\top M\boldsymbol{\xi}^* + (\boldsymbol{\xi}^*)^\top\lambda\widetilde{G}\boldsymbol{\eta}^* + \frac{1}{2}\boldsymbol{\zeta}^\top M\boldsymbol{\zeta} + \mu\boldsymbol{\zeta}^\top\mathbf{1} \\
&\geq \frac{1}{2}(\boldsymbol{\xi}^*)^\top M\boldsymbol{\xi}^* + (\boldsymbol{\xi}^*)^\top\lambda\widetilde{G}\boldsymbol{\eta}^*,
\end{aligned}$$

where in the last step we have used that M is positive definite and that $\boldsymbol{\zeta}^\top\mathbf{1} = 0$. Therefore $\boldsymbol{\xi}^*$ minimizes (3.12) in the class $\mathscr{X}_{\mathrm{det}}(X_0, \mathbb{T})$ for $\boldsymbol{\eta} = \boldsymbol{\eta}^*$. In the same way, one shows that $\boldsymbol{\eta}^*$ minimizes $\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,]$ over $\boldsymbol{\eta} \in \mathscr{X}_{\mathrm{det}}(X_0, \mathbb{T})$. $\square$

3.1.3 Effects of transaction costs

We now turn toward a qualitative analysis of the equilibrium strategies. By means of numerical simulations and the analysis of a particular example, Schöneborn [2008, Section 9.3] observes that the equilibrium strategies may exhibit strong oscillations when $\theta = 0$; see Figure 3.1 for illustrations. In Schöneborn [2008], these oscillations are interpreted as a way of hedging against predatory trading by the other agent. The subsequent proposition implies that such oscillations will always occur in a Nash equilibrium with $X_0 = -Y_0$ and $\theta = 0$ when the trading frequency is sufficiently high. Note that this Nash equilibrium is completely determined by the vector $\boldsymbol{w}$. Throughout this subsection, we will concentrate on *equidistant time grids*,

$$\mathbb{T}_N := \Big\{ \frac{kT}{N} \,\Big|\, k = 0, 1, \dots, N \Big\}, \qquad N \in \mathbb{N}. \tag{3.14}$$

Proposition 3.1.12. *Suppose that ρ, λ, $T > 0$ and $N \in \mathbb{N}$ are fixed, and $\mathbb{T}_N$ is as in* (3.14).

1. *For $\theta = 0$, there exists $N_0 \in \mathbb{N}$ such that for $N \geq N_0$ the entries of the vector $\boldsymbol{w} = (w_1, \dots, w_{N+1})$ are nonzero and have alternating signs: $w_k w_{k+1} < 0$ for $k = 1, \dots, N$.*

2. *For $N \geq N_0$ there exists $\delta > 0$ such that for $0 \leq \theta < \delta$ the entries of the vector $\boldsymbol{w} = (w_1, \dots, w_{N+1})$ are nonzero and have alternating signs.*

Proof. We need to compute the inverse of the matrix $(\lambda G - \lambda \widetilde{G} + 2\theta\,\mathrm{Id})$. Setting $\kappa := 2\theta/\lambda + \frac{1}{2}$, this matrix is equal to $\frac{1}{\lambda}(G - \widetilde{G} + (\kappa - \frac{1}{2})\,\mathrm{Id})^{-1}$. Setting furthermore $a := e^{-\rho T}$, we have

$$\lambda G - \lambda \widetilde{G} + \lambda\Big(\kappa - \frac{1}{2}\Big)\,\mathrm{Id} = \lambda \begin{pmatrix} \kappa & a^{\frac{1}{N}} & a^{\frac{2}{N}} & \cdots & a^{\frac{N-1}{N}} & a \\ 0 & \kappa & a^{\frac{1}{N}} & \cdots & a^{\frac{N-2}{N}} & a^{\frac{N-1}{N}} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \kappa & a^{\frac{1}{N}} \\ 0 & \cdots & \cdots & \cdots & 0 & \kappa \end{pmatrix}.$$

It is easy to verify that the inverse of this matrix is given by

$$\Gamma_N := \frac{1}{\lambda} \begin{pmatrix} \frac{1}{\kappa} & \frac{-a^{\frac{1}{N}}}{\kappa^2} & \frac{-a^{\frac{2}{N}}(\kappa-1)}{\kappa^3} & \cdots & \frac{-a^{\frac{N-1}{N}}(\kappa-1)^{N-2}}{\kappa^N} & \frac{-a^{\frac{N}{N}}(\kappa-1)^{N-1}}{\kappa^{N+1}} \\ 0 & \frac{1}{\kappa} & \frac{-a^{\frac{1}{N}}}{\kappa^2} & \cdots & \frac{-a^{\frac{N-2}{N}}(\kappa-1)^{N-3}}{\kappa^{N-1}} & \frac{-a^{\frac{N-1}{N}}(\kappa-1)^{N-2}}{\kappa^N} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{1}{\kappa} & \frac{-a^{\frac{1}{N}}}{\kappa^2} \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{1}{\kappa} \end{pmatrix}. \tag{3.15}$$

Let us denote by $\boldsymbol{u} = (u_1, u_2, \dots, u_{N+1}) \in \mathbb{R}^{N+1}$ the vector $\lambda\Gamma_N\mathbf{1}$. Then we have $u_{N+1} = \frac{1}{\kappa}$

and, for $n = 1, \ldots, N$, $u_n = u_{n+1} - a^{(N+1-n)/N}(\kappa-1)^{N-n}/\kappa^{N+2-n}$. That is,

$$\begin{aligned} u_n &= \frac{1}{\kappa} - \frac{a^{\frac{1}{N}}}{\kappa^2}\sum_{m=n}^{N}\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{N-m} = \frac{1}{\kappa} - \frac{a^{\frac{1}{N}}}{\kappa^2}\sum_{k=0}^{N-n}\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{k} \\ &= \frac{1}{\kappa}\bigg[1 - \frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}})+a^{\frac{1}{N}}} + (-1)^{N+1-n}\frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}})+a^{\frac{1}{N}}}\Big(\frac{a^{\frac{1}{N}}(1-\kappa)}{\kappa}\Big)^{N+1-n}\bigg]. \end{aligned} \tag{3.16}$$

When $\theta = 0$, we have

$$u_n = 2\bigg[1 - \frac{2a^{\frac{1}{N}}}{1+a^{\frac{1}{N}}} + (-1)^{N+1-n}\frac{2a^{\frac{N+2-n}{N}}}{1+a^{\frac{1}{N}}}\bigg]. \tag{3.17}$$

Since $a < 1$, we have

$$0 \le 1 - \frac{2a^{\frac{1}{N}}}{1+a^{\frac{1}{N}}} < 1 - a^{\frac{1}{N}} \longrightarrow 0, \qquad \text{as } N \uparrow \infty.$$

On the other hand, we have

$$\frac{2a^{\frac{N+2-n}{N}}}{1+a^{\frac{1}{N}}} \ge a^{\frac{N+2-n}{N}} \ge a^{\frac{N+1}{N}} \longrightarrow a, \qquad \text{as } N \uparrow \infty.$$

Therefore, the signs of u_n will alternate as soon as N is large enough to have $1-a^{\frac{1}{N}} < a^{\frac{N+1}{N}}$. This proves part (a). As for part (b), since the expression (3.16) is continuous in κ, the signs of u_n will still alternate if, for fixed $N \ge N_0$, we take κ slightly larger than $1/2$. (Note however that the term $(1-\kappa)^N/\kappa^N$ tends to zero faster than $1 - a^{\frac{1}{N}}$, so we cannot get this result uniformly in N). □

We refer to the right-hand panel of Figure 3.1 for an illustration of the oscillations of the vector $\boldsymbol{w}$. As shown in the left-hand panel of the same figure, similar oscillations occur for the vector $\boldsymbol{v}$. The mathematical analysis for $\boldsymbol{v}$, however, is much harder than for $\boldsymbol{w}$, and at this time we are not able to prove a result that could be an analogue of Proposition 3.1.12 for the vector $\boldsymbol{v}$.

Remark 3.1.13. Alfonsi et al. [2012] discover similar oscillations for the trade execution strategies of a *single* trader under transient price impact when price impact does not decay as a convex function of time. These oscillations, however, result from an attempt to exploit the delay in market response to a large trade, and they disappear when price impact decays as a convex function of time, see Alfonsi et al. [2012, Theorem 1]. In particular, when there is just one agent active in our market impact model then for each time grid $\mathbb{T}$ there exists a unique optimal strategy, which consists either of all buy trades or of all sell trades. As a matter of fact, there is an explicit formula for this strategy when $\theta = 0$; see Alfonsi et al. [2008] and Figure 3.2.

We can now turn to present one of the main results in this chapter. It is concerned with the oscillations of both $\boldsymbol{v}$ and $\boldsymbol{w}$ when the parameter θ increases. Intuitively it is clear that increased transaction costs will penalize trading oscillations and thus lead to a smoothing of the equilibrium strategies. The following theorem shows that there exists a

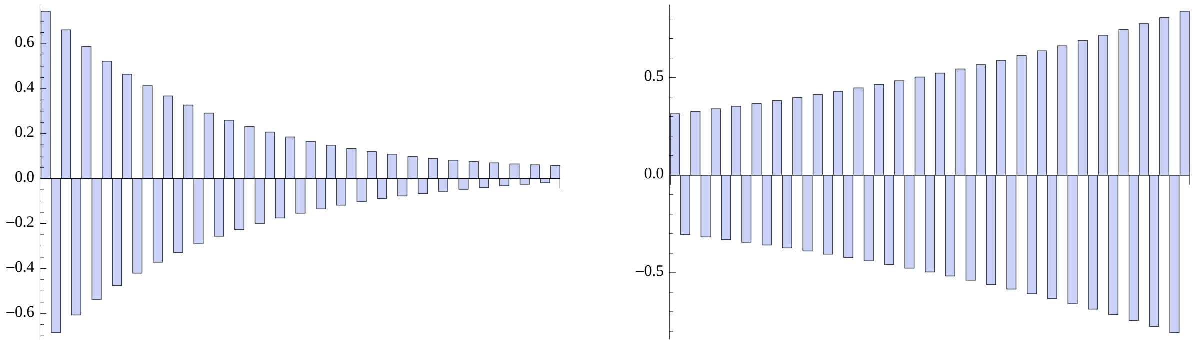


Figure 3.1: Vectors $\boldsymbol{v}$ (left) and $\boldsymbol{w}$ (right) for the equidistant time grid $\mathbb{T}_{50}$ and parameters $\lambda = \rho = 1$ and $\theta = 0$. By (3.6), $(\boldsymbol{v}, \boldsymbol{v})$ is the equilibrium for $X_0 = Y_0 = 1$, and $(\boldsymbol{w}, -\boldsymbol{w})$ is the equilibrium for $X_0 = -Y_0 = 1$. Yet, some individual components of both $\boldsymbol{v}$ and $\boldsymbol{w}$ exceed in either direction 60% of the sizes of the initial positions X_0 and Y_0.

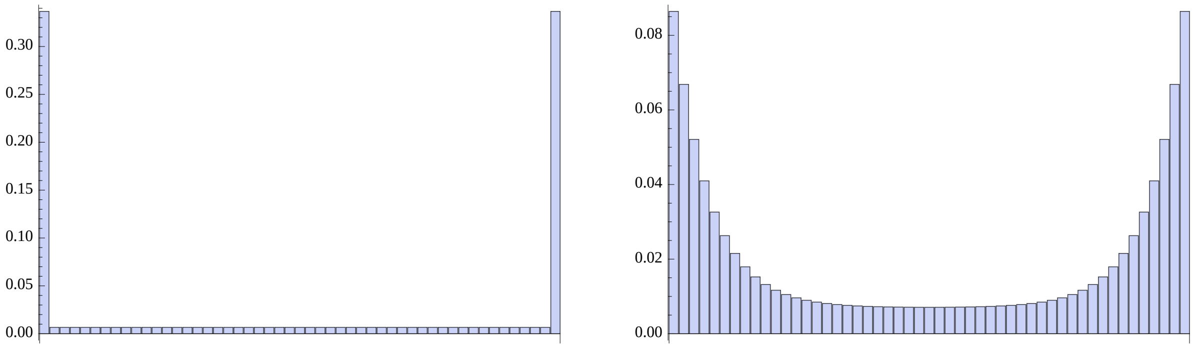


Figure 3.2: Optimal strategies with parameters $\theta = 0$ (left) and $\theta = 0.25$ (right) for the case in which there is just one agent active in the market. We have used the equidistant time grid $\mathbb{T}_{50}$, the initial condition $Z_0 = 1$, and parameters $\lambda = \rho = 1$ and $\theta = 0$.

critical value θ^* at which all oscillations of $\boldsymbol{v}$ and $\boldsymbol{w}$ disappear but below which oscillations are present. That is, for $\theta \geq \theta^*$ all equilibrium strategies for $X_0 = Y_0$ or $X_0 = -Y_0$ consist exclusive of all buy trades or of all sell trades. For $\theta < \theta^*$, the corresponding equilibrium strategies will contain both buy and sell trades as soon as resilience between two trades is sufficiently small. We can even determine the precise critical value θ^* at which the transition between oscillation and monotonicity occurs: it is given by $\theta^* = \lambda/4$.

Theorem 3.1.14. *Suppose that λ, $T > 0$ are fixed and $\mathbb{T}_N$ is as in* (3.14). *Then the following conditions are equivalent.*

(a) For every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{v}$ are nonnegative.

(b) For every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{w}$ are nonnegative.

(c) $\theta \geq \theta^* = \lambda/4$.

Proof of (c)$\Leftrightarrow$(b) *in Theorem 3.1.14.* Recall from the proof of Proposition 3.1.12 the notations $a := e^{-\rho T}$ and $\kappa = 2\theta/\lambda + \frac{1}{2}$ and the definition of the vector $\boldsymbol{u} = (u_1, \ldots, u_{N+1})$, which has only nonnegative entries if and only if $\boldsymbol{w}$ has only nonnegative entries. Our

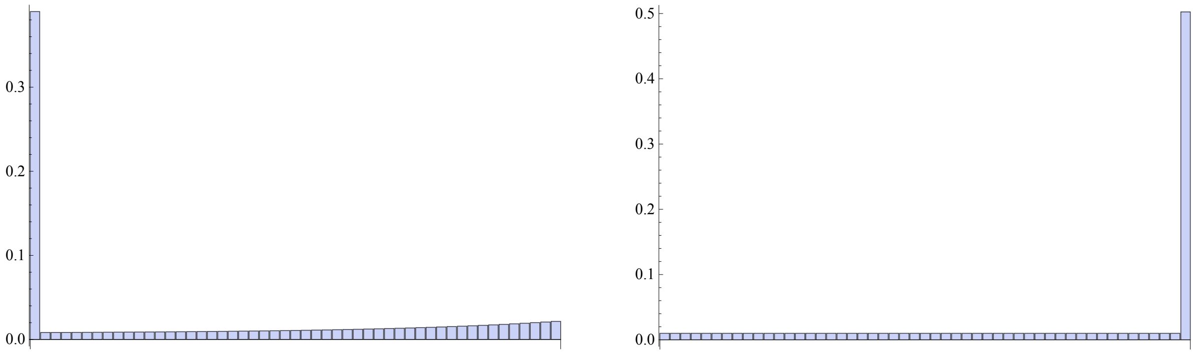


Figure 3.3: Vectors $\boldsymbol{v}$ (left) and $\boldsymbol{w}$ (right) for the equidistant time grid $\mathbb{T}_{50}$ and parameters $\lambda = \rho = 1$ and $\theta = 0.25$. In contrast to the equilibrium strategies in Figure 3.1, these strategies consist of only buy trades.

condition (c) is equivalent to $\kappa - 1 \geq 0$. Therefore, when (c) holds, we get with (3.16) that

$$u_n = \frac{1}{\kappa} - \frac{a^{\frac{1}{N}}}{\kappa^2} \sum_{m=0}^{N-n} \Big(\frac{a^{\frac{1}{N}}(\kappa - 1)}{\kappa} \Big)^m \geq \frac{1}{\kappa} - \frac{1}{\kappa^2} \sum_{m=0}^{\infty} \Big(\frac{\kappa - 1}{\kappa} \Big)^m = 0.$$

This establishes (c)⇒(b).

For the proof of the converse implication, we observe that $\kappa < 1$ and $\lim_{N \uparrow +\infty} a^{\frac{1}{N}} = 1$ imply that

$$u_N = \frac{1}{\kappa} - \frac{a^{\frac{1}{N}}}{\kappa^2}$$

must be strictly negative for sufficiently large N or sufficiently small ρ. This shows that we cannot have (b) without (c). □

Proof of (a)⇒(c) *in Theorem 3.1.14.* We consider the case $N = 2$ and show that $v_2 < 0$ for $\theta < \lambda/4$ when ρ is sufficiently small. Setting $a := e^{-\rho T}$ and $\tau := 3/2 + 2\theta/\lambda$, we have

$$M := \frac{1}{\lambda}\big(\lambda G + \lambda \widetilde{G} + 2\theta \,\mathrm{Id}\big) = \begin{pmatrix} \tau & \sqrt{a} & a \\ 2\sqrt{a} & \tau & \sqrt{a} \\ 2a & 2\sqrt{a} & \tau \end{pmatrix}.$$

A straightforward calculation verifies that

$$M^{-1} = \frac{1}{\det M} \begin{pmatrix} \tau^2 - 2a & \sqrt{a}(2a - \tau) & a - a\tau \\ 2\sqrt{a}(a - \tau) & \tau^2 - 2a^2 & \sqrt{a}(2a - \tau) \\ -2a(\tau - 2) & 2\sqrt{a}(a - \tau) & \tau^2 - 2a \end{pmatrix},$$

where

$$\det M = \tau^3 - (4a + 2a^2)\tau + 6a^2. \tag{3.18}$$

Therefore, the second component of the vector $(\lambda G + \lambda \widetilde{G} + 2\theta \,\mathrm{Id})^{-1}\mathbf{1}$ is given by

$$\frac{1}{\lambda \det M}\big(\tau^2 - 3\tau + 2\big). \tag{3.19}$$

It is easy to see from (3.18) that $\det M$ is increasing in τ for $\tau \geq 3/2$. Hence,

$$\det M \geq \Big(\frac{3}{2}\Big)^3 - 6a + 3a^2 \geq \Big(\frac{3}{2}\Big)^3 - 3 = 0.375 > 0.$$

One therefore sees that (3.19) is strictly negative for all $\tau \in [3/2, 2)$, which shows that we cannot have $v_2 \geq 0$ when $\theta < \lambda/4$. □

We will now prepare for the proof of the implication (c)⇒(a) in Theorem 3.1.14. This proof relies results on so-called M-matrices given in the book Berman and Plemmons [1994]. We first introduce some notations. When A is a matrix or vector, we will write

1. $A \geq 0$ if each entry of A is nonnegative;
2. $A > 0$ if $A \geq 0$ and at least one entry is strictly positive;
3. $A \gg 0$ if each entry of A is strictly positive.

Definition 3.1.15 (Definition 1.2 in Chapter 6 of Berman and Plemmons [1994])**.** A matrix $A \in \mathbb{R}^{n\times n}$ is called a nonsingular *M-matrix* if it is of the form $A = s\,\mathrm{Id} - B$, where the matrix $B \in \mathbb{R}^{n\times n}$ satisfies $B \geq 0$ and the parameter $s > 0$ is strictly larger than the spectral radius of B.

Also recall that a matrix $A \in \mathbb{R}^{n\times n}$ is called a *Z-matrix* if all its off-diagonal elements are nonpositive. Berman and Plemmons [1994] give 50 equivalent characterizations of the fact that a given Z-matrix is a nonsingular M-matrix. We will need two of them here and summarize them in the following statement.

Theorem 3.1.16 (From Theorem 2.3 in Chapter 6 of Berman and Plemmons [1994])**.** *For a Z-matrix $A \in \mathbb{R}^{n\times n}$, the following conditions are equivalent.*

1. *A is a nonsingular M-matrix.*
2. *All the leading principal minors of A are positive.*
3. *A is inverse-positive; that is, A^{-1} exists and $A^{-1} \geq 0$.*
4. *$A + \alpha\,\mathrm{Id}$ is nonsingular for all $\alpha \geq 0$.*

Proof of (c)⇒(a) *in Theorem 3.1.14.* In the first step, we show that the matrix

$$M := (\lambda G)^{-1}(2\theta\,\mathrm{Id} + \lambda\widetilde{G})$$

is a nonsingular M-matrix when $\theta \geq \frac{\lambda}{4}$. To this end, will also once again use the shorthand notations $a := e^{-\rho T}$ and $\kappa = 2\theta/\lambda + \frac{1}{2}$ and we will use the explicit form of G^{-1}, which has been derived in Alfonsi et al. [2008, Theorem 3.4]:

$$G^{-1} = \frac{1}{1-a^{\frac{2}{N}}}\begin{pmatrix} 1 & -a^{\frac{1}{N}} & 0 & \cdots & \cdots & 0 \\ -a^{\frac{1}{N}} & 1+a^{\frac{2}{N}} & -a^{\frac{1}{N}} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a^{\frac{1}{N}} & 1+a^{\frac{2}{N}} & -a^{\frac{1}{N}} \\ 0 & \cdots & \cdots & 0 & -a^{\frac{1}{N}} & 1 \end{pmatrix}.$$

Thus, $M = (\lambda G)^{-1}(2\theta\,\mathrm{Id} + \lambda\widetilde{G})$ is equal to

$$G^{-1}\cdot\begin{pmatrix} \kappa & 0 & \cdots & \cdots & \cdots & 0 \\ a^{\frac{1}{N}} & \kappa & 0 & \cdots & \cdots & 0 \\ a^{\frac{2}{N}} & a^{\frac{1}{N}} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 & \vdots \\ a^{\frac{N-1}{N}} & a^{\frac{N-2}{N}} & \ddots & \ddots & \kappa & 0 \\ a^{\frac{N}{N}} & a^{\frac{N-1}{N}} & \cdots & \cdots & a^{\frac{1}{N}} & \kappa \end{pmatrix}$$

$$= \frac{1}{1-a^{\frac{2}{N}}}\begin{pmatrix} \kappa - a^{\frac{2}{N}} & -\kappa a^{\frac{1}{N}} & 0 & \cdots & \cdots & 0 \\ -a^{\frac{1}{N}}(\kappa-1) & a^{\frac{2}{N}}(\kappa-1)+\kappa & -\kappa a^{\frac{1}{N}} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & -a^{\frac{1}{N}}(\kappa-1) & a^{\frac{2}{N}}(\kappa-1)+\kappa & -\kappa a^{\frac{1}{N}} \\ 0 & \cdots & \cdots & 0 & -a^{\frac{1}{N}}(\kappa-1) & \kappa \end{pmatrix}.$$

Our condition (c) is equivalent to $\kappa \geq 1$, which by the preceding matrix identity is in turn equivalent to the fact that M is a Z-matrix. By Theorem 3.1.16, M will thus be an M-matrix if and only if all its leading principal minors $M_{[n]}$, $n \in \{1, 2, \cdots, N+1\}$, are positive. Using the following identity for $M_{[n]}$, which will be shown below in Lemma 3.1.18, we obtain that for $n \in \{1, 2, \ldots, N\}$

$$\begin{aligned} M_{[n]} &= \frac{1}{(1-a^{\frac{2}{N}})^n}\Big(\kappa^n - \sum_{k=0}^{n-1}\kappa^{n-1-k}(\kappa-1)^k a^{\frac{2(k+1)}{N}}\Big) \\ &> \frac{1}{(1-a^{\frac{2}{N}})^n}\Big(\kappa^n - \sum_{k=0}^{n-1}\kappa^{n-1-k}(\kappa-1)^k\Big) \\ &= \frac{(\kappa-1)^n}{(1-a^{\frac{2}{N}})^n} \\ &\geq 0. \end{aligned}$$

And $M_{[N+1]} = \frac{\kappa^{N+1}}{(1-a^{\frac{2}{N}})^N} > 0$. Therefore, $M = (\lambda G)^{-1}(2\theta\,\mathrm{Id} + \lambda\widetilde{G})$ is a nonsingular M-matrix. It can hence be written as $s\,\mathrm{Id} - B$, where B is non-negative and s is strictly larger than the spectral radius, $\rho(B)$, of B. It follows that the matrix

$$\mathrm{Id} + (\lambda G)^{-1}(2\theta\,\mathrm{Id} + \lambda\widetilde{G}) = (s+1)\,\mathrm{Id} - B$$

is also a nonsingular M-matrix. By Theorem 3.1.16, every nonsingular M-matrix is inverse-positive. Since

$$(\lambda G)^{-1}\mathbf{1} = \frac{1}{\lambda(1+a^{\frac{1}{N}})}\Big(1, 1-a^{\frac{1}{N}}, \ldots, 1-a^{\frac{1}{N}}, 1\Big)^T \gg 0,$$

we thus have

$$\big(\lambda G + 2\theta\,\mathrm{Id} + \lambda\widetilde{G}\big)^{-1}\mathbf{1} = \Big(\mathrm{Id} + (\lambda G)^{-1}(2\theta\,\mathrm{Id} + \lambda\widetilde{G})\Big)^{-1}(\lambda G)^{-1}\mathbf{1} \geq 0.$$

This concludes the proof. $\square$

Lemma 3.1.17. *For $n \geq 3$, let T_n be the $n \times n$-tridiagonal matrix*

$$T_n = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & \cdots & 0 \\ c_1 & a_2 & b_2 & 0 & \ddots & 0 \\ 0 & c_2 & a_3 & b_3 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & c_{n-2} & a_{n-1} & b_{n-1} \\ 0 & \cdots & \cdots & 0 & c_{n-1} & a_n \end{pmatrix}.$$

Its determinant can be computed recursively as follows

$$\det T_n = a_n \det T_{n-1} - b_{n-1}c_{n-1} \det T_{n-2}.$$

Proof. We develop the determinant first via the cofactors of the last row and then by the cofactors of the last column to get

$$\begin{aligned} \det T_n = (-1)^{2n-1} c_{n-1} & \begin{vmatrix} a_1 & b_1 & 0 & \cdots & \cdots & 0 \\ c_1 & a_2 & b_2 & 0 & \ddots & 0 \\ 0 & c_2 & a_3 & b_3 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & c_{n-3} & a_{n-2} & 0 \\ 0 & \cdots & \cdots & 0 & c_{n-2} & b_{n-1} \end{vmatrix} \\ + (-1)^{2n} a_n & \begin{vmatrix} a_1 & b_1 & 0 & \cdots & \cdots & 0 \\ c_1 & a_2 & b_2 & 0 & \ddots & 0 \\ 0 & c_2 & a_3 & b_3 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & c_{n-3} & a_{n-2} & b_{n-2} \\ 0 & \cdots & \cdots & 0 & c_{n-2} & a_{n-1} \end{vmatrix} \\ &= (-1)^{2n-1}(-1)^{2(n-1)} b_{n-1}c_{n-1} \det T_{n-2} + a_n \det T_{n-1} \\ &= a_n \det T_{n-1} - b_{n-1}c_{n-1} \det T_{n-2}. \end{aligned}$$

□

Lemma 3.1.18. *The leading principal minors of the matrix $M = (\lambda G)^{-1}(2\theta \operatorname{Id} + \lambda \widetilde{G})$ are given by*

$$M_{[n]} = \begin{cases} \dfrac{1}{(1 - a^{\frac{2}{N}})^n} \Big(\kappa^n - \displaystyle\sum_{k=0}^{n-1} \kappa^{n-1-k} (\kappa - 1)^k a^{\frac{2(k+1)}{N}} \Big), & \text{for } n \in \{1, 2, \ldots, N\}, \\ \dfrac{\kappa^{N+1}}{(1 - a^{\frac{2}{N}})^N}, & \text{for } n = N + 1. \end{cases} \tag{3.20}$$

Proof. We proceed by induction on n. First, we clearly have

$$
\begin{aligned}
M_{[1]} &= \frac{\kappa - a^{\frac{2}{N}}}{1 - a^{\frac{2}{N}}}, \\
M_{[2]} &= \frac{1}{(1 - a^{\frac{2}{N}})^2}\Big((\kappa - a^{\frac{2}{N}})\big(a^{\frac{2}{N}}(\kappa - 1) + \kappa\big) - \kappa a^{\frac{2}{N}}(\kappa - 1)\Big) \\
&= \frac{1}{(1 - a^{\frac{2}{N}})^2}\Big(\kappa^2 - \kappa a^{\frac{2}{N}} - (\kappa - 1)a^{\frac{4}{N}}\Big), \\
M_{[3]} &= \frac{1}{(1 - a^{\frac{2}{N}})^3}\Big(\kappa^3 - \kappa^2 a^{\frac{6}{N}} - \kappa^2 a^{\frac{4}{N}} - \kappa^2 a^{\frac{2}{N}} + 2a^{\frac{6}{N}}\kappa + \kappa a^{\frac{4}{N}} - a^{\frac{6}{N}}\Big) \\
&= \frac{1}{(1 - a^{\frac{2}{N}})^3}\Big(\kappa^3 - \kappa^2 a^{\frac{2}{N}} - \kappa(\kappa - 1)a^{\frac{4}{N}} - (\kappa - 1)^2 a^{\frac{6}{N}}\Big).
\end{aligned}
$$

These formulas coincide with (3.20) for $n = 1, 2, 3$.

Now we assume that for some $3 \le n \le N - 1$,

$$
M_{[m]} = \frac{1}{(1 - a^{\frac{2}{N}})^m}\Big(\kappa^m - \sum_{k=0}^{m-1} \kappa^{m-1-k}(\kappa - 1)^k a^{\frac{2(k+1)}{N}}\Big), \qquad \text{where } m = n - 1, n.
$$

By Lemma (3.1.17), $M_{[n+1]}$ can be computed via $M_{[n]}$ and $M_{[n-1]}$:

$$
\begin{aligned}
M_{[n+1]} &= \frac{a^{\frac{2}{N}}(\kappa - 1) + \kappa}{1 - a^{\frac{2}{N}}} M_{[n]} - \frac{\kappa a^{\frac{2}{N}}(\kappa - 1)}{(1 - a^{\frac{2}{N}})^2} M_{[n-1]} \\
&= \frac{a^{\frac{2}{N}}(\kappa - 1) + \kappa}{(1 - a^{\frac{2}{N}})^{n+1}}\Big(\kappa^n - \sum_{k=0}^{n-1} \kappa^{n-1-k}(\kappa - 1)^k a^{\frac{2(k+1)}{N}}\Big) \\
&\quad - \frac{\kappa a^{\frac{2}{N}}(\kappa - 1)}{(1 - a^{\frac{2}{N}})^{n+1}}\Big(\kappa^{n-1} - \sum_{k=0}^{n-2} \kappa^{n-2-k}(\kappa - 1)^k a^{\frac{2(k+1)}{N}}\Big) \\
&= \frac{1}{(1 - a^{\frac{2}{N}})^{n+1}}\Bigg(\kappa(\kappa - 1)a^{\frac{2}{N}} + \kappa^{n+1} - \sum_{k=0}^{n-1} \kappa^{n-1-k}(\kappa - 1)^{k+1} a^{\frac{2(k+2)}{N}} \\
&\quad - \sum_{k=0}^{n-1} \kappa^{n-k}(\kappa - 1)^k a^{\frac{2(k+1)}{N}} - \kappa^n a^{\frac{2}{N}}(\kappa - 1) + \sum_{k=0}^{n-2} \kappa^{n-1-k}(\kappa - 1)^{k+1} a^{\frac{2(k+2)}{N}}\Bigg) \\
&= \kappa^{n+1} - (\kappa - 1)^n a^{\frac{2(n+1)}{N}} - \sum_{k=0}^{n-1} \kappa^{n-k}(\kappa - 1)^k a^{\frac{2(k+1)}{N}} \\
&= \kappa^{n+1} - \sum_{k=0}^{n} \kappa^{n-k}(\kappa - 1)^k a^{\frac{2(k+1)}{N}},
\end{aligned}
$$

which also coincides with (3.20).

In the last step, we compute $M_{[N+1]}$. To this end, we use once again Lemma 3.1.17:

$$
\begin{aligned}
M_{[N+1]} &= \frac{\kappa}{1-a^{\frac{2}{N}}} M_{[N]} - \frac{\kappa a^{\frac{2}{N}}(\kappa-1)}{(1-a^{\frac{2}{N}})^2} M_{[N-1]} \\
&= \frac{\kappa}{(1-a^{\frac{2}{N}})^{N+1}} \Big(\kappa^N - \sum_{k=0}^{N-1} \kappa^{N-1-k}(\kappa-1)^k a^{\frac{2(k+1)}{N}} \Big) \\
&\qquad - \frac{\kappa a^{\frac{2}{N}}(\kappa-1)}{(1-a^{\frac{2}{N}})^{n+1}} \Big(\kappa^{N-1} - \sum_{k=0}^{N-2} \kappa^{N-2-k}(\kappa-1)^k a^{\frac{2(k+1)}{N}} \Big) \\
&= \frac{\kappa}{(1-a^{\frac{2}{N}})^{N+1}} \Big(\kappa^{N+1} - \sum_{k=0}^{N-1} \kappa^{N-k}(\kappa-1)^k a^{\frac{2(k+1)}{N}} \\
&\qquad - \kappa^N a^{\frac{2}{N}}(\kappa-1) + \sum_{k=0}^{N-2} \kappa^{N-1-k}(\kappa-1)^{k+1} a^{\frac{2(k+2)}{N}} \Big) \\
&= \frac{\kappa}{(1-a^{\frac{2}{N}})^{N+1}} \Big(\kappa^{N+1} - \sum_{k=0}^{N-1} \kappa^{N-k}(\kappa-1)^k a^{\frac{2(k+1)}{N}} \\
&\qquad - \kappa^N a^{\frac{2}{N}}(\kappa-1) + \sum_{k=0}^{N-1} \kappa^{N-k}(\kappa-1)^k a^{\frac{2(k+1)}{N}} \Big) \\
&= \frac{\kappa}{(1-a^{\frac{2}{N}})^{N+1}} \Big(\kappa^{N+1} - \kappa^N a^{\frac{2}{N}} - \kappa^{N+1} a^{\frac{2}{N}} + \kappa^N a^{\frac{2}{N}} \Big) \\
&= \frac{\kappa^{N+1}}{(1-a^{\frac{2}{N}})^N}.
\end{aligned}
$$

This completes the proof. □

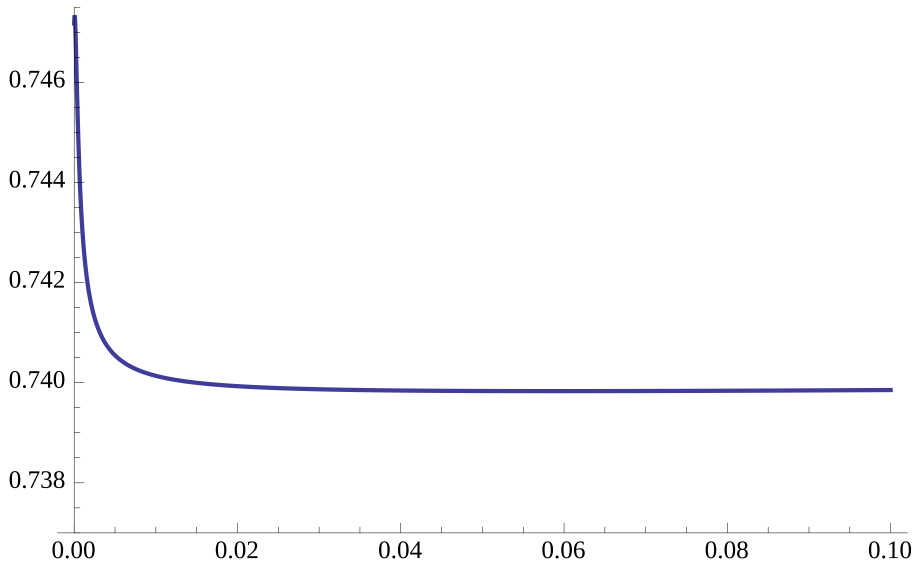


Figure 3.4: Expected costs $\mathbb{E}[\mathcal{C}_{\mathbb{T}_{1000}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)] = \mathbb{E}[\mathcal{C}_{\mathbb{T}_{1000}}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)]$ as a function of θ. The costs decrease steeply from the value of 0.7472 for $\theta = 0$ to 0.7398 for $\theta = 0.06$. From then on there is a moderate and almost linear increase with, e.g., a value of 0.7403 at $\theta = 0.5$. We take the equidistant time grid $\mathbb{T}_{1000}$, initial values $X_0 = Y_0 = 1$, and $\lambda = \rho = 1$.

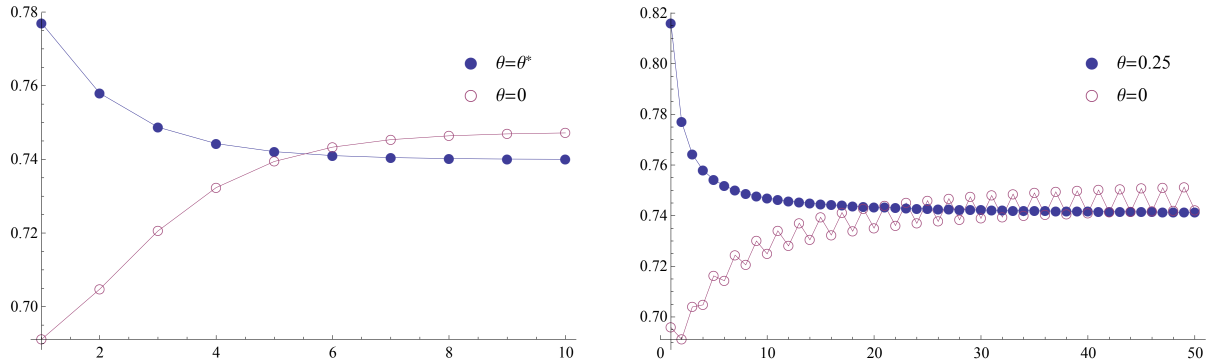


Figure 3.5: Expected costs $\mathbb{E}[\,\mathcal{C}_{\mathbb{T}_{2^N}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,] = \mathbb{E}[\,\mathcal{C}_{\mathbb{T}_{2^N}}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)\,]$ as a function of N with the equidistant time grids $\mathbb{T}_{2^N}$ (left) and $\mathbb{T}_N$ (right) and parameters $\lambda = \rho = 1$, $X_0 = Y_0 = 1$.

Due to our explicit formulas (3.5) and (3.6), it is easy to analyze the Nash equilibrium numerically. These numerical simulations exhibit a striking effect: it is possible that the individual costs for each agent are *lower* in a model with *nonzero* transaction costs than in the corresponding model without transaction costs. Figure 3.4 shows the expected costs $\mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,] = \mathbb{E}[\,\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)\,]$ for $X_0 = Y_0$ as a function of θ. When θ increases from the value $\theta = 0$, these costs start to decline steeply toward a minimum, after which there is a slow and steady increase. Even more interesting is the behavior illustrated in Figure 3.5, where the same expected costs are plotted for $\theta = 0$ and $\theta = \theta^*$ as a function of N. For $\theta = \theta^*$ the costs are *decreasing in* N, whereas they are *increasing* for $\theta = 0$ (modulo a sawtooth pattern corresponding to odd and even values of N). In particular the observation that the expected costs are increasing for $\theta = 0$ is very surprising, because a higher trading frequency suggests greater flexibility and the possibility to use more efficient trading strategies.

So why are the costs then increasing in N? The answer is that a higher trading frequency increases also the possibility for the competitor to conduct predatory strategies at the expense of the other agent. In reaction, the other needs to take stronger protective measures against predatory trading. The hot-potato game, thus, is the result of the need of protection against predatory trading by competitors. When transaction costs increase, predatory trading becomes less profitable, and so both agents need less protection against possible assaults by their competitor. As a result, both can choose more efficient strategies than before and thus benefit from an increase in trading frequency. Therefore, the expected costs are decreasing in N when $\theta = \theta^*$. The same effect is responsible for the steep initial decrease of the expected costs as a function of θ, which is illustrated in Figure 3.4.

3.1.4 Analysis of the high-frequency limit

In this subsection, we analyze the possible convergence of the equilibrium strategies when the trading frequency tends to infinity. To this end, we consider the equidistant time grids $\mathbb{T}_N$ as defined in (3.14) for varying $N \in \mathbb{N}$ and write $\boldsymbol{v}^{(N)} = (v_1^{(N)}, \dots, v_{N+1}^{(N)})$ and $\boldsymbol{w}^{(N)} = (w_1^{(N)}, \dots, w_{N+1}^{(N)})$ for the vectors in (3.5) to make the dependence on N explicit. We start with the following proposition, which analyzes the convergence of the individual components of $\boldsymbol{w}^{(N)}$ when $N \uparrow \infty$. By (3.6), a Nash equilibrium with $X_0 = -Y_0$ is completely determined by $\boldsymbol{w}^{(N)}$.

Proposition 3.1.19. *Suppose that n is fixed.*

1. *When $\theta = 0$, we have*

$$\lim_{N\uparrow\infty} w_n^{(2N)} = (-1)^{n+1}\frac{2a}{2\rho T + a + 1} \quad \textit{and} \quad \lim_{N\uparrow\infty} w_n^{(2N+1)} = (-1)^n \frac{2a}{2\rho T - a + 1}, \tag{3.21}$$

as well as

$$\lim_{N\uparrow\infty} w_{2N+1-n}^{(2N)} = (-1)^n\frac{2}{2\rho T + a + 1} \quad \textit{and} \quad \lim_{N\uparrow\infty} w_{2N+2-n}^{(2N+1)} = (-1)^n \frac{2}{2\rho T - a + 1}. \tag{3.22}$$

2. *When $\theta > 0$, we have*

$$\lim_{N\uparrow\infty} w_n^{(N)} = 0,$$

and

$$\lim_{N\uparrow\infty} w_{N+1-n}^{(N)} = \Big(\frac{4\theta-\lambda}{4\theta+\lambda}\Big)^n \frac{2\lambda}{(\rho T+1)(4\theta+\lambda)}. \tag{3.23}$$

Lemma 3.1.20. *Let Γ_N be as in* (3.15) *and let us denote by $\boldsymbol{u}^{(N)} = (u_1^{(N)}, u_2^{(N)}, \dots, u_{N+1}^{(N)}) \in \mathbb{R}^{N+1}$ the vector $\lambda\Gamma_N\mathbf{1}$. When $n \in \{1,\dots,N+1\}$, then*

$$\sum_{m=1}^n u_m^{(N)} = \frac{1}{\kappa}\Bigg[n\Bigg(1-\frac{a^{\frac1N}}{\kappa(1-a^{\frac1N})+a^{\frac1N}}\Bigg)+\frac{a^{\frac1N}}{\kappa(1-a^{\frac1N})+a^{\frac1N}}\Big(\frac{a^{\frac1N}(\kappa-1)}{\kappa}\Big)^{N+1-n}\frac{\big(\frac{a^{\frac1N}(\kappa-1)}{\kappa}\big)^n-1}{\frac{a^{\frac1N}(\kappa-1)}{\kappa}-1}\Bigg].$$

Proof. The assertion follows from (3.16) by noting that

$$\sum_{m=1}^n \Big(\frac{a^{\frac1N}(\kappa-1)}{\kappa}\Big)^{N+1-m} = \Big(\frac{a^{\frac1N}(\kappa-1)}{\kappa}\Big)^{N+1-n}\frac{\big(\frac{a^{\frac1N}(\kappa-1)}{\kappa}\big)^n-1}{\frac{a^{\frac1N}(\kappa-1)}{\kappa}-1}.$$

□

Proof of Proposition 3.1.19. Let Γ_N and $\boldsymbol{u}^{(N)}$ be as in Lemma 3.1.20. We need to normalize the vector $\boldsymbol{u}^{(N)}$ with $\mathbf{1}^\top\lambda\Gamma_N\mathbf{1} = \mathbf{1}^\top\boldsymbol{u}^{(N)}$ to get $\boldsymbol{w}^{(N)}$. Taking $n = N+1$ in Lemma 3.1.20 yields

$$\begin{aligned}\mathbf{1}^\top\lambda\Gamma_N\mathbf{1} &= \sum_{n=1}^{N+1} u_n^{(N)}\\ &= \frac{1}{\kappa}\Bigg[(N+1)\Bigg(1-\frac{a^{\frac1N}}{\kappa(1-a^{\frac1N})+a^{\frac1N}}\Bigg)+\frac{a^{\frac1N}}{\kappa(1-a^{\frac1N})+a^{\frac1N}}\cdot\frac{\big(\frac{a^{\frac1N}(\kappa-1)}{\kappa}\big)^{N+1}-1}{\frac{a^{\frac1N}(\kappa-1)}{\kappa}-1}\Bigg].\end{aligned}$$

Note that

$$(N+1)\left(1-\frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}})+a^{\frac{1}{N}}}\right)\longrightarrow -\kappa\log a=\kappa\rho T \qquad \text{as } N\uparrow\infty. \tag{3.24}$$

Moreover, $\kappa\geq 1/2$ implies that $|\kappa-1|/\kappa\leq 1$ with equality if and only if $\kappa=1/2$. We therefore get that for $\kappa=1/2$, which is the same as $\theta=0$,

$$\begin{aligned}\lim_{N\uparrow\infty}\mathbf{1}^\top\lambda\Gamma_{2N}\mathbf{1}&=\rho T+\frac{1+a}{2\kappa}=\rho T+a+1,\\ \lim_{N\uparrow\infty}\mathbf{1}^\top\lambda\Gamma_{2N+1}\mathbf{1}&=\rho T+\frac{1-a}{2\kappa}=\rho T-a+1.\end{aligned} \tag{3.25}$$

For $\kappa>1/2$, we have

$$\lim_{N\uparrow\infty}\mathbf{1}^\top\lambda\Gamma_N\mathbf{1}=\rho T+1. \tag{3.26}$$

The assertions now follow easily by taking limits in (3.17) and (3.16). □

The preceding proposition gives further background on the oscillations of equilibrium strategies in the regime $\theta<\lambda/4$. In particular, (3.21) shows that in a Nash equilibrium with $X_0=-Y_0$ and with $\theta=0$ the trades of both agents asymptotically oscillate between $\pm$*const* and that the sign of each trade also depends on whether N is odd or even. Moreover, (3.23) implies that, for $K\in\mathbb{N}$ fixed and $N\uparrow\infty$, the terminal K trades in an equilibrium strategy asymptotically oscillate between $\pm$*const* if and only if $\theta<\lambda/4$.

Now we consider the Nash equilibrium $(\boldsymbol{\xi}^{*,(N)},\boldsymbol{\eta}^{*,(N)})$ with initial positions X_0, Y_0 and time grid $\mathbb{T}_N$. We define the asset positions of the two agents via

$$X_t^{(N)}:=X_0-\sum_{k=1}^{\lceil\frac{Nt}{T}\rceil}\xi_k^{*,(N)},\qquad Y_t^{(N)}:=Y_0-\sum_{k=1}^{\lceil\frac{Nt}{T}\rceil}\eta_k^{*,(N)},\qquad t\geq 0. \tag{3.27}$$

In the case $X_0=-Y_0$, we have $X^{(N)}=-Y^{(N)}=X_0W^{(N)}$, where

$$W_t^{(N)}=1-\sum_{k=1}^{\lceil\frac{Nt}{T}\rceil}w_k^{(N)},\qquad t\geq 0. \tag{3.28}$$

Proposition 3.1.21. *When $\theta>0$, we have for $t<T$*

$$\lim_{N\uparrow\infty}W_t^{(N)}=\frac{\rho(T-t)+1}{\rho T+1} \tag{3.29}$$

and $W_t=0$ for $t>T$.

Proof. Let $n_t:=\lceil Nt/T\rceil$. Then, with the notation introduced in the proof of Proposition 3.1.19,

$$W_t^{(N)}=1-\frac{1}{\mathbf{1}^\top\lambda\Gamma_N\mathbf{1}}\sum_{k=1}^{n_t}u_k^{(N)}.$$

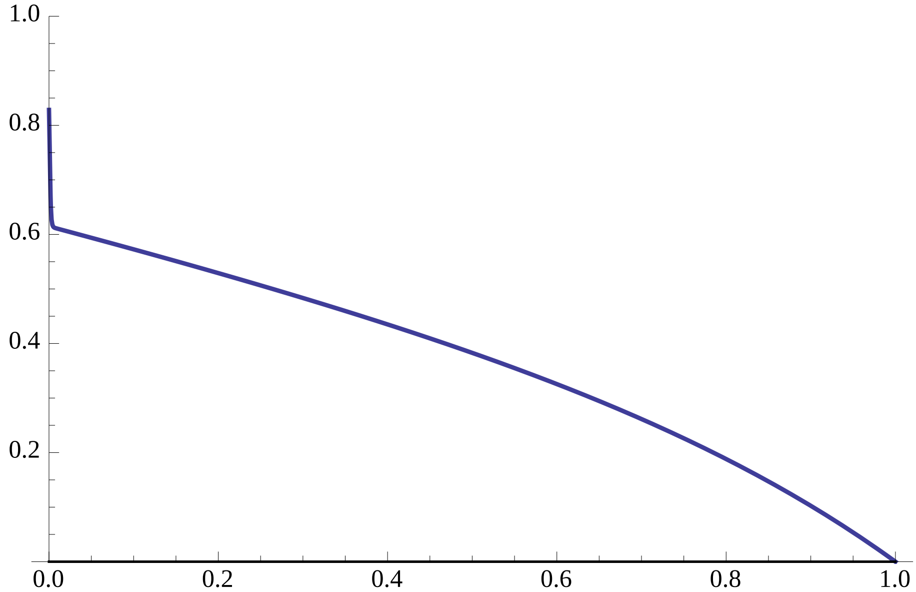


Figure 3.6: The function $V^{(N)}$ for $N = 1000$ with parameters $\theta = 0.5$ and $\lambda = \rho = 1$.

For $\theta > 0$ and $t < T$ it follows from Lemma 3.1.20 that

$$\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{N+1-n_t} \longrightarrow 0 \qquad \text{as } N \uparrow \infty.$$

Therefore, with (3.24),

$$\lim_{N\uparrow\infty} \sum_{k=1}^{n_t} u_k^{(N)} = \frac{1}{\kappa} \lim_{N\uparrow\infty} n_t \bigg(1 - \frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}}) + a^{\frac{1}{N}}}\bigg) = \rho t.$$

The assertion now follows with (3.26). □

Note that the limiting function in (3.29) is independent of θ as long as $\theta > 0$. Let

$$V_t^{(N)} = 1 - \sum_{k=1}^{\lceil \frac{Nt}{T} \rceil} \boldsymbol{v}_k^{(N)}, \qquad t \geq 0. \tag{3.30}$$

The asymptotic analysis for $V^{(N)}$ is much more difficult than for $W^{(N)}$. The numerical simulation in Figure 3.6 suggests that $V^{(N)}$ converges for $\theta \geq \lambda/4$ to a function $V^{(\infty)}$, which has a jump at $t = 0$ and is otherwise a nonlinear function of time.

3.2 Three extensions of the primary model

3.2.1 Incorporation of permanent impact

The first extension of the primary model is an incorporation of *permanent impact*. In comparison with transient impact, permanent impact affects the asset prices during the whole execution period and does not decay. It is accumulated by all transactions until the current transaction time. We assume that when the two financial agents X and Y apply respective strategies $\boldsymbol{\xi} \in \mathscr{X}(X_0, \mathbb{T})$ and $\boldsymbol{\eta} \in \mathscr{X}(Y_0, \mathbb{T})$, the asset prices are given by

$$S_t^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_t^0 - \lambda \sum_{t_k<t} e^{-\rho(t-t_k)}(\xi_k + \eta_k) - \gamma \sum_{t_k<t}(\xi_k + \eta_k). \tag{3.31}$$

As it has been described in Subsection 3.1.1, $S^0 = (S_t^0)_{t\geq 0}$ is a right-continuous martingale on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$, for which $\mathcal{F}_0$ is $\mathbb{P}$-trivial. In this subsection, we still assume that none of the two agents has an advantage in latency over the other. At each execution time, the execution priority is given to that agent who wins an independent coin toss if both agents place nonzero orders at that time. Note that this assumption affects not only the combined liquidity costs caused by transient impact but also the costs caused by permanent impact.

Suppose that just one agent, say X, places a nonzero order at time t_k, then the price is moved from $S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}$ to $S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - (\lambda+\gamma)\xi_k$ and the trade ξ_k incurs the following expenses:

$$\int_{S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}}^{S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}-(\lambda+\gamma)\xi_k} \frac{z}{\lambda+\gamma}\,dz = \frac{\lambda+\gamma}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k.$$

Now suppose that the order η_k of agent Y is executed immediately after the order ξ_k. Then the price is moved from $S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - (\lambda+\gamma)\xi_k$ to $S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - (\lambda+\gamma)(\xi_k+\eta_k)$, and agent Y incurs the expenses

$$\int_{S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}-(\lambda+\gamma)\xi_k}^{S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}-(\lambda+\gamma)(\xi_k+\eta_k)} \frac{z}{\lambda+\gamma}\,dz = \frac{\lambda+\gamma}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + (\lambda+\gamma)\xi_k\eta_k.$$

Therefore, each agent has the combined liquidation costs of $(\lambda+\gamma)\xi_k\eta_k/2$ at time t_k in expectation if they are equally split.

Furthermore, for each trade ξ_k and η_k the quadratic transaction costs of $\theta\xi_k^2$ and $\theta\eta_k^2$ are still imposed, where θ is nonnegative. Now we give the definition of the liquidation costs by each agent in this extended model.

Definition 3.2.1. Suppose that $\mathbb{T} = \{t_0, t_1, \ldots, t_N\}$, X_0 and Y_0 are given. Let furthermore $(\varepsilon_i)_{i=0,1,\ldots}$ be an i.i.d. sequence of Bernoulli $(\frac{1}{2})$-distributed random variables that are independent of $\sigma(\bigcup_{t\geq 0}\mathcal{F}_t)$. Then the *liquidation costs* of $\boldsymbol{\xi} \in \mathscr{X}(X_0, \mathbb{T})$ given $\boldsymbol{\eta} \in \mathscr{X}(Y_0, \mathbb{T})$ are defined as

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}) = X_0 S_0^0 + \sum_{k=0}^{N} \Big(\frac{\lambda+\gamma}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k + \varepsilon_k(\lambda+\gamma)\xi_k\eta_k + \theta\xi_k^2\Big), \tag{3.32}$$

and the *liquidation costs of* $\boldsymbol{\eta}$ *given* $\boldsymbol{\xi}$ are

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}) = Y_0 S_0^0 + \sum_{k=0}^{N} \Big(\frac{\lambda+\gamma}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + (1-\varepsilon_k)(\lambda+\gamma)\xi_k\eta_k + \theta\eta_k^2\Big). \tag{3.33}$$

We assume that both agents X and Y have full knowledge of the other's trading strategy and minimize the expected costs of their strategies accordingly. The optimality for both agents is described through a Nash equilibrium defined in Definition 3.1.5.

To simplify our notation, we define the $(N+1)\times(N+1)$-matrix H and its lower triangular matrix $\widetilde{H}$ for a fixed time grid $\mathbb{T}=\{t_0,\dots,t_N\}$ by

$$H_{ij}=1 \text{ for all } i,j; \qquad \widetilde{H}_{ij}=\begin{cases}1 & \text{if } i>j,\\ \frac{1}{2} & \text{if } i=j,\\ 0 & \text{otherwise.}\end{cases}$$

Lemma 3.2.2. *The expected liquidation costs of an admissible strategy $\boldsymbol{\xi}\in\mathcal{X}(X_0,\mathbb{T})$ given another admissible strategy $\boldsymbol{\eta}\in\mathcal{X}(Y_0,\mathbb{T})$ are*

$$\mathbb{E}\Big[\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\Big]=\mathbb{E}\Big[\frac{1}{2}\boldsymbol{\xi}^\top(\lambda G+2\theta\,\mathrm{Id})\boldsymbol{\xi}+\boldsymbol{\xi}^\top(\lambda\widetilde{G}+\gamma\widetilde{H})\boldsymbol{\eta}\Big]+\frac{\gamma}{2}X_0^2, \tag{3.34}$$

and the expected liquidation costs of $\boldsymbol{\eta}$ given $\boldsymbol{\xi}$ are

$$\mathbb{E}\Big[\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi})\Big]=\mathbb{E}\Big[\frac{1}{2}\boldsymbol{\eta}^\top(\lambda G+2\theta\,\mathrm{Id})\boldsymbol{\eta}+\boldsymbol{\eta}^\top(\lambda\widetilde{G}+\gamma\widetilde{H})\boldsymbol{\xi}\Big]+\frac{\gamma}{2}Y_0^2. \tag{3.35}$$

Proof. We refer to the proof of Lemma 3.2.11, which shows a more general case. □

Theorem 3.2.3. *Given $\rho>0$, $\lambda>0$ and $\theta>0$. For any time grid $\mathbb{T}$ and initial asset positions $X_0,Y_0\in\mathbb{R}$, there exists a unique Nash equilibrium $(\boldsymbol{\xi}^*,\boldsymbol{\eta}^*)\in\mathcal{X}(X_0,\mathbb{T})\times\mathcal{X}(Y_0,\mathbb{T})$. The optimal strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ are deterministic.*

Proof. We refer to the proof of Theorem 3.2.12, which shows a more general case. □

Remark 3.2.4. To derive the optimal strategies in Theorem 3.2.3 explicitly, one can use the method stated in Remark 3.1.11. By Lemma 3.2.2 and the method of Lagrange multipliers, a necessary condition for $(\boldsymbol{\xi}^*,\boldsymbol{\eta}^*)$ to be a Nash equilibrium in $\mathcal{X}_{\det}(X_0,\mathbb{T})\times\mathcal{X}_{\det}(Y_0,\mathbb{T})$ is the existence of $\alpha,\beta\in\mathbb{R}$, such that

$$\begin{cases}P\boldsymbol{\xi}^*+Q\boldsymbol{\eta}^*=\alpha\mathbf{1},\\ P\boldsymbol{\eta}^*+Q\boldsymbol{\xi}^*=\beta\mathbf{1},\\ \mathbf{1}^\top\boldsymbol{\xi}^*=X_0,\\ \mathbf{1}^\top\boldsymbol{\eta}^*=Y_0,\end{cases} \tag{3.36}$$

where

$$P=\lambda G+2\theta\,\mathrm{Id},\quad Q=\lambda\widetilde{G}+\gamma\widetilde{H}.$$

As the existence and uniqueness of optimal strategies are guaranteed by Theorem 3.2.3, the linear equation system(3.36) has a unique solution.

Moreover, if the matrices $P\pm Q$ are invertible and $\mathbf{1}^\top(P\pm Q)^{-1}\mathbf{1}\neq 0$, we define the following two vectors

$$\begin{aligned}\boldsymbol{v}&=\frac{(P+Q)^{-1}\mathbf{1}}{\mathbf{1}^\top(P+Q)^{-1}\mathbf{1}}=\frac{(\lambda G+2\theta\,\mathrm{Id}+\lambda\widetilde{G}+\gamma\widetilde{H})^{-1}\mathbf{1}}{\mathbf{1}^\top(\lambda G+2\theta\,\mathrm{Id}+\lambda\widetilde{G}+\gamma\widetilde{H})^{-1}\mathbf{1}},\\ \boldsymbol{w}&=\frac{(P-Q)^{-1}\mathbf{1}}{\mathbf{1}^\top(P-Q)^{-1}\mathbf{1}}=\frac{(\lambda\widetilde{G}^\top+2\theta\,\mathrm{Id}-\gamma\widetilde{H})^{-1}\mathbf{1}}{\mathbf{1}^\top(\lambda\widetilde{G}^\top+2\theta\,\mathrm{Id}-\gamma\widetilde{H})^{-1}\mathbf{1}}.\end{aligned} \tag{3.37}$$

In this case, by the same method in Remark 3.1.11, the optimal strategies in a Nash equilibrium are given by

$$\begin{aligned}\boldsymbol{\xi}^* &= \frac{1}{2}\Big((X_0+Y_0)\boldsymbol{v} + (X_0-Y_0)\boldsymbol{w}\Big),\\ \boldsymbol{\eta}^* &= \frac{1}{2}\Big((X_0+Y_0)\boldsymbol{v} - (X_0-Y_0)\boldsymbol{w}\Big).\end{aligned} \tag{3.38}$$

Moreover, since $H(\boldsymbol{\xi} \pm \boldsymbol{\eta}) = (X_0 \pm Y_0)\mathbf{1}$ for all admissible strategies $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathcal{X}(X_0, \mathbb{T}) \times \mathcal{X}(Y_0, \mathbb{T})$, the conditions (3.36) can be written as

$$\begin{cases} (P+Q-\gamma H)(\boldsymbol{\xi}^* + \boldsymbol{\eta}^*) = (\alpha - X_0 - Y_0)\mathbf{1},\\ (P-Q+\gamma H)(\boldsymbol{\xi}^* + \boldsymbol{\eta}^*) = (\alpha + X_0 + Y_0)\mathbf{1},\\ \mathbf{1}^\top \boldsymbol{\xi}^* = X_0,\\ \mathbf{1}^\top \boldsymbol{\eta}^* = Y_0. \end{cases} \tag{3.39}$$

If the matrices $P \pm Q \mp \gamma H$ are invertible and $\mathbf{1}^\top (P \pm Q \mp \gamma H)^{-1}\mathbf{1} \neq 0$, we have

$$\begin{aligned}\boldsymbol{v} &= \frac{(P+Q-\gamma H)^{-1}\mathbf{1}}{\mathbf{1}^\top (P+Q-\gamma H)^{-1}\mathbf{1}} = \frac{\Big(\lambda G + 2\theta\,\mathrm{Id} + \lambda \widetilde{G} - \gamma \widetilde{H}^\top\Big)^{-1}\mathbf{1}}{\mathbf{1}^\top \Big(\lambda G + 2\theta\,\mathrm{Id} + \lambda \widetilde{G} - \gamma \widetilde{H}^\top\Big)^{-1}\mathbf{1}},\\ \boldsymbol{w} &= \frac{(P-Q+\gamma H)^{-1}\mathbf{1}}{\mathbf{1}^\top (P-Q+\gamma H)^{-1}\mathbf{1}} = \frac{\Big(\lambda \widetilde{G}^\top + 2\theta\,\mathrm{Id} + \gamma \widetilde{H}^\top\Big)^{-1}\mathbf{1}}{\mathbf{1}^\top \Big(\lambda \widetilde{G}^\top + 2\theta\,\mathrm{Id} + \gamma \widetilde{H}^\top\Big)^{-1}\mathbf{1}}.\end{aligned} \tag{3.40}$$

□

After deriving the existence and uniqueness of Nash equilibria, we analyze the effects of transaction costs. The aim of imposing transaction costs is to avoid unstable optimal strategies exhibiting oscillations. As stated in Subsection 3.1.3, oscillations disappear completely, if transaction costs are sufficiently large, i.e., $\theta \geq \lambda/4$. In comparison to transient impact, permanent impact affects asset prices constantly during a trading period. In the following theorem, we see that the critical value of transaction costs increases if there exists permanent impact.

Theorem 3.2.5. *Given λ, $T > 0$ and $\mathbb{T}_N := \left\{ \frac{kT}{N} \,\middle|\, k = 0, 1, \ldots, N \right\}$. The following conditions are equivalent.*

(a) For every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{v}$ are nonnegative.

(b) For every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{w}$ are nonnegative.

(c) $\theta \geq \theta^* = (\lambda + \gamma)/4$.

To prove (c)⇒(a) in Theorem 3.2.5, we need the following auxiliary lemmas

Lemma 3.2.6. *A triangular Z-matrix $A \in \mathbb{R}^{n \times n}$ with positive diagonal is an M-matrix.*

Proof. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & a_{nn} \end{pmatrix}$$

be an upper triangular Z-matrix with positive diagonal. Then all of its leading principle minors are positive:

$$A_{[k]} = \prod_{i=1}^{k} a_{ii} > 0, \text{ for } k \in 1, 2, \ldots, N.$$

By Theorem 3.1.16, A is an M-matrix. □

To state out the following lemmas, we define $\widehat{G} = \widetilde{G} + \frac{1}{2}\,\mathrm{Id}$.

Lemma 3.2.7. *For $\alpha \geq 0$, the inverse of the matrix $\widehat{G} + \alpha G$ is equal to*

$$\begin{pmatrix} \omega & -a^{\frac{1}{N}}\mu\omega^2 & -a^{\frac{2}{N}}\mu\omega^3 & \cdots & -a^{\frac{N-1}{N}}\mu\omega^N & -\frac{a\alpha}{1+\alpha}\omega^N \\ -a^{\frac{1}{N}}\omega & (1+(1-a^{\frac{4}{N}})\alpha)\omega^2 & -a^{\frac{1}{N}}\mu\nu\omega^3 & \cdots & -a^{\frac{N-2}{N}}\mu\nu\omega^N & -a^{\frac{N-1}{N}}\mu\omega^N \\ 0 & -a^{\frac{1}{N}}\omega & (1+(1-a^{\frac{4}{N}})\alpha)\omega^2 & \cdots & -a^{\frac{N-3}{N}}\mu\nu\omega^{N-1} & -a^{\frac{N-2}{N}}\mu\omega^{N-1} \\ 0 & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -a^{\frac{1}{N}}\omega & (1+(1-a^{\frac{4}{N}})\alpha)\omega^2 & -a^{\frac{1}{N}}\mu\omega^2 \\ 0 & \cdots & \cdots & 0 & -a^{\frac{1}{N}}\omega & \omega \end{pmatrix},$$

where

$$\omega = \left(1 + (1-a^{\frac{2}{N}})\alpha\right)^{-1}, \qquad \mu = (1-a^{\frac{2}{N}})\alpha, \qquad \nu = (1-a^{\frac{2}{N}})(1+\alpha).$$

Proof. Let the matrix in the statement be denoted by D. We rewrite D as

$$D_{ij} = \begin{cases} \omega, & \text{if } (i=j=1) \vee (i=j=N+1); \\ (1+(1-a^{\frac{4}{N}})\alpha)\omega^2, & \text{if } i=j \in \{2,\ldots,N\}; \\ -a^{\frac{1}{N}}\omega, & \text{if } i-j=1; \\ -a^{\frac{k}{N}}\omega^{k+2}\mu\nu, & \text{if } (j-i =: k \in \{1,\ldots,N-2\}) \wedge (i \neq 1) \wedge (j \neq N+1); \\ -a^{\frac{k}{N}}\omega^{k+1}\mu, & \text{if } (j-i =: k \in \{1,\ldots,N-1\}) \wedge \big((i=1) \vee (j=N+1)\big); \\ -\frac{a\alpha}{1+\alpha}\omega^N, & \text{if } (i=1) \wedge (j=N+1); \\ 0, & \text{if } i \geq j+2. \end{cases}$$

On the other hand, the matrix $\widehat{G} + \alpha G$ can be written as

$$(\widehat{G} + \alpha G)_{ij} = \begin{cases} 1+\alpha, & \text{if } i=j; \\ \alpha a^{\frac{j-i}{N}}, & \text{if } i<j; \\ (1+\alpha)a^{\frac{i-j}{N}}, & \text{if } i>j. \end{cases}$$

Checking

$$\sum_{k=1}^{N+1} D_{ik}(\widehat{G}+\alpha G)_{kj} = \sum_{k=1}^{N+1} (\widehat{G}+\alpha G)_{ik} D_{kj} = \delta_{ij}$$

for each case completes the proof. □

Let $\widehat{H} := \widetilde{H}^\top - \frac{1}{2}\,\mathrm{Id}$.

Lemma 3.2.8. *The matrix $G^{-1}\big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H}\big)$ is a Z-matrix and a nonsingular M-matrix.*

Proof. It was shown in [Alfonsi et al., 2008, Theorem 3.4] that

$$G^{-1} = \frac{1}{1 - a^{\frac{2}{N}}} \begin{pmatrix} 1 & -a^{\frac{1}{N}} & 0 & \cdots & \cdots & 0 \\ -a^{\frac{1}{N}} & 1 + a^{\frac{2}{N}} & -a^{\frac{1}{N}} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a^{\frac{1}{N}} & 1 + a^{\frac{2}{N}} & -a^{\frac{1}{N}} \\ 0 & \cdots & \cdots & 0 & -a^{\frac{1}{N}} & 1 \end{pmatrix}. \tag{3.41}$$

The matrix $\widehat{G} - \frac{\gamma}{\lambda}\widehat{H}$ is equal to

$$\begin{pmatrix} 1 & -\frac{\gamma}{\lambda} & -\frac{\gamma}{\lambda} & \cdots & \cdots & -\frac{\gamma}{\lambda} \\ a^{\frac{1}{N}} & 1 & -\frac{\gamma}{\lambda} & \cdots & \cdots & -\frac{\gamma}{\lambda} \\ a^{\frac{2}{N}} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a^{\frac{1}{N}} & 1 & -\frac{\gamma}{\lambda} \\ a & \cdots & \cdots & a^{\frac{2}{N}} & a^{\frac{1}{N}} & 1 \end{pmatrix}.$$

A straightforward computation now yields that the matrix $(1 - a^{\frac{2}{N}})G^{-1}\big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H}\big)$ is equal to

$$\begin{pmatrix} 1 - a^{\frac{2}{N}} & -a^{\frac{1}{N}} - \frac{\gamma}{\lambda} & -(1 - a^{\frac{1}{N}})\frac{\gamma}{\lambda} & -(1 - a^{\frac{1}{N}})\frac{\gamma}{\lambda} & \cdots & -(1 - a^{\frac{1}{N}})\frac{\gamma}{\lambda} \\ 0 & 1 + a^{\frac{1}{N}}\frac{\gamma}{\lambda} & -a^{\frac{1}{N}} - (1 - a^{\frac{1}{N}} + a^{\frac{2}{N}})\frac{\gamma}{\lambda} & -(1 - a^{\frac{1}{N}})^2\frac{\gamma}{\lambda} & \cdots & -(1 - a^{\frac{1}{N}})^2\frac{\gamma}{\lambda} \\ 0 & 0 & 1 + a^{\frac{1}{N}}\frac{\gamma}{\lambda} & -a^{\frac{1}{N}} - (1 - a^{\frac{1}{N}} + a^{\frac{2}{N}})\frac{\gamma}{\lambda} & \cdots & -(1 - a^{\frac{1}{N}})^2\frac{\gamma}{\lambda} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & -(1 - a^{\frac{1}{N}})^2\frac{\gamma}{\lambda} \\ \vdots & \ddots & \ddots & 0 & 1 + a^{\frac{1}{N}}\frac{\gamma}{\lambda} & -a^{\frac{1}{N}} - (1 - a^{\frac{1}{N}} + a^{\frac{2}{N}})\frac{\gamma}{\lambda} \\ 0 & \cdots & \cdots & 0 & 0 & 1 + a^{\frac{1}{N}}\frac{\gamma}{\lambda} \end{pmatrix},$$

which is an upper triangular Z-matrix with positive diagonal. By Lemma 3.2.6, $G^{-1}\big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H}\big)$ is hence a nonsingular M-matrix.

□

Lemma 3.2.9. *For $\delta \geq 0$ the matrix $M := G^{-1}\big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H}\big) + \delta G^{-1}$ is a nonsingular M-matrix.*

Proof. For $\delta = 0$ this assertion is proved by Lemma 3.2.8. We consider now only for $\delta > 0$. Note that M is a Z-matrix since both $G^{-1}\big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H}\big)$ and G^{-1} are Z matrices by Lemma 3.2.8 and (3.41), respectively. Hence condition (4) of Theorem 3.1.16 will imply that M is a nonsingular M-matrix as soon as we can show that $M + \alpha\,\mathrm{Id}$ is invertible for all $\alpha \geq 0$.

In a first step, we note that taking $\gamma = 0$ in Lemma 3.2.8 yields that $G^{-1}\widehat{G}$ is a nonsingular M-matrix. Hence $(\alpha \operatorname{Id} + G^{-1}\widehat{G})^{-1} \geq 0$ for all $\alpha \geq 0$. It follows that

$$\big(\widehat{G} + \alpha G\big)^{-1}\mathbf{1} = \big(\operatorname{Id} + (\alpha G)^{-1}\widehat{G}\big)^{-1}(\alpha G)^{-1}\mathbf{1} = \Big(\alpha \operatorname{Id} + G^{-1}\widehat{G}\Big)^{-1} G^{-1}\mathbf{1} > 0.$$

By [Alfonsi et al., 2008, Example 3.5], we obtain

$$G^{-1}\mathbf{1} = \frac{1}{1 + a^{\frac{1}{N}}}\Big(1, 1 - a^{\frac{1}{N}}, \ldots, 1 - a^{\frac{1}{N}}, 1\Big)^T \gg 0. \tag{3.42}$$

Since moreover $(\widehat{G} + \alpha G)^{-1}$ is a Z-matrix by Lemma 3.2.7, it follows that $(\widehat{G} + \alpha G)^{-1}$ is a diagonally dominant Z-matrix for all $\alpha \geq 0$.

In the next step, we show that the matrix

$$Q := (\widehat{G} + \alpha G)^{-1}\Big(\delta \operatorname{Id} - \frac{\gamma}{\lambda}\widehat{H}\Big)$$

is a Z-matrix. Denoting $P := (\widehat{G} + \alpha G)^{-1}$, we get

$$Q_{ij} = \delta P_{ij} - \frac{\gamma}{\lambda}\sum_{k=1}^{j-1} P_{ik},$$

with the convention that $\sum_{k=1}^{0} a_k = 0$. It follows that $Q_{ii} \geq 0$ for all i, because $P_{ii} \geq 0$ and $\frac{\gamma}{\lambda}\sum_{k=1}^{i-1} P_{ik} \leq 0$ by the fact that P is a Z-matrix. Since P is diagonally dominant, we have $\sum_{k=1}^{j-1} P_{ik} \geq 0$ for any $j > i$ and hence $Q_{ij} = \delta P_{ij} - \frac{\gamma}{\lambda}\sum_{k=1}^{j-1} P_{ik} \leq 0$ for $j > i$. Using the fact that $P_{ik} = 0$ for $k \leq i - 1$, we get that for $j < i$

$$Q_{ij} = \delta P_{ij} - \frac{\gamma}{\lambda}\sum_{k=1}^{j-1} P_{ik} = \delta P_{ij} \leq 0.$$

This shows that Q is a Z-matrix.

We show next that Q is a nonsingular M-matrix. To this end, we note first that the triangular matrix $\Big(\delta \operatorname{Id} - \frac{\gamma}{\lambda}\widehat{H}\Big)$ is invertible. An easy calculation verifies that its inverse is given by

$$\frac{1}{\delta}\begin{pmatrix} 1 & \sigma & \sigma(1+\sigma) & \cdots & \sigma(1+\sigma)^{N-2} & \sigma(1+\sigma)^{N-1} \\ 0 & 1 & \sigma & \sigma(1+\sigma) & \cdots & \sigma(1+\sigma)^{N-2} \\ 0 & 0 & 1 & \sigma & \cdots & \sigma(1+\sigma)^{N-3} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & \sigma \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \geq 0,$$

where $\sigma := \frac{\gamma}{\lambda\delta} > 0$. Hence,

$$Q^{-1} = \Big(\delta \operatorname{Id} - \frac{\gamma}{\lambda}\widehat{H}\Big)^{-1}(\widehat{G} + \alpha G) \geq 0.$$

So Theorem 3.1.16 (3) shows that Q is a nonsingular M-matrix.

For the final step, we note first that Theorem 3.1.16 (4) implies that $\mathrm{Id} + Q$ is a nonsingular M-matrix. In particular, $(\mathrm{Id} + Q)^{-1}$ exists, and so we can define the matrix

$$\begin{aligned}
&\Big(\mathrm{Id} + (\widehat{G} + \alpha G)^{-1}\Big(\delta\,\mathrm{Id} - \frac{\gamma}{\lambda}\widehat{H}\Big)\Big)^{-1}(\widehat{G} + \alpha G)^{-1}G \\
&= \Big(\delta\,\mathrm{Id} + \widehat{G} + \alpha G - \frac{\gamma}{\lambda}\widehat{H}\Big)^{-1}G \\
&= \Big(\mathrm{Id} + (\delta G^{-1})^{-1}\big(G^{-1}(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H}) + \alpha\,\mathrm{Id}\big)\Big)^{-1}(\delta G^{-1})^{-1} \\
&= \Big(G^{-1}\Big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H}\Big) + \delta G^{-1} + \alpha\,\mathrm{Id}\Big)^{-1} \\
&= (M + \alpha\,\mathrm{Id})^{-1}.
\end{aligned}$$

This completes the proof. □

Lemma 3.2.10. *Let A be an invertible matrix and suppose that $\alpha \in \mathbb{R}$ is such that $A+\alpha H$ is invertible. Then the vector $A^{-1}\mathbf{1}$ is proportional to $(A + \alpha H)^{-1}\mathbf{1}$.*

Proof. Note that $H\boldsymbol{x}$ is proportional to $\mathbf{1}$ for any vector $\boldsymbol{x}$. Hence,

$$(A + \alpha H)A^{-1}\mathbf{1} = (\mathrm{Id} + \alpha HA^{-1})\mathbf{1} = (1 + \beta)\mathbf{1}$$

for some constant β. Applying $(A + \alpha H)^{-1}$ to both sides of this equation yields the result. □

Proof of (c)⇔(a) *in Theorem 3.2.5.* Let

$$\theta \geq \frac{\lambda + \gamma}{4} \qquad \text{and} \qquad \delta := \frac{4\theta - (\lambda + \gamma)}{2\lambda} \geq 0.$$

It is clear that

$$\widetilde{G} + \frac{2\theta}{\lambda}\,\mathrm{Id} - \frac{\gamma}{\lambda}\widetilde{H}^\top = \widehat{G} - \frac{\gamma}{\lambda}\widehat{H} + \delta\,\mathrm{Id}.$$

Lemma 3.2.9 yields that the matrix $\lambda G + 2\theta\,\mathrm{Id} + \lambda\widetilde{G} - \gamma\widetilde{H}^\top$ is invertible, its inverse is given by

$$\begin{aligned}
&\Big(\lambda G + 2\theta\,\mathrm{Id} + \lambda\widetilde{G} - \gamma\widetilde{H}^\top\Big)^{-1} \\
&= \Big(\mathrm{Id} + (\lambda G)^{-1}\Big(2\theta\,\mathrm{Id} + \lambda\widetilde{G} - \gamma\widetilde{H}^\top\Big)\Big)^{-1}(\lambda G)^{-1} \\
&= \Big(\mathrm{Id} + G^{-1}\Big(\widetilde{G} + \frac{2\theta}{\lambda}\,\mathrm{Id} - \frac{\gamma}{\lambda}\widetilde{H}^\top\Big)\Big)^{-1}(\lambda G)^{-1} \\
&= \Big(\mathrm{Id} + G^{-1}\Big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H} + \delta\,\mathrm{Id}\Big)\Big)^{-1}(\lambda G)^{-1}.
\end{aligned}$$

It also follows by Lemma 3.2.9 that the matrix

$$G^{-1}\Big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H} + \delta\,\mathrm{Id}\Big)$$

is a nonsingular M-matrix. By Theorem 3.1.16, we obtain that

$$\Big(\mathrm{Id} + G^{-1}\Big(\widehat{G} - \frac{\gamma}{\lambda}\widehat{H} + \delta\,\mathrm{Id}\Big)\Big)^{-1} \geq 0.$$

As it has been shown in the proof of Theorem 3.1.14 that $G^{-1}\mathbf{1} \gg 0$, it yields that

$$\Big(\lambda G + 2\theta\,\mathrm{Id} + \lambda\widetilde{G} - \gamma\widetilde{H}^\top\Big)^{-1}\mathbf{1} \geq 0.$$

Therefore, the vector $\boldsymbol{v}$ stated in (3.40) in this case is well defined. This establishes (c)$\Rightarrow$(a).

To show (c)$\Leftarrow$(a), we consider the case $N = 1$. By definition, $\boldsymbol{v}$ is proportional to the vector

$$2\det(\lambda G + 2\theta\,\mathrm{Id} + \widetilde{G}) \cdot (\lambda G + 2\theta\,\mathrm{Id} + \widetilde{G})^{-1}\mathbf{1} = \begin{pmatrix} \lambda(3-2a) + \gamma + 4\theta \\ \lambda(3-4a) - \gamma + 4\theta \end{pmatrix}.$$

Clearly, the first component of this vector is positive for all $a \in (0,1)$ and $\theta \geq 0$. By sending $a \uparrow 1$ one sees, however, that the second component is negative for $\theta < \theta^*$ and a sufficiently close to 1. Thus, we cannot have $\boldsymbol{v} \geq 0$ in this case. $\square$

Proof of (c)$\Leftrightarrow$(b) *in Theorem 3.2.5.* We assume that $\theta \geq (\lambda+\gamma)/4$. Note that

$$(\widetilde{H}^\top + \frac{1}{2}\,\mathrm{Id})^{-1} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Let $\kappa := \frac{1}{2} + \frac{2\theta}{\lambda} - \frac{\gamma}{2\lambda}$. It follows that

$$\begin{aligned} &\frac{\gamma}{\lambda}\Big(\gamma\widetilde{H}^\top + \frac{\gamma}{2}\,\mathrm{Id}\Big)^{-1}\Big(\lambda\widetilde{G}^\top + \Big(2\theta - \frac{\gamma}{2}\Big)\,\mathrm{Id}\Big) \\ &= \Big(\widetilde{H}^\top + \frac{1}{2}\,\mathrm{Id}\Big)^{-1}\Big(\widetilde{G}^\top + \Big(\kappa - \frac{1}{2}\Big)\,\mathrm{Id}\Big) \\ &= \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} \kappa & a^{\frac{1}{N}} & a^{\frac{2}{N}} & \cdots & a^{\frac{N-1}{N}} & a \\ 0 & \kappa & a^{\frac{1}{N}} & \cdots & a^{\frac{N-2}{N}} & a^{\frac{N-1}{N}} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \kappa & a^{\frac{1}{N}} \\ 0 & \cdots & \cdots & \cdots & 0 & \kappa \end{pmatrix}. \end{aligned} \tag{3.43}$$

A straightforward computation yields that (3.43) is equal to

$$\frac{\lambda}{\gamma}\begin{pmatrix} \kappa & a^{\frac{1}{N}}-\kappa & (a^{\frac{1}{N}}-1)a^{\frac{1}{N}} & \cdots & (a^{\frac{1}{N}}-1)a^{\frac{N-2}{N}} & (a^{\frac{1}{N}}-1)a^{\frac{N-1}{N}} \\ 0 & \kappa & a^{\frac{1}{N}}-\kappa & (a^{\frac{1}{N}}-1)a^{\frac{1}{N}} & \cdots & (a^{\frac{1}{N}}-1)a^{\frac{N-2}{N}} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & (a^{\frac{1}{N}}-1)a^{\frac{1}{N}} \\ \vdots & \ddots & \ddots & \ddots & \kappa & a^{\frac{1}{N}}-\kappa \\ 0 & \cdots & \cdots & \cdots & 0 & \kappa \end{pmatrix}.$$

If $\theta \geq (\lambda+\gamma)/4$, we have $\kappa - 1 \geq 0$. Because $a^{\frac{1}{N}} < 1$ for all $N \in \mathbb{N}$, the matrix $(\gamma\widetilde{H}^\top + \frac{\gamma}{2}\operatorname{Id})^{-1}\big(\lambda\widetilde{G}^\top + \Big(2\theta - \frac{\gamma}{2}\Big)\operatorname{Id}\big)$ is a Z-matrix. By Lemma 3.2.6, it is an M-matrix. Therefore, by Theorem 3.1.16, the matrix

$$A := \operatorname{Id} + \Big(\gamma\widetilde{H}^\top + \frac{\gamma}{2}\operatorname{Id}\Big)^{-1}\Big(\lambda\widetilde{G}^\top + \Big(2\theta - \frac{\gamma}{2}\Big)\operatorname{Id}\Big)$$

is also an M-matrix, which is inverse-positive. We have

$$\begin{aligned}\Big(\gamma\widetilde{H}^\top + \frac{\gamma}{2}\operatorname{Id}\Big)A &= \gamma\widetilde{H}^\top + \lambda\widetilde{G}^\top + 2\theta\operatorname{Id} = \lambda G - \lambda\widetilde{G} - \gamma\widetilde{H} + 2\theta\operatorname{Id} + \gamma H \\ &= \lambda\widetilde{G}^\top + 2\theta\operatorname{Id} + \gamma\widetilde{H}^\top.\end{aligned}$$

It follows that the latter matrix is invertible. An application of Lemma 3.2.10 therefore yields that $\boldsymbol{w}$ is proportional to the vector

$$\begin{aligned}&\Big(\lambda\widetilde{G}^\top + 2\theta\operatorname{Id} + \gamma\widetilde{H}^\top\Big)^{-1}\mathbf{1} \\ &= \Big(\operatorname{Id} + \big(\gamma\widetilde{H}^\top + \frac{\gamma}{2}\operatorname{Id}\big)^{-1}\big(\lambda\widetilde{G}^\top + \Big(2\theta - \frac{\gamma}{2}\Big)\operatorname{Id}\big)\Big)^{-1}\Big(\gamma\widetilde{H}^\top + \frac{\gamma}{2}\operatorname{Id}\Big)^{-1}\mathbf{1} \\ &= \frac{1}{\gamma}\Big(\operatorname{Id} + \big(\gamma\widetilde{H}^\top + \frac{\gamma}{2}\operatorname{Id}\big)^{-1}\big(\lambda\widetilde{G}^\top + \Big(2\theta - \frac{\gamma}{2}\Big)\operatorname{Id}\big)\Big)^{-1}\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{\gamma}\Big(\operatorname{Id} + \big(\gamma\widetilde{H}^\top + \frac{\gamma}{2}\operatorname{Id}\big)^{-1}\big(\lambda\widetilde{G}^\top + \Big(2\theta - \frac{\gamma}{2}\Big)\operatorname{Id}\big)\Big)^{-1} \cdot (0,0,\dots,0,1)^\top \geq 0.\end{aligned}$$

This establishes (c)$\Rightarrow$(b).

To show (c)$\Leftarrow$(b), we assume $N = 1$ and $\boldsymbol{w} \geq 0$ for all $a \in (0,1)$. It follows that the determinant of the matrix $\lambda\widetilde{G}^\top + 2\theta\operatorname{Id} + \gamma\widetilde{H}^\top$ is given by

$$D := \frac{1}{4}(\gamma + 4\theta + \lambda)^2 > 0.$$

We obtain then

$$(\lambda\widetilde{G}^\top + 2\theta\operatorname{Id} + \gamma\widetilde{H}^\top)^{-1}\mathbf{1} = \frac{1}{2D}\begin{pmatrix}(1-2a)\lambda - \gamma + 4\theta \\ \gamma + 4\theta + \lambda\end{pmatrix} =: \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}.$$

We assume by way of contradiction that $\theta < (\lambda+\gamma)/4$. Then we have

$$\frac{\lambda-\gamma+4\theta}{2\lambda} < 1.$$

For $a \in (\frac{\lambda-\gamma+4\theta}{2\lambda}, 1)$ we obtain $w_2 < 0$. Since $w_1 + w_2 = 8\theta + 2(1-a)\lambda > 0$, it yields that $\boldsymbol{w} = (\frac{w_1}{w_1+w_2}, \frac{w_2}{w_1+w_2})^\top$ is not nonnegative. This completes the proof. $\square$

In Figures 3.7 and 3.8, we see the interaction between permanent impact γ and transaction costs θ in the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, i.e., when $X_0 = Y_0 = 1$ and $X_0 = -Y_0 = 1$. For both vectors, oscillations exist when there exists permanent impact and no transactions costs; on the other hand, oscillations can be completely avoided when $\theta = \theta^* = (\lambda+\gamma)/4 = 0.75$.

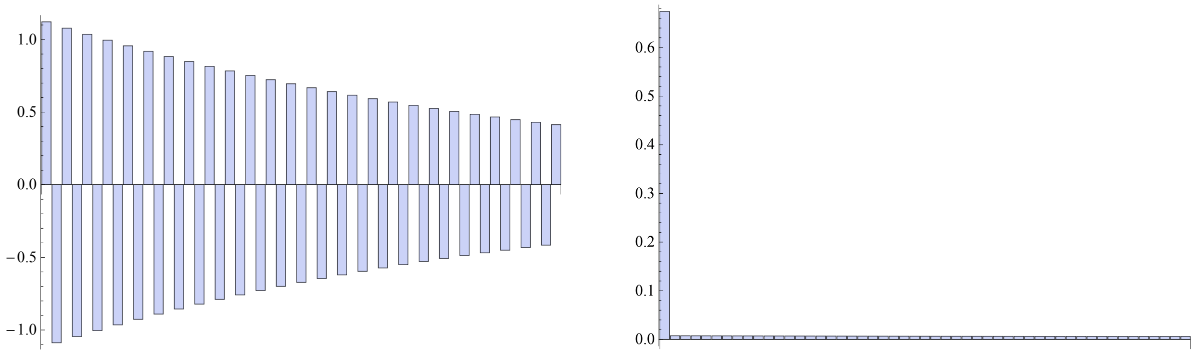


Figure 3.7: Optimal strategies in the equilibrium for the equidistant time grid $\mathbb{T}_{50}$ with parameters $\theta = 0$ (left), $\theta = 0.75$ (right) and $\rho = 1$, $\lambda = 1$, $\gamma = 2$, initial values $X_0 = Y_0 = 1$.

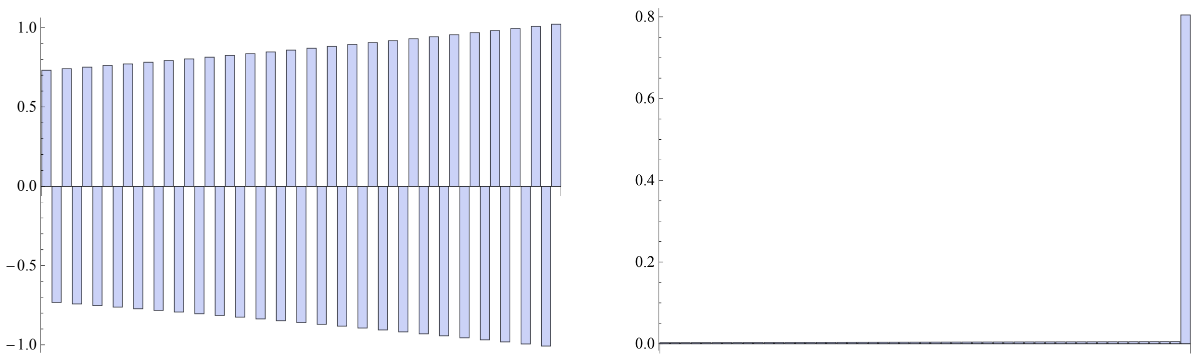


Figure 3.8: Optimal strategies in the equilibrium for the equidistant time grid $\mathbb{T}_{50}$ with parameters $\theta = 0$ (left), $\theta = 0.75$ (right) and $\rho = 1$, $\lambda = 1$, $\gamma = 2$, initial values $X_0 = -Y_0 = 1$.

3.2.2 Splitting of combined liquidation costs

In Subsections 3.1.1 and 3.2.1, we assume none of the two agents has an advantage in latency over the other. Their execution priorities are determined by an i.i.d. sequence of

Bernoulli $(\frac{1}{2})$-distributed random variables. As a result, the additional expected liquidation costs $\mathbb{E}[\lambda\xi_k\eta_k]$ or $\mathbb{E}[(\lambda+\gamma)\xi_k\eta_k]$ occurred jointly by the two agents at time point t_k are split equally for each $t_k \in \mathbb{T}$. Now we assume the two agents have different advantages in latency in our model. The main consequence of such unsymmetrical latency priority is that the combined liquidation costs are not split equally.

In this subsection, we derive the existence and uniqueness of Nash equilibria for unequal splitting of combined liquidity costs. As stated in Subsection 3.2.1, the price process is given by

$$S_t^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_t^0 - \lambda \sum_{t_k<t} e^{-\rho(t-t_k)}(\xi_k+\eta_k) - \gamma \sum_{t_k<t}(\xi_k+\eta_k).$$

To describe the splitting of combined liquidity costs, we assume the execution priorities of the two agents are determined by an independent sequence of Bernoulli (q_i)-distributed random variables $(\varepsilon_i)_{i=0,1,\dots}$ that are independent of $\sigma(\cup_{t\geq 0}\mathcal{F}_t)$ with $\mathbb{E}[\varepsilon_i]=q_i$ for $q_i\in[0,1]$ and each $i\in\mathbb{N}$. Without transaction costs, the liquidation costs of $\boldsymbol{\xi}\in\mathcal{X}(X_0,\mathbb{T})$ given $\boldsymbol{\eta}\in\mathcal{X}(Y_0,\mathbb{T})$ are defined as

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}) = X_0S_0^0 + \sum_{k=0}^{N}\Big(\frac{\lambda+\gamma}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k + \varepsilon_k(\lambda+\gamma)\xi_k\eta_k\Big),$$

and the liquidation costs of $\boldsymbol{\eta}$ given $\boldsymbol{\xi}$ are

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}) = Y_0S_0^0 + \sum_{k=0}^{N}\Big(\frac{\lambda+\gamma}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + (1-\varepsilon_k)(\lambda+\gamma)\xi_k\eta_k\Big).$$

Lemma 3.2.11. *The expected liquidation costs of an admissible strategy $\boldsymbol{\xi}\in\mathcal{X}(X_0,\mathbb{T})$ given another admissible strategy $\boldsymbol{\eta}\in\mathcal{X}(Y_0,\mathbb{T})$ are*

$$\mathbb{E}[\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})] = \mathbb{E}\Big[\frac{1}{2}\boldsymbol{\xi}^\top\lambda G\boldsymbol{\xi} + \boldsymbol{\xi}^\top(\lambda\widetilde{G}_q+\gamma\widetilde{H}_q)\boldsymbol{\eta}\Big] + \frac{\gamma}{2}X_0^2,$$

and the expected liquidation costs of $\boldsymbol{\eta}$ given $\boldsymbol{\xi}$ are

$$\mathbb{E}[\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi})] = \mathbb{E}\Big[\frac{1}{2}\boldsymbol{\eta}^\top\lambda G\boldsymbol{\eta} + \boldsymbol{\eta}^\top\big(\lambda(G-\widetilde{G}_q)^\top+\gamma(H-\widetilde{H}_q)^\top\big)\boldsymbol{\xi}\Big] + \frac{\gamma}{2}Y_0^2,$$

where $\widetilde{G}_q$ is the lower-triangular matrices of G with $(\widetilde{G}_q)_{nn}=q_n$ and $\widetilde{H}_q$ is the lower-triangular matrices of H with $(\widetilde{H}_q)_{nn}=q_n$.

Proof. We note that

$$\begin{aligned}
\mathbb{E}[\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})] &= \mathbb{E}\Big[\sum_{n=0}^{N}\Big(-\xi_nS_{t_n}^{\boldsymbol{\xi},\boldsymbol{\eta}} + \frac{\lambda+\gamma}{2}\xi_n^2 + \varepsilon_n(\lambda+\gamma)\xi_n\eta_n\Big)\Big] + X_0S_0^0 \\
&= \mathbb{E}\Big[\sum_{n=0}^{N}\Big(-\xi_nS_{t_n}^0 + \big(\xi_n\sum_{m=0}^{n-1}\xi_m\lambda e^{-\rho(t_n-t_m)} + \xi_n\sum_{m=0}^{n-1}\gamma\xi_m + \frac{\lambda+\gamma}{2}\xi_n^2\big) \\
&\quad + \big(\xi_n\sum_{m=0}^{n-1}\eta_m\lambda e^{-\rho(t_n-t_m)} + \xi_n\sum_{m=0}^{n-1}\gamma\eta_m + q_n(\lambda+\gamma)\xi_n\eta_n\big)\Big)\Big] + X_0S_0^0.
\end{aligned}$$

By the proof of Lemma 3.1.7, we have

$$\mathbb{E}\Big[\sum_{n=0}^{N}\xi_n S^0_{t_n}\Big] = \mathbb{E}\Big[\sum_{n=0}^{N}\xi_n S^0_T\Big] = X_0\mathbb{E}[\,S^0_T\,] = X_0 S^0_0.$$

Moreover, it holds that

$$\begin{aligned}
&\xi_n\sum_{m=0}^{n-1}\xi_m\lambda e^{-\rho(t_n-t_m)} + \xi_n\sum_{m=0}^{n-1}\gamma\xi_m + \frac{\lambda+\gamma}{2}\xi_n^2\\
&= \boldsymbol{\xi}^\top(\lambda\widetilde{G}+\gamma\widetilde{H})\boldsymbol{\xi} = \frac{1}{2}\boldsymbol{\xi}^\top(\lambda G+\gamma H)\boldsymbol{\xi} = \frac{1}{2}\boldsymbol{\xi}^\top\lambda G\boldsymbol{\xi} + \frac{\gamma}{2}X_0^2,
\end{aligned}$$

and

$$\xi_n\sum_{m=0}^{n-1}\eta_{t_m}\lambda e^{-\rho(t_n-t_m)} + \xi_n\sum_{m=0}^{n-1}\gamma\eta_m + q_n(\lambda+\gamma)\xi_n\eta_n = \boldsymbol{\xi}^\top(\lambda\widetilde{G}_q+\gamma\widetilde{H}_q)\boldsymbol{\eta}.$$

It follows that

$$\mathbb{E}[\mathcal{C}_\mathbb{T}(\boldsymbol{\xi}|\boldsymbol{\eta})] = \mathbb{E}\Big[\frac{1}{2}\boldsymbol{\xi}^\top\lambda G\boldsymbol{\xi} + \boldsymbol{\xi}^\top(\lambda\widetilde{G}_q+\gamma\widetilde{H}_q)\boldsymbol{\eta}\Big] + \frac{\gamma}{2}X_0^2.$$

The proof for $\mathbb{E}[\mathcal{C}_\mathbb{T}(\boldsymbol{\eta}|\boldsymbol{\xi})]$ is analogous. $\square$

Theorem 3.2.12. *Given $\rho>0$, $\lambda>0$ and $\gamma>0$. For any time grid $\mathbb{T}$ and initial values X_0, $Y_0\in\mathbb{R}$, there exists a unique Nash equilibrium $(\boldsymbol{\xi}^*,\boldsymbol{\eta}^*)\in\mathscr{X}(X_0,\mathbb{T})\times\mathscr{X}(Y_0,\mathbb{T})$ in the class of adapted strategies. The optimal strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ are deterministic.*

Proof. First we prove that for each sequence $(q_n)_{n\in\{0,1,\dots,N\}}$ for $q_n\in[0,1]$, there exists at least one Nash equilibrium in the class of deterministic strategies. To this end, we follow the idea of the proof of Rosen [1965, Theorem 1]. Note that the set $\mathscr{X}_{\det}(X_0,\mathbb{T})$ is convex and closed for each $X_0\in\mathbb{R}$. We define the convex and closed set $E_{X_0\times Y_0} := \mathscr{X}_{\det}(X_0,\mathbb{T})\times\mathscr{X}_{\det}(Y_0,\mathbb{T})$ and the function ρ by

$$\rho: E_{X_0\times Y_0}\times E_{X_0\times Y_0}\to\mathbb{R},\quad \rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix},\begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}\right) = \mathcal{C}_\mathbb{T}(\boldsymbol{\xi}|\boldsymbol{y}) + \mathcal{C}_\mathbb{T}(\boldsymbol{\eta}|\boldsymbol{x}). \tag{3.44}$$

It is clear that ρ is continuous in $\binom{\boldsymbol{x}}{\boldsymbol{y}}$ and $\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}$ with respect to Euclidean norm. By the strict convexity of $\mathcal{C}_\mathbb{T}(\boldsymbol{\xi}|\boldsymbol{y})$ and $\mathcal{C}_\mathbb{T}(\boldsymbol{\eta}|\boldsymbol{x})$ in $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, we obtain for $\alpha\in[0,1]$ and $\binom{\boldsymbol{\xi}^0}{\boldsymbol{\eta}^0},\binom{\boldsymbol{\xi}^1}{\boldsymbol{\eta}^1}\in E_{X_0\times Y_0}$ that

$$\begin{aligned}
&\rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix},\alpha\begin{pmatrix}\boldsymbol{\xi}^0\\ \boldsymbol{\eta}^0\end{pmatrix}+(1-\alpha)\begin{pmatrix}\boldsymbol{\xi}^1\\ \boldsymbol{\eta}^1\end{pmatrix}\right)\\
&= \mathcal{C}_\mathbb{T}(\alpha\boldsymbol{\xi}^0+(1-\alpha)\boldsymbol{\xi}^1|\boldsymbol{y}) + \mathcal{C}_\mathbb{T}(\alpha\boldsymbol{\eta}^0+(1-\alpha)\boldsymbol{\eta}^1|\boldsymbol{x})\\
&\le \alpha\mathcal{C}_\mathbb{T}(\boldsymbol{\xi}^0\,|\,\boldsymbol{y}) + (1-\alpha)\mathcal{C}_\mathbb{T}(\boldsymbol{\xi}^1|\boldsymbol{y}) + \alpha\mathcal{C}_\mathbb{T}(\boldsymbol{\eta}^0|\boldsymbol{x}) + (1-\alpha)\mathcal{C}_\mathbb{T}(\boldsymbol{\eta}^1|\boldsymbol{x})\\
&= \alpha\rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix},\begin{pmatrix}\boldsymbol{\xi}^0\\ \boldsymbol{\eta}^0\end{pmatrix}\right) + (1-\alpha)\rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix},\begin{pmatrix}\boldsymbol{\xi}^1\\ \boldsymbol{\eta}^1\end{pmatrix}\right),
\end{aligned}$$

which means ρ is convex in $\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}$. Then we consider the point-to-set function $\Gamma : \binom{\boldsymbol{x}}{\boldsymbol{y}} \in E_{X_0\times Y_0} \mapsto \Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}} \subset E_{X_0\times Y_0}$ with

$$\Gamma\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix} := \left\{\begin{pmatrix}\boldsymbol{\xi}^*\\ \boldsymbol{\eta}^*\end{pmatrix} \in E_{X_0\times Y_0} \,\middle|\, \rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}^*\\ \boldsymbol{\eta}^*\end{pmatrix}\right) = \min_{\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}\in E_{X_0\times Y_0}} \rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}\right)\right\}.$$

We claim that for each $\binom{\boldsymbol{x}}{\boldsymbol{y}} \in E_{X_0\times Y_0}$, $\Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}}$ is a non-empty, convex and closed set in $E_{X_0\times Y_0}$. To this end, first note that again due to the strict convexity of $\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}\,|\,\boldsymbol{y})$ and $\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{x})$ in $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, there exist $\boldsymbol{\xi}^* \in \mathcal{X}_{\mathrm{det}}(X_0, \mathbb{T})$ and $\boldsymbol{\eta}^* \in \mathcal{X}_{\mathrm{det}}(Y_0, \mathbb{T})$ which minimize $\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{y})$ and $\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{x})$ respectively. Hence, $\Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}}$ is non-empty.

Let $\binom{\boldsymbol{\xi}^0}{\boldsymbol{\eta}^0}, \binom{\boldsymbol{\xi}^1}{\boldsymbol{\eta}^1} \in \Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}}$, i.e.,

$$\rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}^0\\ \boldsymbol{\eta}^0\end{pmatrix}\right) = \rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}^1\\ \boldsymbol{\eta}^1\end{pmatrix}\right) = \min_{\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}\in E_{X_0\times Y_0}} \rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}\right).$$

For $\alpha \in [0, 1]$, due to the convexity of ρ in $\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}$, we obtain

$$\begin{aligned}
&\min_{\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}\in E_{X_0\times Y_0}} \rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}\right) \le \rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \alpha\begin{pmatrix}\boldsymbol{\xi}^0\\ \boldsymbol{\eta}^0\end{pmatrix} + (1-\alpha)\begin{pmatrix}\boldsymbol{\xi}^1\\ \boldsymbol{\eta}^1\end{pmatrix}\right)\\
&\le \alpha\rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}^0\\ \boldsymbol{\eta}^0\end{pmatrix}\right) + (1-\alpha)\rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}^1\\ \boldsymbol{\eta}^1\end{pmatrix}\right)\\
&= \min_{\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}\in E_{X_0\times Y_0}} \rho\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}\right),
\end{aligned}$$

which shows $\alpha\binom{\boldsymbol{\xi}^0}{\boldsymbol{\eta}^0} + (1-\alpha)\binom{\boldsymbol{\xi}^1}{\boldsymbol{\eta}^1} \in \Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}}$, i.e., $\Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}}$ is convex.

To show $\Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}}$ is closed, let $\binom{\boldsymbol{\xi}_n}{\boldsymbol{\eta}_n} \in \Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}}$ for each $n \in \mathbb{N}$ and $\binom{\boldsymbol{\xi}_n}{\boldsymbol{\eta}_n} \to \binom{\boldsymbol{\xi}^*}{\boldsymbol{\eta}^*}$ in $E_{X_0\times Y_0}$ for $n \to \infty$. Since ρ is continuous in $\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}$, we get $\binom{\boldsymbol{\xi}^*}{\boldsymbol{\eta}^*} \in \Gamma\binom{\boldsymbol{x}}{\boldsymbol{y}}$.

By the Kakutani fixed point theorem, see Kakutani et al. [1941], there exists a point $\binom{\boldsymbol{x}^*}{\boldsymbol{y}^*} \in E_{X_0\times Y_0}$ such that

$$\begin{pmatrix}\boldsymbol{x}^*\\ \boldsymbol{y}^*\end{pmatrix} \in \Gamma\begin{pmatrix}\boldsymbol{x}^*\\ \boldsymbol{y}^*\end{pmatrix} \text{ i.e., } \rho\left(\begin{pmatrix}\boldsymbol{x}^*\\ \boldsymbol{y}^*\end{pmatrix}, \begin{pmatrix}\boldsymbol{x}^*\\ \boldsymbol{y}^*\end{pmatrix}\right) = \min_{\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}\in E_{X_0\times Y_0}} \rho\left(\begin{pmatrix}\boldsymbol{x}^*\\ \boldsymbol{y}^*\end{pmatrix}, \begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}\right).$$

We now prove that $\binom{\boldsymbol{x}^*}{\boldsymbol{y}^*}$ is a Nash equilibrium in $E_{X_0\times Y_0}$. Suppose by way of contradiction there exists $\widetilde{\boldsymbol{y}} \in \mathcal{X}_{\mathrm{det}}(Y_0, \mathbb{T})$ such that $\mathcal{C}_{\mathbb{T}}(\widetilde{\boldsymbol{y}}|\boldsymbol{x}^*) < \mathcal{C}_{\mathbb{T}}(\boldsymbol{y}^*|\boldsymbol{x}^*)$, it follows that

$$\rho\left(\begin{pmatrix}\boldsymbol{x}^*\\ \boldsymbol{y}^*\end{pmatrix}, \begin{pmatrix}\boldsymbol{x}^*\\ \widetilde{\boldsymbol{y}}\end{pmatrix}\right) = \mathcal{C}_{\mathbb{T}}(\boldsymbol{x}^*|\boldsymbol{y}^*) + \mathcal{C}_{\mathbb{T}}(\widetilde{\boldsymbol{y}}|\boldsymbol{x}^*) < \mathcal{C}_{\mathbb{T}}(\boldsymbol{x}^*|\boldsymbol{y}^*) + \mathcal{C}_{\mathbb{T}}(\boldsymbol{y}^*|\boldsymbol{x}^*) = \rho\left(\begin{pmatrix}\boldsymbol{x}^*\\ \boldsymbol{y}^*\end{pmatrix}, \begin{pmatrix}\boldsymbol{x}^*\\ \boldsymbol{y}^*\end{pmatrix}\right),$$

which contradicts (3.44).

In the second step we show that for each sequence $(q_n)_{n\in\mathbb{N}}$ with $q_n \in [0, 1]$ for all $n \in \mathbb{N}$, there exists at most one Nash equilibrium in the class of adapted strategies. To this end, we use the same method which is used to proved Lemma 3.1.9. We define

$$\boldsymbol{\xi}^\alpha := \alpha\boldsymbol{\xi}^1 + (1-\alpha)\boldsymbol{\xi}^0, \quad \boldsymbol{\eta}^\alpha := \alpha\boldsymbol{\eta}^1 + (1-\alpha)\boldsymbol{\eta}^0, \text{ where } \alpha \in [0, 1];$$

and

$$f(\alpha) := \mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^{\alpha}|\boldsymbol{\eta}^0) + \mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}^{\alpha}|\boldsymbol{\xi}^0) + \mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^{1-\alpha}|\boldsymbol{\eta}^1) + \mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}^{1-\alpha}|\boldsymbol{\xi}^1).$$

Since the matrix λG is strictly positive definite, the cost functional $\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})$ is strictly convex with respect to $\boldsymbol{\xi}$. Hence, we obtain

$$f(\alpha) > f(0) \quad \text{for } \alpha > 0.$$

It follows that

$$\lim_{h\downarrow 0} \frac{f(h) - f(0)}{h} = \frac{df(\alpha)}{d\alpha}\Big|_{\alpha=0+} \geq 0. \tag{3.45}$$

On the other hand, it holds that

$$\begin{aligned}
\frac{d}{d\alpha}\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}^{\alpha}|\boldsymbol{\eta})\Big|_{\alpha=0} &= \mathbb{E}\Big[(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^0)^\top \lambda G \boldsymbol{\xi}^0 + (\boldsymbol{\xi}^1 - \boldsymbol{\xi}^0)^\top (\lambda \widetilde{G}_q + \gamma \widetilde{H}_q)\boldsymbol{\eta}\Big], \\
\frac{d}{d\alpha}\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}^{\alpha}|\boldsymbol{\xi})\Big|_{\alpha=0} &= \mathbb{E}\Big[(\boldsymbol{\eta}^1 - \boldsymbol{\eta}^0)^\top \lambda G \boldsymbol{\eta}^0 + (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^0)^\top \big(\lambda (G - \widetilde{G}_q)^\top + \gamma (H - \widetilde{H}_q)^\top\big)\boldsymbol{\xi}\Big].
\end{aligned}$$

Note that

$$\boldsymbol{\xi}^\top (\lambda \widetilde{G}_q + \gamma \widetilde{H}_q)\boldsymbol{\eta} + \boldsymbol{\eta}^\top \big(\lambda (G - \widetilde{G}_q)^\top + \gamma (H - \widetilde{H}_q)^\top\big)\boldsymbol{\xi} = \boldsymbol{\xi}^\top (\lambda G + \gamma H)\boldsymbol{\eta}.$$

Hence, it follows that

$$\begin{aligned}
&\frac{d}{d\alpha} f(\alpha)\Big|_{\alpha=0+} \\
&= -\mathbb{E}\Big[(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^0)^\top \lambda G (\boldsymbol{\xi}^1 - \boldsymbol{\xi}^0) + (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^0)^\top \lambda G (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^0)\Big] - \mathbb{E}\Big[(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^0)^\top \lambda G (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^0)\Big], \\
&\qquad -\mathbb{E}\Big[(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^0)^\top \lambda H (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^0)\Big] \\
&< -\frac{1}{2}\mathbb{E}\Big[(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^0)^\top \lambda G (\boldsymbol{\xi}^1 - \boldsymbol{\xi}^0) + (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^0)^\top \lambda G (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^0)\Big] \\
&< 0,
\end{aligned}$$

which contradicts (3.45). Therefore, there exists at most one Nash equilibrium in the class of adapted strategies.

In the last step, by the same method which is used to prove Lemma 3.1.10, we show that a Nash equilibrium in the class of deterministic strategies is also a Nash equilibrium in the class of adapted strategies. □

Remark 3.2.13. We explain how to derive explicit equilibrium strategies for the model. By the uniqueness of Nash equilibrium and the fact that a Nash equilibrium in the class of deterministic strategies is also a Nash equilibrium in the class of adapted strategies, we consider only deterministic strategies. We assume that there exists a Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*) \in \mathcal{X}_{\mathrm{det}}(x, \mathbb{T}) \times \mathcal{X}_{\mathrm{det}}(y, \mathbb{T})$, such that

$$\begin{cases}
\frac{1}{2}(\boldsymbol{\xi}^*)^\top \lambda G \boldsymbol{\xi}^* + (\boldsymbol{\xi}^*)^\top (\lambda \widetilde{G}_q + \gamma \widetilde{H}_q)\boldsymbol{\eta}^* \to \min, \\
\frac{1}{2}(\boldsymbol{\eta}^*)^\top \lambda G \boldsymbol{\eta}^* + (\boldsymbol{\eta}^*)^\top \big(\lambda (G - \widetilde{G}_q)^\top + \gamma (H - \widetilde{H}_q)^\top\big)\boldsymbol{\xi}^* \to \min.
\end{cases}$$

By the method of Lagrange multipliers, a necessary condition for $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ is the existence of $\alpha, \beta \in \mathbb{R}$, such that

$$\begin{cases} \lambda G \boldsymbol{\xi}^* + (\lambda \widetilde{G}_q + \gamma \widetilde{H}_q)\boldsymbol{\eta}^* = \alpha \mathbf{1}, \\ \lambda G \boldsymbol{\eta}^* + \big(\lambda (G - \widetilde{G}_q)^\top + \gamma (H - \widetilde{H}_q)^\top\big)\boldsymbol{\xi}^* = \beta \mathbf{1}, \end{cases} \tag{3.46}$$

which can be re-written as

$$\begin{pmatrix} \lambda G & \lambda \widetilde{G}_q + \gamma \widetilde{H}_q \\ \lambda (G - \widetilde{G}_q)^\top + \gamma (H - \widetilde{H}_q)^\top & \lambda G \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi}^* \\ \boldsymbol{\eta}^* \end{pmatrix} = \begin{pmatrix} \alpha \mathbf{1} \\ \beta \mathbf{1} \end{pmatrix}. \tag{3.47}$$

To simplify our notation, we define

$$U := \lambda G, \quad V_q := \lambda \widetilde{G}_q + \gamma \widetilde{H}_q, \quad V_{1-q} := \lambda (G - \widetilde{G}_q)^\top + \gamma (H - \widetilde{H}_q)^\top.$$

If the matrix

$$U - V_{1-q} U^{-1} V_q$$

is invertible, see Bernstein [2005, Page 44], the block matrix on the left side of (3.47) is invertible with the inverse

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} := \begin{pmatrix} U^{-1} + U^{-1} V_q (U - V_{1-q} U^{-1} V_q)^{-1} V_{1-q} U^{-1} & -U^{-1} V_q (U - V_{1-q} U^{-1} V_q)^{-1} \\ -(U - V_{1-q} U^{-1} V_q)^{-1} V_{1-q} U^{-1} & (U - V_{1-q} U^{-1} V_q)^{-1} \end{pmatrix}.$$

Therefore, we obtain

$$\begin{cases} \boldsymbol{\xi}^* = \alpha A \mathbf{1} + \beta B \mathbf{1}, \\ \boldsymbol{\eta}^* = \alpha C \mathbf{1} + \beta D \mathbf{1}. \end{cases}$$

By the constraints that

$$\mathbf{1}^\top \boldsymbol{\xi}^* = X_0, \quad \mathbf{1}^\top \boldsymbol{\eta}^* = Y_0;$$

and if $ad - bc \neq 0$, it follows that

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} X_0 \\ Y_0 \end{pmatrix},$$

where

$$a = \mathbf{1}^\top A \mathbf{1}, \quad b = \mathbf{1}^\top B \mathbf{1}, \quad c = \mathbf{1}^\top C \mathbf{1}, \quad d = \mathbf{1}^\top D \mathbf{1}.$$

□

Now we give numerical simulations of unequal splitting of combined liquidation costs. For all scenarios in this subsection, we set $X_0 = 2$, $Y_0 = 1$, $\rho = \lambda = 1$, $\gamma = 0$ and use the equidistant time grid $\mathbb{T}_{50}$.

In Figure 3.9, we see that if agent X has a complete advantage in latency, i.e., $q_n = 0$ for all execution time points, a large portion will be liquidated at the beginning since agent X's initial value is larger than Y's and all combined costs occurred by the large liquidation have to be paid by agent Y. On the contrary, agent Y will not hesitate to liquidate almost whole position given a complete advantage in latency, i.e., $q_n = 1$ for all execution time points while the agent X has to liquidate a large portion at the end of time grid to minimize the price impact.

Figure 3.10 describes the equilibria when the advantage in latency for one agent increases or decreases as the execution time goes on. If the advantage in latency for agent X decreases, i.e., q_n increases with respect to n, the beginning part of asset positions for both agents is similar to the left part of Figure 3.9, since agent X still has advantage in latency. However, as execution time goes on, due to the increase of advantage in latency for agent Y, our simulation shows that agent X will minimize the liquidation costs by a small round trip near the end of the trading. On the other hand, if the advantage in latency for agent Y decreases, i.e., q_n decreases with respect to n, agent Y will liquidate a larger amount of stock at the beginning of trading in comparison to the right part of Figure 3.9 to have a short position. Note that although agent Y losses the advantage at the end of trading, i.e., $q_{50} = 0$, a large purchase at the time point still reduces the costs because agent X has a sell-order at that time, the "combined costs" turn to be a profit, i.e., $(1 - q_{50})\lambda\xi_{50}\eta_{50} < 0$.

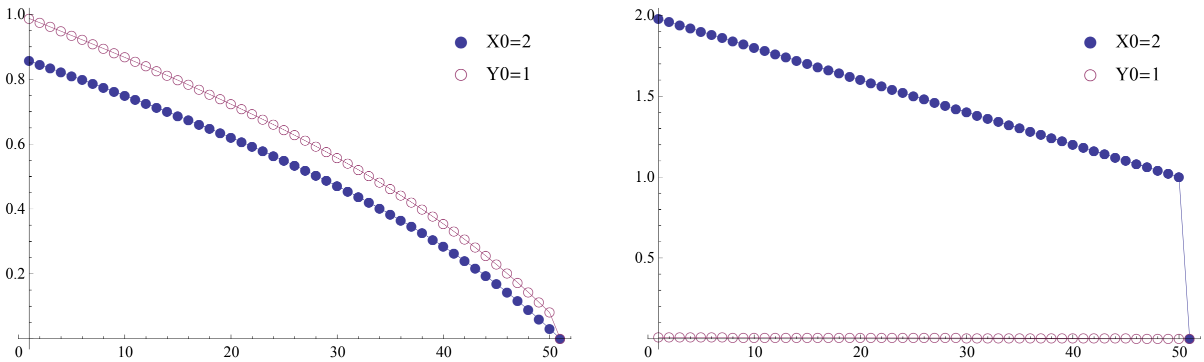


Figure 3.9: Optimal asset positions in the equilibrium with parameters $q_n = 0$ (left) and $q_n = 1$ (right).

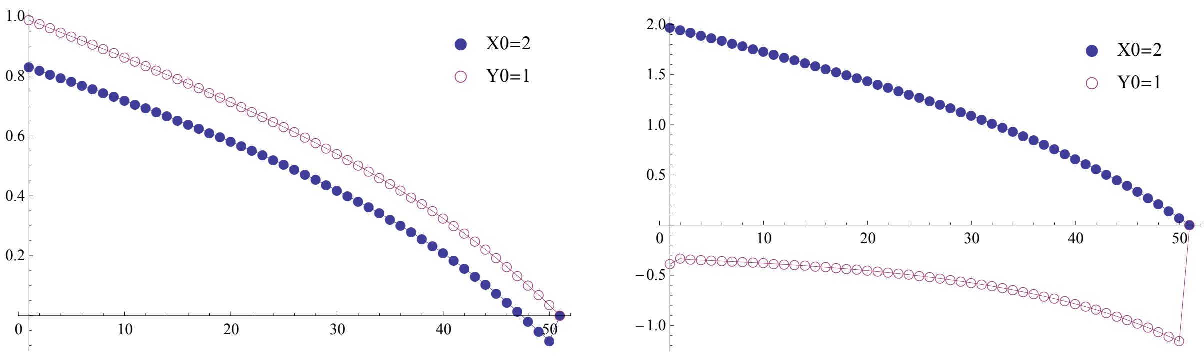


Figure 3.10: Optimal asset positions in the equilibrium with parameters $q_n = (\frac{n}{50})^2$ (left) and $q_n = (1 - \frac{n}{50})^2$ (right).

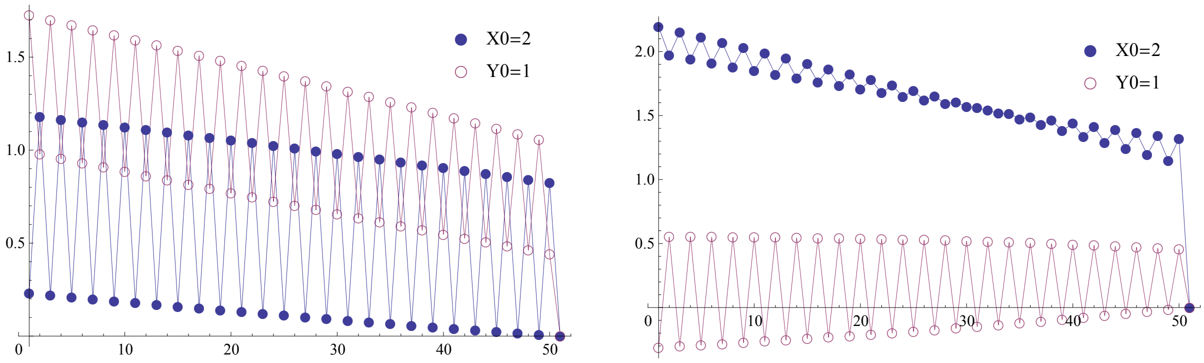


Figure 3.11: Optimal asset positions in the equilibrium $q_n = \frac{1}{2}\big(1 + (-1)^n\big)$ (left) and $q_n = \frac{1}{2}\big(1 + (-1)^{n+1}\big)$ (right).

Figure 3.11 shows a simulation of asset positions in Nash equilibria when the advantage in latency fluctuates. One sees that fluctuations of the advantages in latency lead to fluctuations of asset positions. The design of q_n makes that two agents have complete advantage in latency alternatively. In the left part of the figure, agent X has a complete advantage in latency first, i.e., $q_0 = 0$. Both of the two agents use "hot-potato" strategies to protect themselves. In particular, agent Y holds a larger amount of stock than agent X although agent Y has smaller initial asset position. In contract to the left part, the fluctuation of asset positions is reduced if agent Y has a complete advantage in latency first, i.e., $q_0 = 1$ and agent X always holds a larger asset position during the whole trading.

3.2.3 Closed-loop strategies

So far we have looked at various models of Nash equilibria in open-loop setup. That is, at each execution time, trading strategies do not depend on previous trades of the other trader. In this subsection, we take a quick view of closed-loop strategies. In a closed-loop model, at each execution time t_n all trades of the other agent at $t_1, t_2, \dots, t_{n-1}$ will be considered.

For a general time grind $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$ with $0 = t_0 < t_1 < \cdots < t_N = T$, when the strategies $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ by the agents X and Y are applied, we assume that the price process is given by

$$S_t^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_t^0 - \lambda \sum_{t_k < t} e^{-\rho(t-t_k)}(\xi_k + \eta_k) - \gamma \sum_{t_k < t} (\xi_k + \eta_k).$$

Furthermore, we assume the execution priority at each time is determined by a sequence of Bernoulli (q)-distributed i.i.d. random variables $(\varepsilon_i)_{i=0,1,\dots,N}$ which are independent of $\sigma(\bigcup_{t\ge 0} \mathcal{F}_t)$ with $\mathbb{E}[\varepsilon_i] = q$ for all $\{i = 0, 1, \dots, N\}$, $q \in [0,1]$. We define the liquidation costs of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ as

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}) = X_0 S_0^0 + \sum_{k=0}^{N} \Big(\frac{\lambda+\gamma}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k + \varepsilon_k(\lambda+\gamma)\xi_k\eta_k + \theta\xi_k^2 \Big),$$

$$\mathcal{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}) = Y_0 S_0^0 + \sum_{k=0}^{N} \Big(\frac{\lambda+\gamma}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + (1-\varepsilon_k)(\lambda+\gamma)\xi_k\eta_k + \theta\eta_k^2 \Big),$$

where the transaction costs are denoted by θ. Both of the two agents minimize the expected costs of their strategies accordingly. As it is applied to analyze the optimal closed-loop strategies for a similar case in Schöneborn [2008], we use dynamic programming to derive a unique Nash equilibrium in this closed-loop model.

Theorem 3.2.14. *There exists a unique Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ in the class of closed-loop strategies. At each execution time point t_n, the optimal orders are linear in remained asset positions X_n, Y_n, and transient impact I_n, i.e.,*

$$\xi_n^* = A_n^X X_n + B_n^X Y_n + C_n^X I_n, \quad \eta_n^* = A_n^Y X_n + B_n^Y Y_n + C_n^Y I_n. \tag{3.48}$$

The cost functionals of the Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ for $n \in 0, 1, \ldots, N$ are

$$\begin{aligned}
&J_n^X(X_n, Y_n, I_n) \\
=&a_n^X X_n^2 + b_n^X Y_n^2 + c_n^X I_n^2 + u_n^X X_n Y_n + v_n^X X_n I_n + w_n^X Y_n I_n + \gamma(X_0 + Y_0) X_n, \\
&J_n^Y(X_n, Y_n, I_n) \\
=&a_n^Y X_n^2 + b_n^Y Y_n^2 + c_n^Y I_n^2 + u_n^Y X_n Y_n + v_n^Y X_n I_n + w_n^Y Y_n I_n + \gamma(X_0 + Y_0) Y_n.
\end{aligned} \tag{3.49}$$

All of the coefficients

$$A_n^X, B_n^X, C_n^X, a_n^X, b_n^X, c_n^X, u_n^X, v_n^X, w_n^X, A_n^Y, B_n^Y, C_n^Y, a_n^Y, b_n^Y, c_n^Y, u_n^Y, v_n^Y, w_n^Y \in \mathbb{R}$$

can be recursively computed for each time t_n, $n \in \{0, 1, \ldots, N-1\}$.

Proof. We use a backward induction to prove that the optimal strategies $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ and their cost functionals J^X, J^Y can be expressed in the way stated in the assertion. Let $n = N$, both agents have no choice but liquidate their remained asset positions, since they are at the last execution time point. It follows that the equations (3.48) and (3.49) hold for the following coefficients:

$$\begin{aligned}
A_N^X = 1, B_N^X = C_N^X = 0, a_N^X = \frac{\lambda - \gamma}{2} + \theta, b_N^X = c_N^X = w_N^X = 0, u_N^X = q(\lambda + \gamma) - \gamma, v_N^X = 1; \\
B_N^Y = 1, A_N^Y = C_N^Y = 0, b_N^Y = \frac{\lambda - \gamma}{2} + \theta, a_N^Y = c_N^Y = v_N^Y = 0, u_N^Y = \lambda - q(\lambda + \gamma), w_N^Y = 1.
\end{aligned}$$

Then we assume that the equations (3.48) and (3.49) hold for $n+1$. Given an arbitrary strategy $\boldsymbol{\eta}$, the cost functionals for X starting at t_n should be determined by

$$\begin{aligned}
J_n^X(X_n, Y_n, I_n \,|\, \boldsymbol{\eta}) = &\inf_{\xi_n \in \mathbb{R}} \Big\{ F^X(\xi_n, \eta_n) \Big\} \\
&:= \inf_{\xi_n \in \mathbb{R}} \Big\{ -\xi_n\big(-I_n - \gamma(X_0 - X_n + Y_0 - Y_n)\big) + \big(\frac{\lambda + \gamma}{2} + \theta\big)\xi_n^2 \\
&\quad + q(\lambda + \gamma)\xi_n \eta_n + J_{n+1}^X\big(X_n - \xi_n, Y_n - \eta_n, \big(I_n + \lambda(\xi_n + \eta_n)\big) e^{-\rho(t_{n+1} - t_n)}\big) \Big\},
\end{aligned}$$

which is the minimum of the quadratic function $F^X(\xi_n, \eta_n)$ with respect to ξ_n. Therefore, we obtain the optimal strategy for X

$$\xi_n^* = \varphi_n^X X_n + \psi_n^X Y_n + \mu_n^X I_n + \nu_n^X \eta_n, \tag{3.50}$$

where

$$\begin{aligned}
\varphi_n^X &= \frac{2a_{n+1}^X + \gamma - \lambda v_{n+1}^X e^{-(t_{n+1}-t_n)}}{2\left(a_{n+1}^X + \frac{\gamma+\lambda}{2} + c_{n+1}^X \lambda^2 e^{-2(t_{n+1}-t_n)} - \lambda v_{n+1}^X e^{-(t_{n+1}-t_n)} + \theta\right)}, \\
\psi_n^X &= \frac{u_{n+1}^X + \gamma - \lambda w_{n+1}^X e^{-(t_{n+1}-t_n)}}{2\left(a_{n+1}^X + \frac{\gamma+\lambda}{2} + c_{n+1}^X \lambda^2 e^{-2(t_{n+1}-t_n)} - \lambda v_{n+1}^X e^{-(t_{n+1}-t_n)} + \theta\right)}, \\
\mu_n^X &= \frac{v_{n+1}^X e^{-(t_{n+1}-t_n)} - 2c_{n+1}^X \lambda e^{-2(t_{n+1}-t_n)} - 1}{2\left(a_{n+1}^X + \frac{\gamma+\lambda}{2} + c_{n+1}^X \lambda^2 e^{-2(t_{n+1}-t_n)} - \lambda v_{n+1}^X e^{-(t_{n+1}-t_n)} + \theta\right)}, \\
\nu_n^X &= \frac{\lambda v_{n+1}^X e^{-\rho(t_{n+1}-t_n)} - 2c_{n+1}^X \lambda^2 e^{-2\rho(t_{n+1}-t_n)} + \lambda w_{n+1}^X e^{-\rho(t_{n+1}-t_n)} - q(\lambda+\gamma) - u_{n+1}^X}{2\left(a_{n+1}^X + \frac{\gamma+\lambda}{2} + c_{n+1}^X \lambda^2 e^{-2(t_{n+1}-t_n)} - \lambda v_{n+1}^X e^{-(t_{n+1}-t_n)} + \theta\right)}.
\end{aligned}$$

In the same way, we obtain the optimal strategy for Y given an arbitrary $\boldsymbol{\xi}$,

$$\eta_n^* = \varphi_n^Y X_n + \psi_n^Y Y_n + \mu_n^Y I_n + \nu_n^Y \xi_n. \tag{3.51}$$

Therefore, there exists a Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ if and only if $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ satisfy (3.50) and (3.51), i.e.,

$$\begin{aligned}
\xi_n^* &= \varphi_n^X X_n + \psi_n^X Y_n + \mu_n^X I_n + \nu_n^X \eta_n, \\
\eta_n^* &= \varphi_n^Y X_n + \psi_n^Y Y_n + \mu_n^Y I_n + \nu_n^Y \xi_n^*.
\end{aligned}$$

We then obtain the unique solution for each $n \in 0, 1, \ldots, N-1$,

$$\begin{aligned}
\xi_n^* &= \frac{\mu_n^X I_n + \nu_n^X \mu_n^Y I_n + \varphi_n^X X_n + \nu_n^X \varphi_n^Y X_n + \psi_n^X Y_n + \nu_n^X \psi_n^Y Y_n}{1 - \nu_n^X \nu_n^Y}, \\
\eta_n^* &= \frac{\mu_n^Y I_n + \nu_n^Y \mu_n^X I_n + \varphi_n^Y X_n + \nu_n^Y \varphi_n^X X_n + \psi_n^Y Y_n + \nu_n^Y \psi_n^X Y_n}{1 - \nu_n^X \nu_n^Y}.
\end{aligned}$$

Note that the optimal orders ξ_n^* and η_n^* are linear in X_n, Y_n, I_n for each $n \in \{0, 1, \ldots, N\}$. We define

$$\begin{aligned}
J_n^X(X_n, Y_n, I_n) &:= J_n^X(X_n, Y_n, I_n \,|\, \boldsymbol{\eta}^*) = \inf_{\xi_n \in \mathbb{R}} \left\{ F^X(\xi_n, \eta_n^*) \right\} = F^X(\xi_n^*, \eta_n^*), \\
J_n^Y(X_n, Y_n, I_n) &:= J_n^Y(X_n, Y_n, I_n \,|\, \boldsymbol{\xi}^*) = \inf_{\eta_n \in \mathbb{R}} \left\{ F^Y(\xi_n^*, \eta_n) \right\} = F^Y(\xi_n^*, \eta_n^*).
\end{aligned}$$

$J_n^X(X_n, Y_n, I_n)$ and $J_n^Y(X_n, Y_n, I_n)$ are quadratic in X_n, Y_n, I_n and satisfy (3.49) for each $n \in \{0, 1, \ldots, N\}$. This completes our proof. □

Now we give numerical simulations to compare optimal strategies in Nash equilibria between open-loop and closed-loop models. In all examples we use the equidistant time grid $\mathbb{T}_{50}$ and set $\lambda = \rho = T = 1$.

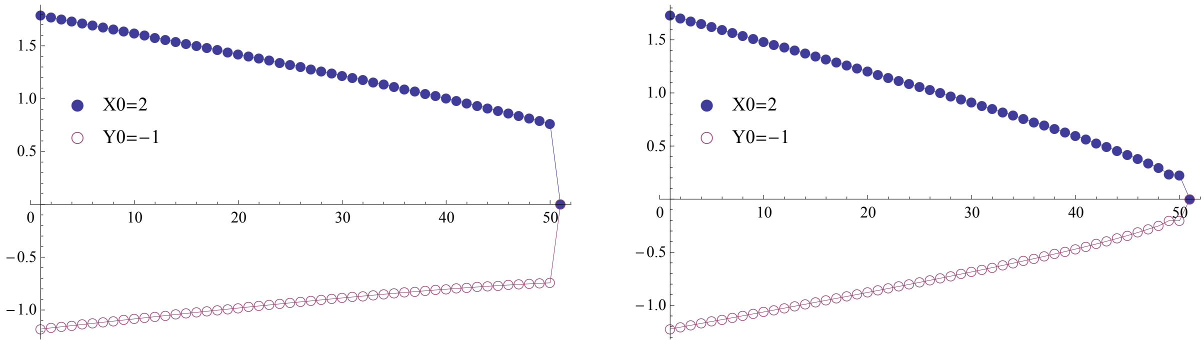


Figure 3.14: Accumulated asset positions of optimal open-loop (left) and closed-loop (right) strategies with parameters $\gamma = 0$, $\theta = \frac{1}{4}$, $q = \frac{1}{2}$ and initial values $X_0 = 2$, $Y_0 = -1$.

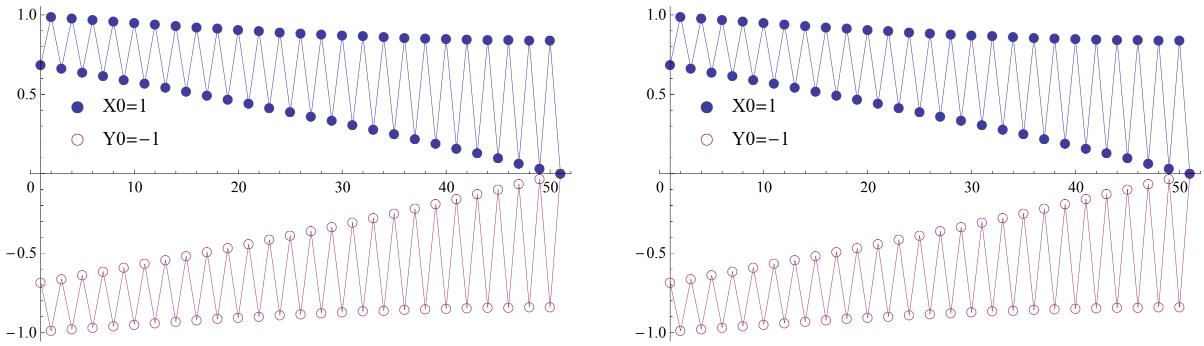


Figure 3.12: Accumulated asset positions of optimal open-loop (left) and closed-loop (right) strategies with parameters $\gamma = 0$, $\theta = 0$, $q = \frac{1}{2}$ and initial values $X_0 = -Y_0 = 1$.

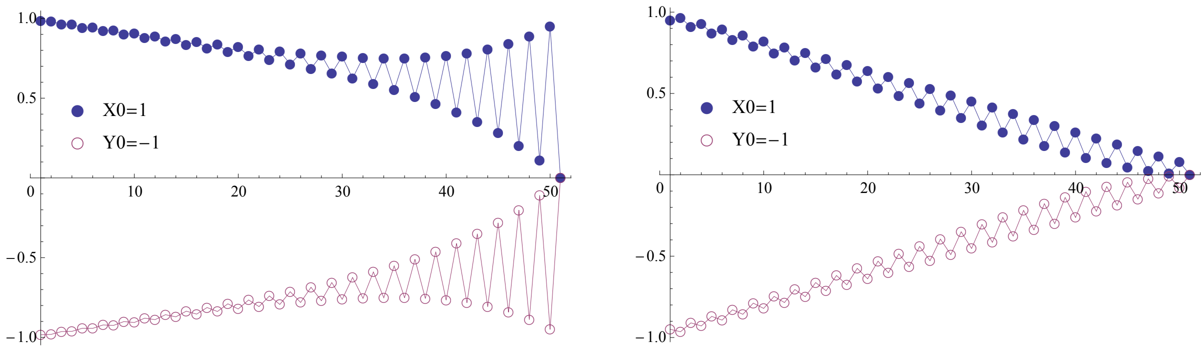


Figure 3.13: Accumulated asset positions of optimal open-loop (left) and closed-loop (right) strategies with parameters $\gamma = 0$, $\theta = 0.01$, $q = \frac{1}{2}$ and initial values $X_0 = -Y_0 = 1$.

Figure 3.12 shows that optimal closed-loop strategies coincide with optimal open-loop strategies in the trivial case for $X_0 = -Y_0 = 1$. The optimal strategies consist of strong oscillations and form a hot-potato game. In this simulation, taking previous trading of each agent into account does not lead to more efficient equilibrium strategies. The absence of transaction costs and equal splitting of combined liquidation costs facilitate oscillatory strategies.

However, once transaction costs are imposed, oscillations in the equilibrium strategies are reduced. We see that in Figure 3.13 there is little oscillation in closed-loop strategies than in open-loop strategies, especially at the end of trading. This phenomenon shows that taking account of previous trading at each execution time helps to improve the efficiency of the trading and leads to certain cooperation. Both agents do not need to use strong oscillations. Figure 3.14 also supports this point of view by showing that the last jump in optimal closed-loop strategies is smaller than in open-loop strategies. In closed-loop model both agents do not need to wait until the last execution time to buy or sell a large amount of asset, they can achieve their goals earlier by smoother strategies.

In Figures 3.15 and 3.16 we compare equilibrium strategies within the closed-loop model. Figure 3.15 shows the situation when the combined liquidation costs are not split equally. As same as in Figure 3.9, if agent X has a complete advantage in latency, a large amount of the asset will be liquidated at the beginning. Having a smaller initial asset position, agent Y holds the position at approximately 0.6 for quite a long time after a large liquidation at the beginning. If agent Y has a complete advantage in latency, a large amount of the asset will be also liquidated and a low position is kept during the trading. Note that in both situations there are cooperations between the two agents at the end of trading. That is, they hold opposite positions before the last trade to obtain a negative price impact at the last trade even if there are transaction costs.

In Figure 3.16, we see that transaction costs play a similar role as they do in open-loop models. That is, there are no oscillations in the equilibrium strategies if the transaction costs are sufficiently large. However, in comparison to open-loop models the critical value of transaction costs for removing oscillations in this simulation is smaller.

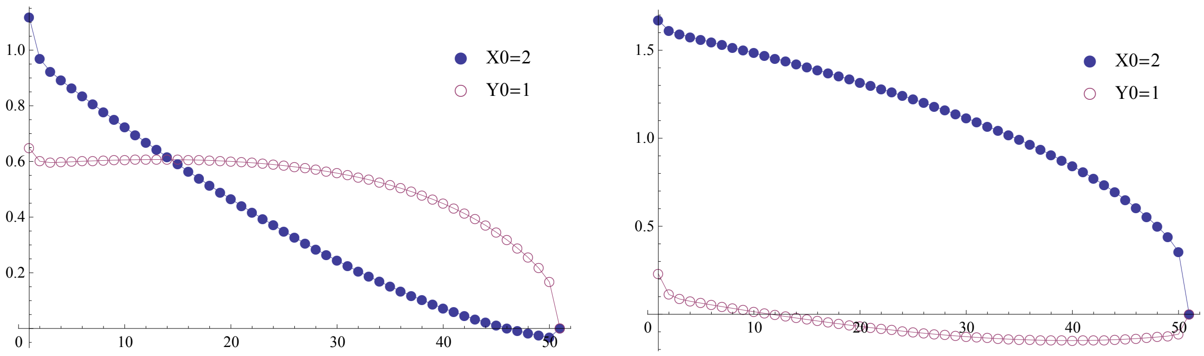


Figure 3.15: Accumulated asset positions of optimal closed-loop strategies with parameters $q = 0$ (left), $q = 1$ (right), $\gamma = 0$, $\theta = \frac{1}{4}$ and initial values $X_0 = 2$, $Y_0 = 1$.

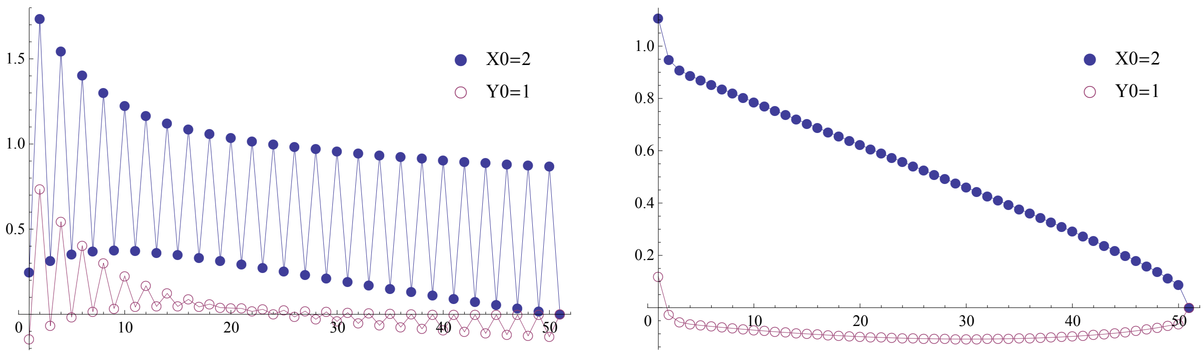


Figure 3.16: Accumulated asset positions of optimal closed-loop strategies with parameters $\theta = 0$ (left), $\theta = \frac{1}{4}$ (right), $\gamma = 1$, $q = \frac{1}{2}$ and initial values $X_0 = 2$, $Y_0 = 1$.

3.3 Continuous-time version of the primary model

3.3.1 Model setup

We consider a continuous-time version of the primary model. The market we consider consists of only one risky asset. We assume there are two financial agents X and Y who can affect the asset prices by their trading. When X and Y are not active, the price process is described by a martingale $(S_t^0)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. The trading strategies used by the two large traders are stochastic processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$, which are defined on the same probability space.

Definition 3.3.1. A strategy $(Z_t)_{t\geq 0-}$ is called *admissible*, if it satisfies the following conditions:

- the function $t \mapsto Z_t$ is right-continuous, adapted and bounded;
- the function $t \mapsto Z_t$ has finite and $\mathbb{P}$-a.s. bounded total variation;
- there exists $T > 0$ such that $Z_t = 0$ $\mathbb{P}$-a.s. for all $t \geq T$.

We denote the class of all admissible strategies Z with an initial asset position $Z_{0-} = z$ in the time horizon $[0, T]$ by $\mathcal{X}(z, [0, T])$. When X and Y are active, the *affected price* $S^{X,Y}$ is assumed to be the sum of the unaffected asset price S^0 and the market impact caused by the two agents:

$$S_t^{X,Y} = S_t^0 + \lambda \int_{[0,t)} e^{-\rho(t-s)}\, dX_s + \lambda \int_{[0,t)} e^{-\rho(t-s)}\, dY_s.$$

This model we describe is an extended version of the continuous-time model introduced in Gatheral et al. [2012] with the exponential decay $G(t) = e^{-\rho t}$ for two agents. This model can also be regarded as the continuous-time version of the model introduced in Section 3.1 that comes from Obizhaeva and Wang [2013] originally.

Now we prepare to define the liquidation costs for X and Y. Through the following two different approaches, we will obtain the same expression of liquidation costs.

First, to derive the liquidation costs that are consistent to our primary model, we consider an approximation from the discrete-time case. To this end, let $X \in \mathcal{X}(x,[0,T])$ and $Y \in \mathcal{X}(y,[0,T])$ be admissible strategies and let $\mathbb{T}_N := \{t_k^N \,|\, t_k^N = \frac{kT}{N}, k \in \{0,1,\dots,N\}\}$ be an equidistant time grid. We define the following discrete trades

$$\begin{aligned}
\xi_0^N &:= X_0 - X_{0-} \quad \text{and} \quad \xi_k^N := X_{t_k^N} - X_{t_{k-1}^N} \text{ for } k = \{1,2,\dots,N\};\\
\eta_0^N &:= Y_0 - Y_{0-} \quad \text{and} \quad \eta_k^N := Y_{t_k^N} - Y_{t_{k-1}^N} \text{ for } k = \{1,2,\dots,N\}.
\end{aligned}$$

Then $\boldsymbol{\xi}^N := (\xi_k^N)$ and $\boldsymbol{\eta}^N := (\eta_k^N)$ are admissible strategies in sense of Definition 3.1.1, i.e., $\boldsymbol{\xi}^N \in \mathcal{X}(x,\mathbb{T}_N)$ and $\boldsymbol{\eta}^N \in \mathcal{X}(y,\mathbb{T}_N)$. Furthermore, for each $N \in \mathbb{N}$ let $(\varepsilon_k^N)_{k\in\{0,1,\dots,N\}}$ be a sequence of i.i.d. Bernoulli $(\frac{1}{2})$-distributed random variables that are independent of $\sigma(\bigcup_{t\geq 0}\mathcal{F}_t)$. According to Definition 3.1.2, we get the liquidation costs for $\boldsymbol{\xi}^N$ and $\boldsymbol{\eta}^N$:

$$\begin{aligned}
\mathcal{C}^N(\boldsymbol{\xi}^N|\boldsymbol{\eta}^N) &:= xS_{0-}^0 + \sum_{k=0}^N \Big(\frac{\lambda}{2}(\xi_k^N)^2 - S_{t_k^N}^{\boldsymbol{\xi}^N,\boldsymbol{\eta}^N}\xi_k^N + \varepsilon_k^N\lambda\xi_k^N\eta_k^N + \theta(\xi_k^N)^2\Big),\\
\mathcal{C}^N(\boldsymbol{\eta}^N|\boldsymbol{\xi}^N) &:= yS_{0-}^0 + \sum_{k=0}^N \Big(\frac{\lambda}{2}(\eta_k^N)^2 - S_{t_k^N}^{\boldsymbol{\xi}^N,\boldsymbol{\eta}^N}\eta_k^N + (1-\varepsilon_k^N)\lambda\xi_k^N\eta_k^N + \theta(\eta_k^N)^2\Big).
\end{aligned} \tag{3.52}$$

In the following lemma we obtain the convergence of the expected liquidation costs.

Lemma 3.3.2. *As $N \uparrow \infty$, we have*

$$\begin{aligned}
\mathbb{E}[\mathcal{C}^N(\boldsymbol{\xi}^N|\boldsymbol{\eta}^N)] \longrightarrow &\mathbb{E}\Big[\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\,dX_s\,dX_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\,dY_s\,dX_t\\
&\quad + \frac{\lambda}{2}\sum_{t\in[0,T]}\Delta X_t\Delta Y_t + \theta\sum_{t\in[0,T]}(\Delta X_t)^2\Big],\\
\mathbb{E}[\mathcal{C}^N(\boldsymbol{\eta}^N|\boldsymbol{\xi}^N)] \longrightarrow &\mathbb{E}\Big[\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\,dY_s\,dY_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\,dX_s\,dY_t\\
&\quad + \frac{\lambda}{2}\sum_{t\in[0,T]}\Delta X_t\Delta Y_t + \theta\sum_{t\in[0,T]}(\Delta Y_t)^2\Big].
\end{aligned} \tag{3.53}$$

Proof. We first note that

$$\begin{aligned}
\mathcal{C}^N(\boldsymbol{\xi}^N|\boldsymbol{\eta}^N) = xS_{0-}^0 &- \sum_{k=0}^N \xi_k^N S_{t_k^N}^0 + \lambda\sum_{k=0}^N\sum_{m=0}^{k-1}\xi_k^N\xi_m^N e^{-\rho(t_k^N - t_m^N)} + \frac{\lambda}{2}\sum_{k=0}^N(\xi_k^N)^2\\
&+ \lambda\sum_{k=0}^N \varepsilon_k^N\xi_k^N\eta_k^N + \lambda\sum_{k=0}^N\sum_{m=0}^{k-1}\xi_k^N\eta_m^N e^{-\rho(t_k^N - t_m^N)} + \theta\sum_{k=0}^N(\xi_k^N)^2.
\end{aligned}$$

The proof of Lorenz and Schied [2012, Lemma 1] yields that, as $N \uparrow \infty$,

$$\begin{aligned}
&\sum_{k=0}^N \xi_k^N S_{t_k^N}^0 \longrightarrow \int_{[0,T]} S_{t-}^0\,dX_t + [S^0, X]_T, \qquad \sum_{k=0}^N(\xi_k^N)^2 \longrightarrow [X]_T,\\
&\sum_{k=0}^N\sum_{m=0}^{k-1}\xi_k^N\xi_m^N e^{-\rho(t_k^N - t_m^N)} \longrightarrow \int_{[0,T]} e^{-\rho t}\int_{[0,t)} e^{\rho s}\,dX_s\,dX_t,
\end{aligned}$$

in probability. Similarly, we have

$$\sum_{k=0}^{N}\sum_{m=0}^{k-1}\xi_k^N\eta_m^N e^{-\rho(t_k^N-t_m^N)} \longrightarrow \int_{[0,T]} e^{-\rho t}\int_{[0,t)} e^{\rho s}\, dY_s\, dX_t, \qquad \text{as } N\uparrow\infty,$$

in probability.

Furthermore, by Protter [2004, Theorem II.6.23], we have $\sum_{k=0}^{N}\xi_k^N\eta_k^N \longrightarrow [X,Y]_T$ in probability. Since ε_k^N is independent of $\sigma(\bigcup_{t\geq 0}\mathcal{F}_t)$ for every N and k, we obtain that

$$\mathbb{E}\Big[\sum_{k=0}^{N}\varepsilon_k^N\xi_k^N\eta_k^N\Big] = \frac{1}{2}\mathbb{E}\Big[\sum_{k=0}^{N}\xi_k^N\eta_k^N\Big] \longrightarrow \frac{1}{2}\mathbb{E}\Big[[X,Y]_T\Big], \qquad \text{as } N\uparrow\infty.$$

As it has been explained in Lorenz and Schied [2012, Remark 3], the integration by parts formula for stochastic integrals can be stated as

$$X_tY_t = X_{0-}Y_{0-} + \int_{[0,t]} X_{s-}\, dY_s + \int_{[0,t]} Y_{s-}\, dX_s + [X,Y]_t,$$

for semi-martingales X and Y. Hence,

$$\mathbb{E}\Big[\int_{[0,T]} S^0_{t-}\, dX_t + [S^0,X]_T\Big] = \mathbb{E}\Big[X_TS^0_T - X_{0-}S^0_{0-} - \int_{[0,T]} X_{t-}\, dS^0_t\Big] = -xS^0_{0-}.$$

Since X and Y are of $\mathbb{P}$-a.s. bounded variation by Definition 3.3.1, we have

$$\mathbb{E}\Big[[X]_T\Big] = \mathbb{E}\Big[\sum_{t\in[0,T]}(\Delta X_t)^2\Big] \quad \text{and} \quad \mathbb{E}\Big[[X,Y]_T\Big] = \mathbb{E}\Big[\sum_{t\in[0,T]}\Delta X_t\Delta Y_t\Big].$$

It also holds that

$$\int_{[0,T]} e^{-\rho t}\int_{[0,t)} e^{\rho s}\, dX_s\, dX_t + \frac{1}{2}\sum_{t\in[0,T]}(\Delta X_t)^2 = \frac{1}{2}\int_{[0,T]}\int_{[0,T]} e^{-\rho|t-s|}\, dX_s\, dX_t.$$

Putting everything together yields

$$\begin{aligned}\mathbb{E}[\mathcal{C}^N(\boldsymbol{\xi}^N|\boldsymbol{\eta}^N)] \longrightarrow \mathbb{E}\Big[&\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dY_s\, dX_t \\ &+ \frac{\lambda}{2}\sum_{t\in[0,T]}\Delta X_t\Delta Y_t + \theta\sum_{t\in[0,T]}(\Delta X_t)^2\Big], \qquad \text{as } N\uparrow\infty.\end{aligned}$$

It is similar to prove the convergence of $\mathbb{E}[\mathcal{C}^N(\boldsymbol{\eta}^N|\boldsymbol{\xi}^N)]$. This completes the proof. □

Now we verify the liquidation costs through an intuitive consideration. If $X \in \mathcal{X}(x,[0,T])$ and $Y \in \mathcal{X}(y,[0,T])$ are continuous in t for all $t\in[0,T]$, the total costs for X are

$$\int_{[0,T]} S_t^{X,Y}\, dX_t = \int_{[0,T]} S_t^0\, dX_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s\, dX_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dY_s\, dX_t.$$

If X has several jumps and Y is still continuous, there are quadratic transaction costs and additional liquidation costs incurred by linear transient impact. The total liquidation costs of X become

$$\int_{[0,T]} S_t^{X,Y}\, dX_t + \Big(\theta + \frac{\lambda}{2}\Big) \sum_{t\in[0,T]} (\Delta X_t)^2.$$

Now if there are jumps in both of X and Y, and some of them are at the same time, the liquidation costs are incurred jointly, i.e.,

$$\lambda \sum_{t\in[0,T]} \Delta X_t \Delta Y_t \neq 0.$$

These combined costs will be taken completely by one agent, whose trading is only slightly slower than the other's, if both of them execute "jump trades" at the same time. As same as the primary model, we assume that none of the two agents has an advantage in latency over the other. The execution priority is determined by an independent coin toss if both agents executes jump trades at the same time. In this sense, the liquidation costs of X and Y can be defined as

$$\begin{aligned}
&\widetilde{\mathcal{C}}(X|Y)\\
&= xS^0_{0-} + \int_{[0,T]} S_t^{X,Y}\, dX_t + \lambda \sum_{t\in[0,T]} \varepsilon_t(\Delta X_t \Delta Y_t) + \frac{\lambda}{2} \sum_{t\in[0,T]} (\Delta X_t)^2 + \theta \sum_{t\in[0,T]} (\Delta X_t)^2,\\
&\widetilde{\mathcal{C}}(Y|X)\\
&= yS^0_{0-} + \int_{[0,T]} S_t^{X,Y}\, dY_t + \lambda \sum_{t\in[0,T]} (1-\varepsilon_t)(\Delta X_t \Delta Y_t) + \frac{\lambda}{2} \sum_{t\in[0,T]} (\Delta Y_t)^2 + \theta \sum_{t\in[0,T]} (\Delta Y_t)^2;
\end{aligned}$$

where $(\varepsilon_t)_{t\in[0,T]}$ is an i.i.d. sequence of Bernoulli $(\frac{1}{2})$-distributed random variables that are independent of $\sigma(\bigcup_{t\geq 0} \mathcal{F}_t)$. In the following lemma, we see that the two approaches give the same expression of liquidation costs in expectation.

Lemma 3.3.3. *For $X \in \mathcal{X}(x,[0,T])$ and $Y \in \mathcal{X}(y,[0,T])$, it holds that*

$$\begin{aligned}
\mathbb{E}\Big[\widetilde{\mathcal{C}}(X|Y)\Big] &= \mathbb{E}\Big[\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dY_s\, dX_t\\
&\qquad + \frac{\lambda}{2}\sum_{t\in[0,T]} \Delta X_t \Delta Y_t + \theta \sum_{t\in[0,T]} (\Delta X_t)^2\Big],\\
\mathbb{E}\Big[\widetilde{\mathcal{C}}(Y|X)\Big] &= \mathbb{E}\Big[\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dY_s\, dY_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s\, dY_t\\
&\qquad + \frac{\lambda}{2}\sum_{t\in[0,T]} \Delta X_t \Delta Y_t + \theta \sum_{t\in[0,T]} (\Delta Y_t)^2\Big].
\end{aligned}$$

Proof.

$$
\begin{aligned}
&\mathbb{E}\Big[\widetilde{\mathcal{C}}(X|Y)\Big] \\
&= xS^0_{0-} + \mathbb{E}\bigg[\int_{[0,T]} \Big(S^0_t + \lambda \int_{[0,t)} e^{-\rho(t-s)}\, dX_s + \lambda \int_{[0,t)} e^{-\rho(t-s)}\, dY_s\Big)\, dX_t \\
&\quad + \frac{\lambda}{2} \sum_{t\in[0,T]} \Big((\Delta X_t)^2 + (\Delta X_t \Delta Y_t)\Big) + \theta \sum_{t\in[0,T]} (\Delta X_t)^2\bigg] \\
&= xS^0_{0-} + \mathbb{E}\bigg[\int_{[0,T]} S^0_t\, dX_t\bigg] + \mathbb{E}\bigg[\Big(\int_{[0,T]}\int_{[0,t)} e^{-\rho(t-s)}\, dX_s\, dX_t + \frac{\lambda}{2}\sum_{t\in[0,T]} (\Delta X_t)^2\Big) \\
&\quad + \Big(\int_{[0,T]}\int_{[0,t)} e^{-\rho(t-s)}\, dY_s\, dX_t + \frac{\lambda}{2}\sum_{t\in[0,T]} (\Delta X_t \Delta Y_t)\Big) + \theta \sum_{t\in[0,T]} (\Delta X_t)^2\bigg].
\end{aligned}
$$

Gatheral et al. [2012, Lemma 2.3] yields that

$$
xS^0_{0-} + \mathbb{E}\bigg[\int_{[0,T]} S^0_t\, dX_t\bigg] = 0.
$$

Moreover it holds that

$$
\mathbb{E}\bigg[\int_{[0,T]}\int_{[0,t)} e^{-\rho(t-s)}\, dX_s\, dX_t + \frac{\lambda}{2}\sum_{t\in[0,T]} (\Delta X_t)^2\bigg] = \mathbb{E}\bigg[\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t\bigg].
$$

Therefore, we obtain

$$
\begin{aligned}
\mathbb{E}\Big[\widetilde{\mathcal{C}}(X|Y)\Big] \;=\; &\mathbb{E}\bigg[\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dY_s\, dX_t \\
&+ \frac{\lambda}{2}\sum_{t\in[0,T]} \Delta X_t \Delta Y_t + \theta \sum_{t\in[0,T]} (\Delta X_t)^2\bigg].
\end{aligned}
$$

It is analogous to show the equation of $\mathbb{E}[\widetilde{\mathcal{C}}(Y|X)]$. This completes the proof. □

Following the two approaches discussed above, we define the liquidation costs.

Definition 3.3.4. Given initial asset positions $x, y \in \mathbb{R}$ and $T > 0$. The *liquidation costs* of $X \in \mathcal{X}(x, [0,T])$ given $Y \in \mathcal{X}(y, [0,T])$ are defined as

$$
\begin{aligned}
\mathcal{C}(X|Y) = &\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dY_s\, dX_t \\
&+ \frac{\lambda}{2}\sum_{t\in[0,T]} \Delta X_t \Delta Y_t + \theta \sum_{t\in[0,T]} (\Delta X_t)^2,
\end{aligned}
$$

and the *liquidation costs* of $Y \in \mathcal{X}(y, [0,T])$ given $X \in \mathcal{X}(x, [0,T])$ are defined as

$$
\begin{aligned}
\mathcal{C}(Y|X) = &\frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dY_s\, dY_t + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s\, dY_t \\
&+ \frac{\lambda}{2}\sum_{t\in[0,T]} \Delta X_t \Delta Y_t + \theta \sum_{t\in[0,T]} (\Delta Y_t)^2.
\end{aligned}
$$

Then we assume the two agents X and Y are competing against each other. There is no cooperation between them and both of them use open-loop strategies. Given initial asset positions, they minimize their expected liquidation costs accordingly. In such a situation, the optimality for both of them is defined again through Nash equilibrium.

Definition 3.3.5. For a given time horizon $[0,T]$ and initial asset positions x, $y \in \mathbb{R}$, a *Nash equilibrium* is a pair (X^*, Y^*) of strategies in $\mathcal{X}(x,[0,T]) \times \mathcal{X}(y,[0,T])$ such that

$$\mathbb{E}[\,\mathcal{C}(X^*|Y^*)\,] = \inf_{X\in\mathcal{X}(x,[0,T])} \mathbb{E}[\,\mathcal{C}(X|Y^*)\,] \quad \text{and} \quad \mathbb{E}[\,\mathcal{C}(Y^*|X^*)\,] = \inf_{Y\in\mathcal{X}(y,[0,T])} \mathbb{E}[\,\mathcal{C}(Y|X^*)\,].$$

3.3.2 Critical value of transaction costs

Now we analyze the existence of Nash equilibria and effects of transaction costs. Lorenz and Schied [2012] find that in a single-agent continuous-time model without transaction costs, there exists an optimal strategy only when the derivative of the price drift is absolutely continuous. On the other hand, we have seen in Section 3.1 that transaction costs help to eliminate oscillations and lead to the convergence of optimal strategies in the primary model. In the following theorem, we see that for any nontrivial case, a unique Nash equilibrium exits if and only if the transaction costs are exactly equal to a critical value.

Theorem 3.3.6. *Given $\rho > 0$, $\lambda > 0$, $T > 0$ and $x, y \in \mathbb{R}$ with $x \neq 0$ or $y \neq 0$. There exists a unique Nash equilibrium (X^*, Y^*) in the class $\mathcal{X}(x,[0,T])\times\mathcal{X}(y,[0,T])$ of adapted strategies if and only if $\theta = \lambda/4$. The optimal strategies X^* and Y^* are deterministic and given by*

$$\begin{aligned} X_t^* &= \frac{1}{2}(x+y)V_t + \frac{1}{2}(x-y)W_t, \\ Y_t^* &= \frac{1}{2}(x+y)V_t - \frac{1}{2}(x-y)W_t, \end{aligned} \tag{3.54}$$

where

$$\begin{aligned} V_t &= \frac{e^{3\rho T}\big(6\rho(T-t)+4\big) - 4e^{3\rho t}}{2e^{3\rho T}(3\rho T + 5) - 1} \quad \text{if } t \in [0,T], \text{ and } V_{0-} = 1, \\ W_t &= \frac{\rho(T-t)+1}{\rho T + 1} \quad \text{if } t \in [0,T), \text{ and } W_{0-} = 1, \quad W_T = 0. \end{aligned} \tag{3.55}$$

We use several lemmas to prove Theorem 3.3.6. Let $\mathcal{X}_{\mathrm{det}}(z,[0,T])$ denote the class of all deterministic strategies in $\mathcal{X}(z,[0,T])$. First we derive in Lemma 3.3.7 an equivalent condition for the optimality of a deterministic strategy $X^* \in \mathcal{X}_{\mathrm{det}}(x,[0,T])$ given another deterministic strategy $Y \in \mathcal{X}_{\mathrm{det}}(y,[0,T])$. This condition helps us to find Nash equilibria in the class $\mathcal{X}_{\mathrm{det}}(x,[0,T]) \times \mathcal{X}_{\mathrm{det}}(y,[0,T])$ of deterministic strategies. Then we find by Lemma 3.3.8 a Nash equilibrium in $\mathcal{X}_{\mathrm{det}}(x,[0,T]) \times \mathcal{X}_{\mathrm{det}}(y,[0,T])$ if the transaction costs θ are exactly equal to $\lambda/4$. In Lemmas 3.3.11 and 3.3.12, we show that this deterministic Nash equilibrium is also a Nash equilibrium in the class of adapted strategies and it is unique. At last we show by way of contradiction that there exists no Nash equilibrium in the class of adapted strategies $\mathcal{X}(x,[0,T]) \times \mathcal{X}(y,[0,T])$ if $\theta \neq \lambda/4$.

To simplify the notations, we define for the rest part of this section that

$$
\begin{aligned}
&C(Y,X) := \mathbb{E}\Big[\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\,dY_s\,dX_t\Big],\\
&C_1(Y,X) := \mathbb{E}\Big[\int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho|t-s|}\,dY_s\,dX_t\Big], \quad C_2(Y,X) := \mathbb{E}\Big[\sum_{t\in[0,T]} \Delta X_t \Delta Y_t\Big].
\end{aligned} \tag{3.56}
$$

Lemma 3.3.7. *Given an admissible deterministic strategy $Y \in \mathcal{X}_{det}(y,T)$, there exists an optimal deterministic strategy $X^* \in \mathcal{X}_{det}(x,T)$ minimizing the liquidation costs $\mathcal{C}(X|Y)$ in $\mathcal{X}_{det}(x,[0,T])$ if and only if there exists a constant $\eta \in \mathbb{R}$ such that for all $t \in [0,T]$,*

$$
\int_{[0,T]} \lambda e^{-\rho|t-s|}\,dX^*_s + \int_{[0,t)} \lambda e^{-\rho(t-s)}\,dY_s + \frac{\lambda}{2}\Delta Y_t + 2\theta \Delta X^*_t = \eta. \tag{3.57}
$$

Proof. We follow the proof of Gatheral et al. [2012, Theorem 2.11]. By the notation defined in (3.56) we rewrite the liquidation costs $\mathcal{C}(X|Y)$ as follows:

$$
\mathcal{C}(X|Y) = \frac{1}{2}C(X) + C_1(Y,X) + \frac{\lambda}{2}C_2(Y,X) + \theta C_2(X,X).
$$

Necessity. We refer to the proof of Lemma 3.3.13, which shows a more general case.

Sufficiency. Assume that $X^* \in \mathcal{X}_{det}(x,[0,T])$ satisfies (3.57). Let $Z \in \mathcal{X}_{\mathrm{det}}(0,[0,T])$ be a round trip. Then we have:

$$
\begin{aligned}
&\mathcal{C}(X^*+Z|Y)\\
&= \frac{1}{2}C(X^*+Z,X^*+Z) + C_1(Y,X^*+Z) + \frac{\lambda}{2}C_2(Y,X^*+Z) + \theta C_2(X^*+Z,X^*+Z)\\
&= \frac{1}{2}C(X^*,X^*) + C_1(Y,X^*) + \frac{\lambda}{2}C_2(Y,X^*) + \frac{1}{2}C(Z,Z) + \theta C_2(X^*,X^*)\\
&\qquad + C(X^*,Z) + C_1(Y,Z) + \frac{\lambda}{2}C_2(Y,Z) + 2\theta C_2(X^*,Z) + \theta C_2(Z,Z).
\end{aligned}
$$

Since

$$
\frac{1}{2}C(X^*,X^*) + C_1(Y,X^*) + \frac{\lambda}{2}C_2(Y,X^*) + \theta C_2(X^*,X^*) = \mathcal{C}(X^*|Y)
$$

and

$$
C(Z,Z) \ge 0, \quad C_2(Z,Z) \ge 0,
$$

it holds that

$$
\begin{aligned}
&\mathcal{C}(X^*+Z|Y)\\
&\ge \mathcal{C}(X^*|Y) + C(X^*,Z) + C_1(Y,Z) + \frac{\lambda}{2}C_2(Y,Z) + 2\theta C_2(X^*,Z)\\
&= \mathcal{C}(X^*|Y) + \int_{[0,T]}\Big(\int_{[0,T]} \lambda e^{-\rho|t-s|}\,dX_t + \int_{[0,s)} \lambda e^{-\rho|t-s|}\,dY_t + \frac{\lambda}{2}\Delta Y_s + 2\theta\Delta X^*_s\Big)\,dZ_s\\
&= \mathcal{C}(X^*|Y) + \eta(Z_T - Z_0) = \mathcal{C}(X^*|Y).
\end{aligned}
$$

Since every strategy $X \in \mathcal{X}_{\mathrm{det}}(x,[0,T])$ can be written as $X^* + Z$ for some round trip Z, this shows the optimality of X^*. □

Lemma 3.3.8. *Given $\rho > 0, \lambda > 0, T > 0$ and initial asset positions $x, y \in \mathbb{R}$. If $\theta = \frac{\lambda}{4}$, there exists a Nash equilibrium $(X^*, Y^*) \in \mathcal{X}_{det}(x, [0,T]) \times \mathcal{X}_{det}(y, [0,T])$ in the class of deterministic strategies. The optimal strategies X^* and Y^* are given by*

$$\begin{aligned} X_t^* &= \frac{1}{2}(x+y)V_t + \frac{1}{2}(x-y)W_t, \\ Y_t^* &= \frac{1}{2}(x+y)V_t - \frac{1}{2}(x-y)W_t, \end{aligned} \tag{3.58}$$

where

$$\begin{aligned} V_t &= \frac{e^{3\rho T}\big(6\rho(T-t)+4\big) - 4e^{3\rho t}}{2e^{3\rho T}(3\rho T+5) - 1} \quad \text{if } t \in [0,T], \text{ and } V_{0-} = 1, \\ W_t &= \frac{\rho(T-t)+1}{\rho T+1} \quad \text{if } t \in [0,T), \text{ and } W_{0-} = 1, W_T = 0. \end{aligned} \tag{3.59}$$

Proof. A straightforward computation yields that, for all $t \in [0,T]$,

$$\begin{aligned} &\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s^* + \int_{[0,t)} \lambda e^{-\rho(t-s)}\, dY_s^* + \frac{\lambda}{2}\Delta Y_t^* + 2\theta \Delta X_t^* \\ &= -\frac{1}{2}\Big(\frac{\lambda(x-y)}{\rho T+1} + \frac{18\lambda(x+y)}{10+6\rho T - e^{-3\rho T}}\Big), \\ &\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dY_s^* + \int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s^* + \frac{\lambda}{2}\Delta X_t^* + 2\theta \Delta Y_t^* \\ &= -\frac{1}{2}\Big(\frac{18\lambda(x+y)}{10+6\rho T - e^{-3\rho T}} - \frac{\lambda(x-y)}{\rho T+1}\Big). \end{aligned}$$

Applying Lemma 3.3.7 completes this proof. □

Remark 3.3.9 (Heuristic derivation of W and V)**.** We explain how we heuristicly derive the formula of W and V in Lemma 3.3.8.

First we show the derivation of W. According to Lemma 3.3.7, there exit $\eta_1, \eta_2 \in \mathbb{R}$ such that for all $t \in [0,T]$,

$$\begin{aligned} &\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s^* + \int_{[0,t)} \lambda e^{-\rho(t-s)}\, dY_s^* + \frac{\lambda}{2}\Delta Y_t^* + 2\theta \Delta X_t^* = \eta_1, \\ &\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dY_s^* + \int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s^* + \frac{\lambda}{2}\Delta X_t^* + 2\theta \Delta Y_t^* = \eta_2. \end{aligned} \tag{3.60}$$

We define W through $X_t^* - Y_t^* = (x-y)W_t$. Then we have $W_{0-} = 1$ and $W_T = 0$. By subtracting the second equation in (3.60) from the first, we obtain

$$\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dW_s - \int_{[0,t)} \lambda e^{-\rho(t-s)}\, dW_s + \Big(2\theta - \frac{\lambda}{2}\Big)\Delta W_t = \eta_1 - \eta_2, \quad \text{for all } t \in [0,T]. \tag{3.61}$$

Setting $t = T$ yields

$$\Big(2\theta + \frac{\lambda}{2}\Big)\Delta W_T = \eta_1 - \eta_2. \tag{3.62}$$

We assume in addition that X^* and Y^* are absolutely continuous in $(0,T)$. Then we have $\Delta W_t = 0$ for $t \in (0,T)$ and

$$
\begin{aligned}
&\int_{[0,T]} \lambda e^{-\rho|t-s|}\,dW_s - \int_{[0,t)} \lambda e^{-\rho(t-s)}\,dW_s + (2\theta - \frac{\lambda}{2})\Delta W_t \\
&= \int_{(t,T]} \lambda e^{-\rho(s-t)}\,dW_s = \int_{(t,T)} \lambda e^{-\rho(s-t)}\,dW_s + \lambda e^{-\rho(T-t)}\Delta W_T \\
&= \lambda\Big(e^{-\rho(s-t)}W_s\Big|_{t+}^{T-} + \rho\int_{(t,T)} W_s e^{-\rho(s-t)}\,ds + e^{-\rho(T-t)}\Delta W_T\Big) \\
&= \lambda\Big(-e^{-\rho(T-t)}\Delta W_T - W_t + \rho\int_{(t,T)} W_s e^{-\rho(s-t)}\,ds + e^{-\rho(T-t)}\Delta W_T\Big) \\
&= \eta_1 - \eta_2.
\end{aligned}
$$

By (3.61) and (3.62), the latter expression must be equal to $\Big(2\theta + \frac{\lambda}{2}\Big)\Delta W_T$. It then follows that

$$
\lambda W_t = \lambda\rho\int_{(t,T)} W_s e^{-\rho(s-t)}\,ds - \Big(2\theta + \frac{\lambda}{2}\Big)\Delta W_T, \quad \text{for all} \quad t \in (0,T). \tag{3.63}
$$

Since

$$
\lim_{t\to T-} W_t = W_{T-} = -\Delta W_T, \qquad \lim_{t\to T-}\int_{(t,T)} W_s e^{-\rho(s-t)}\,ds = 0,
$$

we obtain

$$
\lambda\Delta W_T = \Big(2\theta + \frac{\lambda}{2}\Big)\Delta W_T.
$$

This implies that $\theta = \lambda/4$ or $\Delta W_T = 0$. If $\Delta W_T = 0$, it holds that

$$
W_t = \rho\int_{(t,T)} W_s e^{-\rho(s-t)}\,ds, \quad \text{for all} \quad t \in (0,T). \tag{3.64}
$$

Since W_t is absolutely continuous, we may differentiate both sides of (3.64) with respect to t and get

$$
W_t' = \rho\Big(\rho e^{\rho t}\int_{(t,T)} W_s e^{-\rho s}\,ds - e^{\rho t}W_t e^{-\rho t}\Big) = -\rho(W_t - W_t) = 0,
$$

which means $W_t = c \in \mathbb{R}$, for almost all $t \in (0,T)$. Since $\Delta W_T = 0$, it follows that $c = 0$ and $\Delta W_0 = -1$. However, by setting $t = 0$ in (3.61), we obtain

$$
\Big(2\theta + \frac{\lambda}{2}\Big)\Delta W_0 + \lambda e^{-\rho T}\Delta W_T + \int_{(0,T)} \lambda e^{-\rho s}\,dW_s = \eta_1 - \eta_2 = \Big(2\theta + \frac{\lambda}{2}\Big)\Delta W_T. \tag{3.65}
$$

This implies $\Delta W_0 = 0$ and shows $\Delta W_T \neq 0$. Therefore, we must have $\theta = \lambda/4$.

Now let $\theta = \lambda/4$. Note that

$$
\begin{aligned}
&\int_{(0,T)} e^{-\rho s}\,dW_s \\
&= e^{-\rho s}W_s\Big|_{0+}^{T-} + \rho\int_{(0,T)} W_s e^{-\rho s}\,ds = -e^{-\rho T}\Delta W_T - (\Delta W_0 + 1) + \rho\int_{(0,T)} W_s e^{-\rho s}\,ds.
\end{aligned}
$$

Thus, equation (3.65) becomes

$$\rho \int_{(0,T)} W_s e^{-\rho s}\, ds - 1 = \Delta W_T.$$

Putting this result into (3.63) yields

$$W_t = \rho \int_{(t,T)} W_s e^{-\rho(s-t)}\, ds + 1 - \rho \int_{(0,T)} W_s e^{-\rho s}\, ds, \quad \text{for all} \quad t \in (0,T).$$

Taking the derivative with respect to t yields

$$W_t' = \rho\Big(\rho \int_{(0,T)} W_s e^{-\rho s}\, ds - 1\Big) \in \mathbb{R}. \tag{3.66}$$

Solving (3.66) yields

$$W_t = \frac{\rho(T-t)+1}{\rho T + 1} \quad \text{if } t \in [0,T), \text{ and } W_T = 0.$$

Now we show the derivation of V. Assume that $\theta = \lambda/4$. By Lemma 3.3.7, there exists a constant $\eta \in \mathbb{R}$, such that for all $t \in [0,T]$, it holds

$$\int_{[0,T]} e^{-\rho|t-s|}\, dV_s + \int_{[0,t)} e^{-\rho(t-s)}\, dV_s + \Delta V_t = \eta, \tag{3.67}$$

where V is defined through $X_t^* + Y_t^* = (x+y)V_t$. And we have $V_{0-} = 1$ and $V_T = 0$. We furthermore assume that V_t is absolutely continuous in $(0,T)$. Then (3.67) turns to be

$$2e^{-\rho t}\Delta V_0 + 2\int_{(0,t)} e^{-\rho(t-s)}\, dV_s + \int_{(t,T)} e^{-\rho(s-t)}\, dV_s = \eta.$$

Then we have

$$\begin{aligned}
&\int_{(0,t)} e^{-\rho(t-s)}\, dV_s \\
&= e^{-\rho(t-s)} V_s \Big|_{s=0+}^{s=t-} - \rho \int_{(0,t)} V_s e^{-\rho(t-s)}\, ds = V_t - e^{-\rho t} - e^{-\rho t}\Delta V_0 - \rho \int_{(0,t)} V_s e^{-\rho(t-s)}\, ds, \\
&\int_{(t,T)} e^{-\rho(s-t)}\, dV_s \\
&= e^{-\rho(t-s)} V_s \Big|_{s=t+}^{s=T-} + \rho \int_{(t,T)} V_s e^{-\rho(s-t)}\, ds = -e^{-\rho(T-t)}\Delta V_T - V_t + \rho \int_{(t,T)} V_s e^{-\rho(t-s)}\, ds.
\end{aligned}$$

Hence, we obtain

$$V_t - 2e^{-\rho t} - 2\rho \int_{(0,t)} V_s e^{-\rho(t-s)}\, ds + \rho \int_{(t,T)} V_s e^{-\rho(s-t)}\, ds = \eta, \qquad \text{if } t \in (0,T). \tag{3.68}$$

Similarly,

$$\begin{aligned}
\Delta V_0 - 1 + \rho \int_{(0,T)} V_s e^{-\rho s}\, ds &= \eta, \qquad \text{if } t = 0; \\
-2e^{-\rho T} - 2\rho \int_{(0,T)} V_s e^{-\rho(T-s)}\, ds &= \eta, \qquad \text{if } t = T.
\end{aligned} \tag{3.69}$$

Note that $\lim_{t\to T-} V_t = V_{T-} = -\Delta V_T$ and $\lim_{t\to T-} \int_{(t,T)} V_s e^{-\rho(s-t)}\, ds = 0$. We obtain

$$-\Delta V_T - 2e^{-\rho T} - 2\rho \int_{(0,T)} V_s e^{-\rho(T-s)}\, ds = \eta = -2e^{-\rho T} - 2\rho \int_{(0,T)} V_s e^{-\rho(T-s)}\, ds.$$

This shows $\Delta V_T = 0$.

We assume in addition V is twice differentiable in $(0,T)$. By differentiating (3.68) twice with respect to t, we get

$$V_t'' + \rho^2\Big(- 2e^{-\rho t} - 2\rho \int_{(0,t)} V_s e^{-\rho(t-s)}\, ds + \rho \int_{(t,T)} V_s e^{-\rho(s-t)}\, ds\Big) - 3\rho V_t' + \rho^2 V_t = 0. \quad (3.70)$$

Putting (3.68) into (3.70) yields that

$$V_t'' - 3\rho V_t' + \rho^2 \eta = 0, \quad (3.71)$$

which has the solution

$$V_t = \frac{e^{3\rho t}}{3\rho} c_1 + \frac{\eta \rho t}{3} + c_2, \qquad c_1, c_2 \in \mathbb{R}.$$

Applying the terminal condition $V_T = 0$ yields

$$V_t = \frac{-(T-t)\rho^2\eta + c\Big(e^{3\rho t} - e^{3\rho T}\Big)}{3\rho}, \qquad c \in \mathbb{R}. \quad (3.72)$$

The constants c and η can be derived explicitly by putting (3.72) into (3.69):

$$c = -\frac{12\rho}{6\rho T e^{3\rho T} + 10 e^{3\rho T} - 1}, \qquad \eta = -\frac{18 e^{3\rho T}}{6\rho T e^{3\rho T} + 10 e^{3\rho T} - 1}.$$

Therefore we obtain

$$V_t = \frac{e^{3\rho T}\big(6\rho(T-t) + 4\big) - 4e^{3\rho t}}{2e^{3\rho T}(3\rho T + 5) - 1} \quad \text{if } t \in (0,T], \text{ and } V_0 = 1.$$

In the end, one can validate V by checking

$$\int_{[0,T]} e^{-\rho|t-s|}\, dV_s + \int_{[0,t)} e^{-\rho(t-s)}\, dV_s + \Delta V_t = -\frac{18 e^{3\rho T}}{6\rho T e^{3\rho T} + 10 e^{3\rho T} - 1}, \qquad \text{for all } t \in [0,T].$$

□

Lemma 3.3.10. *Given $T > 0, \rho > 0, \lambda > 0$, $\theta \geq 0$, initial asset positions $x, y \in \mathbb{R}$ and an admissible strategy $Y \in \mathcal{X}(y,[0,T])$, the functional $\mathbb{E}[\mathcal{C}(X|Y)]$ is strictly convex with respect to $X \in \mathcal{X}(x,[0,T])$.*

Proof. Let $\alpha \in (0,1)$ and $X^0, X^1 \in \mathcal{X}(x,[0,T])$ be two distinct admissible strategies. Since the function $t \mapsto \lambda e^{-\rho t}$ is positive definite, we obtain

$$\mathbb{E}\Big[\int_{[0,T]} \int_{[0,T]} \lambda e^{-\rho|s-t|}\, d\big(X_s^1 - X_s^0\big)\, d\big(X_t^1 - X_t^0\big)\Big] > 0.$$

Hence,

$$\begin{aligned}
&C(\alpha X^1+(1-\alpha)X^0, \alpha X^1+(1-\alpha)X^0) \\
&< C(\alpha X^1+(1-\alpha)X^0, \alpha X^1+(1-\alpha)X^0)+\alpha(1-\alpha)C(X^1-X^0, X^1-X^0) \\
&= \alpha^2 C(X^1,X^1)+(1-\alpha)^2 C(X^0,X^0)+2\alpha(1-\alpha)C(X^1,X^0) \\
&\qquad +\alpha(1-\alpha)C(X^1,X^1)-2\alpha(1-\alpha)C(X^1,X^0)+\alpha(1-\alpha)C(X^0,X^0) \\
&= \alpha C(X^1,X^1)+(1-\alpha)C(X^0,X^0).
\end{aligned}$$

By the same way, we have

$$\sum_{t\in[0,T]}\Big(\Delta\big(\alpha X_t^1+(1-\alpha X_t^0)\big)\Big)^2 < \alpha\sum_{t\in[0,T]}\big(\Delta X_t^1\big)^2+(1-\alpha)\sum_{t\in[0,T]}\big(\Delta X_t^0\big)^2.$$

Therefore,

$$\mathbb{E}\Big[\mathcal{C}(\alpha X^1+(1-\alpha)X^0|Y)\Big] < \alpha\mathbb{E}\Big[\mathcal{C}(X^1|Y)\Big]+(1-\alpha)\mathbb{E}\Big[\mathcal{C}(X^0|Y)\Big] \quad \text{for all} \quad \alpha\in(0,1).$$

□

Lemma 3.3.11. *Given $T>0, \rho>0, \lambda>0$, $\theta\geq 0$ and initial asset positions $x, y\in\mathbb{R}$, there exists at most one Nash equilibrium in the class $\mathcal{X}(x,[0,T])\times\mathcal{X}(y,[0,T])$ of adapted strategies.*

Proof. We assume by way of contradiction that there exist two distinct Nash equilibria (X^0, Y^0) and (X^1, Y^1) in $\mathcal{X}(x,[0,T])\times\mathcal{X}(y,[0,T])$. Then we define for $\alpha\in[0,1]$

$$X^\alpha := \alpha X^1+(1-\alpha)X^0 \qquad \text{and} \qquad Y^\alpha := \alpha Y^1+(1-\alpha)Y^0.$$

We furthermore let

$$f(\alpha) := \mathbb{E}\Big[\mathcal{C}(X^\alpha|Y^0)+\mathcal{C}(Y^\alpha|X^0)+\mathcal{C}(X^{1-\alpha}|Y^1)+\mathcal{C}(Y^{1-\alpha}|X^1)\Big].$$

According to Lemma 3.3.10, each term of $f(\alpha)$ is strictly convex in α. Since the two Nash equilibria (X^0, Y^0) and (X^1, Y^1) are distinct, $f(\alpha)$ must also be convex in α and have its unique minimun in $\alpha=0$. That is

$$f(\alpha)\geq f(0) \quad \text{for} \quad \alpha>0.$$

It follows that

$$\lim_{h\downarrow 0}\frac{f(h)-f(0)}{h} = \frac{df(\alpha)}{d\alpha}\Big|_{\alpha=0+}\geq 0. \tag{3.73}$$

On the other hand, we have

$$\begin{aligned}
\frac{d}{d\alpha}\mathbb{E}[\mathcal{C}(X^\alpha|Y^0)]\Big|_{\alpha=0} &= C(X^0, X^1-X^0)+C_1(Y^0, X^1-X^0) \\
&\quad +\frac{\lambda}{2}C_2(Y^0, X^1-X^0)+2\theta C_2(X^0, X^1-X^0).
\end{aligned}$$

In the same way, we take derivatives of $\mathbb{E}[\mathcal{C}(Y^\alpha|X^0)]$, $\mathbb{E}[\mathcal{C}(X^{1-\alpha}|Y^1)]$ and $\mathbb{E}[\mathcal{C}(Y^{1-\alpha}|X^1)]$ with respect to α, set $\alpha = 0$ and then put them together. Then we have

$$\begin{aligned}
&\frac{df(\alpha)}{d\alpha}\Big|_{\alpha=0} \\
&= -C(X^1 - X^0, X^1 - X^0) - C(Y^1 - Y^0, Y^1 - Y^0) - C(Y^1 - Y^0, Y^1 - Y^0) \\
&\quad - 2\theta\Big(C_2(X^1 - X^0, X^1 - X^0) + C_2(Y^1 - Y^0, Y^1 - Y^0)\Big) \\
&< -\frac{1}{2}C(X^1 - X^0, X^1 - X^0) - \frac{1}{2}C(Y^1 - Y^0, Y^1 - Y^0) - \frac{1}{2}C(Y^1 - Y^0, X^1 - X^0) \\
&< 0,
\end{aligned}$$

which contradicts (3.73). Therefore, there exists at most one Nash equilibrium in the class $\mathcal{X}(x, [0,T]) \times \mathcal{X}(y, [0,T])$ of adapted strategies. □

Lemma 3.3.12. *A Nash equilibrium in the class $\mathcal{X}_{det}(x, [0,T]) \times \mathcal{X}_{det}(y, [0,T])$ of deterministic strategies is also a Nash equilibrium in the class $\mathcal{X}(x, [0,T]) \times \mathcal{X}(y, [0,T])$ of adapted strategies.*

Proof. Let $(X^*, Y^*) \in \mathcal{X}_{\mathrm{det}}(x, [0,T]) \times \mathcal{X}(y, [0,T])$ be a Nash equilibrium in the class of deterministic strategies. For any strategy $X \in \mathcal{X}(x, [0,T])$, we have

$$\mathcal{C}(X(\omega)|Y^*)) \geq \mathcal{C}(X^*|Y^*) \quad \text{for almost all } \omega \in \Omega.$$

Therefore we obtain $\mathbb{E}[\mathcal{C}(X|Y^*)] \geq \mathcal{C}(X^*|Y^*)$ with equality if and only if $\mathcal{C}(X|Y^*) = \mathcal{C}(X^*|Y^*)$ $\mathbb{P}$-a.s. This shows the optimality of X^* within the class $\mathcal{X}(x, [0,T])$ of adaptive strategies. Analogous we obtain the optimality of Y^* within the class $\mathcal{X}(y, [0,T])$ of adaptive strategies. This completes the proof. □

Lemma 3.3.13. *Given an adapted strategy $Y \in \mathcal{X}(y, [0,T])$, if there exists an optimal strategy X^* minimizing the expected liquidation costs $\mathbb{E}[\mathcal{C}(X|Y)]$ in $\mathcal{X}(x, [0,T])$, then for any stopping time $\tau \in [0,T]$, there exists an $\mathcal{F}_\tau$-measurable random variable η such that X^* solves the following integral equation $\mathbb{P}$-a.s., for all stopping time $\sigma \in [\tau, T]$,*

$$\mathbb{E}\left[\int_{[0,T]} \lambda e^{-\rho|\sigma - t|}\, dX_t^* + \int_{[0,\sigma)} \lambda e^{-\rho(\sigma - t)}\, dY_t + \frac{\lambda}{2}\Delta Y_\sigma + 2\theta \Delta X_\sigma^* \,\Big|\, \mathcal{F}_\tau\right] = \eta.$$

Proof. We first note that

$$\mathbb{E}[\mathcal{C}(X|Y)] = \frac{1}{2}C(X,X) + C_1(Y,X) + \frac{\lambda}{2}C_2(Y,X) + \theta C_2(X,X).$$

Let $\tau \in [0,T]$, $\sigma \in [\tau, T]$ be stopping time, δ be Dirac measure and $A \in \mathcal{F}_\tau$. We define a round trip Z by

$$dZ_t = \mathbb{1}_A\Big(\delta_\tau(dt) - \delta_\sigma(dt)\Big).$$

For $\alpha \in \mathbb{R}$, we have

$$
\begin{aligned}
&\mathbb{E}[\mathcal{C}(X+\alpha Z|Y)]\\
&= \frac{1}{2}C(X+\alpha Z, X+\alpha Z) + C_1(Y, X+\alpha Z) + \frac{\lambda}{2}C_2(Y, X+\alpha Z) + \theta C_2(X+\alpha Z, X+\alpha Z)\\
&= \frac{1}{2}\big(C(X,X) + 2\alpha C(Z,X) + \alpha^2 C(Z,Z)\big) + C_1(Y,X) + \alpha C_1(Y,Z) + \frac{\lambda}{2}C_2(Y,X)\\
&\quad + \frac{\lambda}{2}\alpha C_2(Y,Z) + \theta\big(C_2(X,X) + 2\alpha C_2(X,Z) + \alpha^2 C_2(Z,Z)\big).
\end{aligned}
$$

We take the derivative with respect to α at $\alpha = 0$. Hence, a necessary condition for the optimality is:

$$
0 = C(Z,X) + C_1(Y,Z) + \frac{\lambda}{2}C_2(Y,Z) + 2\theta C_2(X,Z). \tag{3.74}
$$

We have

$$
\begin{aligned}
C(Z,X) &= \mathbb{E}\Big[\mathbb{1}_A\Big(\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dZ_s\, dX_t\Big)\Big] = \mathbb{E}\Big[\mathbb{1}_A\Big(\int_{[0,T]} \big(\lambda e^{-\rho|\tau-t|} - \lambda e^{-\rho|\sigma-t|}\big)\, dX_t\Big)\Big],\\
C_1(Y,Z) &= \mathbb{E}\Big[\int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho|t-s|}\, dY_s\, dZ_t\Big] = \mathbb{E}\Big[\mathbb{1}_A\Big(\int_{[0,\tau)} \lambda e^{-\rho|\tau-t|}\, dY_t - \int_{[0,\sigma)} \lambda e^{-\rho|\sigma-t|}\, dY_t\Big)\Big],\\
\frac{\lambda}{2}C_2(Y,Z) &= \mathbb{E}\Big[\frac{\lambda}{2}\int_{[0,T]} \Delta Y_t\, dZ_t\Big] = \mathbb{E}\Big[\frac{\lambda}{2}\mathbb{1}_A\Big(\Delta Y_\tau - \Delta Y_\sigma\Big)\Big],\\
2\theta C_2(X,Z) &= \mathbb{E}\Big[2\theta \mathbb{1}_A\Big(\Delta X_\tau - \Delta X_\sigma\Big)\Big].
\end{aligned}
$$

Then (3.74) becomes

$$
\begin{aligned}
0 =&\mathbb{E}\Big[\mathbb{1}_A\Big(\int_{[0,T]} \big(\lambda e^{-\rho|\tau-t|} - \lambda e^{-\rho|\sigma-t|}\big)\, dX_t + \int_{[0,\tau)} \lambda e^{-\rho|\tau-t|}\, dY_t - \int_{[0,\sigma)} \lambda e^{-\rho|\sigma-t|}\, dY_t\\
&+ \frac{\lambda}{2}\left(\Delta Y_\tau - \Delta Y_\sigma\right) + 2\theta\left(\Delta X_\tau - \Delta X_\sigma\right)\Big)\Big].
\end{aligned}
$$

This implies that for all $A \in \mathcal{F}_\tau$ and for all $\sigma \geq \tau$,

$$
\begin{aligned}
&\mathbb{E}\Big[\mathbb{1}_A\Big(\int_{[0,T]} \lambda e^{-\rho|\sigma-t|}\, dX_t^* + \int_{[0,\sigma)} \lambda e^{-\rho(\sigma-t)}\, dY_t + \frac{\lambda}{2}\Delta Y_\sigma + 2\theta \Delta X_\sigma^*\Big)\Big]\\
=&\mathbb{E}\Big[\mathbb{1}_A\Big(\int_{[0,T]} \lambda e^{-\rho|\tau-t|}\, dX_t^* + \int_{[0,\tau)} \lambda e^{-\rho(\tau-t)}\, dY_t + \frac{\lambda}{2}\Delta Y_\tau + 2\theta \Delta X_\tau^*\Big)\Big].
\end{aligned}
$$

By the definition of conditional expectation, there exists a $\mathcal{F}_\tau$-measurable random variable η, such that X^* solves the following integral equation for all stopping time $\sigma \in [\tau, T]$

$$
\mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\sigma-t|}\, dX_t^* + \int_{[0,\sigma)} \lambda e^{-\rho(\sigma-t)}\, dY_t + \frac{\lambda}{2}\Delta Y_\sigma + 2\theta \Delta X_\sigma^* \Big|\, \mathcal{F}_\tau\Big] = \eta, \ \mathbb{P}\text{-a.s.}
$$

□

Proof of Theorem 3.3.6. Putting Lemmas 3.3.8, 3.3.12 and 3.3.11 together shows that if $\theta = \lambda/4$, the strategies $(X^*, Y^*) \in \mathcal{X}_{\mathrm{det}}(x, [0,T]) \times \mathcal{X}_{\mathrm{det}}(y, [0,T])$ stated in (3.58) form a unique Nash equilibrium in the class $\mathcal{X}(x, [0,T]) \times \mathcal{X}(y, [0,T])$ of adapted strategies.

Now we show that if $\theta \neq \lambda/4$, there exists no Nash equilibrium in the class of adapted strategies. To this end, we suppose by contradiction that there exists a Nash equilibrium $(X^*, Y^*) \in \mathcal{X}(x, [0,T]) \times \mathcal{X}(y, [0,T])$ and $\theta \neq \lambda/4$. By Lemma 3.3.13, for any stopping time $\tau \in [0,T]$, there exists an $\mathcal{F}_\tau$-measurable random variable η, such that X^* solves the following integral equation for all stopping time $\sigma \in [\tau, T]$

$$\mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\sigma-s|}\, dX_s^* + \int_{[0,\sigma)} \lambda e^{-\rho(\sigma-s)}\, dY_s^* + \frac{\lambda}{2}\Delta Y_\sigma^* + 2\theta \Delta X_\sigma^* \,\Big|\, \mathcal{F}_\tau\Big] = \eta. \tag{3.75}$$

We define a stopping time $\tau := \inf\{t \mid \Delta X_t^* \neq 0\}$. If $\tau < T$, let $(\tau_n)_{n\in\mathbb{N}}$ be a sequence of stopping time with $\Delta X_{\tau_n}^* = 0$ and $\tau_n \downarrow \tau$ as $n \uparrow \infty$. This sequence $(\tau_n)_{n\in\mathbb{N}}$ is well defined because by Definition 3.3.1 admissible strategies are of bounded total variation and there are at most only countable many jumps in $[0,T]$.

Since $\mathbb{E}[\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s^*]$ is continuous with respect to t, we have

$$\lim_{n\uparrow\infty} \mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\tau_n-s|}\, dX_s^*\Big] = \mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\tau-s|}\, dX_s^*\Big]. \tag{3.76}$$

Furthermore, it holds that for all $n \in \mathbb{N}$

$$\Big|\int_{[0,\tau_n)} \lambda e^{-\rho(\tau_n-s)}\, dY_s^*\Big| \leq \lambda|y| < \infty.$$

By dominated convergence theorem, we have

$$\mathbb{E}\Big[\lim_{n\uparrow\infty} \int_{[0,\tau_n)} \lambda e^{-\rho(\tau_n-s)}\, dY_s^*\Big] = \mathbb{E}\Big[\int_{[0,\tau)} \lambda e^{-\rho(\tau-s)}\, dY_s^*\Big]. \tag{3.77}$$

Putting (3.75), (3.76) and (3.77) together yields that

$$\begin{aligned} \eta &= \lim_{n\to\infty} \mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\tau_n-s|}\, dX_s^* + \int_{[0,\tau_n)} \lambda e^{-\rho(\tau_n-s)}\, dY_s^* + \frac{\lambda}{2}\Delta Y_{\tau_n}^* + 2\theta\Delta X_{\tau_n}^* \,\Big|\, \mathcal{F}_\tau\Big] \\ &= \mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\tau-s|}\, dX_s^* + \int_{[0,\tau]} \lambda e^{-\rho(\tau-s)}\, dY_s^* \,\Big|\, \mathcal{F}_\tau\Big]. \end{aligned} \tag{3.78}$$

On the other hand, setting $\sigma = \tau$ in (3.75) yields that

$$\eta = \mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\tau-s|}\, dX_s^* + \int_{[0,\tau)} \lambda e^{-\rho(\tau-s)}\, dY_s^* + \frac{\lambda}{2}\Delta Y_\tau^* + 2\theta\Delta X_\tau^* \,\Big|\, \mathcal{F}_\tau\Big]. \tag{3.79}$$

By subtracting (3.79) from (3.78), we obtain

$$\mathbb{E}\Big[\frac{\lambda}{2}\Delta Y_\tau^* + 2\theta\Delta X_\tau^* \,\Big|\, \mathcal{F}_\tau\Big] = \frac{\lambda}{2}\Delta Y_\tau^* + 2\theta\Delta X_\tau^* = 0. \tag{3.80}$$

By the same way, we obtain

$$\frac{\lambda}{2}\Delta X_\tau^* + 2\theta\Delta Y_\tau^* = 0. \tag{3.81}$$

Combining (3.80) and (3.81) yields

$$\begin{aligned}&\Big(\frac{\lambda}{2}-2\theta\Big)(\Delta X^*_\tau-\Delta Y^*_\tau)=0,\\&\Big(\frac{\lambda}{2}+2\theta\Big)(\Delta X^*_\tau+\Delta Y^*_\tau)=0.\end{aligned}\tag{3.82}$$

Since $\theta\neq\lambda/4$ and $\lambda>0$, we must have

$$\Delta X^*_\tau=\Delta Y^*_\tau=0.$$

This contradicts the definition of τ. Therefore, X^*_t must be continuous for $t\in[0,T)$. Similarly, Y^*_t must be also continuous for $t\in[0,T)$.

In the next step we show $\Delta X^*_T=\Delta Y^*_T=0$ if $\theta\neq\lambda/4$. To this end, let $(\tau_n)_{n\in\mathbb{N}}$ be a sequence of stopping time in $[0,T)$ with $\lim_{n\uparrow\infty}\tau_n=T$, $\mathbb{P}$-a.s. For each $n\in\mathbb{N}$, let $\big(\sigma^{(n)}_m\big)_{m\in\mathbb{N}}$ be a sequence of stopping time in $[\tau_n,T)$ with $\lim_{m\uparrow\infty}\sigma^{(n)}_m=T$, $\mathbb{P}$-a.s. By Lemma 3.3.13, for each $n\in\mathbb{N}$ there exists an $\mathcal{F}_{\tau_n}$-measurable random variable η_n such that X^* solves the following integral equation for all $\big(\sigma^{(n)}_m\big)_{m\in\mathbb{N}}$,

$$\begin{aligned}\eta_n&=\mathbb{E}\bigg[\int_{[0,T]}\lambda e^{-\rho|\sigma^{(n)}_m-s|}\,dX^*_s+\int_{[0,\sigma^{(n)}_m)}\lambda e^{-\rho(\sigma^{(n)}_m-s)}\,dY^*_s+\frac{\lambda}{2}\Delta Y^*_{\sigma^{(n)}_m}+2\theta\Delta X^*_{\sigma^{(n)}_m}\,\bigg|\,\mathcal{F}_{\tau_n}\bigg]\\&=\mathbb{E}\bigg[\int_{[0,T]}\lambda e^{-\rho|\sigma^{(n)}_m-s|}\,dX^*_s+\int_{[0,\sigma^{(n)}_m)}\lambda e^{-\rho(\sigma^{(n)}_m-s)}\,dY^*_s\,\bigg|\,\mathcal{F}_{\tau_n}\bigg].\end{aligned}$$

Since X^*_t and Y^*_t are continuous for $t\in[0,T)$, we have

$$\begin{aligned}\eta_n&=\lim_{m\uparrow\infty}\mathbb{E}\bigg[\int_{[0,T]}\lambda e^{-\rho|\sigma^{(n)}_m-s|}\,dX^*_s+\int_{[0,\sigma^{(n)}_m)}\lambda e^{-\rho(\sigma^{(n)}_m-s)}\,dY^*_s\,\bigg|\,\mathcal{F}_{\tau_n}\bigg]\\&=\mathbb{E}\bigg[\int_{[0,T]}\lambda e^{-\rho|T-s|}\,dX^*_s+\int_{[0,T)}\lambda e^{-\rho(T-s)}\,dY^*_s\,\bigg|\,\mathcal{F}_{\tau_n}\bigg].\end{aligned}\tag{3.83}$$

On the other hand, we also have

$$\eta_n=\mathbb{E}\bigg[\int_{[0,T]}\lambda e^{-\rho|T-s|}\,dX^*_s+\int_{[0,T)}\lambda e^{-\rho(T-s)}\,dY^*_s+\frac{\lambda}{2}\Delta Y^*_T+2\theta\Delta X^*_T\,\bigg|\,\mathcal{F}_{\tau_n}\bigg].\tag{3.84}$$

Comparing (3.83) and (3.84) yields that

$$\mathbb{E}\Big[\frac{\lambda}{2}\Delta Y^*_T+2\theta\Delta X^*_T\,\Big|\,\mathcal{F}_{\tau_n}\Big]=0.\tag{3.85}$$

Analogously, we also obtain that

$$\mathbb{E}\Big[\frac{\lambda}{2}\Delta X^*_T+2\theta\Delta Y^*_T\,\Big|\,\mathcal{F}_{\tau_n}\Big]=0.\tag{3.86}$$

Note that

$$\sigma\Big(\bigcup_{n=0}^{\infty}\mathcal{F}_{\tau_n}\Big)\subset\mathcal{F}_{T-},\ \ \Delta X^*_T=-X^*_{T-},\ \ \Delta Y^*_T=-Y^*_{T-}\text{ are }\mathcal{F}_{T-}\text{-measurable.}$$

Since

$$\Big|\frac{\lambda}{2}\Delta Y_T^* + 2\theta \Delta X_T^*\Big| < \infty, \qquad \Big|\frac{\lambda}{2}\Delta X_T^* + 2\theta \Delta Y_T^*\Big| < \infty,$$

martingale convergence theorem, see Theorem 7.23 in Kallenberg [2002], yields that

$$\begin{aligned}\lim_{n\uparrow\infty} \mathbb{E}\Big[\frac{\lambda}{2}\Delta Y_T^* + 2\theta\Delta X_T^* \,\Big|\, \mathcal{F}_{\tau_n}\Big] &= \mathbb{E}\Big[\frac{\lambda}{2}\Delta Y_T^* + 2\theta\Delta X_T^* \,\Big|\, \mathcal{F}_{T-}\Big] = \frac{\lambda}{2}\Delta Y_T^* + 2\theta\Delta X_T^* = 0,\\ \lim_{n\uparrow\infty} \mathbb{E}\Big[\frac{\lambda}{2}\Delta X_T^* + 2\theta\Delta Y_T^* \,\Big|\, \mathcal{F}_{\tau_n}\Big] &= \mathbb{E}\Big[\frac{\lambda}{2}\Delta X_T^* + 2\theta\Delta Y_T^* \,\Big|\, \mathcal{F}_{T-}\Big] = \frac{\lambda}{2}\Delta X_T^* + 2\theta\Delta Y_T^* = 0.\end{aligned} \tag{3.87}$$

Since $\theta \neq \lambda/4$, by the same argument of (3.82), we must have

$$\Delta X_T^* = \Delta Y_T^* = 0.$$

Therefore, X^* and Y^* must be continuous in $[0, T]$.

By Lemma 3.3.13 again and choosing the same stopping time $\tau \in [0, T)$ for X^* and Y^*, there exist $\mathcal{F}_\tau$-measurable random variables η_1 and η_2 such that X^* and Y^* solve the following integral equations for all stopping time $\sigma \in [\tau, T]$,

$$\begin{aligned}\mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\sigma-t|}\, dX_t^* + \int_{[0,\sigma)} \lambda e^{-\rho(\sigma-t)}\, dY_t^* \,\Big|\, \mathcal{F}_\tau\Big] &= \eta_1,\\ \mathbb{E}\Big[\int_{[0,T]} \lambda e^{-\rho|\sigma-t|}\, dY_t^* + \int_{[0,\sigma)} \lambda e^{-\rho(\sigma-t)}\, dX_t^* \,\Big|\, \mathcal{F}_\tau\Big] &= \eta_2.\end{aligned} \tag{3.88}$$

Subtracting one equation of (3.88) from the other yields

$$\mathbb{E}\Big[\int_{[\sigma,T]} \lambda e^{-\rho(t-\sigma)}\, d(X_t^* - Y_t^*) \,\Big|\, \mathcal{F}_\tau\Big] = \eta_1 - \eta_2.$$

Setting $\sigma = T$ and the continuity of X^* and Y^* yield that $\eta_1 - \eta_2 = 0$. Hence,

$$\mathbb{E}\Big[\int_{[\sigma,T]} \lambda e^{-\rho(t-\sigma)}\, d(X_t^* - Y_t^*) \,\Big|\, \mathcal{F}_\tau\Big] = 0 \text{ for all } \sigma \geq \tau.$$

Setting $\sigma = \tau$ yields

$$\mathbb{E}\Big[\int_{[\tau,T]} \lambda e^{-\rho t}\, d(X_t^* - Y_t^*) \,\Big|\, \mathcal{F}_\tau\Big] e^{\rho\tau} = 0.$$

Hence,

$$\mathbb{E}\Big[\int_{[\tau,T]} e^{-\rho t}\, d(X_t^* - Y_t^*) \,\Big|\, \mathcal{F}_\tau\Big] = 0 \text{ for all } \tau \in [0, T].$$

Since $X^* - Y^*$ is continuous and $e^{-\rho t} > 0$, we have $X_t^* = Y_t^*$ $\mathbb{P}$-a.s. for all $t \in [0, T]$. Hence, X^* and Y^* are indistinguishable and it must hold that $x = y$. Therefore, it suffices to consider only the deterministic case, i.e., $X^*, Y^* \in \mathcal{X}_{\text{det}}(x, [0, T])$. Note that (3.88) becomes

$$\int_{[0,T]} \lambda e^{-\rho|s-t|}\, dX_t^* + \int_{[0,s)} \lambda e^{-\rho(s-t)}\, dX_t^* = \eta, \tag{3.89}$$

where $s \in [0, T]$ is arbitrary and $\eta \in \mathbb{R}$. Hence, for all $u, s \in [0, T]$ with $u > s$, we have

$$\int_{[0,T]} \frac{\lambda\big(e^{-\rho|u-t|} - e^{-\rho|s-t|}\big)}{u-s}\, dX_t^* + \frac{\int_{[0,u)} \lambda e^{-\rho(u-t)}\, dX_t^* - \int_{[0,s)} \lambda e^{-\rho(s-t)}\, dX_t^*}{u-s} = 0. \tag{3.90}$$

Since

$$\left| \int_{[0,T]} \frac{\lambda\big(e^{-\rho|u-t|} - e^{-\rho|s-t|}\big)}{u-s}\, dX_t^* \right| \le K, \qquad \text{for some } K > 0,$$

the proof of Gatheral et al. [2012, Theorem 2.23] yields that

$$\begin{aligned} &\lim_{s\downarrow 0} \lim_{u\downarrow s} \int_{[0,T]} \frac{\lambda\big(e^{-\rho|u-t|} - e^{-\rho|s-t|}\big)}{u-s}\, dX_t^* \\ &= \lim_{s\downarrow 0} \left(\int_{[0,s]} -\lambda\rho e^{-\rho(s-t)}\, dX_t^* + \int_{(s,0]} \lambda\rho e^{-\rho(t-s)}\, dX_t^* \right) = \int_{[0,T]} \lambda\rho e^{-\rho t}\, dX_t^* < \infty. \end{aligned} \tag{3.91}$$

Furthermore, the function $s \mapsto \int_{[0,s)} \lambda e^{-\rho(s-t)}\, dX_t^*$ is continuous in $[0, T]$. It follows that

$$Q := \lim_{s\downarrow 0} \lim_{u\downarrow s} \frac{\int_{[0,u)} \lambda e^{-\rho(u-t)}\, dX_t^* - \int_{[0,s)} \lambda e^{-\rho(s-t)}\, dX_t^*}{u-s} < \infty. \tag{3.92}$$

Putting (3.91) and (3.92) into (3.90) yields that

$$\int_{[0,T]} \lambda\rho e^{-\rho t}\, dX_t^* + Q = 0. \tag{3.93}$$

On the other hand, (3.89) can be written as

$$2\int_{[0,T]} \lambda e^{-\rho|s-t|}\, dX_t^* - \int_{[s,T]} \lambda e^{-\rho(t-s)}\, dX_t^* = \eta.$$

Since

$$\frac{-\int_{[u,T]} \lambda e^{-\rho(t-u)}\, dX_t^* + \int_{[s,T]} \lambda e^{-\rho(t-s)}\, dX_t^*}{u-s} = \frac{\int_{[0,u)} \lambda e^{-\rho(u-t)}\, dX_t^* - \int_{[0,s)} \lambda e^{-\rho(s-t)}\, dX_t^*}{u-s},$$

by the same method used above, we have

$$2\int_{[0,T]} \lambda\rho e^{-\rho t}\, dX_t^* + Q = 0. \tag{3.94}$$

Comparing (3.93) and (3.94) yields

$$\int_{[0,T]} \lambda\rho e^{-\rho t}\, dX_t^* = 0. \tag{3.95}$$

Setting $s = 0$ in (3.89) yields

$$\int_{[0,T]} \lambda e^{-\rho t}\, dX_t^* = \eta,$$

then we have $\eta = 0$ by (3.95). This implies for any $s \in [0,T]$, it holds that

$$\int_{[0,T]} \lambda e^{-\rho|s-t|}\, dX_t^* + \int_{[0,s)} \lambda e^{-\rho(s-t)}\, dX_t^* = 0.$$

It then follows that

$$\begin{aligned}
&\mathcal{C}(X^*|Y^*) = \mathcal{C}(Y^*|X^*)\\
&= \frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s^*\, dX_t^* + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s^*\, dX_t^*\\
&\leq \int_{[0,T]} \left(\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s^* + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s^* \right) dX_t^*\\
&= 0.
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
&\mathcal{C}(X^*|Y^*) = \mathcal{C}(Y^*|X^*)\\
&= \frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s^*\, dX_t^* + \int_{[0,T]}\int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s^*\, dX_t^*\\
&= \int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s^*\, dX_t^* \geq 0.
\end{aligned}$$

This implies $\mathcal{C}(X^*|Y^*) = \mathcal{C}(Y^*|X^*) = 0$, which holds if and only if $X_t^* = Y_t^* = 0$ for all $t \in [0,T]$. The continuity of X^* and Y^* implies then $x = y = 0$. This completes the proof. □

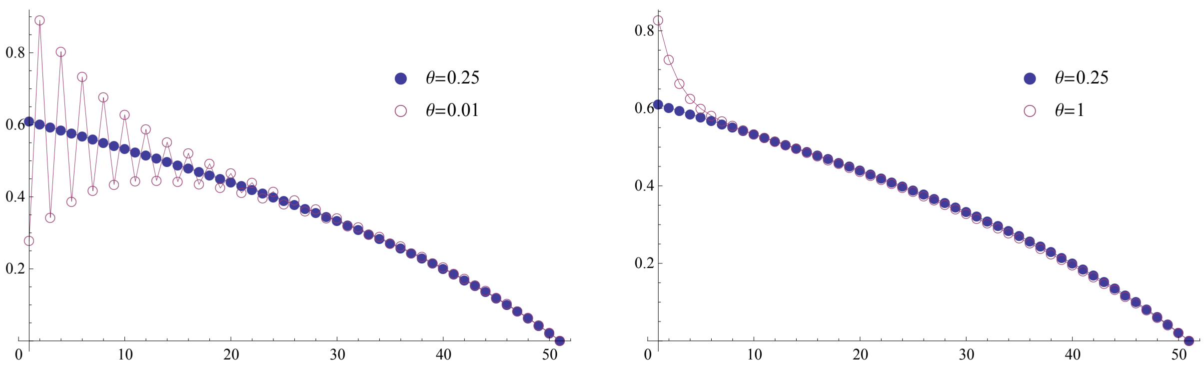


Figure 3.17: Comparison of optimal strategies in discrete-time model with parameters $\theta = 0.01$, $\theta = 0.25$ (left) and $\theta = 1$, $\theta = 0.25$ (right). We have used the equidistant time grid $\mathbb{T}_{50}$, the initial condition $x = y = 1$, and parameters $\lambda = \rho = 1$. The optimal strategies with $\theta \neq \lambda/4$ consist growing jumps near $t = 0$, while the optimal strategy with $\theta = \lambda/4$ coincides with the function V stated in Theorem 3.3.6.

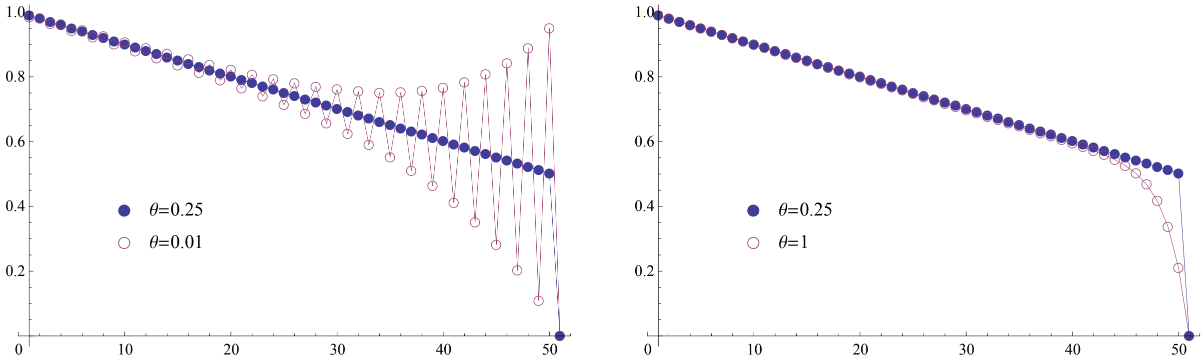


Figure 3.18: Comparison of optimal strategies in discrete-time model with parameters $\theta = 0.01$, $\theta = 0.25$ (left) and $\theta = 1$, $\theta = 0.25$ (right). We have used the equidistant time grid $\mathbb{T}_{50}$, the initial condition $x = -y = 1$, and parameters $\lambda = \rho = 1$. The optimal strategies with $\theta \neq \lambda/4$ consist of growing jumps near $t = T$, while the optimal strategy with $\theta = \lambda/4$ is linear and coincides with the function W stated in Theorem 3.3.6.

Remark 3.3.14. Figures 3.17 and 3.18 of the primary model suggest an intuitive explanation for the nonexistence of Nash equilibia when $\theta \neq \theta^*$: when $\theta < \theta^*$, oscillatory strategies are not convergent to a continuous-time strategy; when $\theta > \theta^*$, numerical computations suggest that the components $\boldsymbol{v}$ and $\boldsymbol{w}$ defined in (3.5) are convergent to the functions V and W defined in (3.55). However, strategies consisted of V and W as stated in (3.54) form a Nash equilibrium if and only if $\theta = \theta^*$. □

3.3.3 Review of single-agent models

In this subsection we analyze the effects of transaction costs in single-agent models. First we impost the transaction costs in the discrete-time model of Obizhaeva and Wang [2013]. Given a discrete time grid $\mathbb{T} = \{t_0, t_1, \ldots, t_N\}$, where $0 = t_0 < t_1 < \cdots < T_N = T$, an admissible trading strategy for $\mathbb{T}$ and an initial asset position $x \in \mathbb{R}$ used by the signal agent X is described by a vector[2] $\boldsymbol{\xi} = (\xi_0, \ldots, \xi_N) \in \mathbb{R}^{N+1}$ with $\sum_{k=0}^{N} \xi_k = x$. If the agent X is active, the affected price process S_t is given by

$$S_t = S_t^0 - \lambda \sum_{t_k < t} e^{-\rho(t-t_k)} \xi_k,$$

where the unaffected price process S_t^0 is defined as same as at the beginning of Subsection 3.3.1. For each execution ξ_k, $k \in \{0, 1 \ldots, N\}$, there are additional transaction costs $\theta \xi_k^2$ for $\theta \geq 0$. This model is a single-agent version of the model stated in Section 3.1. The liquidation costs are defined as

$$\mathcal{C}(\boldsymbol{\xi}) = \frac{1}{2} \boldsymbol{\xi}^\top \lambda (G + 2\theta \, \mathrm{Id}) \boldsymbol{\xi},$$

where

$$G_{i+1,j+1} = e^{-\rho|t_i - t_j|}, \qquad i, j = 0, \ldots, N.$$

[2] We consider only deterministic strategies since one can verify that the unique optimal strategy in the class of adapted strategies is deterministic by Lemma 3.3.12.

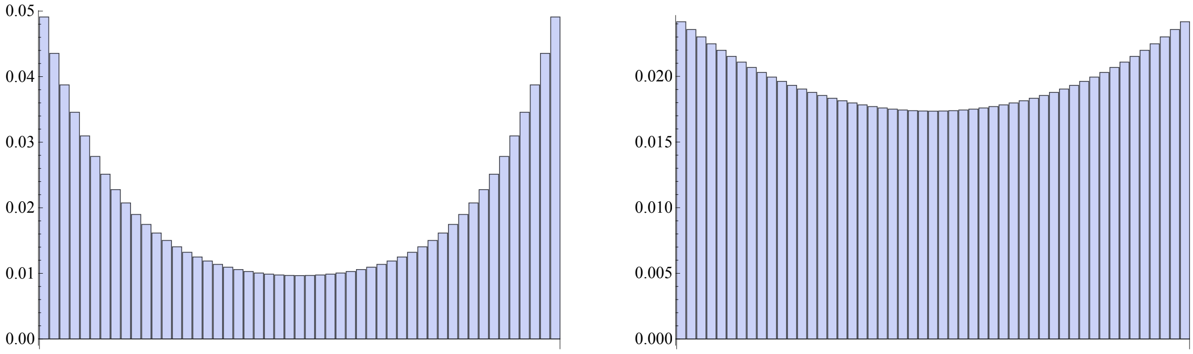


Figure 3.19: Optimal strategies with parameters $\theta = 1$ (left) and $\theta = 10$ (right). We use the equidistant time grid $\mathbb{T}_{50}$, the initial value $x = 1$, and parameters $\lambda = \rho = 1$. The optimal strategy tends to have a constant trading rate as the transaction costs increase.

For this model, the unique optimal strategy can be obtained by following Alfonsi et al. [2008, Theorem 3.1]

$$\boldsymbol{\xi}^* = x \frac{\Big(\lambda(G + 2\theta\,\mathrm{Id})\Big)^{-1}\mathbf{1}}{\mathbf{1}^T\Big(\lambda(G + 2\theta\,\mathrm{Id})\Big)^{-1}\mathbf{1}}, \qquad \text{for all } \theta \geq 0.$$

The addition of transaction costs does not affect the existence of optimal strategy, see Figure 3.19. As the transaction costs increase, the optimal strategy tends to be more similar to the trivial strategy.

We turn to analyze the effects of transaction costs in the continuous-time version of the single-agent model. We assume that if the agent X is active, the affected price S_t is

$$S_t = S_t^0 + \int_{[0,t)} \lambda e^{-\rho(t-s)}\, dX_s.$$

If there are no transaction costs, the liquidation costs are defined as

$$\mathcal{C}(X) := \frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t = \frac{1}{2}\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t, \tag{3.96}$$

which is a special case in Gatheral et al. [2012]. This model is a continuous-time variant of the simplified version of the limit order book model of Obizhaeva and Wang [2013], which is introduced by Alfonsi et al. [2008] and Alfonsi et al. [2010]. It has been shown in Obizhaeva and Wang [2013] that the unique optimal strategy $X^* \in \mathcal{X}(x, T)$ is given by

$$dX_s^* = \frac{-x}{\rho T + 2}\Big(\delta_0(ds) + \rho\, ds + \delta_T(ds)\Big).$$

Now we assume that there are quadratic transaction costs associated with jumps of strategies. Then the liquidation costs are

$$\mathcal{C}(X) := \frac{1}{2}\bigg(\int_{[0,T]}\int_{[0,T]} \lambda e^{-\rho|t-s|}\, dX_s\, dX_t + 2\theta \sum_{t\in[0,T]} \big(\Delta X_t\big)^2\bigg), \qquad \theta > 0.$$

This formula can be regarded as the costs with respected to a decay kernel $g(t)$ which has a jump at $t = 0$, i.e.,

$$\mathcal{C}(X) := \frac{1}{2}\left(\int_{[0,T]}\int_{[0,T]} g(|t-s|)\,dX_s\,dX_t\right), \qquad g(t) = \begin{cases} \lambda e^{-\rho t}, & \text{if } t \in (0,T]; \\ \lambda + 2\theta, & \text{for } t = 0. \end{cases}$$

Gatheral et al. [2012, Theorem 2.11] gives an equivalent condition for the existence of an optimal strategy minimizing the liquidation costs

$$\frac{1}{2}\int_{[0,T]}\int_{[0,T]} h(|t-s|)\,dX_s\,dX_t,$$

for a continuous decay kernel $h : [0,\infty) \to [0,\infty)$. In the following proposition we find that this equivalent condition does not hold if the decay kernel is not continuous at zero.

Proposition 3.3.15. *Let $h : [0,\infty) \to [0,\infty)$ be a positive definite decay kernel. If $h(t)$ is continuous in $(0,\infty)$ but $h(0) \neq \lim_{t\downarrow 0} h(t)$, then for all strategies $X \in \mathcal{X}_{det}(x,[0,T])$ there is no constant $\eta \in \mathbb{R}$, such that X solves the generalized Fredholm integral equation,*

$$\int_{[0,T]} h(|t-s|)\,dX_s = \eta.$$

Proof. Suppose by way of contradiction that there exits a constant $\eta \in \mathbb{R}$ and a strategy $X^* \in \mathcal{X}_{\mathrm{det}}(x,[0,T])$ that solves the generalized Fredholm integral equation

$$\int_{[0,T]} h(|t-s|)\,dX^*_s = \eta, \qquad \text{for all } t \in [0,T].$$

We assume

$$X^* = X^c + X^d$$

where X^c denotes the continuous part of X^* and $X^d := X^* - X^c$. Since X^*_t has finite and $\mathbb{P}$-a.s. bounded total variation, the support $J := \operatorname{supp} dX^d$ must be countable. Furthermore, we define

$$\widetilde{h}(t) = \begin{cases} h(t), & \text{if } t \in (0,T]; \\ \lim_{t\downarrow 0} h(t), & \text{for } t = 0; \end{cases}$$

as a continuous version of $h(t)$. It holds that for all $t \in [0,T]$,

$$\begin{aligned} &\int_{[0,T]} h(|t-s|)\,dX^*_s \\ &= \int_{[0,T]} h(|t-s|)\,dX^c_s + \int_{[0,T]} \widetilde{h}(|t-s|)\,dX^d_s + \big(h(0) - \widetilde{h}(0)\big)\int_{[0,T]} \mathbb{1}_{\{t\}}(s)\,dX^d_s = \eta. \end{aligned}$$

Now let $t_0 \in J$ and a sequence $(t_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} t_n = t_0$ and $t_n \notin J$ for all $n \in \mathbb{N}$. Since the terms

$$\int_{[0,T]} h(|t-s|)\,dX^c_s \qquad \text{and} \qquad \int_{[0,T]} \widetilde{h}(|t-s|)\,dX^d_s$$

are continuous with respect to t. We have

$$
\begin{aligned}
&\int_{[0,T]} h(|t_0-s|)\,dX^*_s \\
&= \lim_{n\to\infty}\Big(\int_{[0,T]} h(|t_n-s|)\,dX^c_s + \int_{[0,T]} \widetilde{h}(|t_n-s|)\,dX^d_s + \big(h(0)-\widetilde{h}(0)\big)\int_{[0,T]} \mathbb{1}_{\{t_n\}}(s)\,dX^d_s\Big) \\
&= \lim_{n\to\infty}\Big(\int_{[0,T]} h(|t_n-s|)\,dX^c_s + \int_{[0,T]} \widetilde{h}(|t_n-s|)\,dX^d_s\Big) \\
&= \int_{[0,T]} h(|t_0-s|)\,dX^c_s + \int_{[0,T]} \widetilde{h}(|t_0-s|)\,dX^d_s = \eta.
\end{aligned}
$$

On the other hand, it holds that

$$
\begin{aligned}
&\int_{[0,T]} h(|t_0-s|)\,dX^*_s \\
&\quad = \int_{[0,T]} h(|t_0-s|)\,dX^c_s + \int_{[0,T]} \widetilde{h}(|t_0-s|)\,dX^d_s + \big(h(0)-\widetilde{h}(0)\big)\Delta X^d_{t_0} = \eta,
\end{aligned}
$$

with $h(0)-\widetilde{h}(0)\neq 0$. This shows that $J=\varnothing$. For a continuous strategy X, its liquidation costs become

$$
\mathcal{C}(X) = \frac{1}{2}\int_{[0,T]}\int_{[0,T]} h(|t-s|)\,dX_s\,dX_t = \frac{1}{2}\int_{[0,T]}\int_{[0,T]} \widetilde{h}(|t-s|)\,dX_s\,dX_t.
$$

However, Theorem 2.23 in Gatheral et al. [2012] asserts that an optimal strategy minimizing the liquidation costs with a continuous decay kernel must have jumps at $t\in\{0,T\}$. This shows there is no optimal strategy in the class $\mathcal{X}_{\text{det}}(x,[0,T])$ of deterministic strategies and completes the proof. □

Corollary 3.3.16. *Given $\lambda>0$, $\rho>0$, $T>0$, $\theta>0$ and an initial asset position $x\neq 0$. There is no optimal strategy in the class $\mathcal{X}_{det}(x,[0,T])$ of deterministic strategies.*

Proof. This assertion is obtained directly through Proposition 3.3.15. □

We see the key point for the nonexistence of optimal strategies is that the decay kernel $g(t)$ is not continuous at $t=0$. Due to this discontinuity, the term

$$
\int_{[0,T]} g(|t-s|)\,dX_s
$$

is not a continuous function with respect to t. Therefore, the optimal cost functional $\mathcal{C}(\boldsymbol{\xi}^{(N)})$ in discrete-time model does not converge to a cost functional $\mathcal{C}(X^*)$ with respect to the weak topology as $N\uparrow\infty$, where $\boldsymbol{\xi}^{(N)}$ denotes the optimal strategy with a time grid $\mathbb{T}_N$ consisting of N points and $X^*\in\mathcal{X}_{\text{det}}(x,[0,T])$.

Intuitively, in the continuous-time model one would try to use continuous strategies to approximate jumps infinitesimally to avoid the quadratic transaction costs incurred by jumps. However it has been shown in Gatheral et al. [2012] that the optimal strategy $X^*\in\mathcal{X}_{\text{det}}(x,[0,T])$ with $x\neq 0$ must have jumps at $t=0$ and $t=T$ in case $\theta=0$. These two conflicting situations lead to the nonexistence of optimal strategy.

Chapter 4

Outlook

In this thesis, we combine market impact models and Nash equilibria to analyze the competing behavior of two or more large traders in a financial market. In Chapter 2 we consider n risk-averse agents who compete for liquidity in an Almgren–Chriss market impact model. In Chapter 3 motivated by the observation that high-frequency traders may use oscillatory trading strategies, we propose quadratic transaction costs in order to make the market more stable and efficient. Moreover, we extend this model in three aspects and introduce a continuous-time version of the model.

Possible further research could be carried out based on the two main concepts of this thesis. From the side of market impact models, we assume in Chapter 2 that the unaffected price process follows a Bachelier model. A more general assumption, such as martingale or even semi-martingale, is expected. Schied [2011] analyzes robust optimal strategies in the Almgren–Chriss framework for a single trader. This work provides a basis to study Nash equilibria for two or more traders when the unaffected price process is defined in a broader class. Note that one difficulty may be the deviation of the optimal value function. In Chapter 3 we analyze the effects of quadratic transaction costs. It is plausible to consider transaction costs which are proportional to the size of trades. A technical problem is to deal with absolute value, since transaction costs can be regarded as a nonnegative penalty. This kind of problem is related to the optimization with l_1-regularization of the following form

$$\inf_{X\in\mathcal{X}} f(X) = C(X) + \theta\|X\|_1.$$

Such optimization problem is usually solved by numerical methods. Another open question is the optimal closed-loop strategies in continuous-time models. As we state in Chapter 3, the explicit formula of optimal strategies obtained by dynamic programming is very complex. It is difficult to analyze its high-frequency limit directly. Moreover, in Section 3.3 we show that optimal open-loop strategies exist only for $\theta = \theta^*$. Whether this result still holds for closed-loop strategies is an interesting question.

On the other hand, there are also much possible further work with respect to Nash equilibria. A main open question in this thesis is the uniqueness of Nash equilibria in the class of adapted strategies under the model setting of Chapter 2. Moreover, as it is mentioned at the beginning of this thesis, Nash equilibrium is a concept for non-cooperative game. In some cases, Nash equilibrium is not the “best” solution if cooperation is allowed. A classical example is *prisoner’s dilemma*. Therefore, to consider an alternative optimality for a cooperative game is desirable. We suggest *Pareto optimality* as one candidate.

Bibliography

A. Alfonsi and A. Schied. Optimal trade execution and absence of price manipulations in limit order book models. *SIAM J. Financial Math.*, 1:490–522, 2010.

A. Alfonsi, A. Fruth, and A. Schied. Constrained portfolio liquidation in a limit order book model. *Banach Center Publications*, 83:9–25, 2008.

A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quant. Finance*, 10:143–157, 2010.

A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM J. Financial Math.*, 3:511–533, 2012.

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

R. Almgren and N. Chriss. Value under liquidation. *Risk*, 12:61–63, 1999.

R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–39, 2000.

A. Berman and R. J. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. ISBN 0-89871-321-8. URL `http://dx.doi.org/10.1137/1.9781611971262`.

D. S. Bernstein. Matrix mathematics: theory, facts, and formulas with application to linear system theory, 2005.

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

J.-P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quant. Finance*, 4: 176–190, 2004.

M. K. Brunnermeier and L. H. Pedersen. Predatory trading. *Journal of Finance*, 60(4): 1825–1863, August 2005.

B. I. Carlin, M. S. Lobo, and S. Viswanathan. Episodic liquidity crises: cooperative and predatory trading. *Journal of Finance*, 65:2235–2274, 2007.

R. Carmona and K. Webster. A belief-driven order book model. 2012.

R. A. Carmona and J. Yang. Predatory trading: a game on volatility and liquidity. *Preprint*, 2011. URL `http://www.princeton.edu/~rcarmona/download/fe/PredatoryTradingGameQF.pdf`.

M. Cecchi, M. Marini, and P. L. Zezza. Linear boundary value problems for systems of ordinary differential equations on noncompact intervals. *Ann. Mat. Pura Appl. (4)*, 123:267–285, 1980. ISSN 0003-4622.

CFTC-SEC. Findings regarding the market events of May 6, 2010. Report, 2010.

J. Gatheral. No-dynamic-arbitrage and market impact. *Quant. Finance*, 10:749–759, 2010.

J. Gatheral and A. Schied. Dynamical models of market impact and algorithms for order execution. In J.-P. Fouque and J. Langsam, editors, *Handbook on Systemic Risk*, pages 579–602. Cambridge University Press, 2013.

J. Gatheral, A. Schied, and A. Slynko. Transient linear price impact and Fredholm integral equations. *Math. Finance*, 22:445–474, 2012.

S. Kakutani et al. A generalization of brouwers fixed point theorem. *Duke mathematical journal*, 8(3):457–459, 1941.

O. Kallenberg. *Foundations of modern probability.* springer, 2002.

A. A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun. The flash crash: The impact of high frequency trading on an electronic market. *Manuscript, U of Maryland*, 2010. URL `http://ssrn.com/abstract=1686004`.

J. Kurzweil. *Ordinary differential equations*, volume 13 of *Studies in Applied Mechanics*. Elsevier Scientific Publishing Co., Amsterdam, 1986. ISBN 0-444-99509-9. Introduction to the theory of ordinary differential equations in the real domain, Translated from the Czech by Michal Basch.

A. Lachapelle, J.-M. Lasry, C.-A. Lehalle, and P.-L. Lions. Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis. *arXiv preprint arXiv:1305.6323*, 2013.

O. Lebedeva, E. Maug, and C. Schneider. Trading strategies of corporate insiders. *Available at SSRN 2118110*, 2012.

C.-A. Lehalle. Market microstructure knowledge needed to control an intra-day trading process. In J.-P. Fouque and J. Langsam, editors, *Handbook on Systemic Risk*. Cambridge University Press, 2013.

C. Lorenz and A. Schied. Drift dependence of optimal order execution strategies under transient price impact. To appear in *Finance and Stochastics*, 2012. URL `http://ssrn.com/abstract=1993103`.

J. Lorenz and R. Almgren. Mean-variance optimal adaptive execution. *Appl. Math. Finance*, 18(5):395–422, 2011. ISSN 1350-486X.

C. C. Moallemi, B. Park, and B. V. Roy. Strategic execution in the presence of an uninformed arbitrageur. *Journal of Financial Markets*, 15(4):361 – 391, 2012. ISSN 1386-4181. doi: http://dx.doi.org/10.1016/j.finmar.2011.11.002.

J. Nash. Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295, 1951.

J. F. Nash et al. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.

A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16:1–32, 2013.

S. Predoiu, G. Shaikhet, and S. Shreve. Optimal execution in a general one-sided limit-order book. *SIAM J. Financial Math.*, 2:183–212, 2011. ISSN 1945-497X.

P. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, New York, second edition, 2004.

J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, pages 520–534, 1965.

A. Schied. Robust strategies for optimal order execution in the Almgren–Chriss framework. Forthcoming in: *Applied Mathematical Finance*, 2011. URL `http://ssrn.com/paper=1991097`.

A. Schied and T. Schöneborn. Risk aversion and the dynamics of optimal trading strategies in illiquid markets. *Finance Stoch.*, 13:181–204, 2009.

A. Schied and T. Zhang. A hot-potato game under transient price impact and some effects of a transaction tax. *arXiv preprint arXiv:1305.4013*, 2013a.

A. Schied and T. Zhang. A state-constrained differential game arising in optimal portfolio liquidation. *arXiv preprint arXiv:1312.7360*, 2013b.

A. Schied, T. Schöneborn, and M. Tehranchi. Optimal basket liquidation for CARA investors is deterministic. *Applied Mathematical Finance*, 17:471–489, 2010.

T. Schöneborn. Trade execution in illiquid markets. Optimal stochastic control and multi-agent equilibria. Ph.D. thesis, TU Berlin, 2008.

T. Schöneborn and A. Schied. Liquidation in the face of adversity: stealth vs. sunshine trading. Preprint, 2009.