Effect of Transaction Tax on Optimal Portfolio Liquidation and Predatory Trading

Enkhjargal Lkhagvajav
Advisor: Prof. Sergey Nadtochiy
REU Summer, University of Michigan

August 1, 2014

Abstract

Schied et al. showed the existence and uniqueness of a Nash Equilibrium among n-risk averse agents who compete to liquidate their portfolio in an Almgren-Chriss market impact model. In this work, we introduce the transaction tax on each traded share in the same model to discourage predatory trading which distresses liquidators and may contribute to a financial crisis. Our original problem takes an absolute value tax function, which is a non-smooth function. However, this non-smoothness is a complication when we try to derive a computationally tractable description of a Nash equilibrium by maximizing the Lagrangian. To go around this problem, we approximate the absolute value function by a smooth function which converges to the absolute value function. We proved the uniqueness and conjectured the existence of a Nash equilibrium in the case of a smooth tax function. We further show the convergence of mean-variance criterion in a non-smooth tax function game case and that the equilibrium strategies of a smooth tax-function game produce approximate equilibrium in the non-smooth tax-function game.

We compute the Nash equilibrium strategies in 2-player case using Matlab. From the numerical experiment, we can see that the transaction tax discourages and eliminates the predatory trading by putting cost on a predatory trader. Moreover, we show that, in certain cases, the benefit for a liquidator with large initial positions increases as the amount of tax increases up to a certain value.

Contents

1 Introduction

In this paper we introduce transaction tax on stock trade and analyze its effect on Nash equilibrium game and benefit for market players. Our main framework is based on Schied & Zhang(2013)[2] paper which shows existence and uniqueness of Nash equilibrium in a state-constrained differential game arising in optimal portfolio liquidation. By introducing transaction tax on each trade, we are seeking to eliminate and discourage predatory trading and protect liquidators from predatory traders. Predatory trading is a trading strategy where predatory trader exploits a liquidator's need of liquidating his position in a fixed period of time. The strategy includes the period of short-selling same type of shares the liquidator is selling and buying them back in the second period where the price of the share is overshot due to selling by two players. By taxing each transaction, whether it is selling or buying, we are seeking to put cost on predatory trading to discourage and, moreover, eliminate the trading strategy. Absence of predatory trading cuts the cost occurred by a liquidator partly due to price overshooting created by a predatory trader. Moreover, we can avoid the risk of financial crisis which could be fueled by predatory trading.

The paper is organized as follows. In Section 2.1 we introduce Almgren-Chriss model and background material on portfolio liquidation. In Section 2.2, we show the uniqueness and conjecture the existence of a Nash equilibrium. The convergence of strategies in a smooth tax function game to that of a smooth tax function game is shown in Section 2.3. We will look at some simulation of strategies in 2-player game in section 2.4. All proofs are shown in Section 3.

2 Nash equilibrium and convergence of smooth game strategies

2.1 Background

Our paper employs a standard continuous-time Almgren-Chriss(2000) [1] framework for investors who need to liquidate their holdings in a fixed time period [0,T]. An investor is assumed to have initial holding of x shares and be required to have zero shares at time T. An investor's trading strategy is denoted by $X = (X(T))_{t\in[0,T]}$. This strategy should have following conditions of admissibility:

- X satisfies the liquidation constraint $X(T) = 0$;
- X is deterministic in the sense that it is not subject to chance

- X is absolutely continuous in the sense that there exists a progressively measurable process $(\dot{X}(t))_{t\in[0,T]}$ such that $\int_0^T (\dot{X}(t))^2 dt < \infty$ and

$$X(t) = X(0) + \int_0^t \dot{X}(s)ds, t \in [0,T]$$

- there exists a constant $c \geq 0$ such that $|X(t)| \leq c$ for all t..

The class of all strategies that follow above conditions and such that $X(0) = x$ for given $x \in \mathbb{R}$ will be denoted by $\mathfrak{X}(x,T)$ We consider players maximizing a mean-variance functional criterion in a continuous-time Almgren and Chris(2000) model. In this model, 'unaffected price' is price determined by two exogenous factors: volatility and drift, which are assumed to be the result of market forces that occur randomly and independently of trading. The 'unaffected price' S^0 follows a Bachelier model with some drift describing current price trends:

$$S^0(t) = S_0 + \sigma W(t) + \int_0^t b(s)ds \tag{1}$$

S_0-constant,W-standard Brownian motion, $\sigma \geq 0$, b continuous

When the trader starts trading, it will influence the price of the asset it is trading. Again in Almgren-Chriss model, we assume the price influenced by the agent's trading, 'affected price' $S^X(t)$, is given by

$$S^X(t) := S^0(t) + \gamma(X(t) - X(0)) + \lambda\dot{X}(t), t \in [0,T], \tag{2}$$

where the constants $\gamma \geq 0$ and $\lambda > 0$ describe the respective permanent and temporary price impact. At time $t \in [0,T]$, $-\dot{X}(t)dt$ shares are sold at price $S^X(t)$. Total revenues generated by the strategy $X \in \mathfrak{X}(x,T)$ for an investor after introducing transaction tax of fixed rate of μ on each traded share are given by:

$$\mathcal{R}(X) := -\int_0^T \dot{X}(t)S^X(t)dt - \int_0^T |\dot{X}(t)|\mu dt \tag{3}$$

Here, we have used an absolute value tax function, which is a non-smooth function. However, we will modify the tax function to make the problem mathematically more feasible since we will need to take derivative from a tax function. Therefore, we take a smooth approximating function of the absolute value function as a tax function:

$$f(\dot{X}) = \sqrt{(\dot{X}(t))^2 + \epsilon^2} - \epsilon \approx |\dot{X}(t)| \text{ for sufficiently small } \epsilon > 0 \tag{4}$$

A new tax function converges to an absolute value function as $\epsilon \to 0$

The total revenues after introducing a smooth transaction tax function are given by:

$$\mathcal{R}(X) := -\int_0^T \dot{X}(t)S^X(t)dt - \int_0^T (\sqrt{(\dot{X}(t))^2+\epsilon^2}-\epsilon)\mu dt \tag{5}$$

An agent seeks to maximize expected revenue taking into account the volatility risk arising from late execution. We consider that an agent maximizes the following mean-variance functional criterion:

$$\mathbb{E}[\mathcal{R}(X)] - \frac{\alpha}{2}\operatorname{Var}(\mathcal{R}(X)); \tag{6}$$

2.2 Nash Equilibrium

If we have n agents with the respective strategies $X_1, ..., X_n$ in the market, then 'affected price' will be given by modifying (2):

$$S^{X_1,...,X_n}(t) := S^0(t) + \gamma\sum_{j=1}^{n}(X_j(t)-X_j(0)) + \sum_{j=1}^{n}\lambda\dot{X}_j(t), t\in[0,T], \tag{7}$$

Let us denote by $\mathbf{X_{-i}}$ the collection $\mathbf{X_{-i}} := \{X_1, ..., X_{i-1}, X_{i+1}, ..., X_n\}$, then player i's revenue is given by:

$$\mathcal{R}(X_i|\mathbf{X}_{-i}) = -\int_0^T \dot{X}_i(t)S^{X_1,...,X_n}(t)dt - \mu\int_0^T \sqrt{(\dot{X}_i(t))^2+\epsilon^2}-\epsilon dt \tag{8}$$

Each player seeks to maximize the functional (6). Our paper looks at whether there is a Nash equilibrium and if it is unique. First, we define a Nash equilibrium.

Definition. *A Nash equilibrium for mean-variance optimization consists of a collection $X_1^*, ..., X_n^*$ of strategies such that for each i and $\boldsymbol{X}_{-i}^* = \{X_1^*, ..., X_{i-1}^*, X_{i+1}^*, ..., X_n^*\}$ the strategy X_i^* maximizes the mean-variance functional*

$$\mathbb{E}[\mathcal{R}(X|\boldsymbol{X}_{-i}^*)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}(X|\boldsymbol{X}_{-i}^*)) \tag{9}$$

over all X

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

$$\alpha_i\sigma^2X_i(t) - 2\lambda\ddot{X}_i(t) - \mu\frac{\ddot{X}_i(t)\epsilon^2}{(\sqrt{(\dot{X}_i(t))^2+\epsilon^2})^3} = b(t) + \gamma\sum_{j\neq i}\dot{X}_j(t) + \lambda\sum_{j\neq i}\ddot{X}_j(t), \tag{10}$$

with two-point boundary conditions

$$X_i(0) = x_i \text{ and } X_i(T) = 0, i = 1, ..., n \tag{11}$$

2.3 From 'smooth' game to 'non-smooth' game

So far, we have been discussing the result in the 'smooth' game in which we have employed a smooth tax function. Let R^0 be a revenue function when transaction tax function is non-smooth i.e absolute value function and let's call it 'non smooth' game. Let R^ϵ be a revenue function when transaction tax function is smooth i.e approximating smooth function with parameter ϵ and let's call it 'ϵ smooth' game. Let $Y_T^{i,\epsilon}$ be a strategy of Player i in a Nash equilibrium corresponding to ϵ and time T i.e

$$\begin{aligned}\sup_{Y\in\mathfrak{X}(y,T)}(\mathbb{E}[\mathcal{R}^\epsilon(Y_T^i|\mathbf{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^\epsilon(Y_T^i|\mathbf{X}_{-i}^\epsilon))) = \\ \mathbb{E}[\mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon))\end{aligned} \tag{12}$$

Theorem(Approximate Equilibrium). *A mean-variance criterion for* **Player i** *in a 'non smooth' game converges to to its maximum as $\epsilon \to 0$ when it employs a strategy that maximizes its mean-variance criterion in an 'ϵ smooth' game given that it knows other players' an 'ϵ smooth' game strategies i.e*

$$\begin{aligned}\Big|(\mathbb{E}[\mathcal{R}^0(Y_T^{i,\epsilon}|\boldsymbol{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^0(Y_T^{i,\epsilon}|\boldsymbol{X}_{-i}^\epsilon)))- \\ - \sup_{Y\in\mathfrak{X}(y,T)}(\mathbb{E}[\mathcal{R}^0(Y_T^i|\boldsymbol{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^0(Y_T^i|\boldsymbol{X}_{-i}^\epsilon)))\Big| \to 0 \ \textit{as}\ \epsilon \to 0\end{aligned} \tag{13}$$

In other words, the theorem says that if every player plays according to his 'ϵ smooth' game strategy in a 'non smooth' game, the maximum gain in a player's mean-variance he can obtain by deviating from the 'ϵ smooth' equilibrium strategy vanishes as the $\epsilon \to 0$.

Conjecture. *There exists a unique set of equilibrium strategies in a 'non smooth' game. Moreover,the equilibrium strategies of n players in an 'ϵ smooth' game converge to those of n players in the 'non smooth' game.*

2.4 2-player game simulation

In this section, we will show how introduction of transaction tax affects Nash equilibrium strategies. We are particularly interested in 2-player case, where we assume that first player is a liquidator and second player is a predator. The following game simulations were created by solving the boundary value problem for the system of non-linear ODEs given in (10) for $n = 2$ and zero drift in MATLAB.

When there isn't any transaction tax at all, we get trading strategies plotted in Figure 1. From the figure, we can see a trading strategy pursued by Player 2, who started from zero position, and bought and sold shares during the period and ended up with zero position. This strategy is predatory trading because he is not liquidating anything overall and has no incentive to trade since there isn't price drift or trend in the market to make profit out of. Player 2 is trading, in this case, only to make profit by manipulating Player 1's need of liquidating. By using this strategy, Player 2 maximizes his mean-variance criterion and profits while Player 1's mean-variance criterion is less than in the case where Player 2 doesn't employ predatory trading: when the parameters take the values in the Figure 1, Player 1 has mean-variance of 222.3 in the absence of predatory trading while it has mean-variance of 220.1 when Player 2 does predatory trading.

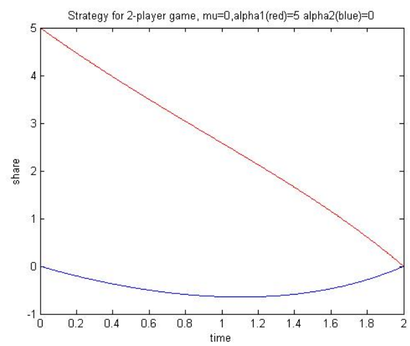

Figure 1: Strategies for two players (Player 1 red, Player 2 blue) for $\mu = 0, x_1 = 5, x_2 = 0, \alpha_1 = 5, \alpha_2 = 0, \sigma = 0.26, \gamma = 1, \lambda = 1, \epsilon = 0.01$ with no drift

The figures show how the introduction of transaction tax reduces the size of predatory trading and essentially eliminates the trading strategy of Player 2.

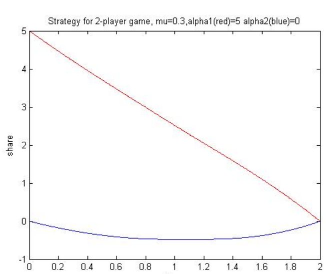

Figure 2: $\mu = 0.3$

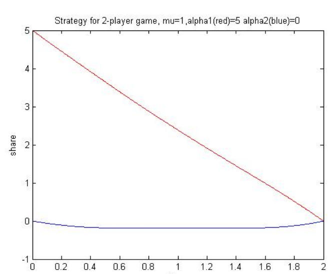

Figure 3: $\mu = 1$

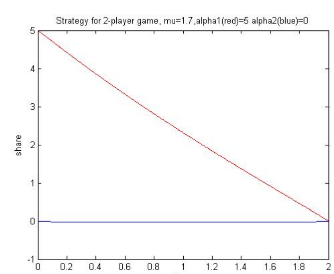

Figure 4: $\mu = 1.7$

We can see that when each transaction is taxed, both of a liquidator and a predatory trader now incur cost on each trade whether they are buying or selling share. Therefore, the predatory trader has less incentive to trade. Because of less amount of share sold by the predatory trader, the price of the share is now higher than in the absence of tax. Hence, an effect of tax on the price benefits the liquidator while the cost he incurs to trade diminishes the benefit of tax. Therefore, there could be either cases depending on the amount of tax and the parameters where mean-variance of a liquidator increases until certain point or decreases as the amount of tax increases. We further plot the mean-variance criterion for a liquidator as an increasing function of μ for some values of parameters.

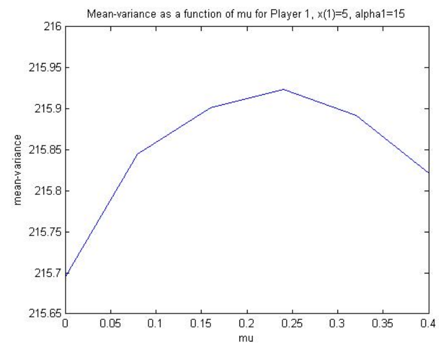

Figure 5: Mean-variance for Player 1 increases as μ increases until certain value. $x_1 = 5, x_2 = 0, T = 100, \alpha_1 = 15, \alpha_2 = 0, \sigma = 0.26, \gamma = 1, \lambda = 1, \epsilon = 0.01$ with no drift

3 Proof

3.1 Nash Equilibrium

Let admissible strategies $X_i \in \mathcal{X}(x_i, T)$ be given and write $\mathbf{X}_{-i} := \{X_1, ..., X_{i-1}, X_{i+1}, ..., X_n\}$ for $i = 1, ..., n$. For $Y \in \mathfrak{X}(y, T)$ we note first that, after some calculation we get

$$\begin{aligned} \mathcal{R}(Y|\mathbf{X}_i) &= yS_0 - \frac{\gamma}{2}y^2 + \mu\epsilon T + \int_0^T Y(t)(b(t) + \gamma \sum_{j \neq i} \dot{X}_j(t))dt \\ &- \lambda \sum_{j \neq i} \int_0^T \dot{Y}(t)\dot{X}_j(t)dt - \lambda \int_0^T \dot{Y}(t)^2 dt + \sigma \int_0^T Y(t)dW(t) - \mu \int_0^T \sqrt{\dot{X}(t)^2 + \epsilon^2}dt \end{aligned} \tag{14}$$

Moreover,

$$\mathbb{E}[\mathcal{R}(Y|\mathbf{X}_{-i})] - \frac{\alpha}{2}\operatorname{Var}(\mathcal{R}(X|\mathbf{X}_{-i})) = c + \int_0^T \mathcal{L}^i(t, Y(t), \dot{Y}(t)|\mathbf{X}_i)dt \tag{15}$$

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

$$\mathcal{L}^i(t,q,p|\mathbf{X}_i) = q(b(t) + \gamma \sum_{j\neq i} \dot{X}_j(t)) - \frac{\alpha_i \sigma^2}{2} q^2 - \lambda p(\sum_{j\neq}) \dot{X}_j(t) + p) - \mu\sqrt{p^2+\epsilon^2} \quad (16)$$

Lemma 1. *In the context of Theorem, there exists at most one Nash equilibrium for mean-variance optimization.*

Proof. Let's prove by sake of contradiction. Let's assume that $X_1^0, ..., X_n^0$ and $X_1^1, ..., X_n^1$ are two distinct Nash equilibrium with $X_i^k \in \mathfrak{X}(x_i, T)$ for $i - 1, ..., n$ and $k = 0, 1$. For $\beta \in [0,1]$ let $X_i^\beta := \beta X_i^1 + (1-\beta)X_i^0$ and define

$$f(\beta) := \sum_{i=1}^n \int_0^T (\mathcal{L}^i(t, X_i^\beta(t), \dot{X}_i^\beta(t)|\mathbf{X}_{-i}^0) + \mathcal{L}^i(t, X_i^{1-\beta}(t), \dot{X}_i^{1-\beta}(t)|\mathbf{X}_{-i}^1))dt \quad (17)$$

By assumption, the strategy X_i^k maximizes the functional $Y \mapsto \int_0^T \mathcal{L}^i(t, Y(t), \dot{Y}(t)|\mathbf{X}_{-i}^k)dt$ within the class $\mathfrak{X}(x_i, T)$ for $k = 0, 1$. Therefore, $f(\beta) \leq f(0)$ for $\beta > 0$

$$\Rightarrow \left.\frac{d}{d\beta}\right|_{\beta=0} f(\beta) \leq 0 \quad (18)$$

On the other hand, since both of our Lagrangian and its partial derivative respect to β are continuous in β and t due to our assumptions on admissible strategies and due to the form of the Lagrangian, we can interchange differentiation and integration:

$$\begin{aligned}
&\left.\frac{d}{d\beta}\right|_{\beta=0} f(\beta) \\
&= \sum_{i=1}^n \int_0^T \Bigg[\gamma(X_i^1(t) - X_i^0(t)) \sum_{j=1}^n (\dot{X}_j^0(t) - \dot{X}_j^1(t)) - \gamma(X_i^1(t) - X_i^0(t))(\dot{X}_i^0(t) - \dot{X}_i^1(t)) \\
&+ \alpha_i\sigma^2(X_i^1(t) - X_i^0(t))^2 + \lambda(\dot{X}_i^1(t) - \dot{X}_i^0(t)) \sum_{j=1}^n (\dot{X}_j^1(t) - \dot{X}_j^0(t)) + \lambda(\dot{X}_i^0(t) - \dot{X}_i^1(t))^2 \\
&- \mu(\dot{X}_i^1(t) - \dot{X}_i^0(t)) \left(\frac{\dot{X}_i^0(t)}{\sqrt{\dot{X}_i^0(t)^2 + \epsilon^2}} - \frac{\dot{X}_i^1(t)}{\sqrt{\dot{X}_i^1(t)^2 + \epsilon^2}} \right) \Bigg] dt \quad (19)
\end{aligned}$$

Here, by integration by part

$$\int_0^T (X_i^1(t) - X_i^0(t))(\dot{X}_i^0(t) - \dot{X}_i^1(t))dt = -\left.(X_i^1(t) - X_i^0(t))^2\right|_0^T - \int_0^T (X_i^1(t) - X_i^0(t))(\dot{X}_i^0(t) - \dot{X}_i^1(t))dt$$

$$\Rightarrow \int_0^T (X_i^1(t) - X_i^0(t))(\dot{X}_i^0(t) - \dot{X}_i^1(t))dt = \frac{1}{2}(X_i^1(0) - X_i^0(0))^2 - \frac{1}{2}(X_i^1(T) - X_i^0(T))^2 = 0 \tag{20}$$

since $X_i^1(T) = X_i^0(T)$ and $X_i^1(0) = X_i^0(0)$. Similarly,

$$\int_0^T (X_i^1(t) - X_i^0(t))(\dot{X}_j^0(t) - \dot{X}_j^1(t))dt = -\int_0^T (\dot{X}_i^1(t) - \dot{X}_i^0(t))(X_j^0(t) - X_j^1(t))dt, \tag{21}$$

therefore,

$$\sum_{i=1}^n \sum_{j=1}^n \int_0^T (X_i^1(t) - X_i^0(t))(\dot{X}_j^0(t) - \dot{X}_j^1(t))dt = 0. \tag{22}$$

Hence, we have

$$\begin{aligned} \left.\frac{d}{d\beta}\right|_{\beta=0} f(\beta) = \int_0^T & \left[\alpha_i\sigma^2 \sum_{i=1}^n (X_i^1(t) - X_i^0(t))^2 + \lambda \sum_{i=1}^n (\dot{X}_i^0(t) - \dot{X}_i^1(t))^2 \right. \\ & \left. + \lambda \left(\sum_{i=1}^n (\dot{X}_i^0(t) - \dot{X}_i^1(t))\right)^2\right] dt \\ + \int_0^T & \left[\mu \sum_{i=1}^n \left((\dot{X}_i^0(t) - \dot{X}_i^1(t))\left(\frac{\dot{X}_i^0(t)}{\sqrt{\dot{X}_i^0(t)^2 + \epsilon^2}} - \frac{\dot{X}_i^1(t)}{\sqrt{\dot{X}_i^1(t)^2 + \epsilon^2}}\right)\right)\right] dt \end{aligned} \tag{23}$$

Here, the first term is strictly positive since the two Nash equilibrium $X_1^0, ..., X_n^0$ and $X_1^1, ..., X_n^1$ are distinct. For the last term,

$$(\dot{X}_i^0(t) - \dot{X}_i^1(t))\left(\frac{\dot{X}_i^0(t)}{\sqrt{\dot{X}_i^0(t)^2 + \epsilon^2}} - \frac{\dot{X}_i^1(t)}{\sqrt{\dot{X}_i^1(t)^2 + \epsilon^2}}\right) \geq 0 \tag{24}$$

if one of $\dot{X}_i^0(t)$ and $\dot{X}_i^1(t)$ is negative and other one is positive. The case where both are negative is essentially same as they are both positive. When they are both positive, without loss of generality, let's assume that $\dot{X}_i^0(t) \geq \dot{X}_i^1(t)$. Then we can clearly see that

$$\dot{X}_i^0(t)\sqrt{\dot{X}_i^1(t)^2+\epsilon^2} \geq \dot{X}_i^1(t)\sqrt{\dot{X}_i^0(t)^2+\epsilon^2} \Rightarrow$$

$$\frac{\dot{X}_i^0(t)}{\sqrt{\dot{X}_i^0(t)^2+\epsilon^2}} - \frac{\dot{X}_i^1(t)}{\sqrt{\dot{X}_i^1(t)^2+\epsilon^2}} \geq 0 \quad (25)$$

Hence

$$\mu \sum_{i=1}^{n} \left((\dot{X}_i^0(t) - \dot{X}_i^1(t)) \left(\frac{\dot{X}_i^0(t)}{\sqrt{\dot{X}_i^0(t)^2+\epsilon^2}} - \frac{\dot{X}_i^1(t)}{\sqrt{\dot{X}_i^1(t)^2+\epsilon^2}} \right) \right) \geq 0, \quad (26)$$

so the last term is also strictly positive. We have shown that $\frac{d}{d\beta}\Big|_{\beta=0} f(\beta) > 0$ which contradicts with (18). By contradiction, we have shown that there exists at most one Nash equilibrium for mean-variance optimization. □

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

$$\begin{cases} \alpha_i\sigma^2 Y(t) - 2\lambda\ddot{Y}(t) - \mu\frac{\ddot{Y}(t)\epsilon^2}{(\sqrt{(\dot{Y}(t))^2+\epsilon^2})^3} = b(t) + \gamma\sum_{j\neq i}\dot{X}_j(t) + \lambda\sum_{j\neq i}\ddot{X}_j(t), \\ Y(0) = yY(T) = 0 \end{cases} \quad (27)$$

Proof. We can see that the Lagrangian $\mathcal{L}^i$ is a strict concave function of p and q and that the set $\mathcal{X}(y,T)$ is convex. Therefore, there could be at most one maximizer in $\mathfrak{X}(y,T)$.

We sketch the main idea of the proof of the existence of a maximizer. The Euler-Lagrange equation, $\mathcal{L}_q^i(t, Y(t), \dot{Y}(t)|\mathbf{X}_i) = \frac{d}{dt}\mathcal{L}_p^i(t, Y(t), \dot{Y}(t)|\mathbf{X}_{-i})$ gives us following equation:

$$\alpha_i\sigma^2 Y(t) - 2\lambda\ddot{Y}(t) - \mu\frac{\ddot{Y}(t)\epsilon^2}{(\sqrt{(\dot{Y}(t))^2+\epsilon^2})^3} = b(t) + \gamma\sum_{j\neq i}\dot{X}_j(t) + \lambda\sum_{j\neq i}\ddot{X}_j(t) \quad (28)$$

with boundary value condition $Y(0) = y$ and $Y(T) = 0$. If there is a solution for this boundary value problem, let $Y^* \in \mathfrak{X}(y,T) \cap C^2[0,T]$ denote the solution. Now, we will show that Y^* maximizes the Lagrangian. Let's take $\forall Y \in \mathfrak{X}(0,T)$. By the concavity of $(q,p) \mapsto \mathcal{L}^i(t,q,p|\mathbf{X}_{-1})$ and the fact that Y^* solves the Euler-Lagrangian equation

(28), we get

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

Therefore,

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

where we used the fact that $Y^*(0) = Y(0)$ and $Y^*(T) = Y(T)$. This shows that $Y^*(t)$ maximized the Lagrangian in $\mathfrak{X}(y, T)$. □

Next, we will give a proof for the main theorem.

Proof. We have shown that there exists at most one Nash Equilibrium in Lemma 1. Now we are left to show the existence of a Nash equilibrium $X_1^*, ..., X_n^*$ such that each strategy X_i^* belongs to $\mathfrak{X}(x_i, T) \cap C^2[0, T]$. Since each strategy X_i^* maximizes the corresponding Lagrangian, each strategy X_i^* solves the second-order nonlinear differential equation

$$
\alpha_i \sigma^2 X_i(t) - 2\lambda \ddot{X}_i(t) - \mu \frac{\ddot{X}_i(t)\epsilon^2}{(\sqrt{(\dot{X}_i(t))^2 + \epsilon^2})^3} = b(t) + \gamma \sum_{j \neq i} \dot{X}_j(t) + \lambda \sum_{j \neq i} \ddot{X}_j(t), \quad (29)
$$

with boundary conditions $X_i(0) = x_i$ and $X_i(T) = 0$. If we combine the n differential equations for n agents, we get a system of n coupled second-order nonlinear ODEs. If this system has a solution, that solution must be a unique solution to our Nash equilibrium by Lemma 2. □

3.2 Convergence result

Proof. We write (13) as follows

$$
\begin{aligned}
&\Big|(\mathbb{E}[\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)))\\
&\quad - \sup_{Y\in\mathfrak{X}(y,T)}(\mathbb{E}[\mathcal{R}^0(Y_T^{i}|\mathbf{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^0(Y_T^{i}|\mathbf{X}_{-i}^\epsilon)))\Big|\\
&\quad = \Big|\Big(\mathbb{E}[\mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) + (\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon))]\\
&\quad - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)) - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)\\
&\quad - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)) - \alpha_i\operatorname{Cov}(\mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon), \mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)\\
&\quad - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon))\Big) - \sup_{Y\in\mathfrak{X}(y,T)}\Big(\mathbb{E}[\mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon) + (\mathcal{R}^0(Y_T^{i}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon))]\\
&\quad - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon)) - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^0(Y_T^{i}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon))\\
&\quad - \alpha_i\operatorname{Cov}(\mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon), \mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon))\Big)\Big|\\
&\quad = \Big|\big(\mathbb{E}[\mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon))\big)\\
&\quad + \mathbb{E}[(\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon))] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)\\
&\quad - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)) - \alpha_i\operatorname{Cov}(\mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon), \mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon))\\
&\quad - \sup_{Y\in\mathfrak{X}(y,T)}\big(\mathbb{E}[\mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon))\big)\\
&\quad - \sup_{Y\in\mathfrak{X}(y,T)}\big(\mathbb{E}[\mathcal{R}^0(Y_T^{i}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon)] - \frac{\alpha_i}{2}\operatorname{Var}(\mathcal{R}^0(Y_T^{i}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon))\\
&\quad - \alpha_i\operatorname{Cov}(\mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon), \mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon))\big)\Big|
\end{aligned}
$$

Here, we note that

$$
\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) = \mu\int_0^T(\sqrt{(\dot{Y}_T^{i,\epsilon}(t))^2+\epsilon^2} - \epsilon)dt - \mu\int_0^T|\dot{Y}_T^{i,\epsilon}(t)|dt
$$

Then,

$$
\begin{aligned}
\operatorname{Var}(\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon)) &= 0\\
\operatorname{Cov}(\mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon), \mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i}|\mathbf{X}_{-i}^\epsilon)) &= 0
\end{aligned}
$$

Then using above and (18), we can write (13) as

$$\Big|\mathbb{E}[(\mathcal{R}^0(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon) - \mathcal{R}^\epsilon(Y_T^{i,\epsilon}|\mathbf{X}_{-i}^\epsilon))] - \sup_{Y\in\mathfrak{X}(y,T)} \big(\mathbb{E}[\mathcal{R}^\epsilon(Y_T^i|\mathbf{X}_{-i}^\epsilon)]\big)\Big|$$

$$= \Big|(\mu\int_0^T(\sqrt{(\dot{Y}_T^{i,\epsilon}(t))^2+\epsilon^2}-\epsilon)dt - \mu\int_0^T|\dot{Y}_T^{i,\epsilon}(t)|dt)$$

$$-\sup_{Y\in\mathfrak{X}(y,T)}(\mu\int_0^T(\sqrt{(\dot{Y}_T^{i,\epsilon}(t))^2+\epsilon^2}-\epsilon)dt - \mu\int_0^T|\dot{Y}_T^{i,\epsilon}(t)|dt)\Big|$$

which goes to 0 as $\epsilon \to 0$ □

References

[1] I. Almgren and N. Chriss. Optimal Execution of Portfolio Transactions. *Journal of Risk*, 3:5–39, 2000.

[2] A. Schied and T. Zhang. A state-constrained differential game arising in optimal portfolio liquidation. *ArXiv e-prints*, December 2013.