UNIVERSITY OF PIRAEUS

DEPARTMENT OF BANKING AND FINANCIAL MANAGEMENT

◇

Essays on Optimal Investment Strategies in Continuous-Time Models

◇

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

◇

Konstantinos Stefanakis

Supervised by

Michail Anthropelos

◇

April 2026

The research project was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "2nd Call for H.F.R.I. Research Projects to support Faculty Members & Researchers" (Project Number: 3444/2022, Project Title: "*Valuation and Optimal Investment of Pension Plans*").

Dedicated to my family.

Acknowledgements

Writing this thesis has been a long and demanding journey, and I am deeply grateful to the people who helped shape it, both mathematically and personally.

First and foremost, I am grateful to my supervisor, Prof. Michail Anthropelos. His guidance, patience, and confidence in me were decisive throughout my doctoral studies. He gave me the freedom to develop my own ideas, while always being available with sharp advice, careful criticism, and steady encouragement. I have learned a great deal from working with him, and I am sincerely grateful for his guidance.

I am especially grateful to Prof. Constantinos Kardaras. Our collaboration has been one of the most formative parts of my PhD. His mathematical clarity and precision have helped me develop considerably as a researcher. I have benefited a great deal from his comments, suggestions, and discussions, which helped me understand several ideas more deeply.

My sincere gratitude also goes to Prof. Gordan Žitković, Prof. Thaleia Zariphopoulou, Prof. Mihai Sîrbu, and Prof. Scott Robertson. Each of them, in different ways, showed genuine interest in my progress and offered guidance, encouragement, and mathematical insight at important moments. Their support meant a great deal to me and helped me see my work within a broader research community.

I am also grateful to the other members of my doctoral committee: Prof. Nikolaos Kourogenis, Prof. Michail Loulakis, Prof. Athanasios Yannacopoulos, Prof. Georgios Skoulakis, and Prof. Nikolaos Englezos. Their time, comments, and suggestions are sincerely appreciated, and their feedback helped improve the final form of this thesis.

I thank the Department of Banking and Financial Management at the University of Piraeus for providing a supportive academic environment throughout my doctoral studies.

I gratefully acknowledge the support of the Hellenic Foundation for Research and Innovation under the project "Valuation and Optimal Investment of Pension Plans", which provided support for part of this work.

Finally, I owe my deepest thanks to my family. Their love, patience, and unwavering support have been the constant background of this thesis. Whatever this work represents, it would not have been possible without them.

Contents

I Summary

This thesis consists of three chapters on continuous-time finance, all centred on the problem of optimal investment. A common theme throughout is that the classical benchmark of a frictionless and complete market is often too restrictive for economically relevant applications. The first chapter studies equilibrium returns when investors act strategically and face trading frictions, whereas the latter two chapters analyse optimal investment and utility-based valuation in incomplete semimartingale markets with non-replicable streams.

In particular, Chapter II studies an Itô financial market in which the returns of risky assets are determined endogenously through a market-clearing condition amongst heterogeneous risk-averse investors with quadratic preferences and random endowments. Investors act strategically by taking into account the effect of their orders on the assets' drift. A frictionless market and one with (exogenous) quadratic transaction costs are analysed and compared. In the former, we derive the unique Nash equilibrium at which investors' demand processes reveal hedging needs that differ from their true ones, resulting in a deviation of the Nash equilibrium from its competitive counterpart. In the presence of price impact and transaction costs, we characterise the Nash equilibrium via the unique solution of a system of FBSDEs and obtain a closed-form expression for the corresponding equilibrium returns. We show that, under common risk aversion and in the absence of noise traders, transaction costs and price impact do not affect equilibrium returns. When noise traders are present, comparing each market structure with its own frictionless benchmark shows that the transaction cost

component of the Nash equilibrium coincides with the competitive one for two investors, but is dampened when more than two investors participate. Comparing instead the full frictional competitive and Nash equilibria, the return gap is shaped jointly by the level of noise trader demand and the trading rate pressure induced by transaction costs.

Chapter III considers an incomplete financial market with general continuous semimartingale dynamics in which a log-utility investor receives, in addition to an initial capital, units of a non-replicable endowment process. Using duality methods, we derive a fourth-order expansion of the primal value function with respect to the number ϵ of endowment units held by the investor. This expansion, in turn, provides the basis for a second-order approximation of the corresponding optimal wealth process. The key quantities governing these asymptotics are expressed through Kunita–Watanabe projections, in close analogy with the structure that appears in lower-order results of the same kind. The treatment covers finite and infinite horizons within a unified framework.

Chapter IV turns to the valuation problem associated with a small non-replicable liability stream in the same general semimartingale setting—under log-utility preferences. Therein the notion of utility-based certainty equivalent is explored, namely the cash amount that offsets the investor's utility loss from bearing such a position. Since explicit valuation formulas are typically unavailable in incomplete markets, the analysis develops a fourth-order asymptotic expansion of this quantity with respect to the size of said stream. The resulting formula makes precise how market incompleteness affects prices beyond the leading terms and shows that higher-order effects can become economically relevant even for moderate positions. The chapter also treats the infinite-horizon case. Beyond its independent theoretical interest, this offers a convenient approximation mechanism for investors with long maturities. In that regime, the infinite-horizon formulation yields a simpler and more tractable benchmark for utility-based pricing, while still capturing the main economic effects of long-dated non-replicable cash flows.

II Continuous-time Equilibrium Returns in Markets with Price Impact and Transaction Costs

1 Introduction

Discussion

It is an undeniable fact that many contemporaneous financial markets feature large institutional investors whose trading strategies affect asset prices (see, among others, the empirical studies in [KY19] and [HHLT21]). Motivated by the related statistical evidence, a large body of theoretical research on price impact modelling has been developed both in static and dynamic settings (for comprehensive literature reviews on price impact models we refer to [RY22] and [Web23]).

The majority of the existing models in this area, however, do not take into account a prevalent force in the markets: the exogenous transaction costs. Especially when a market is thin, i.e., the average trading volume is low, transaction costs burden investors in several forms, such as broker and trading platform fees ([FIM18]) and search costs for counterparties ([DGP05]). These transaction costs can be viewed as an additional layer of market friction, compounding the challenges already posed by the market's inherent non-competitiveness.

In this chapter, we analyse a continuous-time market populated by strategically behaving investors who account for their price impact while incurring transaction costs on their trades. We consider risky assets following Itô dynamics, with their drifts endogenously determined at equilibrium. The model accommodates the presence of noise traders, while the primary market participants are mean-variance investors with heterogeneous random income processes. Driven by their hedging needs, these investors trade risky assets, with the market-clearing condition determining the equilibrium drift of asset returns. Unlike much of the existing literature (see, among others, [CC98] and [Bas00]), we assume—as in [RW15]—that all investors are strategic, in the sense that each one takes into account her impact on market equilibrium when submitting an order. However, unlike [RW15], we impose (as in [BFHMK18]) quadratic transaction costs on investors' trading, which we argue provides a more realistic market setting for thin financial markets. To the best of our knowledge, this is the first study to model equilibrium returns in a continuous-time market with heterogeneous investors, while simultaneously accounting for both endogenous price impact and transaction costs.

Contributions

This chapter analyses and compares two main market settings: one without exogenous transaction costs (the frictionless case) and one with them (the frictional case). We use the frictionless market as a benchmark to highlight the effect of transaction costs on market equilibrium and, at the same time, to introduce more transparently the way we model investors' price impact. For this, we write the market's equilibrium return as a function of each investor's demand given that the rest of the investors act optimally (similarly to the static models of [Ant17] and [AK17]). This yields a best-response function which in turn leads to a continuous-time Nash equilibrium for market returns. Although our analysis formulates equilibrium in terms of asset returns rather than prices, the notion of price impact is economically unchanged. Because asset prices are directly linked to cumulative returns, any change in the latter translates into a corresponding shift in equilibrium pricing. This interpretation is consistent with the relevant literature, following [Kyl85] and [Vay99]. Hence, even though prices are "normalised out" of the equilibrium representation, the terminology price impact captures the economic mechanism by which investors internalise how their trades affect future price dynamics.

We derive these equilibrium returns in closed-form and verify that investors' price impact always alters the market's drift when noise traders are present (even under common risk aversion). In fact, at the equilibrium trading, price impact makes the investors reveal different hedging needs than their true ones. In particular, this leads to lower non-competitive equilibrium returns (and hence higher equilibrium prices) when noise traders buy the risky assets (from equally risk tolerant investors). Consistent with the related literature ([MR17] and [AK24]), we also show that Nash equilibrium yields higher utility for investors with sufficiently low risk aversion, compared with the associated competitive equilibrium.

We then apply this price impact concept in a market with quadratic transaction costs. For such an extension, we have to deal with a number of challenges. It is well known that equilibrium models with transaction costs are often intractable, because trading costs complicate the investors' individual optimisation objectives. In our case, the difficulty is further compounded by the absence of a general explicit expression for competitive equilibrium returns, making it challenging to directly identify investors' price impacts. For this, we restrict the class of admissible strategies and model parameters and assume common risk tolerance, which in turn leads to the (frictional) best-response strategy satisfying an explicitly solvable coupled but linear FBSDE. Having said that, it should be noted that the assumption of homogeneous risk preferences can be relaxed in the case of two investors. Building on the above, we characterise the investors' Nash equilibrium demands through an associated system of linear FBSDEs, which admits an explicit solution. We then derive a closed-form expression for the (frictional) Nash equilibrium returns.

The closed-form formulas allow us to compare the market's returns under different types of equilibria in two directions: with and without transaction costs, and with and without price impact. The impact on investors' frictional objectives comes both in their linear and quadratic parts. The linear part depends solely on the direction of noise traders' demand, while the quadratic part reflects the negative effect of transaction costs on investors' strategic behaviour. We also highlight the importance of homogeneous risk preferences for equilibrium comparisons. In particular, when investors have common risk aversion and there are no noise traders in the market, we show that the frictional Nash equilibrium reduces to its frictionless competitive counterpart, a result that also holds under a competitive market structure; moreover, as exogenous transaction costs tend to zero, it converges to the corresponding frictionless Nash equilibrium.

We further interpret these comparisons by separating two effects. The first concerns the incremental contribution of transaction costs relative to each market structure's own frictionless benchmark. The key mechanism is demand shading: a strategic investor moderates her price-taking demand because she internalises the effect of her order on equilibrium returns. When only one investor behaves strategically, this makes the residual market effectively thinner from her perspective and hence amplifies the transaction cost component relative to the competitive market structure. When all investors behave strategically, however, all demands are moderated simultaneously, so these strategic distortions partly offset through market clearing. In the two-investor case, the strategic interaction is purely bilateral: the moderation of one investor's demand is matched by the opposite position of the other. Hence, the transaction cost component remains the same as in the competitive market structure. With more than two investors, the interaction becomes multilateral, since the adjustment is spread across several strategic counterparties. The offsetting effect is then only partial, and the transaction cost component becomes smaller than in the competitive benchmark.

The second comparison concerns the frictional Nash and competitive equilibrium returns. In the frictionless case, their difference is governed only by the level of noise trader demand: when strategic investors moderate their position, the equilibrium returns adjust so that they are willing to absorb the resulting imbalance. In the frictional case, however, an additional effect appears through the trading rate pressure generated by transaction costs. Thus, the comparison between the two frictional returns is no longer driven only by the accumulated noise trader position, but also by the way this position changes over time. Since these two effects need not move in the same direction, the difference between the frictional Nash and competitive returns may change sign in the scalar case, and more generally depends on their vector interaction.

Related literature

This work is linked with two strands of the literature: equilibrium models with price impact and models of optimal investment and equilibrium pricing under transaction costs.

There are different ways to categorise price impact models (static versus dynamic settings, exogenous versus endogenous price impact, and symmetric versus asymmetric investors). On the one hand, there is a vast literature that assumes exogenous price impact, e.g. [CC98], [AC00], [ATHL05], [HW04] and [SZ19]. Therein, specific assumptions are used to justify the nature of the price impact that could be permanent or temporary. In contrast to this approach, we follow the literature that starts from the seminal work of [Kyl85] and considers an endogenously derived price impact. In fact, similarly to the static models of [Ant17], [Viv11] and [MR17] and the time-dynamic ones in [Vay99] and [RW15], all investors, except the noise ones, have symmetric bargaining power and internalise their ability to affect the equilibrium returns. Related price impact mechanisms in optimal investment and derivative demand are studied in [ARS21], while [AR] analyses strategic informed trading and the value of private information. More recently, a continuous-time model for non-competitive markets has been studied in [CLS21], where in contrast to our setting, each investor's optimisation problem considers the surplus of wealth above a stochastic target at the terminal time. This model with the presence of linear transaction costs has been analysed by [NW22]. This distinction is important for the economic interpretation of transaction costs. In the proportional cost setting of Noh and Weston, transaction costs determine how long agents are willing to trade and can make them cease trading before the end of the trading period. In our quadratic cost setting, transaction costs have a different effect. Since the marginal cost of increasing the trading rate is proportional to the trading rate itself, it vanishes at zero trading speed. Hence, transaction costs do not induce early stopping in our linear-quadratic formulation. Instead, they dampen trading continuously: investors adjust their portfolios more gradually, and the optimal trading rate tapers off as the terminal date approaches.

There are two main advantages of imposing endogenous price impact: the first being that there is no need to impose an exogenous parametrisation of impact and the second that we can analyse the equilibrium price impact in terms of investors' characteristics such as their risk exposures, risk preferences and/or even their population. The broader equilibrium pricing perspective is also connected to the partial-equilibrium approach of [AŽ10a, AŽ10b] and to equilibrium pricing under ambiguity in [AS24].

Our work also contributes to equilibrium market models with transaction costs incorporating a non-competitive market structure. In the same spirit of thin market strategic interaction, [AKV20] shows how market thinness affects effective risk aversion in risk-sharing markets, while [AB20] studies Nash equilibria in optimal reinsurance bargaining. Under that perspective, we partially extend the equilibrium models of [BFHMK18] and [HMKP21] and show that price impact not only can be taken into account, but also the equilibrium outcomes can be qualitatively compared. Equilibrium models with dynamic trading and transaction costs, however without assuming price impact, have been developed both in discrete-time settings (e.g. [BD19]) and under continuous-time (e.g. [Vay98]). Equilibrium continuous-time models with transaction costs are also developed in [VV99] and [Wes18], but under simplified settings with deterministic asset prices. In this context, [GP16] studies the equilibrium returns in a model with a single agent and noise traders—however, the demand is exogenously given. In the same work, the authors also consider optimal allocation under mean-variance preferences as well as temporary and persistent price impact—which, unlike our work, is also given exogenously. These results are extended for the case of exponential preferences in [CDJ20].

On the other hand, Nash equilibria under price impact and transaction costs in a dynamic setting have been studied in [SZ19] and [LS19] and more recently extended in [CL24], under the notion of price-impact game. In contrast to our setting, these works consider exogenous price impact, while the trading occurs in discrete times.

Structure of the chapter

This chapter is organised as follows: Section 2 presents the model's ingredients and the simple case of the competitive equilibrium with no transaction costs. In Section 3, we introduce the price impact modelling in a frictionless market, derive the closed-form expression of the Nash equilibrium returns and comment on its intuition. In Section 4, we generalise the model by including quadratic transaction costs and solve the system of FBSDEs that characterises the corresponding Nash equilibrium returns. The importance of such results is summarised through a discussion on the effect of price impact, and transaction costs on the returns of the market. Finally, in Appendix A.1 we provide the proofs of Section 3, while the proofs of Section 4 are given in Appendix A.2.

2 Setup

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with an augmented filtration $\mathbb{F} := (\mathcal{F}(t))_{t\in[0,T]}$ generated by a standard d-dimensional Brownian motion W, and a finite time horizon $T > 0$. We also fix a constant $r \geq 0$, which will stand for the time discount rate.

2.1 The market

We consider a market with $d+1 \in \mathbb{N}_{>1}$ assets and their cumulative return processes $R = (R^i : 0 \leq i \leq d)$. The first one is exogenous and holds no risk, while the remaining assets are risky with returns modelled by the following $\mathbb{R}^d$-valued Itô process:

$$\begin{bmatrix} \mathrm{d}R^1(t) \\ \vdots \\ \mathrm{d}R^d(t) \end{bmatrix} = \begin{bmatrix} \nu^1(t) \\ \vdots \\ \nu^d(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \sigma_{1,1} & \cdots & \sigma_{1,d} \\ \vdots & & \vdots \\ \sigma_{d,1} & \cdots & \sigma_{d,d} \end{bmatrix} \begin{bmatrix} \mathrm{d}W^1(t) \\ \vdots \\ \mathrm{d}W^d(t) \end{bmatrix}, \quad R^i(0) = 0,\ i \in \mathcal{D}, \tag{2.2.1}$$

where $\mathcal{D} := \{1, ..., d\}$. Here, the $\mathbb{R}^d$-valued (local) return process $\nu = (\nu^i : i \in \mathcal{D})$ is going to be endogenously determined at the market equilibrium, while the constant $\mathbb{R}^{d\times d}$-valued infinitesimal covariance matrix $\Sigma := \sigma\sigma'$ is given exogenously and is assumed to be (symmetric) positive definite. In the above representation W can be seen as the vector of market risk factors that add noise to the assets' returns. Through (2.2.1) we also characterise the dynamics of the risky assets' price processes $S = (S^i : i \in \mathcal{D})$ as:

$$\mathrm{d}S^i(t) = S^i(t)\mathrm{d}R^i(t), \quad i \in \mathcal{D}, \tag{2.2.2}$$

together with the specified initial values $(S^i(0) : i \in \mathcal{D})$ which are strictly positive. Regarding the riskless asset we require it pays and charges no interest, i.e. we set $R^0 = 0$ and $S^0 = 1$. This means that we can consider the wealth and endowment processes in discounted terms[1].

2.2 Investment opportunities

We consider $N \in \mathbb{N}_{>1}$ investors with quadratic preferences over their wealth at terminal time T and risk tolerance coefficients denoted by $\delta_m > 0$ for each $m \in \mathcal{I} := \{1, ..., N\}$. They

[1]The main economic assumption that is imposed regarding the riskless asset in our framework is that its price is not cleared within the equilibrium. In other words, we consider the investors who trade and clear the market of a bundle of risky assets, while they do not impact or change the value of the riskless asset. On the other hand, we assume that $R^0(0) = 0$ for clarity, since we would arrive at the same results by taking R^0 to be an arbitrary adapted, continuous finite-variation process with $R^0(0) = 0$, S^0 satisfying: $\mathrm{d}S^0(t) = S^0(t)\mathrm{d}R^0(t)$ for $S^0(0) = 1$ and considering $R^i - R^0,\ S^i/S^0,\ i \in \mathcal{D}$, as our respective return and price processes, which once more satisfy (2.2.2). We can summarise the aforementioned as normalising the riskless asset price process to one, and "expressing everything else in its units", through discounting.

continuously trade in the market in order to hedge their exogenously given risks. We assume that each investor's random endowment process stems from her exposure to the market's risk factors. In particular, we consider the following cumulative endowment process:

$$\mathrm{d}Y_m(t) = (\zeta_m(t))' \sigma \mathrm{d}W(t), \quad m \in \mathcal{I}, \tag{2.2.3}$$

where $\zeta_m = (\zeta_m^i : i \in \mathcal{D})$ is an optional process that describes the exposure of investor m to the risk factor. The stochasticity of ζ_m for each $m \in \mathcal{I}$ implies that this exposure changes over time and hence investors' hedging needs change accordingly. Note that time-changing exposure to risk factors implies that even if the equilibrium allocation among investors is Pareto optimal, continuous changes in endowments allow for mutually beneficial trading.

Remark 2.2.1. Considering only the diffusion part of the investors' random endowment in (2.2.3) comes without loss of generality. Indeed, as stated in [HMKP21], the solution to the investors' objective functional (see (2.2.5) below) is invariant to any additional finite-variation part or to shocks that are orthogonal to the market. In other words, under suitable integrability assumptions, (2.2.3) can be generalised to:

$$\mathrm{d}Y_m(t) = \mathrm{d}A_m(t) + (\zeta_m(t))' \sigma \mathrm{d}W(t) + \mathrm{d}M_m^{\perp}(t), \qquad m \in \mathcal{I},$$

where the $\mathbb{R}$-valued, finite-variation process A_m models the cumulative cashflows not spanned by the assets and the $\mathbb{R}$-valued orthogonal martingale $M_m^{\perp}$ stands for any unhedgeable shocks in the investors' endowments.

The amount of money that each investor holds in the risky assets is given by the $\mathbb{R}^d$-valued processes $\phi_m = (\phi_m^i : i \in \mathcal{D})$, for each $m \in \mathcal{I}$, which shall be called portfolios or demands. Combining trading in the market with the endowment process and assuming no additional withdrawal or deposit of money until time T, we get that the dynamics of each investor's total wealth is of the following form:

$$\mathrm{d}V_m(t) = (\phi_m(t))' \mathrm{d}R(t) + \mathrm{d}Y_m(t), \quad 1 \leq m \leq N. \tag{2.2.4}$$

Although it is not necessary for the model, we consider the potential participation of noise liquidity providers, hereafter called noise traders, whose demand is exogenously given and denoted by the $\mathbb{R}^d$-valued process ψ.

We will use the following class of processes for any $1 \leq p < \infty$:

$$\mathcal{H}^p := \left\{ X : [0,T] \times \Omega \to \mathbb{R}^d : X \text{ optional s.t. } \|X\|_{\mathcal{H}^p} := \mathbb{E}\left[\int_0^T \|X(t)\|_2^p dt \right]^{1/p} < \infty \right\},$$

and impose the following standing assumption:

Assumption 2.2.2. ψ, ϕ_m, ν and $\zeta_m \in \mathcal{H}^2$, for all $m \in \mathcal{I}$.

Remark 2.2.3. Note that we could also represent the investors' holdings in units of assets as: $\theta_m^i := \phi_m^i / S^i$ for each $i \in \mathcal{D}$, $m \in \mathcal{I}$. That is, by the associativity of the semimartingale integral we have that (2.2.4) can also be written as:

$$\mathrm{d}V_m(t) = (\theta_m(t))' \mathrm{d}S(t) + \mathrm{d}Y_m(t), \quad m \in \mathcal{I}.$$

2.3 Equilibrium returns in a competitive market with no transaction costs

We start our equilibrium study with the simplest case of a competitive market without transaction costs. Although this equilibrium is not our main focus, we use it as a benchmark for equilibrium comparisons that helps the related intuition discussion. Herein, the objective of investor m is to choose a portfolio ϕ through the maximisation of the following quadratic functional $\mathcal{F}_m : \mathcal{H}^2 \to \mathbb{R}$ defined as:

$$\mathcal{F}_m(\phi) := \mathbb{E}\left[\int_0^T e^{-rt} \left((\phi(t))' \mathrm{d}R(t) + \mathrm{d}Y_m(t) - \frac{1}{2\delta_m} \mathrm{d}\left[\int_0^{\cdot} (\phi)' \mathrm{d}R(s) + Y_m \right](t) \right) \right], \quad m \in \mathcal{I},$$

where the wealth dynamics are given by (2.2.4). Hence, the optimisation problem is written as:

$$\sup_{\phi \in \mathcal{H}^2} \mathbb{E}\left[\int_0^T e^{-rt} \Big((\phi(t))' \nu(t) - \frac{1}{2\delta_m} (\phi(t) + \zeta_m(t))' \Sigma (\phi(t) + \zeta_m(t)) \Big) \mathrm{d}t \right], \quad m \in \mathcal{I}. \quad (2.2.5)$$

Remark 2.2.4. The presence of the discount factor e^{-rt} is optional, since we could simply set $r = 0$. A strictly positive r implies the presence of a penalising term on investors' objective functional when the wealth is utilised later in time.

By the strict concavity of $\mathcal{F}_m(\phi)$ the unique solution of (2.2.5), denoted by $\widehat{\phi}_m$, is derived via a simple pointwise optimisation, and has the following representation:

$$\mathcal{H}^2 \ni \widehat{\phi}_m := \delta_m \Sigma^{-1} \nu - \zeta_m, \quad \forall m \in \mathcal{I}. \tag{2.2.6}$$

The above investment strategy shall be called *the competitive frictionless optimal portfolio* of investor m. Note that the first term of (2.2.6) is the classic quadratic optimal portfolio (also called Merton's portfolio), while the second reflects the investor's hedging needs.

Under competitive market structure with no transaction costs, the local return process ν of (2.2.1) is determined simply by the market-clearing condition, where a zero-net supply is imposed:

$$\widehat{\phi}_1 + \cdots + \widehat{\phi}_N + \psi = 0. \tag{2.2.7}$$

Without transaction costs and when all investors are price takers, we readily get from (2.2.6) and (2.2.7), the explicit form for the equilibrium return process, henceforth called *frictionless competitive equilibrium returns*:

$$\mu := \frac{\Sigma(\zeta - \psi)}{\delta}, \tag{2.2.8}$$

where the aggregate investors' risk exposure and risk tolerance are defined as:

$$\zeta := \sum_{m=1}^{N} \zeta_m \quad \text{and} \quad \delta := \sum_{m=1}^{N} \delta_m. \tag{2.2.9}$$

Due to Assumption 2.2.2, we directly verify that $\mu \in \mathcal{H}^2$. Equilibrium (2.2.8) is indicative for the discussion that follows. Say $d = 1$, for simplicity. Assuming that the aggregate investors' exposure to market shocks is higher than the corresponding noise demand, the equilibrium returns increase when aggregate risk tolerance decreases and the asset's volatility increases. This is intuitive: lower risk tolerance and higher risk means that investors decrease their demand for the asset and hence equilibrium returns increase. Note that higher equilibrium returns essentially imply that investors *require* higher expected returns in order to invest on the asset, a situation that is consistent with lower demand. On the other hand, higher exposure to market shocks increases investors' demand and hence decreases the required expected returns at the equilibrium.

3 Price Impact in a Market with no Transaction Costs

We are now ready to introduce the way we model investors' *price impact*. For this, we first take the position of the n-th investor and assume (for now) that she acts strategically, while the rest of the investors behave as price takers. This set-up gives rise to a best-response function, which ultimately leads to the Nash equilibrium when all investors adopt the same strategy. Although treating all but one investor as price takers is a preliminary step towards deriving the Nash equilibrium, this analysis is insightful in its own right. It elucidates the intuition behind the price impact that is proposed in our model and also its distinction from the competitive equilibrium. Moreover, the assumption of a single strategic investor can be realistic in certain cases, such as when a large investor influences the market solely due to her size.[2]

Consider again the market-clearing condition (2.2.7) and assume that all the investors, besides the n-th one, act optimally through (2.2.6). In the spirit of [Ant17], we readily observe that the market's returns can be viewed as a function of investor n's demand. Indeed, in this context the clearing condition is written as:

$$\phi_n + \sum_{\substack{m=1 \\ m\neq n}}^{N} \left(\delta_m \Sigma^{-1}\nu - \zeta_m\right) + \psi = 0, \quad \forall \phi_n \in \mathcal{H}^2. \tag{2.3.1}$$

[2]A prominent example of such a framework is provided in [CC98], where a large investor accounts for her impact on assets' returns when optimising her policy, while other investors remain price takers. Similar to our model, the equilibrium assets' returns in [CC98] are affected by the large investor's individual optimisation. However, the price impact on returns in their model is imposed exogenously. In a way, our model builds on [CC98] by proposing a specific endogenous price impact process, which could lead to a Nash equilibrium when all investors act strategically.

Now, solving for ν we get how any demand ϕ_n submitted by the investor n affects the equilibrium returns (letting the rest of the investors act optimally as price-takers). In other words, we introduce the *price impact process* of investor n, $\nu_n(\phi) \in \mathcal{H}^2$ defined as:

$$\nu_n(\phi) = \frac{\Sigma(\zeta_{-n} - \psi - \phi)}{\delta_{-n}}, \quad \forall \phi \in \mathcal{H}^2, \tag{2.3.2}$$

where $\zeta_{-n} := \sum_{m=1, m \neq n}^{N} \zeta_m$ and $\delta_{-n} := \sum_{m=1, m \neq n}^{N} \delta_m$ denote respectively the aggregate exposure to market shocks and the aggregate risk tolerance of the rest of the investors. With a slight abuse of notation, we distinguish between the price impact process of investor n, and the local market's return process ν—which is to be determined at the equilibrium.

Note that the functional $\nu_n(\phi)$ indicates that when investor n submits positive demand for the asset, the equilibrium return decreases at a rate of Σ/δ_{-n}. In turn, if investors collectively require less compensation for a given level of risk—that is, if the equilibrium expected return falls—then the asset must be more expensive relative to its future payoffs.

Assuming that investor n acts strategically, we essentially consider her optimal demand problem (2.2.5) to take into account her impact on equilibrium returns through (2.3.2). In other words, her adjusted objective functional for investor n, $\widetilde{\mathcal{F}}_n : \mathcal{H}^2 \to \mathbb{R}$, is given as:

$$\widetilde{\mathcal{F}}_n(\phi) := \mathbb{E}\Big[\int_0^T e^{-rt}\Big((\phi(t))'\nu_n(t;\phi) - \frac{1}{2\delta_n}(\phi(t) + \zeta_n(t))'\Sigma(\phi(t) + \zeta_n(t))\Big)\mathrm{d}t\Big]. \tag{2.3.3}$$

Recalling (2.3.2) and rearranging the terms, we get the following optimisation problem:

$$\sup_{\phi \in \mathcal{H}^2} \mathbb{E}\Bigg[\int_0^T e^{-rt}\Big((\phi(t))'\Sigma\Big(\frac{\zeta_{-n}(t) - \psi(t)}{\delta_{-n}} - \frac{\zeta_n(t)}{\delta_n}\Big) - \frac{k_n}{2}(\phi(t))'\Sigma\phi(t) - \frac{1}{2\delta_n}(\zeta_n(t))'\Sigma\zeta_n(t)\Big)\mathrm{d}t\Bigg], \tag{2.3.4}$$

where:

$$k_n := \frac{1}{\delta_{-n}} + \frac{1}{2\delta_n}.$$

To simplify the formulas, we introduce the notation for each investor's relative risk tolerance:

$$\lambda_m := \frac{\delta_m}{\delta} \quad \text{and} \quad \lambda_{-m} := 1 - \lambda_m, \quad m \in \mathcal{I}.$$

The following proposition deals with the solution of problem (2.3.4).

Proposition 2.3.1. Under Assumption 2.2.2, the optimisation problem (2.3.4) has the following unique solution:

$$\mathcal{H}^2 \ni \widetilde{\phi}_n := \frac{\lambda_n(\zeta_{-n} - \psi) - \lambda_{-n}\zeta_n}{\lambda_n + 1}. \tag{2.3.5}$$

Henceforth, we also refer to (2.3.5) as the *best-response strategy* of investor n in the frictionless market.

Remark 2.3.2. Comparing the best-response strategy of investor n with the respective competitive optimal strategy (2.2.6) at the competitive equilibrium we directly get:

$$\widetilde{\phi}_n = \frac{1}{\lambda_n + 1}\widehat{\phi}_n.$$

Hence, investor n has an incentive to decrease the volume of her demand (positive or negative) when compared to her price-taking demand. Intuitively, this adjustment reflects the investor's awareness that increasing her demand lowers the expected return, which in turn raises the asset prices. In other words, by demanding more of a risky asset, the investor bids up its future price, making subsequent gains less attractive. Anticipating this adverse effect on her own payoff, she optimally moderates her position. This behaviour is actually what generates a downward-sloping aggregate demand schedule in equilibrium.

On the other hand, we have that the equilibrium returns when only investor n acts strategically, μ_n, take the following form:

$$\mathcal{H}^2 \ni \frac{\lambda_n}{\lambda_n + 1}\mu_{-n} + \frac{1}{\lambda_n + 1}\mu, \tag{2.3.6}$$

where $\mu_{-n} := \Sigma(\zeta_{-n} - \psi)/\delta_{-n}$ is the competitive equilibrium without the presence of the strategic investor n and μ is the competitive equilibrium when all investors are present (given in (2.2.8)). This implies that the equilibrium returns are the average return of the competitive equilibrium μ for all the market and the competitive equilibrium without investor n, meaning that the impact of investor n's strategy drives the equilibrium return between μ_{-n} and μ.

3.1 Investors' revealed risk exposure

By applying strategy (2.3.5), investor n drives the equilibrium returns to (2.3.6), i.e. to a different equilibrium than the competitive one. As mentioned in the introductory section, the main reason for investors' trading in this market is to hedge their risk exposure (the other reason is to offset noise traders' demand). When investor n applies strategy (2.3.5), she essentially reveals different hedging needs than her true ones. In order to see that, recall the price-taking demand (2.2.6), where the risk exposure is the intercept point of the affine demand function. Under price impact, investor n submits a different demand function and hence the implied risk exposure is shifted. This creates the very useful and intuitive notion of the *revealed risk exposure* that is formally defined below.

Definition 2.3.3 (Revealed risk exposure process)**.** For any given demand process $\phi \in \mathcal{H}^2$ of investor m and local returns $\nu \in \mathcal{H}^2$ pinned down at the equilibrium, the *revealed exposure* of investor m is defined as:

$$\zeta_m(\phi) := \delta_m \Sigma^{-1}\nu - \phi, \quad \forall m \in \mathcal{I}. \tag{2.3.7}$$

With a slight abuse of notation, we distinguish between the revealed risk exposure process $\zeta_m(\cdot)$ and the true exposure process, ζ_m of each investor m. Departing from the above definition, we derive the following result:

Proposition 2.3.4. Impose Assumption 2.2.2. The revealed exposure from the equilibrium returns μ_n, in (2.3.6), and the demand $\widetilde{\phi}_n$ of investor n, in (2.3.5), is given as:

$$\zeta_n(\widetilde{\phi}) = \frac{\lambda_n^2}{1-\lambda_n^2}(\zeta_{-n} - \psi) + \frac{1}{1+\lambda_n}\zeta_n. \tag{2.3.8}$$

Remark 2.3.5. From the representation of the revealed exposure (2.3.8), we get two key observations. First, the revealed exposure of the strategic investor is written as a function of the other investors' exposure. Aiming at the Nash equilibrium, we could, therefore, apply the best-response function in terms of investors' risk exposure, rather than investors' demands. As we will see below, this formulation is quite convenient.

Second, we directly get an interesting intuition: The coefficient of the second term in (2.3.8) is positive and less than one. Whence, investor n's revealed risk exposure is a fraction of her true one and a part of the net aggregate exposure $\zeta_{-n} - \psi$. The latter can be called residual exposure, since it stands for the aggregate exposure to market of the rest of the investors (after the offsetting of the noise traders' demand). In other words, strategic investor n on the one hand lessens the true exposure of her endowment and on the other supplements it by a fraction $\propto (\zeta_{-n} - \psi)$. We could think of the above strategy as investor n declaring only a part of her true exposure to the rest of the investors, while exploiting her impact on the equilibrium by submitting exposure similar to her counterparties.

A reasonable question that arises is when does the strategic demand of an investor coincide with the competitive one. As stated in the following corollary, this coincidence never happens, as long as the investor's demand at competitive equilibrium is non-zero.

Corollary 2.3.6. Under Assumption 2.2.2, it holds that:

$$\widehat{\phi}_n = \widetilde{\phi}_n, \qquad \text{if and only if} \qquad \widehat{\phi}_n = 0.$$

Note that it is rather standard to have $\widehat{\phi}_n \neq 0$. Indeed, from (2.2.6) and (2.2.8), we get that condition $\widehat{\phi}_n = 0$ holds if and only if $\lambda_n(\zeta_{-n} - \psi) = (1 - \lambda_n)\zeta_n$. Even if the investors' endowments have a specific linear dependence, the presence of noise traders makes the condition extremely rare.

3.2 The Nash equilibrium returns in a market with no transaction costs

We now consider a market where all the investors act strategically by applying the strategy that is determined by problem (2.3.3). Recalling (2.3.7), we define the Nash equilibrium returns using the best-response of the revealed risk exposure. The formal definition is given below:

Definition 2.3.7. The processes $(\breve{\phi}_m \in \mathcal{H}^2 : m \in \mathcal{I})$ are called investors' *Nash optimal demands* if for each $m \in \mathcal{I}$, $\breve{\phi}_m$ solves (2.3.4), where the exposure of the other investors is shifted to $\zeta_i(\breve{\phi}),\ i \in \mathcal{I} \setminus \{m\}$. In turn, the return process induced by the market-clearing condition when all the investors act according to their Nash optimal demands, is called *Nash equilibrium returns*.

In view of (2.3.5) and (2.3.7), investors' optimal demands $(\breve{\phi}_m \in \mathcal{H}^2 : m \in \mathcal{I})$ solve the following linear system:

$$\breve{\phi}_m = \frac{\lambda_m(\zeta_{-m}(\breve{\phi}) - \psi) - \lambda_{-m}\zeta_m}{\lambda_m + 1}, \quad m \in \mathcal{I}, \tag{2.3.9}$$

where $\zeta_{-m}(\breve{\phi}) := \sum_{i=1, i\neq m}^{N} \zeta_i(\breve{\phi})$. The fact that solutions to the above belong in $\mathcal{H}^2$ is guaranteed by Assumption 2.2.2 and system's linearity. In other words, Nash equilibrium demands have the form of the best-response strategy (2.3.5), but with the rest of the investors being strategic too. Armed with the above, we give the following characterisation:

Theorem 2.3.8. Under Assumption 2.2.2, the unique Nash equilibrium returns in a market without transaction costs take the following form:

$$\mathcal{H}^2 \ni \breve{\mu} := \frac{\Sigma}{\delta}\frac{(\zeta - \psi) - \sum_{m=1}^{N}\lambda_m\zeta_m}{1 - \sum_{m=1}^{N}\lambda_m^2}. \tag{2.3.10}$$

As Corollary 2.3.6 implies, Nash equilibrium is generally different from the competitive one. Indeed, using (2.2.8) and (2.3.10) we readily get that:

$$\breve{\mu} - \mu = \frac{\Sigma}{\delta}\frac{1}{1 - \sum_{m\in\mathcal{I}}\lambda_m^2}\sum_{m\in\mathcal{I}}\lambda_m\widehat{\phi}_m, \tag{2.3.11}$$

where we recall that $\widehat{\phi}_n$ is the investor n's optimal demand at the competitive equilibrium, given in (2.2.6). The difference of equilibrium returns with and without price impact given in (2.3.11) is also called *liquidity premium* created by the presence of price impact. Similarly to the condition $\widehat{\phi}_n = 0$, the weighted sum $\sum_{m\in\mathcal{I}}\lambda_m\widehat{\phi}_m$ rarely equals zero. For example, when investors have the same risk tolerance, i.e. $\delta_m = \bar{\delta}$, $\forall m \in \mathcal{I}$, the liquidity premium equals to $-(\Sigma/\bar{\delta}N(N-1))\psi$, which vanishes if and only if there is no noise demand in the market. In other words, there is no liquidity premium due to price impact only when investors' risk tolerances are equal and there are no noise traders.

On the other hand, even if $\psi = 0$, then $\breve{\mu} = \mu$ holds if and only if $\zeta = \sum_{m\in\mathcal{I}} \lambda_m \zeta_m / \sum_{m\in\mathcal{I}} \lambda_m^2$, which is quite a rare condition.

The effect of price impact on the equilibrium returns has also a monotonicity property. In particular, we get from (2.3.11) that when $\delta_m = \bar{\delta}$, $\forall m \in \mathcal{I}$: $(\breve{\mu} - \mu)'\psi \leq 0$, which means that price impact creates liquidity premium on the assets' required return in the opposite direction to noise traders' demand. This is apparent when $d = 1$: A positive (resp. negative) noise demand decreases (resp. increases) the equilibrium required asset return, or equivalent increases (resp. decreases) its price. This implies that the noise traders will buy (resp. sell) the asset at a higher (resp. lower) price (an effect that benefits the investors at the aggregate level).

Another issue that stems from equilibria comparison is whether the utility gain (or utility surplus) that is created by the trading is higher or lower for each investor due to price impact. In general, competitive equilibrium is consistent with Pareto optimal allocation and hence another equilibrium allocation decreases the aggregate gain of utility. However, at the individual level, Nash equilibrium may be beneficial for some investor(s). This feature appears in several recent non-competitive settings such as [AK24] and [MR17].

To this end, we define the utility surplus for each asset return $\nu \in \mathcal{H}^2$ and each admissible strategy $\phi \in \mathcal{H}^2$ of the investor m as $\mathbb{US}_m^{\nu}(\phi) := \mathcal{F}_m(\phi) - \mathcal{F}_m(0)$, for each $m \in \mathcal{I}$.

In view of the related literature we get the following limiting result for investor $n = 1$ (which can be directly generalised to any investor by symmetry):

Proposition 2.3.9. Impose Assumption 2.2.2 and fix $(\delta_2, ..., \delta_N) \in (0, \infty)^{N-1}$. Then, we get:

$$\lim_{\delta_1 \to \infty} \mathbb{US}_1^{\mu}(\widehat{\phi}) = 0$$

and on the other hand:

$$\lim_{\delta_1 \to \infty} \mathbb{US}_1^{\breve{\mu}}(\breve{\phi}) = \mathbb{E}\left[\int_0^T e^{-rt} \frac{1}{4\delta_{-1}} (\zeta_{-1}(t) - \psi(t))'\Sigma(\zeta_{-1}(t) - \psi(t))\mathrm{d}t\right] \geq 0.$$

Note that $\delta_1 = \infty$ refers to risk-neutral investor (vanishing risk aversion). Due to positive definiteness of Σ, $\lim_{\delta_1\to\infty} \mathbb{US}_1^{\breve{\mu}}(\breve{\phi})$ is normally strictly positive and becomes zero if and only if the noise traders' demand exactly matches (in a $dt \otimes d\mathbb{P}$ sense) the hedging demands of the rest of the investors. Therefore, the risk-neutral investor benefits from the price impact as long as noise traders' demand is different from the rest of the investors' risk exposure. In fact, her utility gain essentially stems from the aggregate difference $\zeta_{-1} - \psi$.

4 Price Impact in a Market with Transaction Costs

In this section we aim to generalise the concept of the Nash equilibrium under the presence of exogenous market transaction costs. As in the frictionless case, we build our price impact equilibrium upon the corresponding competitive one. For this, we begin our analysis with a brief introduction on the competitive equilibrium when each investor faces transaction costs upon trading.

4.1 Competitive market with transaction costs

As mentioned in the introductory section, we build on the setting of [BFHMK18] and consider the market's transaction costs—which are of quadratic form on the rate of trading. Hence, the frictional (i.e. with transaction costs) analogue of the competitive functional, denoted by $\mathcal{F}_{\Lambda,m}$, is defined as:

$$\mathcal{F}_{\Lambda,m}(\phi) := \mathbb{E}\left[\int_0^T e^{-rt}\Big((\phi(t))'\nu(t) - \frac{1}{2\delta_m}(\phi(t)+\zeta_m(t))'\Sigma(\phi(t)+\zeta_m(t)) - (\dot{\phi}(t))'\Lambda\dot{\phi}(t)\Big)\mathrm{d}t\right], \tag{2.4.1}$$

for each $m \in \mathcal{I}$, where $\dot{\phi}_m(t) := \mathrm{d}\phi_m(t)/\mathrm{d}t$ stands for the investor's trading rate and Λ is an $\mathbb{R}^{d\times d}$-valued constant diagonal and positive definite matrix. In words, investors face the same costs that are proportional to the square of their trading rate. Furthermore, since the frictional competitive functional involves the investor's trading rate, a natural assumption to avoid infinite transaction costs is to impose $\dot{\phi}_m \in \mathcal{H}^2$, for each $m \in \mathcal{I}$.

Focusing on demand functions of the form $\phi_m = \int_0^\cdot \dot{\phi}_m(t)\mathrm{d}t,\ m \in \mathcal{I}$, we can state our maximisation objective in optimal-control terms. More precisely, for each $m \in \mathcal{I}$, ϕ_m can be viewed as a state variable which is gradually evolved by adjusting to the control $\dot{\phi}_m$. Hence, the objective of investor m for a competitive market with transaction costs could be written as

$$\sup_{\dot{\phi}_m \in \mathcal{H}^2} \mathcal{F}_{\Lambda,m}(\dot{\phi}). \tag{2.4.2}$$

From [BFHMK18, Lemma 4.1], we get that (2.4.2) has a unique solution which is characterised by the following coupled and linear Forward-Backward SDE (FBSDE) for each $m \in \mathcal{I}$:

$$\begin{aligned} &\mathrm{d}\phi_m(t) = \dot{\phi}_m(t)\mathrm{d}t, \quad \phi_m(0) = 0, \\ &\mathrm{d}\dot{\phi}_m(t) = \mathrm{d}M_m(t) + \frac{\Lambda^{-1}\Sigma}{2\delta_m}(\phi_m(t) - \widehat{\phi}_m(t))\mathrm{d}t + r\dot{\phi}_m(t)\mathrm{d}t, \quad \dot{\phi}_m(T) = 0, \end{aligned} \tag{2.4.3}$$

where $M_m = (M_{i,m} : i \in \mathcal{D})$ is a $\mathbb{R}^d$-valued, square-integrable (continuous) martingale determined as part of the solution, and $\widehat{\phi}_n$ is the frictionless competitive optimal given by (2.2.6).

4.2 The price impact of a single investor in a market with transaction costs

We return to the concept of price impact, now imposing transaction costs on the market. As in the frictionless equilibrium, we first consider the simplified case, where only one investor (say investor n) acts strategically, while the rest remain price takers. Although there is a characterisation of the unique equilibrium in the competitive setting, there is no general explicit formula for the equilibrium returns. This comes in sharp contrast to the frictionless case and makes the determination of the impact of investor n on the equilibrium return (and hence her best-response) quite challenging. However, as we shall show below, the following assumption allows us to go beyond a characterisation of the market's equilibrium returns and deal both with the best-response and the Nash equilibrium under transaction costs.

Assumption 2.4.1. Investors have the same risk tolerances: $\delta_m = \bar{\delta}$, for each $m \in \mathcal{I}$.

Assumption 2.4.1 allows us to derive an explicit formula for the price impact process under transaction costs (recall (2.3.2) for the frictionless case). It can be relaxed, however, when we consider the case of two investors (i.e. $N = 2$), where we still get similar Nash equilibrium results (see Theorem 2.4.12 below). The assumption $N = 2$ is commonly used in equilibrium models for the sake of tractability (see for example [HMKP21]). While such thin markets may appear stylised, they play a crucial theoretical role by highlighting the relationship between price impact and differences in risk tolerance across investors.

Now, as can be seen by [BFHMK18, Theorem A.4], (2.4.3) admits a (pathwise) unique solution. More precisely, the BSDE part of the solution, assuming $\zeta_m, \nu \in \mathcal{H}^2$ for all $1 \leq m \leq N$, belongs to the following class of continuous semimartingales:

$$\mathcal{A} := \Big\{ X : \mathrm{d}X(t) = a^X(t)\mathrm{d}t + \mathrm{d}M^X(t), \text{ s.t. } X(T) = 0,\, a^X \in \mathcal{H}^2, \text{ square-integrable MG } M^X \Big\}.$$

In turn, we define the set:

$$\mathcal{W}_{\mathcal{A}} := \Big\{ X :\ \mathrm{d}X(t) = \dot{X}(t)\mathrm{d}t,\ t \in [0,T],\ X(0) = 0 \text{ and } \dot{X} \in \mathcal{A} \Big\},$$

which forms the new admissible class of investor n's demand in a market with both transaction costs and price impact. We hence update Assumption 2.2.2 to:

Assumption 2.4.2. $\zeta_m, \nu \in \mathcal{H}^2$ for all $m \in \mathcal{I}$ and $\phi_n, \psi \in \mathcal{W}_{\mathcal{A}}$.

When considering the frictional analogue of the price impact process for investor n introduced in (2.3.2), $\mathcal{A}$ becomes a natural choice for her control. In other words, by observing how the rest of the investors act optimally as price takers at the equilibrium, the strategic investor elicits information from them and informs her choices; leading to an updated optimisation objective. In particular, we have:

Lemma 2.4.3. Impose Assumptions 2.4.1 and 2.4.2. In a market with transaction costs, when only investor n acts strategically the objective functional (2.4.1) becomes $\widetilde{\mathcal{F}}_{\Lambda,n} : \mathcal{A} \to \mathbb{R}$:

$$\widetilde{\mathcal{F}}_{\Lambda,n}(\dot{\phi}) := \mathbb{E}\Big[\int_0^T e^{-rt} \Big((\phi(t))' \Big(\frac{\Sigma(\zeta_{-n}(t) - \phi(t) - \psi(t))}{\bar{\delta}(N-1)} + \frac{2\Lambda}{N-1}(a^{\dot{\psi}}(t) - r\dot{\psi}(t)) \Big)$$

$$-\frac{1}{2\bar{\delta}}(\phi(t)+\zeta_n(t))'\Sigma(\phi(t)+\zeta_n(t))-(\dot{\phi}(t))'\Big(\Lambda+\frac{2\Lambda}{N-1}\Big)\dot{\phi}(t)\Big)\mathrm{d}t\Big]. \tag{2.4.4}$$

Remark 2.4.4. Comparison of (2.4.4) and (2.3.3) directly indicates the impact of the transaction costs on the investor's objective functional. While the term $\Sigma(\zeta_{-n}-\phi_n-\psi)/\bar{\delta}(N-1)$ is common in both cases (assuming equal risk tolerances), transaction costs add two additional terms: one in the linear part which depends on noise demand $2\Lambda(a^{\dot{\psi}}-r\dot{\psi})/(N-1)$ and another in the quadratic part (on top of $\dot{\phi}'\Lambda\dot{\phi}$). In particular, for the first term we get that if noise traders sell at a constant rate, positive investor's demand increases her utility due to price impact. This effect comes solely due to the presence of transaction costs. On the contrary, the second term implies that price impact does increase the effect of the transaction costs, even when there are no noise traders.

Based on a calculus of variations argument, we get the following result for the frictional best-response strategy of investor n:

Proposition 2.4.5. Impose Assumptions 2.4.1 and 2.4.2. The unique solution $(\widetilde{\phi}_{\Lambda,n},\dot{\widetilde{\phi}}_{\Lambda,n})$ of (2.4.4) is characterised by the following FBSDE:

$$\begin{aligned}
&\mathrm{d}\widetilde{\phi}_{\Lambda,n}(t)=\dot{\widetilde{\phi}}_{\Lambda,n}(t)\mathrm{d}t, \quad \widetilde{\phi}_{\Lambda,n}(0)=0,\\
&\mathrm{d}\dot{\widetilde{\phi}}_{\Lambda,n}(t)=\mathrm{d}\widetilde{M}_n(t)+B(\widetilde{\phi}_{\Lambda,n}(t)-\mathbb{TP}_n(t))\mathrm{d}t+r\dot{\widetilde{\phi}}_{\Lambda,n}(t)\mathrm{d}t, \quad \dot{\widetilde{\phi}}_{\Lambda,n}(T)=0,
\end{aligned} \tag{2.4.5}$$

where:

$$\mathbb{TP}_n:=\widetilde{\phi}_n+\frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}}-r\dot{\psi}), \qquad B:=\frac{\Lambda^{-1}\Sigma}{2\bar{\delta}}, \tag{2.4.6}$$

$\widetilde{\phi}_n$ is the best-response strategy in the frictionless case, given in (2.3.5) and $\widetilde{M}_n$ is a continuous square-integrable martingale derived as part of the solution.

The form of (2.4.5) gives some interesting insights when compared to its frictional competitive analogue (2.4.3). In particular, both optimal trading rates adjust at the same rate, but towards a different target portfolio. In other words, both frictional competitive optimal strategy of investor n and its strategic counterpart are "myopic", in the sense that they trade towards their respective frictionless maximisers.

Equation (2.4.5) belongs to the class of linear FBSDEs, which are well-studied in the context of dynamic portfolio choice or even under more general settings (see, among others, the explicit solutions in [BFHMK18] and [BSM17] or a more general discussion in [MY07]). Following along those lines, we get:

Proposition 2.4.6. Impose Assumptions 2.4.1 and 2.4.2. The unique optimal demand $\widetilde{\phi}_{\Lambda,n} \in \mathcal{W}_{\mathcal{A}}$ that satisfies (2.4.5) admits the following explicit form:

$$\widetilde{\phi}_{\Lambda,n} = \int_0^{\cdot} e^{-\int_s^{\cdot} F(u)du} \widetilde{\mathbb{TP}}_n(s)\mathrm{d}s, \tag{2.4.7}$$

where:

$$F(t) := -\Big(\Delta G(t) - \frac{r}{2}\dot{G}(t)\Big)^{-1} B\dot{G}(t), \quad \Delta := B + \frac{r^2}{4} I_d, \quad G(t) := \cosh(\sqrt{\Delta}(T-t))$$
$$\widetilde{\mathbb{TP}}_n(t) := \Big(\Delta G(t) - \frac{r}{2}\dot{G}(t)\Big)^{-1} \mathbb{E}\Big[\int_t^T \Big(\Delta G(s) - \frac{r}{2}\dot{G}(s)\Big) B e^{-\frac{r}{2}(s-t)} \mathbb{TP}_n(s)\mathrm{d}s \Big| \mathcal{F}(t)\Big]$$

and B is given in (2.4.6).

Remark 2.4.7. The explicit representation in Proposition 2.4.6 also shows how transaction costs affect the optimal trading rate near the terminal date. Differentiating (2.4.7), we obtain the feedback form

$$\dot{\widetilde{\phi}}_{\Lambda,n}(t) = \widetilde{\mathbb{TP}}_n(t) - F(t)\widetilde{\phi}_{\Lambda,n}(t).$$

Thus, the investor trades gradually towards the forward-looking smoothed target $\widetilde{\mathbb{TP}}_n$, rather than jumping to the frictionless target portfolio. In particular,

$$\widetilde{\mathbb{TP}}_n(t) = \Big(\Delta G(t) - \frac{r}{2}\dot{G}(t)\Big)^{-1} e^{\frac{r}{2}t} \Bigg(\mathbb{E}\left[\int_0^T \Big(\Delta G(s) - \frac{r}{2}\dot{G}(s)\Big) B e^{-\frac{r}{2}s} \mathbb{TP}_n(s)\mathrm{d}s \middle| \mathcal{F}(t)\right]$$
$$- \int_0^t \Big(\Delta G(s) - \frac{r}{2}\dot{G}(s)\Big) B e^{-\frac{r}{2}s} \mathbb{TP}_n(s)\mathrm{d}s \Bigg).$$

Since G and $\dot{G}$ are continuous up to T, with $G(T) = I_d$ and $\dot{G}(T) = 0$, we have $F(t) \to 0$ as $t \uparrow T$. The preceding expression also gives $\widetilde{\mathbb{TP}}_n(t) \to 0$ as $t \uparrow T$. Hence, from the feedback form, $\dot{\widetilde{\phi}}_{\Lambda,n}(t) \to 0$ as $t \uparrow T$. Thus, quadratic transaction costs make the optimal trading rate taper off smoothly as the terminal date approaches.

This provides a useful comparison with Noh and Weston [NW22], who study price impact and proportional transaction costs. In their model, proportional transaction costs can generate a stopping effect: agents may cease trading before the end of the trading period. In the present quadratic cost setting, the analogous economic force appears differently. Since the marginal cost of increasing the trading rate is proportional to the trading rate itself, it vanishes at zero trading speed. Thus, unlike proportional costs, quadratic costs do not generate a fixed marginal cost wedge around zero trading. In the present linear-quadratic setting, the effect takes the form of a continuous dampening of the trading rate, which converges to zero at the terminal date.

Based on the above characterisation, we are able to provide a closed-form expression for the equilibrium returns under transaction costs when only investor n acts strategically and the rest of the investors are price takers.

Corollary 2.4.8. Impose Assumptions 2.4.1 and 2.4.2. In a market with transaction costs, when only investor n acts strategically, the unique equilibrium returns in this context are given by:

$$\mathcal{H}^2 \ni \mu_{\Lambda,n} := \mu_n + \frac{2N\Lambda}{N^2-1}(a^{\dot{\psi}} - r\dot{\psi}), \tag{2.4.8}$$

where μ_n is the corresponding frictionless equilibrium given in (2.3.6).

Remark 2.4.9. A direct consequence of (2.4.8) is that equal risk tolerances result in $\mu_{\Lambda,n}$ being equal to the corresponding frictionless equilibrium if there are no noise traders (i.e. no liquidity premium is created). This is in line with the price-taking equilibrium. Indeed, as shown in [BFHMK18], quadratic transaction costs do not change the equilibrium returns as long as there is homogeneous risk aversion and absence of noise traders. This also holds when, as we shall verify below, all investors act strategically. From (2.4.8), we also verify that $\mu_{\Lambda,n}$ converges to μ_n (i.e. the liquidity premium vanishes) when Λ goes to zero.

4.3 The Nash equilibrium returns in a market with transaction costs

We are now ready to generalise the concepts presented in §3.2 in a market with transaction costs. To this end, we extend Assumption 2.4.2 to each $\phi_m,\ m \in \mathcal{I}$ and begin with a characterisation of the investors' demands when all of them act strategically. More precisely, recalling that the frictional best-response strategy of each investor trades towards its frictionless counterpart, we shift the investors' exposure in accordance with Definition 2.3.7. Thus, the linear system (2.3.9) becomes the following system of linear FBSDEs, under the same boundary value conditions as (2.4.5), for $m \in \mathcal{I}$:

$$\begin{aligned} &\mathrm{d}\breve{\phi}_{\Lambda,m}(t) = \dot{\breve{\phi}}_{\Lambda,m}(t)\mathrm{d}t, \\ &\mathrm{d}\dot{\breve{\phi}}_{\Lambda,m}(t) = \mathrm{d}\breve{M}_m(t) + B(\breve{\phi}_{\Lambda,m}(t) - \breve{\mathbb{TP}}_m(t))\mathrm{d}t + r\dot{\breve{\phi}}_{\Lambda,m}(t)\mathrm{d}t, \end{aligned} \tag{2.4.9}$$

where:

$$\breve{\mathbb{TP}}_m := \frac{\overbrace{\sum_{i=1\,,i\neq m}^{N} \left(\bar{\delta}\Sigma^{-1}\nu - \breve{\phi}_{\Lambda,i}\right)}^{\zeta_{-m}(\breve{\phi}_\Lambda)} - \psi - (N-1)\zeta_m}{N+1} + \frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}_n} - r\dot{\psi}).$$

With the above in hand, we are ready to derive the point at which the frictional market equilibrates when all the agents act strategically.

Theorem 2.4.10. Impose Assumptions 2.4.1 and 2.4.2 (for all investors). In a market with transaction costs, the unique Nash equilibrium returns are given by:

$$\mathcal{H}^2 \ni \breve{\mu}_\Lambda := \breve{\mu} + \frac{2\Lambda}{N(N-1)}(a^{\dot{\psi}} - r\dot{\psi}), \tag{2.4.10}$$

where $\breve{\mu}$ is the corresponding frictionless Nash equilibrium, for equal risk tolerances, and B is defined in (2.4.6).

Remark 2.4.11. Commenting on the above, there are a couple of immediate things to note. Firstly, when the investors have equal risk tolerances and there are no noise traders in the market, (2.4.10) reduces to its frictionless competitive counterpart in (2.2.8). Secondly, when Λ tends to zero the frictional Nash equilibrium simplifies to its frictionless analogue. This is in line with Remark 2.4.9.

4.4 The case of two investors

Interestingly, as previously discussed, Assumption 2.4.1 in Theorem 2.4.10 can be dropped when considering a market of only two investors, i.e. $N = 2$.

Theorem 2.4.12. Impose Assumption 2.4.2 (for all investors) and let $N = 2$. In a market with transaction costs, the unique Nash equilibrium returns are given by:

$$\mathcal{H}^2 \ni \hat{\mu}_\Lambda := \breve{\mu} + \Lambda(a^{\dot{\psi}} - r\dot{\psi}), \tag{2.4.11}$$

where $\breve{\mu}$ is the corresponding frictionless Nash equilibrium.

Remark 2.4.13. Note that (2.4.11) coincides with (2.4.10) when $\delta_1 = \delta_2$. Now, rewriting (2.4.11) as:

$$\hat{\mu}_\Lambda = \frac{\Sigma}{\delta} \frac{\lambda_2\zeta_1 + \lambda_1\zeta_2 - \psi}{2\lambda_1\lambda_2} + \Lambda(a^{\dot{\psi}} - r\dot{\psi}),$$

we get that as $\delta_1 \to \infty$, the first term converges to $\Sigma(\zeta_2 - \psi)/2\delta_2$. In words, for a two-agent model with price impact and transaction costs, as the first investor approaches risk neutrality her (direct) effect on equilibrium returns vanishes. On the other hand, the second term—which is related to transaction costs—remains unaffected. By symmetry, the analogous conclusion holds for $\delta_2 \to \infty$.

4.5 Discussion

Our results enable a deeper understanding of the interplay between price impact and transaction costs. The starting point is the effect of price impact: As expected, a strategic investor does not submit the same demand she would submit as a price taker, but rather, internalisation of the effect of her order on the equilibrium returns alters her submitted demand schedule. In the frictionless single strategic investor case, this follows directly from Remark 2.3.2, since the best response satisfies:

$$\widetilde{\phi}_n = \frac{1}{1+\lambda_n}\widehat{\phi}_n,$$

where we recall that $\widehat{\phi}_n$ denotes the corresponding price-taking demand, and λ_n the investor's relative risk tolerance. Since:

$$\frac{1}{1+\lambda_n} < 1,$$

the strategic investor “shades" her price-taking demand. In other words, she reveals only a fraction of her true hedging needs (e.g. if she buys at the equilibrium, she buys less, anticipating that a lower demand reduces prices). Thus, price impact distorts the relation between the investor's true hedging needs and the demand she submits to the market.

At the same time, by the market clearing condition, the strategic investors must collectively absorb the noise trader position, while each one acts optimally. The effect of this mechanism, in the context of exogenous transaction costs, can be summarised through two separate comparisons: The first comparison isolates the incremental effect of the transaction costs, while the second one contrasts the frictional Nash equilibrium with the associated frictional competitive equilibrium. Keeping these two comparisons separate aims to capture different economic effects.

For the sake of clarity we focus on the case where $d = 1$ for the rest of the discussion, although similar results hold for more risky assets.

First comparison: Recalling [BFHMK18, Corollary 5.4], under common risk tolerance, the competitive liquidity premium gives:

$$\mu_\Lambda - \mu = \frac{2\Lambda}{N}\big(a^{\dot{\psi}} - r\dot{\psi}\big),$$

where μ_Λ denotes the frictional competitive equilibrium returns. Likewise, when only investor n behaves strategically, equation (2.4.8) gives:

$$\mu_{\Lambda,n} - \mu_n = \frac{2N\Lambda}{N^2-1}\big(a^{\dot{\psi}} - r\dot{\psi}\big).$$

Therefore,

$$\frac{\mu_{\Lambda,n} - \mu_n}{\mu_\Lambda - \mu} = \frac{N^2}{N^2-1} > 1,$$

whenever the common factor $a^{\dot{\psi}} - r\dot{\psi}$ is non-zero. Thus, the presence of a single strategic investor amplifies the transaction cost component relative to the competitive benchmark.

The reason is that the strategic investor faces a residual market consisting of the remaining $N-1$ price-taking investors. From her perspective, the relevant "market depth" is therefore $(N-1)\bar{\delta}$, rather than $N\bar{\delta}$, making the market effectively thinner. A reduction in her demand is then absorbed by counterparties with a smaller aggregate risk-bearing capacity, so a larger movement in the equilibrium returns is needed to clear the market.

By contrast, when all investors behave strategically and have common risk tolerance, (2.4.10) yields:

$$\breve{\mu}_\Lambda - \breve{\mu} = \frac{2\Lambda}{N(N-1)}\big(a^{\dot{\psi}} - r\dot{\psi}\big).$$

Hence,

$$\frac{\breve{\mu}_\Lambda - \breve{\mu}}{\mu_\Lambda - \mu} = \frac{1}{N-1},$$

again whenever $a^{\dot{\psi}} - r\dot{\psi}$ is non-zero. Consequently, for $N = 2$ the transaction cost premium of the (fully) strategic market structure coincides with the competitive one, whereas for $N > 2$ it becomes lower. As previously stated, the single strategic investor case is asymmetric: one investor internalises price impact, while the remaining investors are price takers. In the full Nash equilibrium, however, all investors internalise their impact. These strategic distortions partially offset in the market clearing. As a result, the transaction cost component that "remains" in equilibrium is dampened by the factor $1/(N-1)$ relative to the competitive equilibrium.

The case $N = 2$ is special. Nash equilibrium (2.4.11) gives, without imposing common risk tolerance:

$$\widehat{\mu}_\Lambda - \breve{\mu} = \Lambda\big(a^{\dot{\psi}} - r\dot{\psi}\big).$$

This coincides with the competitive liquidity premium. Any strategic moderation of one investor's demand is matched by the other investor's opposite position through the clearing condition. Hence, the contribution to the incremental transaction cost coefficient stemming from price impact cancels out, due to the purely bilateral strategic interaction.

Second comparison: The previous comparison isolates the transaction cost component relative to each model's own frictionless benchmark. A different question is whether the frictional Nash return is higher or lower than its associated frictional competitive equilibrium return.

Recall that under common risk tolerance, we have:

$$\mu = \frac{\Sigma}{N\bar{\delta}}(\zeta - \psi), \qquad \breve{\mu} = \frac{\Sigma}{N\bar{\delta}}\left(\zeta - \frac{N}{N-1}\psi\right).$$

Thus,

$$\breve{\mu} - \mu = -\frac{\Sigma}{N\bar{\delta}(N-1)}\psi.$$

This term reflects the frictionless price-impact effect. For example, if noise traders sell the risky asset, then $\psi < 0$, and therefore $\breve{\mu} - \mu > 0$. Strategic investors reduce their demand, so the equilibrium return has to increase in order for them to absorb the noise trader supply. On the other hand, it holds that:

$$\breve{\mu}_\Lambda - \mu_\Lambda = (\breve{\mu} - \mu) + \left(\frac{2\Lambda}{N(N-1)} - \frac{2\Lambda}{N}\right)(a^{\dot{\psi}} - r\dot{\psi}) = -\frac{\Sigma}{N\bar{\delta}(N-1)}\psi - \frac{2\Lambda(N-2)}{N(N-1)}(a^{\dot{\psi}} - r\dot{\psi}).$$

In particular, the first term depends on the level of noise trader demand ψ, while the second term depends on the trading rate pressure $a^{\dot{\psi}} - r\dot{\psi}$. Thus, the full comparison between the frictional Nash and frictional competitive returns depends on their interplay. To illustrate this point, suppose that noise traders sell at a constant rate:

$$\dot{\psi} = -q \qquad \text{and} \qquad q > 0.$$

Then

$$\psi(t) = -qt \qquad \text{and} \qquad a^{\dot{\psi}} - r\dot{\psi} = rq.$$

This means that

$$\breve{\mu}_\Lambda(t) - \mu_\Lambda(t) = \frac{q}{N(N-1)}\left(\frac{\Sigma}{\bar{\delta}}t - 2\Lambda r(N-2)\right).$$

In the frictionless case, the Nash return is higher than the competitive return for all $t > 0$, since

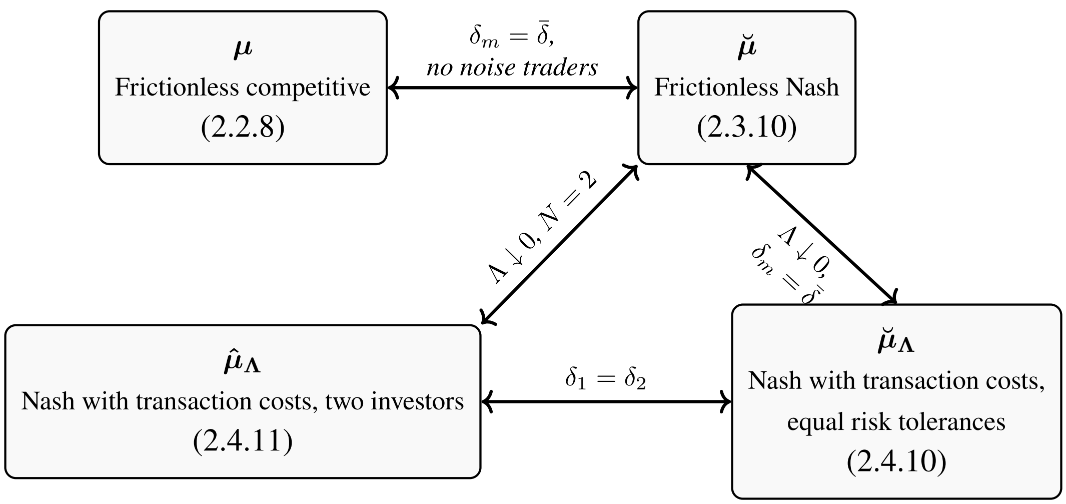
$$\breve{\mu}(t) - \mu(t) = \frac{\Sigma q t}{N\bar{\delta}(N-1)} > 0.$$

However, with transaction costs, the ordering depends on time. At early times, the accumulated imbalance $-\psi(t) = qt$ is low, and the transaction cost effect may dominate. The accumulated inventory imbalance grows over time, and strategic investors require a higher return to absorb the increasing noise traders' supply. In turn, this may change the comparison between the frictional Nash and competitive returns. In particular, the threshold in the scalar case is:

$$t^* = \frac{2\Lambda r(N-2)\bar{\delta}}{\Sigma},$$

and for $t < t^*$, the competitive frictional return can exceed the Nash frictional return, whereas for $t > t^*$, the price impact effect on ψ dominates.

We conclude with an indicative figure that depicts the connections among our different equilibrium return specifications.



Connections among different equilibrium specifications.

III Log-optimality with small liability stream

1 Introduction

Discussion

A central problem in financial economics involves an investor allocating initial wealth across an array of assets, with the goal of maximising the expected utility of terminal wealth. This optimal investment problem in continuous-time settings was initially studied in [Mer69, Mer71], using dynamic programming techniques to derive a non-linear partial differential equation characterising the value function.

A major conceptual advancement came with the development of the theory of equivalent martingale measures in [Ros76], [HK79] and [HP81], which enabled the application of martingale and duality methods to such optimisation problems. Under the assumption of market completeness, this duality approach was further developed in [Pli86], [KLS87], and [CH89, CH91]. The more intricate case of incomplete markets was addressed in the foundational works [HP91a, HP91b] and [KLSX91]. Building on these contributions, [KS99, KS03] established minimal conditions on both the utility function and the financial market under which the core results of the theory remain valid.

In the context of incomplete markets, a natural extension of the problem involves maximising expected utility when the investor receives an additional exogenous random endowment. Typical examples of this include pension funds. A related discrete time treatment of valuation and investment for pension plans in incomplete markets is given in [AB23]. In complete markets, endowments can be perfectly replicated using traded assets, effectively reducing the problem to one with augmented initial wealth and no random endowment. However, as noted among others in [HH07], real-world markets are typically incomplete, with perfect replication impeded by frictions such as transaction costs, non-replicable assets, and portfolio constraints. In such settings, assets are associated with a range of arbitrage-free prices, and the risk of holding them cannot be fully hedged through market trading alone.

Consequently, analysing the value function and its solution in an incomplete market setting is undeniably more challenging. Notable contributions addressing this challenge include [CSW01], who characterised the optimal terminal wealth in a general semimartingale model via a dual formulation, and [KŽ03] extending this framework to account for intertemporal consumption. In [HK04], both the initial capital and the units held in the endowment are treated as optimisation variables, hence not requiring the use of finitely additive measures. [OŽ09] studies the case of unbounded random endowments and utility functions defined over the entire real line, providing necessary and sufficient conditions for the existence of a solution to the primal problem. [Mos15] obtains necessary and sufficient conditions in the general framework of an incomplete financial model with a stochastic field utility and intermediate consumption, occurring according to some stochastic clock, while [Mos17] generalises the former by also incorporating a random endowment process into the model. Staying within the context of intermediate consumption, [CCFM17] show that the key conclusions of the utility maximisation theory hold under the assumptions of No Unbounded Profit with Bounded Risk (NUPBR) and of the finiteness of both primal and dual functions.

Contributions

In an incomplete financial market with general continuous semimartingale dynamics, we model an investor with log-utility preferences who, in addition to an initial capital, receives units of a non-replicable endowment process. As explicit solutions to the associated utility maximisation problem are generally unavailable, even for simple model specifications, we assume that the payoff from the non-replicable endowment is small relative to the investor's total wealth. Within this framework, our contributions to the literature are as follows:

1. *Fourth-order expansion and nearly optimal strategies:* Using duality techniques, we derive the fourth-order expansion of the primal value function with respect to the number of units ϵ held in the non-replicable endowment. In turn, this lays the foundation for expanding the optimal wealth process, in this context, up to second order w.r.t. ϵ. To the best of our knowledge, this is the first result in this direction, extending the work of [KS06], [KS07] for the case of log-utility. Interestingly enough, the key processes underpinning the aforementioned results are given in terms of Kunita-Watanabe projections, mirroring the case of lower-order expansions of similar nature.

2. *Long-horizon asymptotics:* Our model also accommodates the case of "infinite time horizons". This is a non-trivial addition which allows for the original problem to be understood in "myopic terms" for investors with distant maturities. The upshot is that one can obtain explicit results in a larger array of models, as the dependence on the horizon is eliminated.

These results have two key implications. First, they allow for a better understanding of log-optimal behaviour in the incomplete setting, in the presence of a non-replicable endowment. This stems from the fact that the first-order approximation of the optimal wealth process with non-replicable endowment, with respect to ϵ, is actually optimal, assuming market completeness. This provides unique insight on how market incompleteness affects asset allocation. Second, considering the case of arbitrarily large time horizons, i.e. the infinite-horizon setting, comes with its own merits. Namely it enables analysing assets which do not have a certain pre-specified maturity. A prominent example of such a situation arises in the context of a pension fund's liabilities.

Related literature

The existing literature on optimal investment is too vast in order to give a complete overview. Instead, we focus on the specific area of utility-based hedging and pricing, with an emphasis on the former, which is closely aligned to our work.

Exponential utility provides a particularly tractable framework for the study of utility-based hedging. In this direction, [Dav06] used duality techniques to derive an explicit form for the optimal hedging strategy, with related developments also appearing in [Mon13].

Given the scarcity of explicit results, various asymptotic approaches have been proposed for hedging in incomplete markets. For example, [Mon04, Mon07] work within a Black-Scholes framework with basis risk and approximate the hedging strategy in powers of $1-\rho^2$, where ρ denotes the correlation between traded and non-replicable assets. In [Hen02] and [HH02] the authors consider the case of power utility and derive the second-order expansions of the investor's value function with respect to a small position in the contingent claim, thereby approximating the optimal hedging strategy.

A major advance came with [KS07]. Using techniques developed in their earlier work, the authors derive a first-order approximation of the utility-based hedging strategy with respect to the number of units held in the non-replicable endowment. They also show how this approximation is related to quadratic hedging. Similar asymptotic results are found in [Mon10], which considers hedging in the presence of parameter uncertainty under exponential utility and partial information.

In the same spirit, [KR11] analyses utility-based hedging under exponential preferences for a vanishing risk aversion. [KMKV14], focusing on exponential Lévy models, presents alternative representations of the results in [KS07] for power utility functions that avoid the need for a change of numeraire.

There is also an associated strand of literature which conducts sensitivity analysis of the utility maximisation problem with respect to other model perturbations. For example [MS19] studies the sensitivity of the expected utility maximisation problem in a continuous semimartingale market w.r.t. small changes in the market price of risk. [Mos20] investigates the behaviour of the expected utility maximisation problem under small perturbations of the numeraire, while [MS24] study the response of the optimal investment problem to small changes of the stock price dynamics.

Structure of the chapter

This chapter is organised as follows: in §2 the setup of the model is given. Therein, dynamics for the financial market, investment opportunities and the illiquid stream are explained. In §3 we perform the fourth-order expansion for the value function $u(\epsilon)$ of the optimal investment problem with a random endowment. In fact, this lays the foundation to derive the second-order expansion for the optimal wealth X^{ϵ}, i.e. the unique solution of $u(\epsilon)$. In §4 we continue with a discussion of derived results as well as a concrete example. In §3, the finite- and infinite-horizon cases are treated simultaneously. Note that for the sake of readability, all proofs are deferred to Appendix B.

2 Setup

2.1 The market

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathcal{F}(\cdot)$ satisfying the usual conditions, s.t. $\mathcal{F}(0)$ is trivial a.s. We set:

$$\mathcal{F}(\infty) := \sigma\Bigg(\bigcup_{t \in \mathbb{R}_{\geq 0}} \mathcal{F}(t)\Bigg) \subseteq \mathcal{F},$$

and work within the infinite-horizon setting, i.e. $\mathbb{R}_{\geq 0} = [0, \infty)$, since any finite time horizon can be naturally embedded in it. Unless otherwise explicitly stated, identities or inequalities involving random variables are interpreted as being valid almost everywhere under $\mathbb{P}$. Similarly, comparisons between processes will be understood up to indistinguishability. Moreover for $f : \Omega \to \mathbb{R}$, let $f^+ := \max(f, 0)$, $f^- := \max(-f, 0)$ and for $1 \leq p \leq \infty$, denote the class of Lebesgue p-integrable functions in $(\Omega, \mathcal{F}, \mathbb{P})$ by $\mathbb{L}_p$. In cases where it is important to underline that the sigma-algebra is different than $\mathcal{F}$, it will be stated explicitly.

Consider an agent that trades in $1 + d \in \mathbb{N}_{>0}$ assets; a baseline asset with stochastic exponential price, driven by a continuous finite-variation process, and d risky assets. Discounting by the baseline, the market is driven by a continuous, positive, $\mathbb{R}^d$-valued semimartingale $\widetilde{S} \equiv (\widetilde{S}_i; 1 \leq i \leq d)$. In particular, we have:

$$\frac{\mathrm{d}\widetilde{S}_i(t)}{\widetilde{S}_i(t)} = \mathrm{d}\widetilde{A}_i(t) + \mathrm{d}\widetilde{M}_i(t), \qquad 1 \leq i \leq d.$$

Here, each real-valued component $\widetilde{A}_i$ of the vector-valued $\widetilde{A} \equiv (\widetilde{A}_i; 1 \leq i \leq d)$ is a continuous finite-variation process with $\widetilde{A}_i(0) = 0$, whereas each $\widetilde{M}_i$ in $\widetilde{M} \equiv (\widetilde{M}_i; 1 \leq i \leq d)$ is a continuous local martingale with $\widetilde{M}_i(0) = 0$. In turn, we introduce the continuous, nondecreasing scalar process:

$$\widetilde{O} := \sum_{i=1}^{d} \left(\int_0^{\cdot} |\mathrm{d}\widetilde{A}_i(t)| + [\widetilde{M}_i, \widetilde{M}_i] \right),$$

with $\int_0^K |\mathrm{d}\widetilde{A}_i(t)|$ denoting the total variation of $\widetilde{A}_i$ on the interval $[0, K]$, for $K > 0$. This scalar process $\widetilde{O}$ plays the role of an "operational clock" for the market of risky assets: all processes $\widetilde{A}_i$ and $[\widetilde{M}_i, \widetilde{M}_j]$ for $1 \leq i \leq d$ and $1 \leq j \leq d$ are absolutely continuous w.r.t. this clock (i.e. the respective induced measures). Hence, there exist predictable processes $\widetilde{a} \equiv (\widetilde{a}_i; 1 \leq i \leq d)$ and $\widetilde{c} \equiv (\widetilde{c}_{ij}; 1 \leq i, j \leq d)$, s.t.:

$$\widetilde{A} = \int_0^{\cdot} \widetilde{a}(t)\mathrm{d}\widetilde{O}(t), \qquad [\widetilde{M}, \widetilde{M}] = \int_0^{\cdot} \widetilde{c}(t)\mathrm{d}\widetilde{O}(t).$$

By altering it on a $(\mathbb{P} \otimes \widetilde{O})$-null set if necessary, we shall assume throughout that the process $\widetilde{c}$ takes values in the space of symmetric, nonnegative-definite, $d \times d$ matrices.

2.2 Investment opportunities

In the market $(1, \widetilde{S})$, an initial capital $x \in \mathbb{R}$ and a choice of a self-financing strategy $\theta \equiv (\theta_i; 1 \leq i \leq d)$ (assumed to be predictable and $\widetilde{S}$-vector integrable) result in a *wealth process*:

$$\widetilde{X} := x + \int_0^{\cdot} \theta(t) \mathrm{d}\widetilde{S}(t) \equiv \widetilde{X}(\cdot; x, \theta), \tag{3.2.1}$$

where the above is understood as a vector stochastic integral. To avoid the so-called doubling strategies, we restrict the above class as follows: a wealth process $\widetilde{X}$ with initial capital $\widetilde{X}(0) = x$ is called *admissible* if:

$$\widetilde{X} > 0.$$

We denote this subset of wealth processes by $\widetilde{\mathcal{X}}(x)$. The union $\cup_{x>0}\widetilde{\mathcal{X}}(x)$ is denoted by $\widetilde{\mathcal{X}}$.

The market's viability is intimately connected to the following condition, as shown in [KK21, Theorem 2.31]:

$$\widetilde{c} \text{ is non-singular, } (\mathbb{P} \otimes \widetilde{O})\text{-a.e. and } \int_0^K \widetilde{a}(t)' \widetilde{c}(t)^{-1} \widetilde{a}(t) \mathrm{d}\widetilde{O}(t) < \infty, \forall K > 0.\,[1] \tag{3.2.2}$$

In particular, the above condition implies the existence of the local martingale numeraire; denote it by $\widetilde{X}_\nu := \mathcal{E}(\widetilde{R}_\nu)$, where $\widetilde{R}_\nu := \int_0^{\cdot} \nu(t)' \mathrm{d}\widetilde{S}(t)/\widetilde{S}(t)$ and $\nu := \widetilde{c}^{-1}\widetilde{a}$ (refer to §2 in [KK21] for further details on this concept). Note that $\widetilde{X}_\nu$ can be used as a new numeraire, under which each $\widetilde{X} \in \widetilde{\mathcal{X}}$ becomes a local martingale. In fact by considering the auxiliary market, under the local martingale numeraire:

$$S := \left(S_0 := \frac{1}{\widetilde{X}_\nu}, S := \frac{\widetilde{S}}{\widetilde{X}_\nu} \right),$$

we have for each wealth process $\widetilde{X}$ in $(1, \widetilde{S})$:

$$\widetilde{X}/\widetilde{X}_\nu = x + \int_0^{\cdot} \sum_{i=1}^d \theta_i(t) \mathrm{d}S_i(t) + \int_0^{\cdot} \left(\widetilde{X}(t)/\widetilde{X}_\nu(t) - \sum_{i=1}^d \theta_i(t) S_i(t) \right) \frac{\mathrm{d}S_0(t)}{S_0(t)}, \tag{3.2.3}$$

[1] A bit more generally it is necessary that $\widetilde{a}$ is in the range of $\widetilde{c}$, $(\mathbb{P} \otimes \widetilde{O})$-a.e. Defining the "pseudo-inverse" matrix-valued process $\widetilde{c}^\dagger := \lim_{m\to\infty}((\widetilde{c} + (1/m)I_d)^{-2}\widetilde{c})$, where I_d is the identity operator on $\mathbb{R}^d$; we also require $\int_0^K \widetilde{a}(t)' \widetilde{c}(t)^\dagger \widetilde{a}(t) \mathrm{d}\widetilde{O}(t) < \infty, \forall K > 0$.

where S is a $\mathbb{R}^{1+d}$-valued local martingale. In fact, we have:

$$\begin{aligned}\frac{\mathrm{d}S_i(t)}{S_i(t)} &= \mathrm{d}R_i(t), \qquad 0 \le i \le d,\\ R_0 &:= -\int_0^{\cdot} \nu(t)' \mathrm{d}\widetilde{M}(t),\\ R_i &:= \widetilde{M}_i - \int_0^{\cdot} \nu(t)' \mathrm{d}\widetilde{M}(t), \qquad 1 \le i \le d.\end{aligned}$$

2.3 Illiquid stream

Besides trading in the financial market, the agent holds $\epsilon \in \mathbb{R}$ units of an exogenous, non-replicable (at stopping time T) cumulative stream $\widetilde{\Lambda}$ that is a right-continuous, adapted finite-variation process, s.t. $\mathrm{d}\tilde{\Lambda}(t)$ is already discounted by the baseline asset at each t. The only regularity assumption on $\widetilde{\Lambda}$ is that it can be super and subreplicated using the traded asset $\widetilde{S}$. In other words, we use the following standing assumption:

$$\{\widetilde{X} \in \widetilde{\mathcal{X}} : |\widetilde{\Lambda}(T)| \le \widetilde{X}(T)\} \neq \emptyset, \tag{3.2.4}$$

where T is any stopping time (where $\widetilde{X}(T)$ exists).

Remark III.1. Note that the non-replicable stream $\widetilde{\Lambda}$, expressed under the numeraire $\widetilde{X}_\nu$, is given by $\Lambda := \widetilde{\Lambda}/\widetilde{X}_\nu$. Now, integration by parts yields:

$$\Lambda = \int_0^{\cdot} \widetilde{\Lambda}(t-)\mathrm{d}S_0(t) + F, \tag{3.2.5}$$

$F := \int_0^{\cdot} S_0(t)\mathrm{d}\widetilde{\Lambda}(t)$. In turn, we have for a wealth process $\widetilde{X}$, with initial capital x, that:

$$\widetilde{X}/\widetilde{X}_\nu - \int_0^{\cdot} \widetilde{\Lambda}(t-)\mathrm{d}S_0(t),$$

is simply (3.2.3) with a shifted position in S_0. In other words, when considering the illiquid stream discounted by the local martingale numeraire, the non-replicable part of Λ comes solely from F.

2.4 Utility maximisation and hedging

Fix a possibly infinite-valued stopping time T, that will serve as the maturity of the agent. We focus on the model (S, F) in the sequel. Recalling (3.2.3) (under the self-financing condition), a wealth process in S: $X \equiv X(\cdot; x, \theta)$ is given via:

$$X = x + \int_0^{\cdot} \sum_{i=1}^{d} \theta_i(t) \mathrm{d}S_i(t) + \int_0^{\cdot} \left(X(t) - \sum_{i=1}^{d} \theta_i(t) S_i(t) \right) \frac{\mathrm{d}S_0(t)}{S_0(t)}. \tag{3.2.6}$$

Hence, all processes of the form shown in (3.2.3) are in fact wealth processes in S. The subset of admissible processes will be defined by:

$$X > 0 \text{ on } [\![0, T]\!] := \{(\omega, t) \in \Omega \times \mathbb{R}_{\geq 0} : t \leq T(\omega)\},$$

which shall be denoted by $\mathcal{X}(x)$ and its union $\cup_{x>0} \mathcal{X}(x)$ by $\mathcal{X}$. In this context, following [DS97], a process $X \in \mathcal{X}$ is called *maximal* if for each $X^{'} \in \mathcal{X}$ s.t. $X^{'}(T) \geq X(T)$ and $X^{'}(0) = X(0)$, we necessarily have $X^{'} = X$. Lastly, a wealth process in S, X is called *acceptable* if there exists a maximal process $X^{'} \in \mathcal{X}$ s.t. $X + X^{'} \in \mathcal{X}$. Note that when the horizon is infinite, for each $X \in \mathcal{X}$ we have $X(\infty) := \lim_{t \to \infty} X(t)$ exists a.s., by the fact that X is a (continuous) positive local martingale. Hence for any acceptable process X, $\lim_{t \to \infty} X(t)$ exists a.s. in that case since $X = X^{'} - X^{''}$, where $X^{'} \in \mathcal{X}$ and $X^{''}$ is maximal.

Assuming:

$$\begin{aligned} &F(T) := \lim_{t \to T} F(t) \text{ exists a.s., and} \\ &\{X \in \mathcal{X} : |F(T)| \leq X(T)\} \neq \emptyset, \end{aligned} \tag{3.2.7}$$

we define the following class:

$$\mathcal{X}(x, \epsilon) := \{X \text{ a wealth process in } S : X \text{ is acceptable, } X(0) = x \text{ and } X(T) - \epsilon F(T) > 0\}.$$

From the definition of acceptable processes, we also deduce that:

$$\mathcal{X}(x, 0) = \mathcal{X}(x), \qquad x > 0.$$

The set of points $(x, \epsilon) \in (0, \infty) \times \mathbb{R}$ where $\mathcal{X}(x, \epsilon)$ is not empty is a closed convex cone in $\mathbb{R}^2$. Denote its interior by $\mathcal{K} := \{(x, \epsilon) : \mathcal{X}(x, \epsilon) \neq \emptyset\}^{\mathrm{o}}$ and note, as shown in [HK04], that (3.2.7) implies $(x, 0) \in \mathcal{K}$, $x > 0$. In turn consider:

$$u(x, \epsilon) := \sup_{X \in \mathcal{X}(x,\epsilon)} \mathbb{E}[\log(X(T) - \epsilon F(T))],$$

i.e. the log-utility optimisation problem in (S, F). To ensure well-posedness, assume:

We have $u(x, \epsilon) < \infty$, for some $(x, \epsilon) \in \mathcal{K}$.

In fact the above should always hold for the case of log-utility in this context. To see that note $(x, 0) \in \mathcal{K}$ for any $x > 0$ under (3.2.7), as shown in [HK04]. In particular for any such point we have $\mathcal{X}(x, 0) = \mathcal{X}(x)$. Using the inequality $\log(x) \leq x - 1,\ x > 0$ and the fact that $\mathbb{E}[X(T)] \leq x$ for $X \in \mathcal{X}(x)$ shows the claim. The concavity of $u(x, \epsilon)$ on the open set $\mathcal{K}$ finally gives that $u(x, \epsilon) < \infty$ for all $(x, \epsilon) \in \mathcal{K}$.

In turn, following [HK04], since the log-utility satisfies the reasonable asymptotic elasticity condition, i.e. for $U(x) := \log(x)$ we have $\limsup_{x \to \infty} xU^{'}(x)/U(x) < 1$, S trivially satisfies the condition of NFLVR and we also have (3.2.7); the solution to $u(x, \epsilon)$ exists and is unique. We focus on:

$$u(\epsilon) := u(1, \epsilon) = \mathbb{E}[\log(1 + X^{\epsilon}(T) - \epsilon F(T))], \tag{3.2.8}$$

where $X^{\epsilon}(T)$ denotes the shifted counterpart of the solution of $u(1, \epsilon)$ s.t. it starts from zero.

Remark III.2. Note that working with the optimisation problem in (S, F) doesn't come with any loss of generality in this context, since by using the numeraire invariance of log-utility, along with $(\log(\widetilde{X}_{\nu}(T)))^{-} \in \mathbb{L}_1$ as well as (3.2.4) and (3.2.7) yields that the unique solution to the respective optimisation problem of $u(\epsilon)$ in $(1, \widetilde{S}, \widetilde{\Lambda})$, denoted by $\widetilde{X}^{\epsilon}$, is given as:

$$\widetilde{X}^{\epsilon}(T) = \widetilde{X}_{\nu}(T) \left(1 + X^{\epsilon}(T) + \epsilon \int_0^T \widetilde{\Lambda}(t-) \mathrm{d}S_0(t)\right),$$

Even if the "original" utility maximisation problem in $(1, \widetilde{S}, \widetilde{\Lambda})$ is not well-defined for potentially infinite T, due to it taking infinite value, it is still well-defined for the numeraire-discounted market. Hence it is preferable to work with the latter for long time (and even infinite) horizon settings.

3 Fourth-order expansion of value function & Second-order expansion of the optimal wealth

For $p \geq 1$ denote the class of (local) martingales s.t. for each M we have $\mathbb{E}[\overline{M}(T)^p] < \infty$ by $\mathcal{H}^p$; where $\overline{M} := \sup_{t\in[0,\cdot]} |M(t)|$. Likewise, we denote $\mathcal{H}_0^p := \{M \in \mathcal{H}^p : M(0) = 0\}$. Using the above, define:

$$\mathcal{M}_p := \left\{ M \in \mathcal{H}_0^p : M = \int_0^{\cdot} \theta(t)\mathrm{d}S(t) \right\}.$$

A crucial assumption for our exposition is the following:

$$\begin{aligned} &F(T) := \lim_{t\to T} F(t) \text{ exists a.s., and} \\ &\exists x > 0 \text{ and } M \in \mathcal{M}_4 \text{ s.t. } |F(T)| \leq x + M(T). \end{aligned} \tag{3.3.1}$$

Likewise, as discussed in [KS06], asymptotic results regarding the value function up to second order require the bounding martingale to be in $\mathcal{M}_2$ instead of $\mathcal{M}_4$.

Moving forward, the key tools for the fourth-order expansion of $u(\epsilon)$ are two Kunita-Watanabe (K-W) decompositions. Denote by $\mathcal{F}(T)$ the "stopped" sigma-algebra at T. In turn, as $F(T) \in \mathbb{L}_2(\mathcal{F}(T))$ it admits an orthogonal projection on the space of stochastic integrals w.r.t. S that start at zero. In turn, this gives rise to the following K-W decomposition:

$$\mathbb{E}[F(T)|\mathcal{F}(\cdot)] = \Delta + N,$$

where Δ is a stochastic integral w.r.t. S, while N is strongly orthogonal to it and $N(0) = \mathbb{E}[F(T)]$. Applying similar reasoning to the above, we have:

$$\mathbb{E}[N(T)^2|\mathcal{F}(\cdot)] = \Gamma + P,$$

where Γ is a stochastic integral w.r.t. S, while P is strongly orthogonal to it and $P(0) = \mathbb{E}[N(T)^2]$. In fact, by (3.3.1) we have $\Delta, N \in \mathcal{H}^4$ and $\Gamma, P \in \mathcal{H}^2$.

We also denote by $\mathcal{N}_2$ the orthogonal complement to $\mathcal{M}_2$ in $\mathcal{H}_0^2$. That is:

$$\mathcal{N}_2 := \left\{ N \in \mathcal{H}_0^2 : NM \text{ is a martingale for all } M \in \mathcal{M}_2 \right\}.$$

Furthermore, setting:

$$\mathcal{N}_\infty := \{ N \in \mathcal{N}_2 : N \text{ is bounded} \},$$

and following very similar steps as in [KS06, Lemma 4] we have that, under 3.2.2, and 3.3.1 (a sufficient condition for $\mathcal{N}_2$ to be non-trivial), the closure of $\mathcal{N}_\infty$ in $\mathcal{H}_0^2$ coincides with $\mathcal{N}_2$. This will be a useful result for the following theorem.

Theorem 3.3.1. Assume (3.2.2) and (3.3.1); then we have:

$$u(\epsilon) + \epsilon\mathbb{E}[F(T)] + \frac{\epsilon^2}{2}\mathbb{E}[N(T)^2] + \frac{\epsilon^3}{3}\mathbb{E}[N(T)^3] + \frac{\epsilon^4}{4}\mathbb{E}[N(T)^4] - \frac{\epsilon^4}{2}\mathbb{E}[\Gamma(T)^2] = o(\epsilon^4), \quad (3.3.2)$$

Furthermore, we have that:

$$X^\epsilon(T) = \epsilon\Delta(T) + \epsilon^2\Gamma(T) + o_{\mathbb{P}}(\epsilon^2). \quad (3.3.3)$$

4 Discussion

Before considering a concrete example, let us comment on the usefulness of the results derived in the previous sections. In particular, besides the obvious increase of accuracy in $X^\epsilon(T)$'s approximation by also considering $\Gamma(T)$ (compared to first-order results); there is another arguably more important upside. One that provides some added intuition on the way the investor behaves optimally in an incomplete vs. complete market setting. In order to underline this, we claim that $X^\epsilon(T) = \epsilon\Delta(T)$ in a complete market setting, i.e. the first-order approximation discussed in §3 is actually optimal. Indeed, in this case we have that $\Delta(T) = F(T) - \mathbb{E}[F(T)]$, since $N = \mathbb{E}[F(T)]$. This implies that

$1+\epsilon\Delta(T)-\epsilon F(T) = 1-\epsilon\mathbb{E}[F(T)]$. But then, we get for sufficiently small ϵ s.t. $1-\epsilon\mathbb{E}[F(T)] > 0$ and $\forall X \in \mathcal{X}(1,\epsilon)$:

$$\mathbb{E}\left[\frac{X(T)-1-X^\epsilon(T)}{1+X^\epsilon(T)-\epsilon F(T)}\right] = \mathbb{E}\left[\frac{X(T)-1-\epsilon\Delta(T)}{1-\epsilon\mathbb{E}[F(T)]}\right] = \frac{\mathbb{E}[X(T)]-1}{1-\epsilon\mathbb{E}[F(T)]} \leq 0,$$

which shows that $X^\epsilon(T) = \epsilon\Delta(T)$ in view of the first-order conditions. Hence, the term $\epsilon^2\Gamma(T)$ in the approximation of $X^\epsilon(T)$ allows us to get a sense of how market incompleteness directly affects optimal behaviour in a context where a non-replicable endowment is present.

Secondly, the fact that the model accommodates infinite-horizon markets is a non-trivial addition. In order to better grasp the added benefit of such a result, note that while we have a characterisation of Δ, Γ; their dependence on the maturity makes explicit results not possible in many cases. On the other hand the "myopic case" of $T=\infty$ provides a more tractable tool to approximate optimal behaviour. Informally, the message being that in the case of distant horizons asset allocation can be further approximated by the projections of $\mathbb{E}[F(\infty)|\mathcal{F}(\cdot)]$ and $\mathbb{E}[N(\infty)^2|\mathcal{F}(\cdot)]$ on the space of S-integrals.

4.1 Examples

Concretely, let us consider a space that supports a two dimensional standard Brownian motion (W^1, W^2), where the (single) risky asset $\widetilde{S}$ is driven by $\int_0^\cdot \widetilde{\sigma}(t)dW^1(t)$. Furthermore, consider the following factor process Z:

$$Z = Z(0) + \int_0^\cdot \mu(Z(t))\mathrm{d}t + \int_0^\cdot \kappa(Z(t))\mathrm{d}B(t),$$

and non-replicable stream:

$$\widetilde{\Lambda} = \int_0^\cdot \exp\left(-\int_0^t \widetilde{r}(s)\mathrm{d}s\right)\widetilde{\lambda}(t)\mathrm{d}t,$$

where $\int_0^\cdot \widetilde{r}(t)\mathrm{d}t$ is the finite-variation process driving the baseline asset, $B := \rho W^1 + \sqrt{1-\rho^2}W^2$, $\rho \in (-1,1)$ and we define $\beta := \kappa^2$. With a slight abuse of notation, we assume that all $\widetilde{r}$, $\widetilde{a}$, $\widetilde{\sigma}$, $\widetilde{\lambda}$ are functions of the factor ($\widetilde{c} = \widetilde{\sigma}^2$). Explicit calculations regarding Δ

require us to have a sense of the following form:

$$\mathbb{E}[F(\infty)|\mathcal{F}(t)] = \int_0^t S_0(s)\exp\left(-\int_0^s \widetilde{r}(u)\mathrm{d}u\right)\widetilde{\lambda}(s)\mathrm{d}s + S_0(t)\exp\left(-\int_0^t \widetilde{r}(s)\mathrm{d}s\right)\psi(Z(t)),$$

where ψ satisfies the following second-order ODE:

$$\widetilde{\lambda} - \widetilde{r}\psi + (\mu - \nu\widetilde{\sigma}\kappa\rho)\partial_z\psi + \frac{1}{2}\beta\partial_{zz}\psi = 0, \tag{3.4.1}$$

where $\nu = \widetilde{a}/\widetilde{c}$. On the contrary, similar calculations for finite horizons would require us to solve a PDE, since a time variable would also be present.

Example III.3. [Linear payoff on a factor model and an infinite horizon] In a market with one riskless and one risky asset as well as $T = \infty$, we consider a standard two dimensional Brownian motion (W^1, W^2) generating the augmented filtration $\mathcal{F}(\cdot)$, and formulate the following problem for $B := \rho W^1 + \sqrt{1-\rho^2}W^2,\ \rho \in (-1, 1)$:

$$\begin{aligned}
&\widetilde{\Lambda} = \int_0^{\cdot} e^{-\widetilde{r}t} Z(t)\mathrm{d}t, \qquad \widetilde{r} > 0,\\
&\mathrm{d}Z(t) = k(\theta - Z(t))\mathrm{d}t + b\mathrm{d}B(t), \qquad k, b > 0;\ Z(0), \theta \in \mathbb{R},\\
&\frac{\mathrm{d}\widetilde{S}(t)}{\widetilde{S}(t)} = \widetilde{a}\mathrm{d}t + \widetilde{\sigma}\mathrm{d}W^1(t), \qquad \widetilde{a} \in \mathbb{R};\ \widetilde{\sigma} > 0.
\end{aligned}$$

In turn, in the above model, the portfolio that gives rise to the local martingale numeraire is constant and in particular we have:

$$\begin{aligned}
\frac{\mathrm{d}S_0(t)}{S_0(t)} &= \overbrace{-\nu\widetilde{\sigma}\mathrm{d}W^1(t)}^{\mathrm{d}R_0(t)}, \qquad S_0(0) = 1,\\
F &= \int_0^{\cdot} e^{-\widetilde{r}t} S_0(t) Z(t)\mathrm{d}t.
\end{aligned}$$

Note that even if a non-replicable (local) stream given as a linear contract on the factor Z isn't sub/super replicable, the continuity (and adaptedness) of the factor process give that Z is locally bounded. Hence the above are still relevant. In fact, define $\tau_n := \inf\{t : |Z(t)| \geq n\}$ (for big enough n s.t. it dominates $Z(0)$). Then $F_n := \int_0^{\cdot} e^{-\widetilde{r}t} S_0(t) Z^{\tau_n}(t)dt$ satisfies the sub/super replicability conditions for all n, since Z^{τ_n} is (uniformly) bounded. In fact there is a concrete connection between (Δ_n, Δ) and (Γ_n, Γ), explored in §B.2.

We move forward with explicit calculations of Δ, Γ. Beginning with the former denote the measure induced by S_0 as $\mathbb{Q}^\nu$ (which is locally equivalent to $\mathbb{P}$) and note that, by denoting the martingale closed by $F(\infty)$ as $\breve{M}$, we have:

$$\breve{M}(t) = \int_0^t e^{-\widetilde{r}s} S_0(s) Z(s) \mathrm{d}s + e^{-\widetilde{r}t} S_0(t) \int_0^\infty e^{-\widetilde{r}s} \mathbb{E}^{\mathbb{Q}^\nu}_{Z(t)}[Z(s)] \mathrm{d}s.$$

Now, under $\mathbb{Q}^\nu$ we have:

$$dZ(t) = (\eta - kZ(t))\mathrm{d}t + b\mathrm{d}B^{\mathbb{Q}^\nu}(t), \qquad \eta := k\theta - b\widetilde{\sigma}\nu\rho.$$

Hence, we get:

$$\breve{M}(t) = \int_0^t e^{-\widetilde{r}s} S_0(s) Z(s) \mathrm{d}s + e^{-\widetilde{r}t} S_0(t) \left(\frac{Z(t) + \eta/\widetilde{r}}{\widetilde{r} + k} \right).$$

Applying the Kunita-Watanabe decomposition on $\breve{M}$ w.r.t. W^1 (which drives both S_0 and S_1), we get:

$$\Delta = \int_0^\cdot \theta^\Delta(t) \mathrm{d}W^1(t), \qquad \theta^\Delta(t) := \frac{(b\rho - \nu\widetilde{\sigma}\eta/\widetilde{r})}{\widetilde{r} + k} e^{-\widetilde{r}t} S_0(t) - \frac{\nu\widetilde{\sigma}}{\widetilde{r} + k} e^{-\widetilde{r}t} S_0(t) Z(t).$$

Continuing with Γ, denote the martingale closed by $[N, N](\infty)$ as $\breve{K}$, then:

$$\begin{aligned}
\breve{K}(t) &= \mathbb{E}\left[\int_0^\infty \frac{b^2(1-\rho^2)}{(\widetilde{r}+k)^2} e^{-2\widetilde{r}s} S_0(s)^2 \mathrm{d}s \middle| \mathcal{F}(t) \right] = \int_0^t \frac{b^2(1-\rho^2)}{(\widetilde{r}+k)^2} e^{-2\widetilde{r}s} S_0(s)^2 \mathrm{d}s \\
&\quad + \mathbb{E}\Bigg[\int_t^\infty \frac{b^2(1-\rho^2)}{(\widetilde{r}+k)^2} e^{-(2\widetilde{r}-(\nu\widetilde{\sigma})^2)s} \mathcal{E}(2R_0)(s) \mathrm{d}s \Bigg| \mathcal{F}(t) \Bigg] \\
&= \int_0^t \frac{b^2(1-\rho^2)}{(\widetilde{r}+k)^2} e^{-2\widetilde{r}s} S_0(s)^2 \mathrm{d}s + \frac{b^2(1-\rho^2)}{(\widetilde{r}+k)^2(2\widetilde{r} - (\nu\widetilde{\sigma})^2)} e^{-(2\widetilde{r}-(\nu\widetilde{\sigma})^2)t} \mathcal{E}(2R_0)(t),
\end{aligned}$$

for sufficiently big $\widetilde{r}$. Hence:

$$\Gamma = \int_0^\cdot \theta^\Gamma(t) \mathrm{d}W^1(t), \qquad \theta^\Gamma(t) := -\frac{2b^2(1-\rho^2)\nu\widetilde{\sigma}}{(\widetilde{r}+k)^2(2\widetilde{r} - (\nu\widetilde{\sigma})^2)} e^{-(2\widetilde{r}-(\nu\widetilde{\sigma})^2)t} \mathcal{E}(2R_0)(t),$$

Now, it is direct to note that if ρ tends to either 1 or -1, Γ vanishes.

Example III.4. [Polynomial payoff on a factor model and an infinite horizon, for unit-linked liabilities] Borrowing the setup from Example III.3 we can generalise our results for any polynomial payoff, i.e. for $m \in \mathbb{N}_{>0}$. In particular, for added notational clarity we focus on the case where the liability stream is unit-linked to the local martingale numeraire $\widetilde{X}_\nu$, i.e.:

$$F(\cdot; m) = \int_0^{\cdot} e^{-\widetilde{r}t}(t) Z(t)^m \mathrm{d}t,$$

while the more cumbersome case of non-unit-linked liabilities is left for the appendix, in B.3. In turn, following similar steps as for (3.4.1), we have:

$$\mathbb{E}[F(\infty; m)|\mathcal{F}(t)] = \int_0^t e^{-\widetilde{r}s} Z(s)^m \mathrm{d}s + e^{-\widetilde{r}t}\psi_m(Z(t)),$$

where ψ_m solves:

$$(\widetilde{r} - \mathcal{L})\psi_m(z) = z^m, \qquad \mathcal{L}f(z) := k(\theta - z)\partial_z f(z) + \frac{b^2}{2}\partial_{zz} f(z).$$

In fact, by Feynman-Kac (under the implied transversality condition) we have:

$$\psi_m(z) = \mathbb{E}_z\left[\int_0^\infty e^{-\widetilde{r}t} Z(t)^m \mathrm{d}t\right].$$

Furthermore, note that:

$$(\widetilde{r} - \mathcal{L})z^m = (\widetilde{r} + mk)z^m - mk\theta z^{m-1} - \frac{b^2}{2}m(m-1)z^{m-2}.$$

Using $(\widetilde{r} - \mathcal{L})\psi_{m-1}(z) = z^{m-1}$, $(\widetilde{r} - \mathcal{L})\psi_{m-2}(z) = z^{m-2}$ and rearranging the above, we get:

$$(\widetilde{r} - \mathcal{L})\left(\frac{z^m}{\widetilde{r} + mk} + \frac{mk\theta}{\widetilde{r} + mk}\psi_{m-1}(z) + \frac{b^2 m(m-1)}{2(\widetilde{r} + mk)}\psi_{m-2}(z)\right) = z^m,$$

which implies that ψ_m satisfies the following recursive relationship:

$$\begin{aligned}
&\psi_0(z) = \frac{1}{\widetilde{r}}; \qquad \psi_1(z) = \frac{z + k\theta/\widetilde{r}}{\widetilde{r} + k},\\
&\psi_m(z) = \frac{z^m}{\widetilde{r} + mk} + \frac{mk\theta}{\widetilde{r} + mk}\psi_{m-1}(z) + \frac{b^2 m(m-1)}{2(\widetilde{r} + mk)}\psi_{m-2}(z), \qquad m \geq 2.
\end{aligned}$$

In particular, denoting the martingale closed by $F(\infty; m)$ as $\breve{M}_m$, we have:

$$\mathrm{d}\breve{M}_m(t) = e^{-\widetilde{r}t} b\rho\partial_z\psi_m(Z(t))\mathrm{d}W^1(t) + e^{-\widetilde{r}t} b\sqrt{1-\rho^2}\partial_z\psi_m(Z(t))\mathrm{d}W^2(t).$$

So, by the Kunita-Watanabe decomposition, we get:

$$\Delta_m = \int_0^{\cdot} e^{-\widetilde{r}t} b\rho \partial_z \psi_m(Z(t)) \mathrm{d}W^1(t),$$
$$\mathrm{d}N_m(t) = e^{-\widetilde{r}t} b\sqrt{1-\rho^2} \partial_z \psi_m(Z(t)) \mathrm{d}W^2(t),$$

where Δ_m, N_m are the respective forms of Δ, N from Example III.3. Note that—for the case of unit-linked liabilities—when $\rho = 0$, the hedging part Δ_m vanishes since the factor is solely driven by the orthogonal part to the market.

Now, denote the martingale closed by $[N_m, N_m](\infty)$ as $\breve{K}_m$. Then:

$$\breve{K}_m(t) = \int_0^t b^2(1-\rho^2) e^{-2\widetilde{r}s} \partial_z \psi_m(Z(s))^2 \mathrm{d}s + e^{-2\widetilde{r}t} \chi_m(Z(t)),$$

where χ_m solves:

$$(2\widetilde{r} - \mathcal{L})\chi_m(z) = b^2(1-\rho^2) \partial_z \psi_m(z)^2.$$

In fact, by Feynman-Kac (under the implied transversality condition) we have:

$$\chi_m(z) = \mathbb{E}_z \left[\int_0^\infty e^{-2\widetilde{r}t} b^2(1-\rho^2) \partial_z \psi_m(Z(t))^2 \mathrm{d}t \right].$$

We need to pin down $\partial_z \psi_m^2$. Differentiating the recursion for ψ_m, we get:

$$\partial_z \psi_m(z) = \frac{m}{\widetilde{r} + mk} z^{m-1} + \frac{mk\theta}{\widetilde{r} + mk} \partial_z \psi_{m-1}(z) + \frac{b^2 m(m-1)}{2(\widetilde{r} + mk)} \partial_z \psi_{m-2}(z), \qquad m \geq 2,$$

where $\partial_z \psi_0(z) = 0$ and $\partial_z \psi_1(z) = 1/(\widetilde{r} + k)$. In particular, $\partial_z \psi_m$ is of the form: $\partial_z \psi_m(z) = \sum_{j=0}^{m-1} d_{m,j} z^j$. Hence:

$$\sum_{j=0}^{m-1} d_{m,j} z^j = \frac{m}{\widetilde{r} + mk} z^{m-1} + \frac{mk\theta}{\widetilde{r} + mk} \sum_{j=0}^{m-2} d_{m-1,j} z^j + \frac{b^2 m(m-1)}{2(\widetilde{r} + mk)} \sum_{j=0}^{m-3} d_{m-2,j} z^j.$$

By the above, and after setting $d_{m,j} := 0$, whenever $j > m-1$ implies the following recursion for $m \geq 2$:

$$d_{m,j} = \begin{cases} \frac{m}{\widetilde{r}+mk}, & j = m-1, \\ \frac{mk\theta}{\widetilde{r}+mk} d_{m-1,j}, & j = m-2, \\ \frac{mk\theta}{\widetilde{r}+mk} d_{m-1,j} + \frac{b^2 m(m-1)}{2(\widetilde{r}+mk)} d_{m-2,j}, & j = 0, \ldots, m-3, \end{cases}$$

and $d_{1,0} = 1/(\widetilde{r}+k)$. In turn, $\partial_z \psi_m^2$ is of the form: $\partial_z \psi_m(z)^2 = \sum_{j=0}^{2m-2} q_{m,j} z^j$, where $q_{m,j} = \sum_{k=\max(0,j-(m-1))}^{\min(j,m-1)} d_{m,k} d_{m,j-k}$. Lastly, in order to obtain χ_m, consider:

$$(2\widetilde{r} - \mathcal{L})\phi_m(z) = z^m.$$

Following a similar process as ψ_m, we get that:

$$\begin{aligned}
\phi_m(z) &= \mathbb{E}_z\left[\int_0^\infty e^{-2\widetilde{r}t} Z(t)^m \mathrm{d}t\right], \\
\phi_0(z) &= \frac{1}{2\widetilde{r}}; \qquad \phi_1(z) = \frac{z + k\theta/2\widetilde{r}}{2\widetilde{r}+k}, \\
\phi_m(z) &= \frac{z^m}{2\widetilde{r}+mk} + \frac{mk\theta}{2\widetilde{r}+mk}\phi_{m-1}(z) + \frac{b^2 m(m-1)}{2(2\widetilde{r}+mk)}\phi_{m-2}(z).
\end{aligned}$$

Lastly, by linearity we get:

$$\chi_m(z) = b^2(1-\rho^2)\sum_{j=0}^{2m-2} q_{m,j}\phi_j(z).$$

One way to read the formula for $\chi_m(z)$, in comparison to the first-order object $\psi_m(z)$, is that the latter solely captures lower moment-type effects generated by the initial monomial input, while the former also propagates forward—incorporating higher-order effects. The functions $\phi_0(z), \phi_1(z), ...$ are precisely the natural building blocks for this representation. In fact, they are essentially the same family as $\{\psi_0(z), \psi_1(z), ...\}$ only with a faster discounting: $\widetilde{r} \mapsto 2\widetilde{r}$. This is one of the advantages of going to second order w.r.t. ϵ: the correction term not only improves accuracy, but is also more informative about the structure of incompleteness.

The above paragraph also underlines that the tractability of the Ornstein–Uhlenbeck framework is not restricted to linear claims. Indeed, as long as the factor remains affine under the relevant measure changes, the associated hedging strategies can be computed recursively for monomial liability streams, and hence for arbitrary polynomials by linearity. This, in turn, provides a natural gateway to more general payoff functions, for example through analytic expansion.

IV Log-pricing with small liability stream

1 Introduction

Discussion

A natural question once an investor faces a non-replicable liability stream is not only how this stream affects optimal investment, but also how it should be valued in monetary terms. In a complete market, this issue is straightforward: every claim can be replicated and therefore has a unique arbitrage-free price. In an incomplete market, however, perfect hedging is typically impossible, so the same liability may admit many arbitrage-free valuations. This makes utility-based pricing a natural framework, since it ties valuation to the investor's preferences and to the actual trading opportunities available in the market.

In the present setting, the object of interest is the so-called utility-based certainty equivalent of a small liability stream, for the case of log-utility. This quantity has a direct economic interpretation: it is the cash amount that offsets the investor's utility loss from bearing the liability stream. Such a formulation is particularly useful when one wants to assign a price to non-traded streams, as in pension-style applications, where the relevant risks cannot be fully offset through dynamic trading. At the same time, explicit formulas are generally unavailable in incomplete markets, even under relatively simple model specifications. This makes asymptotic analysis a natural tool, especially when the liability is small relative to the investor's wealth.

In this context, the chapter focuses on a higher-order approximation of this certainty equivalent. This is not only a matter of improving numerical accuracy. Lower-order approximations capture the leading effect of the unhedgeable liability, but they do not fully reflect how residual risk propagates into valuation once market incompleteness becomes important. By pushing the expansion further, one obtains a more informative description of how the investor prices non-replicable cash flows and how this price departs from the complete-market benchmark. This is particularly relevant over long horizons, where even small nonlinear effects may accumulate and become economically significant.

Contributions

This chapter follows the same general semimartingale framework outlined in Chapter III, but now from a valuation perspective. In this direction, the contributions of the chapter are the following:

1. The main result of the chapter is a fourth-order asymptotic expansion of the log-based certainty equivalent with respect to the size of the liability stream. To the best of our knowledge, this constitutes a novel result in such a general setting. In turn, this result makes explicit how market incompleteness enters the investor's valuation beyond the leading-order term. In particular, lower-order approximations capture only the first nonlinear correction, whereas the third- and fourth-order terms reveal additional effects generated by the residual unhedgeable part of the liability stream. In fact, it is shown that such effects can be pronounced even for moderate positions in the non-traded stream. In this way, the chapter shows that going beyond second order is not merely a matter of numerical refinement, but also of obtaining a more informative picture of how incomplete hedging opportunities affect utility-based prices.

2. A further contribution of the chapter is that the analysis also covers the infinite horizon case. This is useful not only as a theoretical extension, but also as an additional

approximation tool for investors with very distant maturities. In such cases, the infinite-horizon formulation can provide a simpler and more tractable benchmark for utility-based valuation, while still capturing the main economic effects of a long-dated non-replicable liability stream. In fact, as the maturity becomes larger, the discrepancy between the finite- and infinite-horizon approximations becomes smaller, so the latter provides an increasingly accurate proxy for long-horizon liability valuation. This improvement is particularly pronounced when the non-replicable stream is more closely aligned with the traded source of risk.

Related literature

As previously discussed, the literature on utility-based valuation in incomplete markets is too vast to survey exhaustively. We therefore focus only on the strand most closely related to the present work, namely utility-based pricing of non-replicable streams. Since hedging and pricing are tightly connected in this area, many of the relevant papers study both problems simultaneously. Here, however, we emphasise their pricing content.

A major source of tractability in utility-based valuation comes from the use of exponential utility. Prominent works in this context include [Hen02], [MZ04], [GH07], and [LL12]. These studies leverage a linearisation technique—commonly referred to as the Cole-Hopf transformation or distortion power—first introduced in claim valuation by [Zar01], which reduces the resulting non-linear HJB PDE to a linear form, solvable via standard methods. Further generalisations by [FS08] and [FS10] showed that, even in models with general asset dynamics, the exponential utility-based price admits an explicit expression, although these formulas are often less interpretable.

Even within the relatively tractable exponential utility framework, explicit expressions are not always obtainable. For example, in models where the claim depends on the traded asset, [SZ05] derive asymptotic expansions for the utility-indifference price in the context of fast mean-reverting volatility. A multidimensional non-replicable asset model subject to intertemporal default risk is considered in [HL16], where a semigroup approximation using splitting techniques is developed.

Since explicit formulas are rarely available in incomplete markets, a number of asymptotic methods have been developed to approximate utility-based prices. For instance, [Hen02] and [HH02] study the case of power utility and derive second-order expansions of the investor's value function with respect to a small position in the contingent claim, which in turn yield approximations for the reservation price.

These early results were significantly extended in [KS06], where the authors study a general semimartingale framework—under a broad class of utility functions defined on the positive real line. Particularly, they derive the first-order approximation for marginal (utility-based) indifference prices and study their qualitative features. A related analysis in [Kal02] studies first-order marginal price approximations under local utility maximisation.

Along similar lines, [KR11] studies utility-based pricing under exponential utility in the regime of vanishing risk aversion. In a different direction, [KMKV14] considers exponential Lévy models and, for power utility, provides alternative formulations of the results in [KS06] that do not require a change of numeraire. Within the setting of exponential Lévy processes, [MT16] derives a non-asymptotic approximation for the exponential utility-based indifference price. The approach therein extends the earlier small risk-aversion asymptotics and yields a closed-form approximation by treating the Lévy model as a perturbation of the classical Black-Scholes framework.

Structure of the chapter

This chapter is set in the same market and utility-maximisation framework as Chapter III. We therefore retain the notation, assumptions, and objects introduced there, and only recall the ingredients specific to this part of the text. In particular, §3 is dedicated to exploring the notion of utility-based pricing, in the context of deriving the fourth-order expansion of the log-based certainty equivalent. Therein, the finite- and infinite-horizon cases are treated simultaneously. Note that for the sake of readability, the proof of the main result is deferred to Appendix C.2. Lastly, Appendix C.1 is reserved for introducing several key notions related to arbitrage and market completeness. These concepts, in turn, provide the foundation for the utility-based pricing framework developed in this chapter.

2 Setup

2.1 The market

Consider an agent that trades in $1+d \in \mathbb{N}_{>0}$ assets; a baseline asset with stochastic exponential price, driven by a continuous finite-variation process, and d risky assets. Discounting by the baseline, the market is driven by a continuous, positive, $\mathbb{R}^d$-valued semimartingale $\widetilde{S} \equiv (\widetilde{S}_i; 1 \leq i \leq d)$. In particular, we have:

$$\frac{\mathrm{d}\widetilde{S}_i(t)}{\widetilde{S}_i(t)} = \mathrm{d}\widetilde{A}_i(t) + \mathrm{d}\widetilde{M}_i(t), \qquad 1 \leq i \leq d.$$

Here, each real-valued component $\widetilde{A}_i$ of the vector-valued $\widetilde{A} \equiv (\widetilde{A}_i; 1 \leq i \leq d)$ is a continuous finite-variation process with $\widetilde{A}_i(0) = 0$, whereas each $\widetilde{M}_i$ in $\widetilde{M} \equiv (\widetilde{M}_i; 1 \leq i \leq d)$ is a continuous local martingale with $\widetilde{M}_i(0) = 0$.

2.2 Investment opportunities

In the market $(1, \widetilde{S})$, an initial capital $x \in \mathbb{R}$ and a choice of a self-financing strategy $\theta \equiv (\theta_i; 1 \leq i \leq d)$ (assumed to be predictable and $\widetilde{S}$-vector integrable) result in a *wealth process*:

$$\widetilde{X} := x + \int_0^{\cdot} \theta(t)\mathrm{d}\widetilde{S}(t) \equiv \widetilde{X}(\cdot; x, \theta), \tag{4.2.1}$$

where the above is understood as a vector stochastic integral. To avoid the so-called doubling strategies, we restrict the above class as follows: a wealth process $\widetilde{X}$ with initial capital $\widetilde{X}(0) = x$ is called *admissible* if:

$$\widetilde{X} > 0.$$

We denote this subset of wealth processes by $\widetilde{\mathcal{X}}(x)$. The union $\cup_{x>0}\widetilde{\mathcal{X}}(x)$ is denoted by $\widetilde{\mathcal{X}}$.

The market's viability is intimately connected to the following condition, as shown in [KK21, Theorem 2.31]:

$$\widetilde{c} \text{ is non-singular, } (\mathbb{P} \otimes \widetilde{O})\text{-a.e. and } \int_0^K \widetilde{a}(t)'\widetilde{c}(t)^{-1}\widetilde{a}(t)\mathrm{d}\widetilde{O}(t) < \infty, \forall K > 0, \tag{4.2.2}$$

where $\widetilde{a}$, $\widetilde{c}$ are the local drift and covariation rates w.r.t. the clock $\widetilde{O}$ characterised in Chapter III.

In particular, the above condition implies the existence of the local martingale numeraire; denote it by $\widetilde{X}_\nu := \mathcal{E}(\widetilde{R}_\nu)$, where $\widetilde{R}_\nu := \int_0^\cdot \nu(t)'\mathrm{d}\widetilde{S}(t)/\widetilde{S}(t)$ and $\nu := \widetilde{c}^{-1}\widetilde{a}$ (refer to §2 in [KK21] for further details on this concept). Note that $\widetilde{X}_\nu$ can be used as a new numeraire, under which each $\widetilde{X} \in \widetilde{\mathcal{X}}$ becomes a local martingale. In fact by considering the auxiliary market, under the local martingale numeraire:

$$S := \left(S_0 := \frac{1}{\widetilde{X}_\nu}, S := \frac{\widetilde{S}}{\widetilde{X}_\nu}\right),$$

we have for each wealth process $\widetilde{X}$ in $(1, \widetilde{S})$:

$$\widetilde{X}/\widetilde{X}_\nu = x + \int_0^\cdot \sum_{i=1}^d \theta_i(t)\mathrm{d}S_i(t) + \int_0^\cdot \left(\widetilde{X}(t)/\widetilde{X}_\nu(t) - \sum_{i=1}^d \theta_i(t)S_i(t)\right)\frac{\mathrm{d}S_0(t)}{S_0(t)}, \tag{4.2.3}$$

where S is a $\mathbb{R}^{1+d}$-valued local martingale.

2.3 Illiquid stream

Besides trading in the financial market, the agent holds $\epsilon \in \mathbb{R}$ units of an exogenous, non-replicable (at stopping time T) cumulative stream $\widetilde{\Lambda}$ that is a right-continuous, adapted finite-variation process, s.t. $\mathrm{d}\tilde{\Lambda}(t)$ is already discounted by the baseline asset at each t. The only regularity assumption on $\widetilde{\Lambda}$ is that it can be super and subreplicated using the traded asset $\widetilde{S}$.

Remark IV.1. Note that the non-replicable stream $\widetilde{\Lambda}$, expressed under the numeraire $\widetilde{X}_\nu$, is given by $\Lambda := \widetilde{\Lambda}/\widetilde{X}_\nu$. Now, integration by parts yields:

$$\Lambda = \int_0^\cdot \widetilde{\Lambda}(t-)\mathrm{d}S_0(t) + F, \tag{4.2.4}$$

$F := \int_0^\cdot S_0(t)\mathrm{d}\widetilde{\Lambda}(t)$. In turn, we have for a wealth process $\widetilde{X}$, with initial capital x, that:

$$\widetilde{X}/\widetilde{X}_\nu - \int_0^\cdot \widetilde{\Lambda}(t-)\mathrm{d}S_0(t),$$

is simply (4.2.3) with a shifted position in S_0. In other words, when considering the illiquid stream discounted by the local martingale numeraire, the non-replicable part of Λ comes solely from F.

2.4 Utility maximisation and pricing

Fix a possibly infinite-valued stopping time T, that will serve as the maturity of the agent. We focus on the model (S, F) in the sequel. Recalling (4.2.3) (under the self-financing condition), a wealth process in S: $X \equiv X(\cdot; x, \theta)$ is given via:

$$X = x + \int_0^\cdot \sum_{i=1}^d \theta_i(t)\mathrm{d}S_i(t) + \int_0^\cdot \left(X(t) - \sum_{i=1}^d \theta_i(t)S_i(t)\right)\frac{\mathrm{d}S_0(t)}{S_0(t)}. \tag{4.2.5}$$

Hence, all processes of the form shown in (4.2.3) are in fact wealth processes in S.

Now, assuming:

$$\begin{aligned} &F(T) := \lim_{t\to T} F(t) \text{ exists a.s., and} \\ &\{X \in \mathcal{X} : |F(T)| \le X(T)\} \neq \emptyset, \end{aligned} \tag{4.2.6}$$

we define the following class:

$$\mathcal{X}(x, \epsilon) := \{X \text{ a wealth process in } S : X \text{ is acceptable, } X(0) = x \text{ and } X(T) - \epsilon F(T) > 0\}.$$

In turn, focusing on the case where $x = 1$ for notational simplicity, consider:

$$u(\epsilon) := \underbrace{\sup_{X\in\mathcal{X}(1,\epsilon)} \mathbb{E}[\log(X(T) - \epsilon F(T))]}_{u(1,\epsilon)} = \mathbb{E}[\log(1 + X^\epsilon(T) - \epsilon F(T))], \tag{4.2.7}$$

where $X^\epsilon(T)$ denotes the shifted counterpart of the solution of $u(1,\epsilon)$ s.t. it starts from zero. In particular, this is a well-posed problem that admits a unique solution, as discussed in Chapter III.

3 Fourth-order expansion of log-based certainty equivalent

For $p \geq 1$ denote the class of (local) martingales s.t. for each M we have $\mathbb{E}[\overline{M}(T)^p] < \infty$ by $\mathcal{H}^p$; where $\overline{M} := \sup_{t\in[0,\cdot]} |M(t)|$. Likewise, we denote $\mathcal{H}_0^p := \{M \in \mathcal{H}^p : M(0) = 0\}$. Using the above, define:

$$\mathcal{M}_p := \left\{ M \in \mathcal{H}_0^p : M = \int_0^\cdot \theta(t)\mathrm{d}S(t) \right\}.$$

As it was the case for Chapter III, a crucial assumption for our exposition is the following:

$$\begin{aligned} &F(T) := \lim_{t\to T} F(t) \text{ exists a.s., and} \\ &\exists x > 0 \text{ and } M \in \mathcal{M}_4 \text{ s.t. } |F(T)| \leq x + M(T). \end{aligned} \tag{4.3.1}$$

Likewise, as discussed in [KS06], asymptotic results regarding the value function up to second order require the bounding martingale to be in $\mathcal{M}_2$ instead of $\mathcal{M}_4$.

Moving forward, the key tools for our expansion are two Kunita-Watanabe (K-W) decompositions (particularly the orthogonal parts). Denote by $\mathcal{F}(T)$ the "stopped" sigma-algebra at T. In turn, as $F(T) \in \mathbb{L}_2(\mathcal{F}(T))$ it admits an orthogonal projection on the space of stochastic integrals w.r.t. S that start at zero. In turn, this gives rise to the following K-W decomposition:

$$\mathbb{E}[F(T)|\mathcal{F}(\cdot)] = \Delta + N,$$

where Δ is a stochastic integral w.r.t. S, while N is strongly orthogonal to it and $N(0) = \mathbb{E}[F(T)]$. Applying similar reasoning to the above, we have:

$$\mathbb{E}[N(T)^2|\mathcal{F}(\cdot)] = \Gamma + P,$$

where Γ is a stochastic integral w.r.t. S, while P is strongly orthogonal to it and $P(0) = \mathbb{E}[N(T)^2]$. In fact, by (4.3.1) we have $\Delta, N \in \mathcal{H}^4$ and $\Gamma, P \in \mathcal{H}^2$.

The utility-based approach can also be used for the sake of pricing non-replicable streams. To fix ideas we focus, for now, on the case of a finite horizon (i.e. a finite stopping time). We define:

$$v(x) := \sup_{\widetilde{X} \in \widetilde{\mathcal{X}}(x)} \mathbb{E}[\log(\widetilde{X}(T))],$$

and assume:

$$v(x) < \infty, \text{ for some } x > 0, \tag{4.3.2}$$

which in turn implies that $v(x) < \infty$, for all $x > 0$ by the concavity of the logarithm. In particular, the finiteness of $v(x)$ yields $\log(\widetilde{X}(T))^+ \in \mathcal{L}_1$, for all $\widetilde{X} \in \widetilde{\mathcal{X}}(x)$. To see this, note that for $\delta \in (0,1)$, $\widetilde{\mathcal{X}}(x) \ni \widetilde{X}_\delta := \delta + (1-\delta)\widetilde{X}$ is bounded away from zero. Hence $\mathbb{E}[\log(\widetilde{X}(T))]$ is well-defined with values in $\mathbb{R} \cup \{-\infty\}$, since $\mathbb{E}[\log(\widetilde{X}_\delta(T))^+] \geq (1-\delta)\mathbb{E}[\log(\widetilde{X}(T))^+]$.

In fact, $\log(\widetilde{X}_\nu(T))^- \in \mathcal{L}_1$ along with $\log(\widetilde{X}(T))^+ \in \mathcal{L}_1$, and [KK21, Proposition 2.46] imply that the wealth process in $(1, \widetilde{S})$ with initial value $x > 0$, generated by $\theta_i^\nu := x\widetilde{X}_\nu \nu_i / \widetilde{S}_i$ for $1 \leq i \leq d$, is the solution to $v(x)$. Particularly, we get:

$$v(x) = \mathbb{E}\left[\log\left(x + \int_0^T \theta^\nu(t)' d\widetilde{S}(t)\right)\right] = \mathbb{E}\left[\log\left(x\widetilde{X}_\nu(T)\right)\right].$$

By the above, and after recalling the discussion in Remark III.2, we derive the log-based certainty equivalent $c(\epsilon)$ of the position $(1, \epsilon)$ as the solution to the following equation:

$$\mathbb{E}\left[\log\left(1 + c(\epsilon) + \int_0^T \theta^\nu(t)' d\widetilde{S}(t)\right)\right] = \mathbb{E}\left[\log\left(\widetilde{X}^\epsilon(T) - \epsilon\Lambda(T)\right)\right]. \tag{4.3.3}$$

Rearranging and using the fact that the logarithm turns multiplication into addition, we have:

$$c(\epsilon) = e^{u(\epsilon)} - 1, \tag{4.3.4}$$

where $u(\epsilon)$ is given in (4.2.7). In fact, we may take (4.3.4) as an alternative characterisation of the log-based certainty equivalent, and we shall do so henceforth. The advantage is that

(4.3.4) remains valid on an infinite-horizon setting, does not require (4.3.2), and reduces to (4.3.3) whenever the more restrictive assumptions underlying it are in force.

Putting all these together, we get the following result.

Theorem 4.3.1. Assume (4.2.2) and (4.3.1); then we have:

$$\begin{aligned} c(\epsilon) + A_1\epsilon + \frac{A_2 - A_1^2}{2}\epsilon^2 + \left(\frac{A_3}{3} - \frac{A_1 A_2}{2} + \frac{A_1^3}{6}\right)\epsilon^3 \\ + \left(\frac{A_4}{4} + G - \frac{A_1 A_3}{3} - \frac{A_2^2}{8} + \frac{A_1^2 A_2}{4} - \frac{A_1^4}{24}\right)\epsilon^4 = o(\epsilon^4), \end{aligned} \tag{4.3.5}$$

where:

$$\begin{aligned} &A_1 := \mathbb{E}[F(T)], \quad A_2 := \mathbb{E}[N(T)^2], \quad A_3 := \mathbb{E}[N(T)^3], \quad A_4 := \mathbb{E}[N(T)^4], \\ &G := \mathbb{E}[\Gamma(T)^2/2 - N(T)^2\Gamma(T)]. \end{aligned}$$

Remark IV.2. Note that in a complete market setting we have:

$$\mathbb{E}[F(T)|\mathcal{F}(\cdot)] = \mathbb{E}[F(T)] + \Delta,$$

i.e. the quadratic variation of N as well as Γ vanish. In turn, in this context we get: $A_2 = A_1^2$, $A_3 = A_1^3$, $A_4 = A_1^4$, $G = 0$. Applying the above to (4.3.5) we have that its second-, third-, and fourth-order terms vanish, and the log-based certainty equivalent reduces to $-\epsilon\mathbb{E}[F(T)]$. Recalling the discussion in Appendix C.1, this constitutes the unique arbitrage-free price of the stream in the context of a complete market.

4 Discussion & Applications

Theorem 4.3.1 opens up the possibility for several interesting comparisons with already known results, especially for long time horizons, which are particularly relevant for pension funds. One such comparison is with the already established second-order asymptotic of $c(\epsilon)$ w.r.t. ϵ (cf. [KS06]). The main goal being to isolate the effect of third- and fourth-order terms and compare it to the one of lower-order asymptotics. To this end, consider the following setting of a linear payoff on a factor model and an infinite horizon (which was also explored for the case of log-optimality in Chapter III):

In a market with one riskless and one risky asset as well as $T = \infty$, we consider a standard two dimensional Brownian motion (W^1, W^2) generating the augmented filtration $\mathcal{F}(\cdot)$, and formulate the following problem for $B := \rho W^1 + \sqrt{1-\rho^2}W^2,\ \rho \in (-1, 1)$:

$$\begin{aligned}
&\widetilde{\Lambda} = \int_0^{\cdot} e^{-\widetilde{r}t} Z(t)\mathrm{d}t, \qquad \widetilde{r} > 0,\\
&\mathrm{d}Z(t) = k(\theta - Z(t))\mathrm{d}t + b\mathrm{d}B(t), \qquad k, b > 0;\ Z(0), \theta \in \mathbb{R},\\
&\frac{\mathrm{d}\widetilde{S}(t)}{\widetilde{S}(t)} = \widetilde{a}\mathrm{d}t + \widetilde{\sigma}\mathrm{d}W^1(t), \qquad \widetilde{a} \in \mathbb{R};\ \widetilde{\sigma} > 0.
\end{aligned}$$

In turn, in the above model, the portfolio that gives rise to the local martingale numeraire is constant and in particular we have:

$$\begin{aligned}
&\frac{\mathrm{d}S_0(t)}{S_0(t)} = -\nu\widetilde{\sigma}\mathrm{d}W^1(t), \qquad S_0(0) = 1,\\
&F = \int_0^{\cdot} e^{-\widetilde{r}t} S_0(t) Z(t)\mathrm{d}t.
\end{aligned}$$

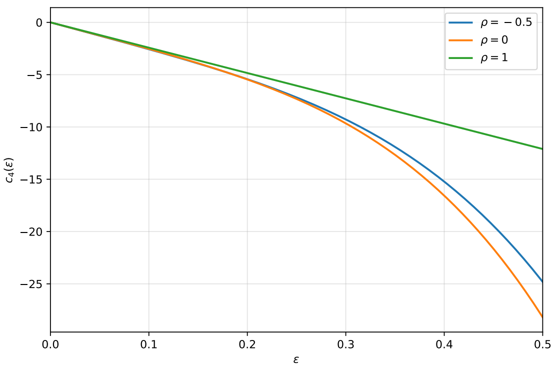
We can directly see what form (4.3.5) takes in the above context. To this end, note that:

$$\begin{aligned}
&A_1 = \frac{Z(0) + \eta/\widetilde{r}}{\widetilde{r} + k}, \qquad \eta := k\theta - b\widetilde{\sigma}\nu\rho,\\
&A_2 = A_1^2 + V_1,\ \ V_1 := \mathbb{E}[[N, N](\infty)] = \frac{b^2(1-\rho^2)}{(\widetilde{r}+k)^2(2\widetilde{r} - (\nu\widetilde{\sigma})^2)},\\
&A_3 = A_1^3 + 3A_1V_1; \qquad G = -V_2\frac{(\nu\widetilde{\sigma})^2}{2\widetilde{r} - (\nu\widetilde{\sigma})^2},\\
&A_4 = A_1^4 + 6A_1^2V_1 + 3V_2; \qquad V_2 := \mathbb{E}[[N, N](\infty)^2] = \frac{b^4(1-\rho^2)^2}{(\widetilde{r}+k)^4(2\widetilde{r} - (\nu\widetilde{\sigma})^2)(2\widetilde{r} - 3(\nu\widetilde{\sigma})^2)},
\end{aligned}$$

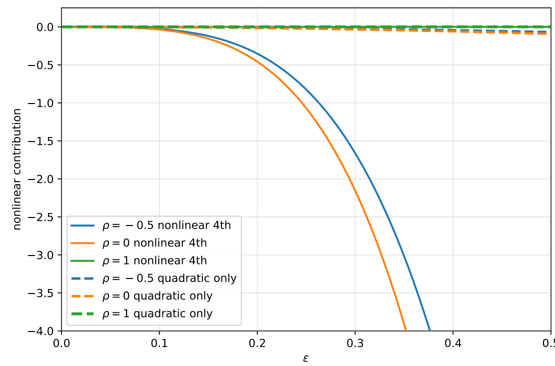
where $\widetilde{r} > 3(\nu\widetilde{\sigma})^2/2$. Substituting the above to (4.3.5), we have:

$$c(\epsilon) = \underbrace{-A_1\epsilon - \frac{V_1}{2}\epsilon^2 - \frac{A_1V_1}{2}\epsilon^3 - \left(\frac{A_1^2V_1}{2} + \frac{3V_1^2}{8} + \frac{V_2}{4}\right)\epsilon^4}_{c_4(\epsilon)} + o(\epsilon^4).$$

More concretely, continuing with Example III.3, say we have a pension fund that trades in the equity market, $\widetilde{S}$, while also holding a non-replicable cumulative liability (Z). Assuming the following parametrisation: $\widetilde{a} = 0.028$, $\widetilde{\sigma} = 0.18$, $k = 0.7$, $b = 0.15$, $Z(0) = 1 = \theta$ (normalised) and $\widetilde{r} = 0.04$ we have:



(a) Fourth-order truncation of $c(\epsilon)$.



(b) Nonlinear part: $c_4(\epsilon) + A_1\epsilon$ (solid) vs $-\frac{V_1}{2}\epsilon^2$ (dashed).

Note that even though ϵ^3, ϵ^4 are dominated by ϵ^2 for $\epsilon \in (0,1)$, in the above example they come with much larger coefficients. In particular, the cubic term scales like $A_1\epsilon$ times the quadratic term and the dominant quartic term scales like $A_1^2\epsilon^2$ times it. In other words, as the magnitude of the (discounted) mean liability stream of the pension fund increases, the difference between the quadratic and the fourth-order expansion of $c(\epsilon)$ becomes more pronounced—for moderate values of ϵ.

Another interesting comparison is between the finite- and infinite-horizon cases. To this end note that for some $T \in \mathbb{R}_{>0}$, the respective forms of A_1, A_2, A_3, A_4, V_1, V_2, and G are:

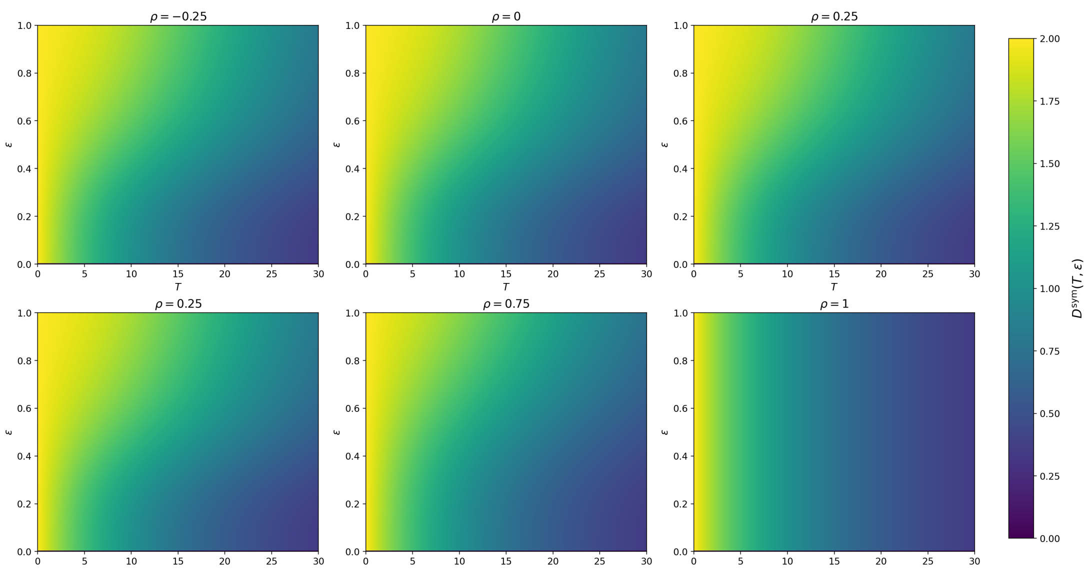
$$
\begin{aligned}
A_1^T &:= \frac{1 - e^{-(\widetilde{r}+k)T}}{\widetilde{r}+k} Z(0) + \frac{\eta}{k}\left(\frac{1-e^{-\widetilde{r}T}}{\widetilde{r}} - \frac{1-e^{-(\widetilde{r}+k)T}}{\widetilde{r}+k}\right), \\
H(t) &:= \int_0^t e^{-(2\widetilde{r}-(\nu\widetilde{\sigma})^2)}\left(1 - e^{-(\widetilde{r}+k)(t-u)}\right)^2 \mathrm{d}u, \\
V_1^T &:= \frac{b^2(1-\rho^2)}{(\widetilde{r}+k)^2} H(T), \\
A_2^T &:= (A_1^T)^2 + V_1^T,
\end{aligned}
$$

$$
\begin{aligned}
A_3^T &:= (A_1^T)^3 + 3A_1^T V_1^T, \\
V_2^T &:= (V_1^T)^2 + 4(\nu\widetilde{\sigma})^2 \left(\frac{b^2(1-\rho^2)}{(\widetilde{r}+k)^2} \right)^2 \int_0^T e^{-(4\widetilde{r}-6(\nu\widetilde{\sigma})^2)t} H(T-t)^2 \mathrm{d}t, \\
G^T &:= -\frac{V_2^T - (V_1^T)^2}{2}, \\
A_4^T &:= (A_1^T)^4 + 6(A_1^T)^2 V_1^T + 3V_2^T,
\end{aligned}
$$

where the above tend to their infinite time horizon analogues when $T \to \infty$, assuming $\widetilde{r} > 3(\nu\widetilde{\sigma})^2/2$. In turn, the respective form of $c_4(\epsilon)$ is:

$$
c_4^T(\epsilon) := -A_1^T \epsilon - \frac{V_1^T}{2}\epsilon^2 - \frac{A_1^T V_1^T}{2}\epsilon^3 - \left(\frac{(A_1^T)^2 V_1^T}{2} + \frac{3(V_1^T)^2}{8} + \frac{V_2^T}{4} \right) \epsilon^4.
$$

With $c_4^T(\epsilon)$ and $c_4(\epsilon)$ in hand, we can determine how the latter approximates the former for different values of (T, ϵ). To this end, note:



Symmetric absolute error $D^{\mathrm{sym}}(T,\epsilon) := \dfrac{2|c_4^T(\epsilon) - c_4(\epsilon)|}{|c_4^T(\epsilon)| + |c_4(\epsilon)|}$ for $(T,\epsilon) \in (0,30] \times (0,1]$.

The above heatmaps indicate that the discrepancy between the finite- and infinite-horizon fourth-order truncations is driven primarily by the maturity T, and decreases steadily as the horizon grows. The dependence on the liability size ϵ is also visible, but is comparatively milder, especially away from very small maturities. Another clear feature is the role of correlation: as ρ approaches 1, the discrepancy becomes uniformly smaller, which is consistent with the fact that the non-replicable component weakens in that regime. Overall, the plots suggest that the infinite horizon truncation provides a much better approximation for long maturities, with the improvement being most pronounced when the factor is more strongly aligned with the traded source of risk.

Another notable feature of the symmetric absolute error heatmaps is the bright region near the lower-left corner. When the liability position is very small, both the finite and infinite horizon truncations are also very small in absolute magnitude. However, for short maturities, the finite-horizon quantity is much closer to zero than its infinite-horizon counterpart, since there is very little time for the liability stream to accumulate. As a result, even though the absolute gap between the two approximations is small there, the gap is still large relative to the size of the quantities being compared.

Appendices

A Supplementary material for Chapter II

A.1 Proof(s) of Section 3

Proof of Proposition 2.3.1. The (strict) concavity of the functional $\widetilde{\mathcal{F}}_n(\phi)$ follows by its quadratic form and the positive definiteness of Σ. Hence, the unique solution to (2.3.4) can be directly seen to satisfy:

$$-\frac{\psi(t)'\Sigma}{\delta_{-n}} - k_n(\phi_n(t))'\Sigma - \frac{(\zeta_n(t))'\Sigma}{\delta_n} + \frac{(\zeta_{-n}(t))'\Sigma}{\delta_{-n}} = 0,$$

in a $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$ sense, which in turn leads to the following representation of the frictionless optimiser under the price impact of investor n:

$$\frac{\zeta_{-n} - \psi - \frac{\delta_{-n}}{\delta_n}\zeta_n}{\delta_{-n}k_n}, \tag{A.1}$$

where $\delta_{-n}k_n = (\delta_n + \delta)/\delta_n$. In turn, (A.1) rearranges to the desired result, which belongs in $\mathcal{H}^2$ by linearity. $\square$

Proof of Proposition 2.3.4. Recalling (2.3.5), (2.3.6), (2.3.7) and evaluating the revealed exposure process of investor n for $(\widetilde{\phi}_n, \mu_n)$, we get:

$$\begin{aligned}\Sigma\zeta_n(\widetilde{\phi}) &= \delta_n \overbrace{\left(\frac{\delta_n\mu_{-n} + \delta\mu}{\delta_n + \delta}\right)}^{\mu_n} - \Sigma\frac{\zeta_{-n} - \psi - \frac{\delta_{-n}}{\delta_n}\zeta_n}{\delta_{-n}k_n} \\ &= \frac{\delta_n\mu_{-n}}{\delta_{-n}k_n} + \frac{\Sigma\zeta_n}{\delta_{-n}k_n} + \frac{\frac{\delta_{-n}}{\delta_n}\Sigma\zeta_n}{\delta_{-n}k_n}.\end{aligned}$$

Therefore, we have:

$$\zeta_n(\widetilde{\phi}) = \frac{\frac{\zeta_{-n}-\psi}{(1/\delta_n)\delta_{-n}} + \zeta_n + \frac{\delta_{-n}}{\delta_n}\zeta_n}{(\delta+\delta_n)/\delta_n}. \tag{A.2}$$

Equivalently, (A.2) can be written as:

$$\begin{aligned}\zeta_n(\widetilde{\phi}) &= \frac{\frac{\zeta-\psi}{(1/\delta_n)\delta_{-n}} - \frac{\zeta_n}{(1/\delta_n)\delta_{-n}} + \zeta_n + \frac{\delta_{-n}}{\delta_n}\zeta_n}{\delta/\delta_n + 1} \\ &= \frac{\lambda_n^2}{1-\lambda_n^2}(\zeta_{-n}-\psi) + \frac{1}{1+\lambda_n}\zeta_n,\end{aligned}$$

which concludes the proof. □

Proof of Corollary 2.3.6. Recalling (2.3.5) and (2.3.7), we get that relation (2.3.8) can be equivalently written as:

$$\widetilde{\phi}_n = \delta_n\Sigma^{-1}\mu_n - \left(\frac{\lambda_n^2}{1-\lambda_n^2}(\zeta_{-n}-\psi) + \frac{1}{1+\lambda_n}\zeta_n\right).$$

Furthermore, evaluating (2.2.6) at μ, yields:

$$\widehat{\phi}_n = \delta_n\Sigma^{-1}\mu - \zeta_n.$$

Thus, for the "if" implication using (2.3.6) and (A.2) we have:

$$\begin{aligned}0 &= \delta_n\Sigma^{-1}\left(\frac{\delta_n\left(\frac{\Sigma(\zeta_{-n}-\psi)}{\delta_{-n}} - \frac{\Sigma(\zeta-\psi)}{\delta}\right)}{\delta_n+\delta}\right) - \zeta_n(\widetilde{\phi}) + \zeta_n \\ &= -\delta_n\Sigma^{-1}\mu - \frac{\delta_{-n}}{\delta_n}\zeta_n + \frac{\delta_{-n}+\delta_n}{\delta_n}\zeta_n \\ &= \delta_n\Sigma^{-1}\mu - \zeta_n.\end{aligned}$$

For the other direction, we have that:

$$\begin{aligned}0 &= \delta_n\Sigma^{-1}\frac{\Sigma(\zeta-\psi)}{\delta} - \zeta_n \\ \zeta_{-n}-\psi &= \frac{\delta_{-n}}{\delta_n}\zeta_n,\end{aligned}$$

which together with (A.1) finishes the proof. □

Proof of Theorem 2.3.8. Let $\mathcal{H}^2 \ni \breve{\phi}_m,\ m \in \mathcal{I}$ be the unique solution set of (2.3.9). Focusing on investor n and evaluating (2.3.7) at (2.3.9), we see that her revealed exposure becomes:

$$\zeta_n(\breve{\phi}) = \delta_n \Sigma^{-1}\nu - \underbrace{\frac{\lambda_n(\zeta_{-n}(\breve{\phi}) - \psi) - \lambda_{-n}\zeta_n}{\lambda_n + 1}}_{(\star)}.$$

We now use the investors' symmetry to derive the aggregate revealed exposure process $\zeta(\breve{\phi}) := \sum_{m=1}^{N} \zeta_m(\breve{\phi})$. Therefore, adding and subtracting $\zeta_n(\breve{\phi})/\delta_{-n}k_n$ and using the equivalent formulation of (A.1) for the term $(\star)$, we have:

$$\underbrace{\frac{\overbrace{(\delta_{-n}k_n - 1)}^{\delta/\delta_n}}{\delta_{-n}k_n}}_{>0} \zeta_n(\breve{\phi}) = \delta_n \Sigma^{-1}\nu - \frac{\zeta(t;\breve{\phi})}{\delta_{-n}k_n} + \frac{\psi}{\delta_{-n}k_n} + \frac{(\delta_{-n}/\delta_n)\zeta_n}{\delta_{-n}k_n}$$

$$\begin{aligned}\sum_{m=1}^{N} \zeta_m(\breve{\phi}) &= \Sigma^{-1}\nu \sum_{m=1}^{N} \frac{1}{(1/\delta_m)^2\delta} + \Sigma^{-1}\nu\delta - \zeta(\breve{\phi}) \sum_{m=1}^{N} \frac{1}{(1/\delta_m)\delta} \\ &\quad + \psi \sum_{m=1}^{N} \frac{1}{(1/\delta_m)\delta} + \frac{1}{\delta}\sum_{m=1}^{N} \delta_{-m}\zeta_m.\end{aligned}$$

Therefore, we have for the aggregate exposure process:

$$\zeta(\breve{\phi}) = \frac{\delta\Sigma^{-1}\nu\Big(1 + \sum_{m=1}^{N} \lambda_m^2\Big) + \sum_{m=1}^{N} \lambda_{-m}\zeta_m + \psi}{2} \tag{A.3}$$

Hence, when all the investors act strategically the market-clearing condition becomes:

$$\breve{\phi}_1 + \cdots + \breve{\phi}_N + \psi = 0. \tag{A.4}$$

Now by using the fact that $\breve{\phi}_m$ for each m (which solves (2.3.9)) can be equivalently written as: $\delta_m\Sigma^{-1}\nu - \zeta_m(\breve{\phi})$ for each m, and using (A.3), we get that (A.4) becomes:

$$\begin{aligned} 0 &= \sum_{m=1}^{N} \Big(\delta_m\Sigma^{-1}\nu - \zeta_m(\breve{\phi})\Big) + \psi \\ \delta\Sigma^{-1}\nu - \delta\Sigma^{-1}\nu \sum_{m=1}^{N} \lambda_m^2 &= \sum_{m=1}^{N} (1 - \lambda_m)\zeta_m - \psi \\ \nu - \nu \sum_{m=1}^{N} \lambda_m^2 &= \frac{\Sigma}{\delta}\Big(\zeta - \psi - \sum_{m=1}^{N} \lambda_m\zeta_m\Big). \end{aligned}$$

Noting that the symmetric matrix $\left(1-\sum_{m=1}^{N}\lambda_m^2\right)I_{d\times d}$ is positive definite, we arrive at the desired result.

The uniqueness of the Nash equilibrium returns is a direct consequence of the uniqueness of the solution set $(\breve{\phi}_1,...,\breve{\phi}_N)$ of the linear system (2.3.9). This stems from the fact that the system's coefficient matrix is invertible, which can be easily determined by checking that its transpose is strictly diagonally dominant. Lastly, by linearity we get that the frictionless Nash equilibrium indeed belongs in $\mathcal{H}^2$. □

Proof of Proposition 2.3.9. We begin with the calculation of the asymptotic surplus at the frictionless Nash equilibrium, i.e.: $\mathbb{US}_1^{\breve{\mu}}(\breve{\phi})$ as $\delta_1\to\infty$. To this end, we first need to derive investor's Nash optimal demand, $\breve{\phi}_1$. This can be readily obtained by noting that (2.3.9) can be decoupled through the use of the equilibrium condition: $\sum_{m=1}^{N}\breve{\phi}_m+\psi=0$. Hence, we have:

$$\begin{aligned}\breve{\phi}_1&=\frac{\lambda_1(\zeta_{-1}(\breve{\phi})-\psi)-\lambda_{-1}\zeta_1}{\lambda_1+1}=\frac{\lambda_1(\delta_{-1}\Sigma^{-1}\nu+\breve{\phi}_1)-\lambda_{-1}\zeta_1}{\lambda_1+1}\\&=\lambda_1\delta_{-1}\Sigma^{-1}\nu-\lambda_{-1}\zeta_1.\end{aligned}\tag{A.5}$$

Furthermore, note that for the frictionless Nash equilibrium returns, $\breve{\mu}$, we have:

$$\lim_{\delta_1\to\infty}\breve{\mu}=\frac{\Sigma(\zeta_{-1}-\psi)}{2\delta_{-1}}.\tag{A.6}$$

Using (A.5) and (A.6), we readily get the following limit for investor's Nash optimal demand at $\breve{\mu}$:

$$\lim_{\delta_1\to\infty}\breve{\phi}_1=\frac{1}{2}(\zeta_{-1}-\psi).$$

We now apply the Dominated Convergence Theorem (DCT) on $dt\otimes d\mathbb{P}$ to justify bringing the limits (w.r.t. δ_1) inside the product integral in $\mathbb{US}_1^{\breve{\mu}}(\breve{\phi})$ (passing between the sequential characterisation of limits and back when necessary). To this end, note that $\psi,\zeta_m\in\mathcal{H}^2$ for each m and:

$$\|\breve{\mu}\|_2=\left\|\frac{\Sigma}{\delta}\frac{(\zeta-\psi)-\sum\limits_{m=1}^{N}\lambda_m\zeta_m}{1-\sum\limits_{m=1}^{N}\lambda_m^2}\right\|_2\le\frac{1}{\delta}\frac{1}{1-\lambda_1^2-\sum\limits_{m=2}^{N}\lambda_m^2}\underbrace{\left(\|\Sigma\|_2\|\zeta-\psi\|_2+\sum_{m=1}^{N}\|\zeta_m\|_2\right)}_{=:C}$$

$$< \frac{1}{\delta \sum_{m=2}^{N} \lambda_m (1-\lambda_m)} C < \frac{1}{\delta_2 \big(1 - \frac{\delta_2}{\delta}\big)} C < \frac{1}{\delta_2 \big(1 - \frac{\delta_2}{\delta_{-1}}\big)} C.$$

With regard to $\breve{\phi}$, we use (A.5) and the above to conclude that $\|\breve{\phi}_1\|_2 \leq \delta_{-1}\|\Sigma^{-1}\|_2\|\breve{\mu}\|_2 + \|\zeta_1\|_2$. Hence, applying DCT gives the desired result for the asymptotic surplus at the frictionless Nash equilibrium.

Following similar arguments, we deal with the competitive case. More precisely:

$$\|\mu\|_2 \leq \frac{1}{\delta_{-1}} \|\Sigma\|_2 \|\zeta - \psi\|_2 \qquad \text{and} \qquad \|\widehat{\phi}_1\|_2 \leq \|\zeta - \psi\|_2 + \|\zeta_1\|_2.$$

Taking into account that $\lim_{\delta_1 \to \infty} \widehat{\phi}_1' \mu = 0$ and applying the DCT once more concludes the proof. □

A.2 Proof(s) of Section 4

Proof of Lemma 2.4.3. We begin by formally generalising (2.3.2). For this, we work similarly as in the frictionless market and introduce the following conditions for all $(\phi_n, \dot{\phi}_n)$:

$$\phi_n + \sum_{\substack{m=1 \\ m \neq n}}^{N} \phi_{\Lambda,m} + \psi = 0, \ \forall \phi_n \in \mathcal{W}_{\mathcal{A}}, \tag{A.7}$$

$$\dot{\phi}_n + \sum_{\substack{m=1 \\ m \neq n}}^{N} \dot{\phi}_{\Lambda,m} + \dot{\psi} = 0, \tag{A.8}$$

where the pair $(\phi_{\Lambda,m}, \dot{\phi}_{\Lambda,m})$, for $m \in \mathcal{I} \setminus \{n\}$ solves (2.4.3). Thus, (A.8) reads as:

$$\begin{aligned} 0 = a^{\dot{\phi}_n}(t)\mathrm{d}t &+ \frac{\Lambda^{-1}}{2} \sum_{\substack{m=1 \\ m \neq n}}^{N} \bigg(\frac{\Sigma \phi_{\Lambda,m}(t)}{\bar{\delta}} - \Big(\nu(t) - \frac{\Sigma \zeta_m(t)}{\bar{\delta}} \Big) \bigg) \mathrm{d}t + r \sum_{\substack{m=1 \\ m \neq n}}^{N} \dot{\phi}_{\Lambda,m}(t)\mathrm{d}t \\ &+ a^{\dot{\psi}}(t)\mathrm{d}t + \mathrm{d}M^{\dot{\phi}_n}(t) + \mathrm{d}M^{\dot{\psi}}(t) + \underbrace{\sum_{\substack{m=1 \\ m \neq n}}^{M} \mathrm{d}M_m(t)}_{=:\mathrm{d}\widehat{M}(t)}. \end{aligned}$$

By using (A.8) again, we have that $r\sum_{\substack{m=1\\m\neq n}}^{N}\dot{\phi}_{\Lambda,m}=-r(\dot{\phi}_n+\dot{\psi})$ which in turn yields:

$$0=a^{\dot{\phi}_n}(t)\mathrm{d}t+\frac{\Lambda^{-1}\Sigma}{2}\sum_{\substack{m=1\\m\neq n}}^{N}\frac{\phi_{\Lambda,m}(t)}{\bar{\delta}}\mathrm{d}t-\frac{\Lambda^{-1}(N-1)}{2}\nu(t)\mathrm{d}t+\frac{\Lambda^{-1}\Sigma}{2}\sum_{\substack{m=1\\m\neq n}}^{N}\frac{\zeta_m(t)}{\bar{\delta}}\mathrm{d}t$$
$$-r(\dot{\phi}_n(t)+\dot{\psi}(t))\mathrm{d}t+a^{\dot{\psi}}(t)\mathrm{d}t+\underbrace{\mathrm{d}M^{\dot{\phi}_n}(t)+\mathrm{d}\widehat{M}(t)+\mathrm{d}M^{\dot{\psi}}(t)}_{=:\mathrm{d}M^S(t)}.$$

Thanks to equal risk tolerances and (A.7) we get that $\sum_{\substack{m=1\\m\neq n}}^{N}\phi_{\Lambda,m}/\bar{\delta}=-(\phi_n+\psi)/\bar{\delta}$, which gives:

$$0=\Big(\frac{2\Lambda}{N-1}a^{\dot{\phi}_n}(t)-\frac{2r\Lambda}{N-1}\dot{\phi}_n(t)+\frac{\Sigma(\zeta_{-n}(t)-\phi_n(t)-\psi(t))}{\bar{\delta}(N-1)}$$
$$+\frac{2\Lambda}{N-1}(a^{\dot{\psi}}(t)-r\dot{\psi}(t))-\nu(t)\Big)\mathrm{d}t+\frac{2\Lambda}{N-1}\mathrm{d}M^S(t).$$

Since the rest of the terms are absolutely continuous, $\mathrm{d}M^S(t)$ vanishes since M^S is a continuous martingale of finite variation (from the stronger absolute continuity). Therefore, we have that the price impact process of investor n, $\nu_n(\dot{\phi})\in\mathcal{H}^2$ in a market with transaction costs is (i.e. the frictional counterpart of (2.3.2)):

$$\nu_n(\dot{\phi})=\Big(\frac{2\Lambda}{N-1}a^{\dot{\phi}_n}-\frac{2r\Lambda}{N-1}\dot{\phi}_n+\frac{\Sigma(\zeta_{-n}-\phi_n-\psi)}{\bar{\delta}(N-1)}+\frac{2\Lambda}{N-1}(a^{\dot{\psi}}-r\dot{\psi})\Big).\tag{A.9}$$

The next step is to incorporate (A.9) into (2.4.1) in order to introduce price impact into the frictional problem (similarly to what we did for the frictionless case). Hence, we get:

$$\mathbb{E}\Big[\int_0^T e^{-rt}(\phi_n(t))'\nu_n(t;\dot{\phi})\mathrm{d}t\Big]=\mathbb{E}\Big[\int_0^T e^{-rt}(\phi_n(t))'\Big(\frac{2\Lambda}{N-1}a^{\dot{\phi}_n}(t)-\frac{2r\Lambda}{N-1}\dot{\phi}_n(t)$$
$$+\frac{\Sigma(\zeta_{-n}(t)-\phi_n(t)-\psi(t))}{\bar{\delta}(N-1)}+\frac{2\Lambda}{N-1}(a^{\dot{\psi}}(t)-r\dot{\psi}(t))\Big)\mathrm{d}t\Big].\tag{A.10}$$

Now note that we have the following equality between (componentwise) integrals:

$$\mathbb{E}\Big[\int_0^T e^{-rt}(\phi_n(t))'\frac{2\Lambda}{N-1}\mathrm{d}\dot{\phi}_n(t)\Big]=\mathbb{E}\Big[\int_0^T e^{-rt}(\phi_n(t))'\frac{2\Lambda}{N-1}a^{\dot{\phi}_n}(t)\mathrm{d}t\Big],\tag{A.11}$$

since $\phi_n, a^{\dot{\phi}_n}\in\mathcal{H}^2$, $M^{\dot{\phi}_n}$ is square-integrable and the local martingales $\int_0^t e^{-rs}\frac{2\Lambda^{i,i}}{N-1}\phi_n^i(s)\mathrm{d}M_i^{\dot{\phi}_n}(s)$ satisfy $\mathbb{E}[[\int_0^\cdot e^{-rs}\frac{2\Lambda^{i,i}}{N-1}\phi_n^i(s)\mathrm{d}M_i^{\dot{\phi}_n}(s)](T)^{1/2}]<\infty$, $\forall i\in\mathcal{D}$ and hence are true martingales by the Burkholder-Davis-Gundy inequality. In turn, applying integration by parts on the left-hand side of (A.11) we get:

$$\mathbb{E}\Big[\int_0^T e^{-rt}(\phi_n(t))'\frac{2r\Lambda}{N-1}\dot{\phi}_n(t) - e^{-rt}(\dot{\phi}_n(t))'\frac{2\Lambda}{N-1}\dot{\phi}_n(t)\mathrm{d}t\Big] \tag{A.12}$$

$$= \mathbb{E}\Big[\int_0^T e^{-rt}(\phi_n(t))'\frac{2\Lambda}{N-1}a^{\dot{\phi}_n}(t)\mathrm{d}t\Big]. \tag{A.13}$$

Thus, taking into account (A.10), (A.12) and adjusting (2.4.1), we get the desired result. □

Proof of Proposition 2.4.5. The concavity of (2.4.4) is readily ensured by its quadratic form. More formally, using a sufficient condition for strict concavity, in this context, through the negativeness of the functional's second Gâteaux differential (see further in [KZ05]), we conclude that for all admissible directions $\dot{\theta}_n$ s.t. $\theta_n \in \mathcal{W}_{\mathcal{A}}$ we have:

$$\Big(D_G^2\widetilde{\mathcal{F}}_{\Lambda,n}(\dot{\phi})\dot{\theta}_n, \dot{\theta}_n\Big) = \mathbb{E}\Bigg[\int_0^T e^{-rt}\Big(- (\theta_n(t))'\frac{2\Sigma}{\bar{\delta}(N-1)}\theta_n(t) - (\dot{\theta}_n(t))'\frac{4\Lambda}{N-1}\dot{\theta}_n(t)$$
$$- (\theta_n(t))'\frac{\Sigma}{\bar{\delta}}\theta_n(t) - (\dot{\theta}_n(t))'2\Lambda\dot{\theta}_n(t)\Big)\mathrm{d}t\Bigg] < 0,$$

for $\bar{\delta}, 1/\bar{\delta}(N-1) > 0$ and positive definite matrices Σ and Λ. By the above, we have that $\widetilde{\mathcal{F}}_{\Lambda,n}(\dot{\phi})$ is strictly concave and therefore the solution of (2.4.4) is characterised by the first order condition through the Gâteaux differential (see further in [ET99]), i.e. for all $\dot{\theta}_n$ as prescribed above we have that:

$$\Big(D_G\widetilde{\mathcal{F}}_{\Lambda,n}(\dot{\phi}), \dot{\theta}_n\Big) = 0.$$

Applying this to (2.4.4), we get:

$$\Big(D_G\widetilde{\mathcal{F}}_{\Lambda,n}(\dot{\phi}), \dot{\theta}_n\Big) = \mathbb{E}\Big[\int_0^T e^{-rt}\Big(\Big(\frac{(\zeta_{-n}(t) - 2\phi_n(t) - \psi(t))'\Sigma}{\bar{\delta}(N-1)} + (a^{\dot{\psi}}(t) - r\dot{\psi}(t))'\frac{2\Lambda}{N-1}$$
$$- (\phi_n(t) + \zeta_n(t))'\frac{\Sigma}{\bar{\delta}}\Big)\theta_n(t) - (\dot{\phi}_n(t))'\Big(\frac{4\Lambda}{N-1} + 2\Lambda\Big)\dot{\theta}_n(t)\Big)\mathrm{d}t\Big]$$
$$= \mathbb{E}\Big[\int_0^T e^{-rt}\Big(\Big(\widetilde{\phi}_n(t) + \frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}_n}(t) - r\dot{\psi}(t)) - \phi_n(t)\Big)'\frac{(N+1)\Sigma}{\bar{\delta}(N-1)}\int_0^t \dot{\theta}_n(s)\mathrm{d}s$$
$$- (\dot{\phi}_n(t))'\frac{2(N+1)\Lambda}{N-1}\dot{\theta}_n(t)\Big)\mathrm{d}t\Big]$$

$$= \mathbb{E}\Big[\int_0^T \Big(\mathbb{E}\Big[\int_0^T f(s)\mathrm{d}s\Big|\mathcal{F}(t)\Big] - \int_0^t f(s)\mathrm{d}s - e^{-rt}(\dot{\phi}_n(t))'\frac{2(N+1)\Lambda}{N-1}\Big)\dot{\theta}_n(t)\mathrm{d}t\Big]$$
$$= 0,$$

where $\widetilde{\phi}_n$ is given in (2.3.5) and

$$f(t) := e^{-rt}\big(\widetilde{\phi}_n(t) + \frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}_n}(t) - r\dot{\psi}(t)) - \phi_n(t)\big)'\frac{(N+1)\Sigma}{\bar{\delta}(N-1)}.$$

The third equality above holds by Fubini-Tonelli and the tower properties of conditional expectation. Now, since the above has to hold for every admissible direction $\dot{\theta}_n$, $D_G\widetilde{\mathcal{F}}_{\Lambda,n}(\dot{\phi}) \in \mathcal{A}$, we have that (2.4.4) has a unique solution if and only if the following stochastic integral equation is uniquely satisfied:

$$\dot{\phi}_n(t) = Be^{rt}\mathbb{E}\Big[\int_t^T e^{-rs}\Big(\widetilde{\phi}_n(s) + \frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}_n}(s) - r\dot{\psi}(s)) - \phi_n(s)\Big)\mathrm{d}s\Big|\mathcal{F}(t)\Big], \quad \text{(A.14)}$$

where $B := \Lambda^{-1}\Sigma/2\bar{\delta}$. In turn, assuming (A.14) admits a (unique) solution and taking the continuous version of the martingale (ensured by martingale representation):

$$\mathbb{E}\Big[\int_0^T e^{-rs}\Big(\widetilde{\phi}_n(s) + \frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}_n}(s) - r\dot{\psi}(s)) - \phi_n(s)\Big)\mathrm{d}s\Big|\mathcal{F}(t)\Big],$$

denoted as $\hat{M}_n$, through integration by parts we can rewrite (A.14) as:

$$\mathrm{d}\dot{\phi}_n(t) = e^{rt}\mathrm{d}\hat{M}_n(t) + B\Big(\phi_n(t) - \widetilde{\phi}(t) - \frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}_n}(t) - r\dot{\psi}(t))\Big)\mathrm{d}t + r\dot{\phi}_n(t)\mathrm{d}t,$$

where by the preservation of the martingale property $\mathrm{d}\widetilde{M}_n(t) := e^{rt}\mathrm{d}\hat{M}_n(t)$ is a square-integrable martingale with continuous paths. Hence, combining the above with $\mathrm{d}\phi_n(t) = \dot{\phi}_n(t)\mathrm{d}t$ leads to the desired FBSDE representation.

Conversely assume that (2.4.5) admits a unique solution, then integration by parts yields:

$$e^{-rt}\dot{\phi}_n(t) = \dot{\phi}_n(0) + \int_0^t e^{-rs}\mathrm{d}\widetilde{M}_n(s) + \int_0^t e^{-rs}B\Big(\phi_n(s) - \widetilde{\phi}(s) - \frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}_n}(s) - r\dot{\psi}(s))\Big)\mathrm{d}s. \quad \text{(A.15)}$$

Now, by the terminal condition $\dot{\phi}_n(T) = 0$, we get:

$$\dot{\phi}_n(0) = -\int_0^T e^{-rs}\mathrm{d}\widetilde{M}_n(s) - \int_0^T e^{-rs}B\Big(\phi_n(s) - \widetilde{\phi}(s) - \frac{2\bar{\delta}\Sigma^{-1}\Lambda}{N+1}(a^{\dot{\psi}_n}(s) - r\dot{\psi}(s))\Big)\mathrm{d}s. \quad \text{(A.16)}$$

Substituting (A.16) into (A.15), rearranging and taking conditional expectations we get (A.14). □

Proof of Proposition 2.4.6. The FBSDE of (2.4.5) is part of larger class of linear FBSDEs that have the following form:

$$\begin{aligned} \mathrm{d}X(t) &= \dot{X}(t)\mathrm{d}t, \quad X(0) = 0, \\ \mathrm{d}\dot{X}(t) &= \mathrm{d}N(t) + B(X(t) - \xi(t))\mathrm{d}t + r\dot{X}(t)\mathrm{d}t, \quad \dot{X}(T) = 0, \end{aligned} \tag{A.17}$$

where B is a real matrix with only positive eigenvalues, $r \geq 0$, $\xi \in \mathcal{H}^2$ and N is a square-integrable martingale. The existence and uniqueness of this class of FBSDEs' solutions is studied in [BFHMK18] (or, for example, [MY07, Chapter 2, §3-5] in general), where it is shown that the (pathwise) unique processes $(X, \dot{X})$ that solve (A.17) satisfy a linear differential equation with the following (unique) solution given through variation of parameters and (primary) matrix functions:

$$X = \int_0^{\cdot} e^{-\int_s^{\cdot} F(u)\mathrm{d}u}\widetilde{\xi}(s)\mathrm{d}s, \tag{A.18}$$

$$\dot{X} = \widetilde{\xi} - FX, \tag{A.19}$$

where:

$$\begin{aligned} &\Delta := B + \frac{r^2}{4}I_d, \\ &G(t) := \cosh(\sqrt{\Delta}(T-t)), \\ &F(t) := -\Big(\Delta G(t) - \frac{r}{2}\dot{G}(t)\Big)^{-1} B\dot{G}(t), \\ &\widetilde{\xi}(t) = \Big(\Delta G(t) - \frac{r}{2}\dot{G}(t)\Big)^{-1}\mathbb{E}\Big[\int_t^T \Big(\Delta G(s) - \frac{r}{2}\dot{G}(s)\Big)Be^{-\frac{r}{2}(s-t)}\xi(s)\mathrm{d}s\Big|\mathcal{F}(t)\Big], \end{aligned}$$

More precisely, for $\widetilde{\xi}$ we have:

$$\begin{aligned} &\Big(\Delta G(t) - \frac{r}{2}\dot{G}(t)\Big)^{-1}\mathbb{E}\Big[\int_t^T \Big(\Delta G(s) - \frac{r}{2}\dot{G}(s)\Big)Be^{-\frac{r}{2}(s-t)}\xi(s)\mathrm{d}s\Big|\mathcal{F}(t)\Big] = \\ &\Big(\Delta G(t) - \frac{r}{2}\dot{G}(t)\Big)^{-1}e^{\frac{r}{2}t}\Big(\bar{M}(t) - \int_0^t \Big(\Delta G(s) - \frac{r}{2}\dot{G}(s)\Big)Be^{-\frac{r}{2}s}\xi(s)\mathrm{d}s\Big), \end{aligned}$$

where $\bar{M}$ is a continuous version of $\mathbb{E}[\int_0^T (\Delta G(s) - \frac{r}{2}\dot{G}(s))Be^{-\frac{r}{2}s}\xi(s)\mathrm{d}s|\mathcal{F}(t)]$, ensured by martingale representation.

In particular, going back to (2.4.5) and applying the above we get that $\widetilde{\phi}_{\Lambda,n} \in \mathcal{H}^2$ as $\mathbb{TP}_n \in \mathcal{H}^2$; which holds since $\mathbb{TP}_n$ is a linear combination of processes in $\mathcal{H}^2$. To see that, initially note that $\widetilde{\mathbb{TP}}_n \in \mathcal{H}^2$ and:

$$\sup_{t\in[0,T]} \|(\Delta G(t) - r/2\dot{G}(t))^{-1}\|_2 < \infty,$$

by the properties of matrix functions (for a detailed exposition, refer to [Hig08]) and the fact that:

$$\inf_{t\in[0,T]} \Big(x\cosh(x(T-t)) + (r/2)x\sinh(x(T-t))\Big) \geq x > 0, \quad \forall x \in (0,\infty).$$

In turn, these yield $\widetilde{\phi}_{\Lambda,n} \in \mathcal{H}^2$ as desired. Furthermore, we also get $\dot{\widetilde{\phi}}_{\Lambda,n} \in \mathcal{H}^2$ by (A.19) and the triangle inequality. Lastly, following similar steps as in [BFHMK18, Theorem A.4], we see that the unique solution of the RDE (A.19) indeed satisfies the FBSDE of (2.4.5). $\square$

Proof of Corollary 2.4.8. Under transaction costs, the market-clearing condition holds for both the investors' demands and trading rates. Now, following similar steps as the ones we did to derive the price impact process (A.9), we may use (2.4.5) for the strategic investor n and get:

$$\nu = \frac{N^2}{N^2-1}\mu - \frac{\Sigma}{\bar{\delta}(N^2-1)}\zeta_n + \frac{2N\Lambda}{N^2-1}(a^{\dot{\psi}} - r\dot{\psi}),$$

where μ is the frictionless competitive equilibrium returns. Uniqueness follows from the uniqueness of the frictional best-response strategy for investor n, as well as from the uniqueness of the competitive frictional strategies for the rest of the investors. Now, recalling (2.3.6) and noting that the above is a linear combination of processes in $\mathcal{H}^2$, we get the desired result. $\square$

Proof of Theorem 2.4.10. Initially note that for:

$$\phi = \begin{bmatrix}\phi_1' & \dots & \phi_N'\end{bmatrix}', \quad \dot{\phi} = \begin{bmatrix}\dot{\phi}_1' & \dots & \dot{\phi}_N'\end{bmatrix}' \quad \text{and} \quad \breve{M} = \begin{bmatrix}\breve{M}_1' & \dots & \breve{M}_N'\end{bmatrix}'$$

the system (2.4.9) can be rewritten as:

$$\begin{aligned}
\mathrm{d}\phi(t) &= \dot{\phi}(t)\mathrm{d}t, \\
\mathrm{d}\dot{\phi}(t) &= \mathrm{d}\breve{M}(t) + \mathbf{B}(\phi(t) - \mathbf{B}^{-1}Z(t))\mathrm{d}t + r\dot{\phi}(t)\mathrm{d}t,
\end{aligned} \tag{A.20}$$

where:

$$\mathbf{B} := \begin{bmatrix} B & \frac{B}{N+1} & \cdots & \frac{B}{N+1} \\ \frac{B}{N+1} & B & \ddots & \vdots \\ \vdots & \ddots & B & \frac{B}{N+1} \\ \frac{B}{N+1} & \cdots & \frac{B}{N+1} & B \end{bmatrix} \in \mathbb{R}^{dN \times dN} \text{ and}$$

$$Z := \left[\left(B \frac{(N-1)\bar{\delta}\Sigma^{-1}\nu - \psi - (N-1)\zeta_i + 2\bar{\delta}\Sigma^{-1}\Lambda(a^{\dot{\psi}} - r\dot{\psi})}{N+1} \right)' \right]'_{i\in\mathcal{I}}.$$

Now note that for:

$$C := \begin{bmatrix} 1 & \frac{1}{N+1} & \cdots & \frac{1}{N+1} \\ \frac{1}{N+1} & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \frac{1}{N+1} \\ \frac{1}{N+1} & \cdots & \frac{1}{N+1} & 1 \end{bmatrix} \in \mathbb{R}^{N \times N},$$

$$\mathbf{B} = C \otimes B,$$

where $\otimes$ here denotes the Kronecker product. The above shows that (A.20) is once more within the class of linear FBSDEs, hence the machinery of Proposition 2.4.6 should apply.

In particular we have that $\mathbf{B}^{-1}$ is well defined as $\det(\mathbf{B}) = \det(C \otimes B) = \det(C)^d \det(B)^N \neq 0$, since C is real, symmetric, strictly diagonally dominant with positive diagonal entries (and hence symmetric positive definite) and B is the product of two symmetric positive definite matrices. Furthermore, the eigenvalues, $(b_i)_{i\in\mathcal{I}}$, of B are all positive (as the product of two symmetric positive definite matrices) and the eigenvalues, $(c_j)_{i\in\mathcal{I}}$ of C are again positive (as a symmetric positive definite matrix). Therefore, the eigenvalues of $\mathbf{B}$ given as: $b_i c_j,\ i,j \in \mathcal{I}$ are also positive. We also have that $\breve{M}$ is square-integrable and $\mathbf{B}^{-1}Z \in \mathcal{H}^2$ as $\nu, \zeta_m \in \mathcal{H}^2$, for each $m \in \mathcal{I}$ and $\psi \in \mathcal{W}_{\mathcal{A}}$. So, (A.20) is of the form of (A.17) (assuming no dependence of ν on $\phi_m, \dot{\phi}_m,\ m \in \mathcal{I}$) and thus it admits a unique solution satisfying the conditions of $\mathcal{W}_{\mathcal{A}}$ (which can be directly shown, following [BFHMK18, Theorem A.4]).

Now, let $(\breve{\phi}_{\Lambda,m}, \dot{\breve{\phi}}_{\Lambda,m}),\ m \in \mathcal{I}$ be the optimal demand and trading rate of (2.4.9) whose existence and uniqueness is shown by the above. In order for the market clearing to hold, the following equilibrium conditions must be satisfied:

$$\sum_{m=1}^{N} \breve{\phi}_{\Lambda,m} + \psi = 0, \tag{A.21}$$

$$\sum_{m=1}^{N} \dot{\breve{\phi}}_{\Lambda,m} + \dot{\psi} = 0. \tag{A.22}$$

In particular, (A.22) develops as:

$$\begin{aligned}
0 &= \sum_{m=1}^{N} \mathrm{d}\breve{M}_m(t) + B\sum_{m=1}^{N} \breve{\phi}_{\Lambda,m}(t)\mathrm{d}t \\
&\quad - B\sum_{m=1}^{N} \frac{\sum_{\substack{i=1\\ i\neq m}}^{N} (\bar{\delta}\Sigma^{-1}\nu(t) - \breve{\phi}_{\Lambda,i}(t)) - \psi(t) - (N-1)\zeta_m(t) + B^{-1}(a^{\dot{\psi}}(t) - r\dot{\psi}(t))}{N+1}\mathrm{d}t \\
&\quad + r\sum_{m=1}^{N} \dot{\breve{\phi}}_{\Lambda,m}(t)\mathrm{d}t + a^{\dot{\psi}}(t)\mathrm{d}t + \mathrm{d}M^{\dot{\psi}}(t).
\end{aligned}$$

Using (A.21) we get:

$$\begin{aligned}
0 &= \overbrace{\sum_{m=1}^{N} \mathrm{d}\breve{M}_m(t) + \mathrm{d}M^{\dot{\psi}}(t)}^{=:\mathrm{d}\breve{M}^S(t)} - B\psi(t)\mathrm{d}t - r\dot{\psi}(t)\mathrm{d}t + a^{\dot{\psi}}(t)\mathrm{d}t \\
&\quad - B\sum_{m=1}^{N} \frac{(N-1)\bar{\delta}\Sigma^{-1}\nu(t) + \breve{\phi}_{\Lambda,m}(t) - (N-1)\zeta_m(t) + B^{-1}(a^{\dot{\psi}}(t) - r\dot{\psi}(t))}{N+1}\mathrm{d}t.
\end{aligned}$$

Using (A.21) once more, we get:

$$\begin{aligned}
0 = \mathrm{d}\breve{M}^S(t) &+ \frac{(N-1)B\zeta(t)}{N+1}\mathrm{d}t - \frac{NB\psi(t)}{N+1}\mathrm{d}t \\
&+ \frac{a^{\dot{\psi}}(t) - r\dot{\psi}(t)}{N+1}\mathrm{d}t - \frac{BN(N-1)\bar{\delta}\Sigma^{-1}\nu(t)}{N+1}\mathrm{d}t.
\end{aligned}$$

Arguing as in the proof of Lemma 2.4.3 for (A.9), $\mathrm{d}\breve{M}^S(t)$ vanishes. The above is rewritten as:

$$\frac{BN(N-1)\bar{\delta}\Sigma^{-1}\nu}{N+1} = \frac{(N-1)B(\zeta-\psi)}{N+1} - \frac{B\psi}{N+1} + \frac{a^{\dot{\psi}} - r\dot{\psi}}{N+1}$$

and by rearranging the terms we get the desired result.

The uniqueness of the frictional Nash equilibrium returns is a direct consequence of the uniqueness of the solution set $(\breve{\phi}_{\Lambda,m}, \dot{\breve{\phi}}_{\Lambda,m}),\ m \in \mathcal{I}$. Lastly, the frictional Nash equilibrium is in $\mathcal{H}^2$ by linearity and the fact that $\zeta_m \in \mathcal{H}^2$, for each $m \in \mathcal{I}$ and $\psi \in \mathcal{W}_{\mathcal{A}}$. □

Proof of Theorem 2.4.12. The steps for this proof are similar to the ones of Lemma 2.4.3, Proposition 2.4.5, Proposition 2.4.6, and Theorem 2.4.10. Thus, we omit part of the analysis and refer to the original proofs for more details.

We first consider the price impact process of investor 1. As in Lemma 2.4.3, the respective form of (A.9) in this case is given as:

$$\nu_1(t;\dot{\phi}) = 2\Lambda a^{\dot{\phi}_1} - 2r\Lambda\dot{\phi}_1 + \frac{\Sigma}{\delta_2}(\zeta_2 - \phi_1 - \psi) + 2\Lambda(a^{\dot{\psi}} - r\dot{\psi}).$$

Using similar arguments as with (A.11) and (A.12) in Lemma 2.4.3, the adjusted frictional objective functional under investor 1's price impact becomes:

$$\begin{aligned}\widetilde{\mathcal{F}}_{\Lambda,1}(\dot{\phi}) := \mathbb{E}\Big[\int_0^T e^{-rt}\Big((\phi_1(t))'\Big(\frac{\Sigma(\zeta_2(t) - \phi_1(t) - \psi(t))}{\delta_2} + 2\Lambda(a^{\dot{\psi}}(t) - r\dot{\psi}(t))\Big)\\ - \frac{1}{2\delta_1}(\phi_1(t) + \zeta_1(t))'\Sigma(\phi_1(t) + \zeta_1(t)) - (\dot{\phi}_1(t))'\Big(\Lambda + 2\Lambda\Big)\dot{\phi}_1(t)\Big)\mathrm{d}t\Big].\end{aligned}$$

Optimising $\widetilde{\mathcal{F}}_{\Lambda,1}(\dot{\phi})$, as in Proposition 2.4.5, we get that the unique solution of the above functional is characterised by the following FBSDE:

$$\begin{aligned}&\mathrm{d}\widetilde{\phi}_{\Lambda,1}(t) = \dot{\widetilde{\phi}}_{\Lambda,1}(t)\mathrm{d}t, \quad \widetilde{\phi}_{\Lambda,1}(0) = 0,\\ &\mathrm{d}\dot{\widetilde{\phi}}_{\Lambda,1}(t) = \mathrm{d}\widetilde{M}_1(t) + B_1(\widetilde{\phi}_{\Lambda,1}(t) - \chi_1(t))\mathrm{d}t + r\dot{\widetilde{\phi}}_{\Lambda,1}(t)\mathrm{d}t, \quad \dot{\widetilde{\phi}}_{\Lambda,1}(T) = 0,\end{aligned}$$

where:

$$\mathcal{H}^2 \ni \chi_1 := \widetilde{\phi}_1 + \frac{2\delta_1\delta_2\Sigma^{-1}\Lambda}{\delta + \delta_1}(a^{\dot{\psi}_n} - r\dot{\psi}), \qquad B_1 := \frac{\delta + \delta_1}{3\delta_1\delta_2}\frac{\Lambda^{-1}\Sigma}{2},$$

$\widetilde{\phi}_1$ is the best-response strategy of investor 1 in the frictionless case and $\widetilde{M}_1$ is a continuous square-integrable MG. In fact, noting that B_1 has only positive eigenvalues and following similar arguments as in Proposition 2.4.6, the above FBSDE admits a unique (pathwise) solution. Using the investors' symmetry and recalling the characterisation of the Nash equilibrium returns in §4.3, we get the following system for the frictional Nash optimal demands of the investors $m = 1, 2$:

$$\begin{aligned} \mathrm{d}\breve{\phi}_{\Lambda,m}(t) &= \dot{\breve{\phi}}_{\Lambda,m}(t)\mathrm{d}t, \quad \breve{\phi}_{\Lambda,m}(0) = 0, \\ \mathrm{d}\dot{\breve{\phi}}_{\Lambda,m}(t) &= \mathrm{d}\breve{M}_m(t) + B_m(\breve{\phi}_{\Lambda,m}(t) - \breve{\chi}_m(t))\mathrm{d}t + r\dot{\breve{\phi}}_{\Lambda,m}(t)\mathrm{d}t, \quad \dot{\breve{\phi}}_{\Lambda,m}(T) = 0, \end{aligned} \tag{A.23}$$

where $B_m := (\delta + \delta_m)\Lambda^{-1}\Sigma/6\delta_m\delta_{-m}$ and:

$$\breve{\chi}_m := \frac{\lambda_m\Big(\overbrace{\sum_{i=1, i\neq m}^{2} (\delta_i\Sigma^{-1}\nu - \breve{\phi}_{\Lambda,i})}^{\zeta_{-m}(\breve{\phi}_\Lambda)} - \psi\Big) - \lambda_{-m}\zeta_m}{\lambda_m + 1} + \frac{2\delta_m\delta_{-m}\Sigma^{-1}\Lambda}{\delta + \delta_m}(a^{\dot{\psi}_n} - r\dot{\psi}).$$

The key ingredient in deriving the frictional Nash equilibrium returns in this setting is to solve (A.23). Note that for $\phi = \begin{bmatrix}\phi_1' & \phi_2'\end{bmatrix}'$, $\dot{\phi} = \begin{bmatrix}\dot{\phi}_1' & \dot{\phi}_2'\end{bmatrix}'$ and $\breve{M} = \begin{bmatrix}\breve{M}_1' & \breve{M}_2'\end{bmatrix}'$ the system of (A.23) can be rewritten as:

$$\begin{aligned} \mathrm{d}\phi(t) &= \dot{\phi}(t)\mathrm{d}t, \\ \mathrm{d}\dot{\phi}(t) &= \mathrm{d}\breve{M}(t) + \breve{\mathbf{B}}(\phi(t) - \breve{\mathbf{B}}^{-1}\breve{Z}(t))\mathrm{d}t + r\dot{\phi}(t)\mathrm{d}t, \end{aligned}$$

where:

$$\breve{\mathbf{B}} := \begin{bmatrix} \frac{\delta+\delta_1}{3\delta_1\delta_2}\frac{\Lambda^{-1}\Sigma}{2} & \frac{\delta_1}{3\delta_1\delta_2}\frac{\Lambda^{-1}\Sigma}{2} \\ \frac{\delta_2}{3\delta_1\delta_2}\frac{\Lambda^{-1}\Sigma}{2} & \frac{\delta+\delta_2}{3\delta_1\delta_2}\frac{\Lambda^{-1}\Sigma}{2} \end{bmatrix} \in \mathbb{R}^{2d\times 2d},$$

$$\breve{Z} := \left[\left(B_1\frac{\delta_1(\delta_2\Sigma^{-1}\nu-\psi)-\delta_2\zeta_1+2\delta_1\delta_2\Sigma^{-1}\Lambda(a^{\dot{\psi}}-r\dot{\psi})}{\delta+\delta_1}\right)' \quad \left(B_2\frac{\delta_2(\delta_1\Sigma^{-1}\nu-\psi)-\delta_1\zeta_2+2\delta_1\delta_2\Sigma^{-1}\Lambda(a^{\dot{\psi}}-r\dot{\psi})}{\delta+\delta_2}\right)'\right]'.$$

Now note that for:

$$\breve{C} := \begin{bmatrix} \delta+\delta_1 & \delta_1 \\ \delta_2 & \delta+\delta_2 \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

$$\breve{\mathbf{B}} = \breve{C} \otimes \frac{\Lambda^{-1}\Sigma}{6\delta_1\delta_2}.$$

But this is the setting studied in Proposition 2.4.5 (assuming that there is not a dependence of ν on $\phi_m, \dot{\phi}_m,\ m \in \mathcal{I}$). Hence, following similar steps we get that (A.23) has a (pathwise) unique solution.

Let $(\breve{\phi}_{\Lambda,m}, \dot{\breve{\phi}}_{\Lambda,m}),\ m \in \mathcal{I}$ be the unique optimal demand and trading rate of (A.23). In order for the market clearing to hold (A.21) and (A.22), as shown in Theorem 2.4.10 must be satisfied at all times. More precisely, using the above, (A.22) reads:

$$\begin{aligned}
0 = &\overbrace{\sum_{i=1}^{2} \mathrm{d}\breve{M}_i(t) + \mathrm{d}M^{\dot{\psi}}(t)}^{=:\mathrm{d}\breve{M}} + B_1(\breve{\phi}_{\Lambda,1}(t) - \breve{\chi}_1(t))\mathrm{d}t + B_2(\breve{\phi}_{\Lambda,2}(t) - \breve{\chi}_2(t))\mathrm{d}t + r(\dot{\breve{\phi}}_{\Lambda,1}(t) \\
&+ \dot{\breve{\phi}}_{\Lambda,2}(t))\mathrm{d}t + a^{\dot{\psi}}(t)\mathrm{d}t.
\end{aligned}$$

Using (A.21), we have:

$$\begin{aligned}
0 = &\mathrm{d}\breve{M}(t) + \frac{\delta\Lambda^{-1}\Sigma}{3\delta_1\delta_2}(\breve{\phi}_{\Lambda,1}(t) + \breve{\phi}_{\Lambda,2}(t))\mathrm{d}t + \frac{\Lambda^{-1}\Sigma}{6\delta_1\delta_2}(\delta_2\zeta_1(t) + \delta_1\zeta_2(t))\mathrm{d}t + \frac{\delta\Lambda^{-1}\Sigma}{6\delta_1\delta_2}\psi(t)\mathrm{d}t \\
&+ \frac{a^{\dot{\psi}}(t) - r\dot{\psi}(t)}{3}\mathrm{d}t - \frac{\Lambda^{-1}\nu(t)}{3}\mathrm{d}t.
\end{aligned}$$

Working as in the proof of Lemma 2.4.3 for (A.9), $\mathrm{d}\breve{M}(t)$ vanishes. Hence, we get:

$$\begin{aligned}
\nu &= \frac{\Sigma}{2\delta_1\delta_2}(\delta_2\zeta_1 + \delta_1\zeta_2) - \frac{\delta\Sigma}{2\delta_1\delta_2}\psi + \Lambda(a^{\dot{\psi}} - r\dot{\psi}) \\
&= \frac{\Sigma}{\delta}\frac{\lambda_2\zeta_1 + \lambda_1\zeta_2}{2\lambda_1\lambda_2} - \frac{\Sigma}{\delta}\frac{\psi}{2\lambda_1\lambda_2} + \Lambda(a^{\dot{\psi}} - r\dot{\psi}).
\end{aligned}$$

Noting that $\zeta - \lambda_1\zeta_1 - \lambda_2\zeta_2 = \lambda_2\zeta_1 + \lambda_1\zeta_2$ and $1 - \lambda_1^2 - \lambda_2^2 = 2\lambda_1\lambda_2$ yields the desired result. The uniqueness of the frictional Nash equilibrium returns is a direct consequence of the uniqueness of the solution set $(\breve{\phi}_{\Lambda,m}, \dot{\breve{\phi}}_{\Lambda,m}),\ m \in \mathcal{I}$. Lastly, the frictional Nash equilibrium is in $\mathcal{H}^2$ by linearity, the fact that $\zeta_1, \zeta_2 \in \mathcal{H}^2$ and $\psi \in \mathcal{W}_{\mathcal{A}}$. □

B Supplementary material for Chapter III

B.1 Proof(s)

Proof of Theorem 3.3.1. We do some general setup that will be useful in deriving both bounds for the value function $u(\epsilon)$. For the remainder of the proof, we may assume without loss of generality that $\epsilon \geq 0$. The case where ϵ takes negative values is analogous.

Let $x + M$ be the martingale implied by (3.3.1) and consider $B_n := P(0) + \widehat{B}_n$, where $\widehat{B}_n \in \mathcal{N}_\infty$ s.t. $\widehat{B}_n(T)$ tends to $P(T) - P(0)$ in $\mathbb{L}_2$. Define:

$$\Xi := x + |\Delta| + M; \qquad \epsilon^L := 1 \wedge \frac{1}{6\Xi(0)} \wedge \frac{1}{2(\|B_n(T)\|_{\mathbb{L}_\infty})^{1/2}},$$

and note that Ξ is a continuous (adapted) process with $\mathbb{E}[\overline{\Xi}(T)^4] < \infty$, that works as an upper bound for $|\Delta|$, $x + M$, as well as $|N|$. Now, for $\epsilon \in (0, \epsilon^L)$, consider the stopping times:

$$\tau^\Gamma_\epsilon := \inf\{t : |\Gamma(t)| \geq 1/6\epsilon^2\}; \qquad \tau^M_\epsilon := \inf\{t : \Xi(t) \geq 1/6\epsilon\},$$

and $\tau_\epsilon := \tau^\Gamma_\epsilon \wedge \tau^\Xi_\epsilon \wedge T$. Now, let ζ be a non-negative random variable with $\mathbb{E}[\zeta^p] < \infty$ for some $p > 0$. Then, it holds that:

$$\lim_{z\to\infty} z^p \mathbb{P}(\zeta > z) = 0,$$

in turn we have:

$$\lim_{\epsilon\to 0+} \frac{1}{\epsilon^4}\mathbb{P}(\tau^\Gamma_\epsilon < T) = 0; \qquad \lim_{\epsilon\to 0+} \frac{1}{\epsilon^4}\mathbb{P}(\tau^\Xi_\epsilon < T) = 0,$$

since $\mathbb{E}[\overline{\Xi}(T)^4] < \infty$ and $\Gamma \in \mathcal{H}^2$. In particular the above also give:

$$\lim_{\epsilon\to 0+} \frac{1}{\epsilon^4}\mathbb{P}(\tau_\epsilon < T) = 0. \tag{B.1}$$

<u>**The lower bound**</u>

Now note that:

$$1 + \epsilon(\Delta^{\tau_\epsilon} + M - M^{\tau_\epsilon}) + \epsilon^2\Gamma^{\tau_\epsilon} \in \mathcal{X}(1, \epsilon).$$

In turn, Taylor expansion and the above give:

$$\begin{aligned}
u(\epsilon) &\geq \mathbb{E}\left[\log\left(1+\epsilon\left(\Delta^{\tau_\epsilon}(T)+M(T)-M^{\tau_\epsilon}(T)-F(T)\right)+\epsilon^2\Gamma^{\tau_\epsilon}(T)\right)\right] \\
&= -\epsilon\mathbb{E}[F(T)]-\frac{1}{2}\mathbb{E}\left[\left(\epsilon\left(\Delta^{\tau_\epsilon}(T)+M(T)-M^{\tau_\epsilon}(T)-F(T)\right)+\epsilon^2\Gamma^{\tau_\epsilon}(T)\right)^2\right] \\
&+\frac{1}{3}\mathbb{E}\left[\left(\epsilon\left(\Delta^{\tau_\epsilon}(T)+M(T)-M^{\tau_\epsilon}(T)-F(T)\right)+\epsilon^2\Gamma^{\tau_\epsilon}(T)\right)^3\right] \\
&-\frac{1}{4}\mathbb{E}\left[\left(\epsilon\left(\Delta^{\tau_\epsilon}(T)+M(T)-M^{\tau_\epsilon}(T)-F(T)\right)+\epsilon^2\Gamma^{\tau_\epsilon}(T)\right)^4\xi^L_\epsilon\right],
\end{aligned}$$

where ξ^L_ϵ is a r.v., s.t. $\lim_{\epsilon\to 0+}\xi^L_\epsilon = 1$ a.s. and $0<\xi^L_\epsilon\leq 2^4$, $\forall\epsilon\in(0,\epsilon^L)$.

Now define $Q^\epsilon := \Delta^{\tau_\epsilon}+M-M^{\tau_\epsilon}\in\mathcal{M}_4$, where $Q^\epsilon(T)\to\Delta(T)$ almost surely as $\epsilon\to 0+$ (similarly $\Gamma^{\tau_\epsilon}(T)\to\Gamma(T)$ a.s.). Collecting terms, we get:

$$\begin{aligned}
u(\epsilon) &\geq -\epsilon\mathbb{E}[F(T)]-\epsilon^2\mathbb{E}[(Q^\epsilon(T)-F(T))^2/2]+\epsilon^3\mathbb{E}[(Q^\epsilon(T)-F(T))^3/3 \\
&-\Gamma^{\tau_\epsilon}(T)(Q^\epsilon(T)-F(T))]-\epsilon^4\mathbb{E}[(Q^\epsilon(T)-F(T))^4\xi^L_\epsilon/4 \\
&-(Q^\epsilon(T)-F(T))^2\Gamma^{\tau_\epsilon}(T)+\Gamma^{\tau_\epsilon}(T)^2/2]+\epsilon^5\mathbb{E}[(Q^\epsilon(T)-F(T))\Gamma^{\tau_\epsilon}(T)^2 \\
&-(Q^\epsilon(T)-F(T))^3\Gamma^{\tau_\epsilon}(T)\xi^L_\epsilon]+\epsilon^6\mathbb{E}[\Gamma^{\tau_\epsilon}(T)^3/3-(3/2)(Q^\epsilon(T)-F(T))^2\Gamma^{\tau_\epsilon}(T)^2\xi^L_\epsilon] \\
&-\epsilon^7\mathbb{E}[(Q^\epsilon(T)-F(T))\Gamma^{\tau_\epsilon}(T)^3\xi^L_\epsilon]-\epsilon^8\mathbb{E}[\Gamma^{\tau_\epsilon}(T)^4\xi^L_\epsilon/4].
\end{aligned}$$

We examine the above terms (w.r.t. to ϵ) one by one, beginning with the highest order.

Eighth order term

$$\begin{aligned}
\epsilon^4\Gamma^{\tau_\epsilon}(T)^4 &= \epsilon^4\Gamma^{\tau_\epsilon}(T)^4\mathbf{1}_{\tau_\epsilon<T}+\epsilon^4\Gamma^{\tau_\epsilon}(T)^4\mathbf{1}_{\tau_\epsilon\geq T} \\
&\leq \frac{1}{(6\epsilon)^4}\mathbf{1}_{\tau_\epsilon<T}+\epsilon^4\overline{\Gamma}(T)^4\mathbf{1}_{\overline{\Gamma}(T)\leq 1/6\epsilon^2}.
\end{aligned}$$

Hence:

$$\epsilon^4\mathbb{E}[\Gamma^{\tau_\epsilon}(T)^4]\leq\frac{1}{(6\epsilon^4)}\mathbb{P}(\tau_\epsilon<T)+\epsilon^4\mathbb{E}\left[\overline{\Gamma}(T)^4\mathbf{1}_{\overline{\Gamma}(T)\leq 1/6\epsilon^2}\right].$$

In turn, $\lim_{\epsilon\to 0+}\epsilon^4\mathbb{E}[\Gamma^{\tau_\epsilon}(T)^4]=0$ follows by (B.1), $\Gamma\in\mathcal{M}_2$ and (B.6) after a simple change of variables. In fact note that the same holds when we multiply with ξ^L due to its boundedness.

Seventh order term

Holder's inequality gives:

$$\epsilon^3\mathbb{E}[|Q^\epsilon(T)-F(T)||\Gamma^{\tau_\epsilon}(T)|^3] \le \left(\mathbb{E}\left[\left(\left|M(T)-F(T)\right| + \sup_{t\in[0,T]}\left|\Delta(t)-M(t)\right|\right)^4\right]\right)^{\frac{1}{4}} \left(\mathbb{E}\left[(\epsilon\Gamma(\tau_\epsilon))^4\right]\right)^{\frac{3}{4}},$$

where the first expectation is finite by (3.3.1) and:

$$\lim_{\epsilon\to 0+}\epsilon^4\mathbb{E}[\Gamma^{\tau_\epsilon}(T)^4]=0,$$

by the previous result on the eighth-order term.

<u>Sixth order terms</u>

The first part follows similarly to the eighth-order term, since:

$$\epsilon^2|\Gamma^{\tau_\epsilon}(T)|^3 \le \frac{1}{6^3\epsilon^4}\mathbf{1}_{\tau_\epsilon<T}+\epsilon^2\overline{\Gamma}(T)^3\mathbf{1}_{\overline{\Gamma}(T)\le 1/6\epsilon^2}.$$

Hence, as before, $\lim_{\epsilon\to 0+}\epsilon^2\mathbb{E}[|\Gamma^{\tau_\epsilon}(T)|^3]=0$ follows by (B.1), $\Gamma\in\mathcal{M}_2$ and (B.6).

For the second part note that by Holder's inequality we have:

$$\epsilon^2\mathbb{E}[(Q^\epsilon(T)-F(T))^2\Gamma^{\tau_\epsilon}(T)^2] \le \left(\mathbb{E}\left[\left(\left|M(T)-F(T)\right| + \sup_{t\in[0,T]}\left|\Delta(t)-M(t)\right|\right)^4\right]\right)^{\frac{1}{2}} \left(\mathbb{E}\left[(\epsilon\Gamma^{\tau_\epsilon}(T))^4\right]\right)^{\frac{1}{2}},$$

where the first expectation is finite by (3.3.1) and:

$$\lim_{\epsilon\to 0+}\epsilon^4\mathbb{E}[\Gamma^{\tau_\epsilon}(T)^4]=0,$$

by the previous result on the eighth-order term.

<u>Fifth order terms</u>

For the first part we have that:

$$\begin{aligned}\epsilon\mathbb{E}[(Q^\epsilon(T)-F(T))\Gamma^{\tau_\epsilon}(T)^2] &= \epsilon\mathbb{E}\left[\mathbb{E}\left[(Q^\epsilon(T)-F(T))\Gamma^{\tau_\epsilon}(T)^2\middle|\mathcal{F}(\tau_\epsilon)\right]\right]\\ &= -\mathbb{E}[\epsilon N^{\tau_\epsilon}(T)\Gamma^{\tau_\epsilon}(T)^2].\end{aligned}$$

Now note that:

$$\mathbb{E}[|\epsilon N^{\tau_\epsilon}(T)|\Gamma^{\tau_\epsilon}(T)^2] \leq \mathbb{E}[|\epsilon N^{\tau_\epsilon}(T)|\overline{\Gamma}(T)^2],$$

where $|N^{\tau_\epsilon}(T)| \leq \Xi^{\tau_\epsilon}(T)$ and $6\epsilon\Xi^{\tau_\epsilon}(T) \leq 1$. In turn, since $\lim_{\epsilon\to 0+}|\epsilon N^{\tau_\epsilon}(T)| = 0$ a.s., we have:

$$\lim_{\epsilon\to 0+} \epsilon\mathbb{E}[(Q^\epsilon(T) - F(T))\Gamma^{\tau_\epsilon}(T)^2] = 0.$$

For the second part,

$$\lim_{\epsilon\to 0+} \epsilon\mathbb{E}[|Q^\epsilon(T) - F(T)|^3|\Gamma^{\tau_\epsilon}(T)|] = 0,$$

follows from Holder's inequality and $\lim_{\epsilon\to 0+}\epsilon^4\mathbb{E}[\Gamma^{\tau_\epsilon}(T)^4] = 0$, similarly to the sixth and seventh order terms.

Fourth-order terms

We have:

$$\begin{aligned}
\lim_{\epsilon\to 0+} \mathbb{E}[\Gamma^{\tau_\epsilon}(T)^2] &= \mathbb{E}[\Gamma(T)^2], \\
\lim_{\epsilon\to 0+} \mathbb{E}[(Q^\epsilon(T) - F(T))^2\Gamma^{\tau_\epsilon}(T)] &= \mathbb{E}[N(T)^2\Gamma(T)], \\
\lim_{\epsilon\to 0+} \mathbb{E}[(Q^\epsilon(T) - F(T))^4\xi^L_\epsilon] &= \mathbb{E}[N(T)^4],
\end{aligned}$$

which all hold by dominated convergence, (3.3.1), $0 < \xi^L_\epsilon \leq 2^4$ and $\lim_{\epsilon\to 0+}\xi^L_\epsilon = 1$ a.s.

Third-order terms

Initially note that:

$$\mathbb{E}[\Gamma^{\tau_\epsilon}(T)(Q^\epsilon(T) - F(T))] = \mathbb{E}\left[\mathbb{E}\left[\Gamma^{\tau_\epsilon}(T)(Q^\epsilon(T) - F(T))\Big|\mathcal{F}(\tau_\epsilon)\right]\right] = 0.$$

For the second part we have by Holder's inequality:

$$\begin{aligned}
\frac{1}{\epsilon}\mathbb{E}[|(Q^\epsilon(T) - F(T))^3 + N(T)^3|] &= \frac{1}{\epsilon}\mathbb{E}[|(Q^\epsilon(T) - F(T))^3 + N(T)^3|\mathbf{1}_{\tau_\epsilon<T}] \\
&\leq 2^{\frac{1}{4}}\Bigg(\mathbb{E}\Bigg[\Bigg(\Bigg|M(T) - F(T)\Bigg|
\end{aligned}$$

$$
\left.\left.+ \sup_{t\in[0,T]}\left|\Delta(t)-M(t)\right|\right)^4 + N(T)^4\right]\Big)^{\frac{3}{4}}
\cdot\left(\frac{1}{\epsilon^4}\mathbb{P}(\tau_\epsilon<T)\right)^{\frac{1}{4}},
$$

where the expectation on the last inequality is finite by (3.3.1), and $\lim_{\epsilon\to 0+}(\epsilon^4)^{-1}\mathbb{P}(\tau_\epsilon < T) = 0$ by (B.1). Hence the second part also tends to zero.

Second-order term

We have:

$$
\lim_{\epsilon\to 0+}\frac{1}{\epsilon^2}\mathbb{E}[|(Q^\epsilon(T)-F(T))^2-N(T)^2|]=0,
$$

following the same process as for the third-order term.

Lastly, combining all the above we have

$$
\limsup_{\epsilon\to 0+}\frac{1}{\epsilon^4}\Big(u(\epsilon)+\epsilon\mathbb{E}[F(T)]+\frac{\epsilon^2}{2}\mathbb{E}[N(T)^2]+\frac{\epsilon^3}{3}\mathbb{E}[N(T)^3]+\frac{\epsilon^4}{4}\mathbb{E}[N(T)^4]
+\epsilon^4\mathbb{E}[\Gamma(T)^2/2-N(T)^2\Gamma(T)]\Big)\geq 0, \tag{B.2}
$$

where $\mathbb{E}[\Gamma(T)^2/2-N(T)^2\Gamma(T)]=-\mathbb{E}[\Gamma(T)^2]/2$.

The upper bound

Note that by Legendre-Fenchel duality we have:

$$
\log(1+X^\epsilon(T)-\epsilon F(T))\leq -1-\log(1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T))
+(1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T))(1+X^\epsilon(T)-\epsilon F(T)).
$$

Now, denoting the probability measure induced by $(1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T))/\mathbb{E}[1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T)]$ as $\mathbb{Q}^\epsilon$ we claim that $1+X^\epsilon$ is a supermartingale under $\mathbb{Q}^\epsilon$ and therefore:

$$
\mathbb{E}^{\mathbb{Q}^\epsilon}[1+X^\epsilon(T)]\leq 1.
$$

Indeed, we have $1+X^\epsilon(T)-\epsilon F(T)>0$ and in particular from (3.3.1) we get $\epsilon|F(T)|\leq \epsilon(x+M(T))$. It follows that $1+X^\epsilon+\epsilon(x+M)$ is a positive stochastic integral w.r.t. S. This, along with the fact that N, P are strongly orthogonal to S yields that $1+X^\epsilon+\epsilon(x+M)$ is a supermartingale under $\mathbb{Q}^\epsilon$. The claim now follows from the fact that $\epsilon(x+M)$ is a martingale under $\mathbb{Q}^\epsilon$. Hence we get:

$$\begin{aligned}\mathbb{E}[(1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T))(1+X^\epsilon(T))] &= \mathbb{E}[1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T)]\mathbb{E}^{\mathbb{Q}^\epsilon}[1+X^\epsilon(T)] \\ &\le \mathbb{E}[1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T)].\end{aligned}$$

Furthermore we have:

$$\begin{aligned}&\mathbb{E}[N^{\tau_\epsilon}(T)F(T)] = \mathbb{E}[N^{\tau_\epsilon}(T)N(T)] = \mathbb{E}[N^{\tau_\epsilon}(T)^2],\\ &\mathbb{E}[B_n(T)F(T)] = \mathbb{E}[B_n(T)N(T)],\end{aligned}$$

given the definitions and properties of N, P. It follows that:

$$\begin{aligned}u(\epsilon) \le& -\epsilon\mathbb{E}[F(T)] - \epsilon^2\mathbb{E}[N^{\tau_\epsilon}(T)^2] - \epsilon^3\mathbb{E}[B_n(T)N(T)]\\ &+ \mathbb{E}[-\log(1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T)) + \epsilon N^{\tau_\epsilon}(T) + \epsilon^2 B_n(T)].\end{aligned}$$

In turn, Taylor expansion yields:

$$\begin{aligned}\mathbb{E}[-\log(1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T)) + \epsilon N^{\tau_\epsilon}(T) + \epsilon^2 B_n(T)] =& \frac{1}{2}\epsilon^2\mathbb{E}[(N^{\tau_\epsilon}(T)+\epsilon B_n(T))^2]\\ &- \frac{1}{3}\epsilon^3\mathbb{E}[(N^{\tau_\epsilon}(T)+\epsilon B_n(T))^3]\\ &+ \frac{1}{4}\epsilon^4\mathbb{E}[(N^{\tau_\epsilon}(T)+\epsilon B_n(T))^4\xi_\epsilon^U],\end{aligned}$$

where the random variables ξ_ϵ^U satisfy:

$$0 < \xi_\epsilon^U \le 2^4,\ \forall \epsilon \in (0,\epsilon^L); \qquad \lim_{\epsilon\to 0+}\xi_\epsilon^U = 1 \text{ a.s.}$$

Collecting terms of same order, we obtain:

$$\begin{aligned}\mathbb{E}[-\log(1+\epsilon N^{\tau_\epsilon}(T)+\epsilon^2 B_n(T)) + \epsilon N^{\tau_\epsilon}(T) + \epsilon^2 B_n(T)] =& \epsilon^2\mathbb{E}\left[\frac{N^{\tau_\epsilon}(T)^2}{2}\right]\\ &+ \epsilon^3\mathbb{E}\left[N^{\tau_\epsilon}(T)B_n(T) - \frac{N^{\tau_\epsilon}(T)^3}{3}\right]\\ &+ \epsilon^4\mathbb{E}\Bigg[\frac{B_n(T)^2}{2} - N^{\tau_\epsilon}(T)^2 B_n(T)\end{aligned}$$

$$
\begin{aligned}
&+ \frac{N^{\tau_\epsilon}(T)^4}{4}\xi_\epsilon^U\Big] \\
&+ \epsilon^5\mathbb{E}[-N^{\tau_\epsilon}(T)B_n(T)^2 \\
&+ N^{\tau_\epsilon}(T)^3B_n(T)\xi_\epsilon^U] \\
&\epsilon^6\mathbb{E}\Big[-\frac{B_n(T)^3}{3} \\
&+ \frac{3}{2}B_n(T)^2N^{\tau_\epsilon}(T)^2\xi_\epsilon^U\Big] \\
&+ \epsilon^7\mathbb{E}[N^{\tau_\epsilon}(T)B_n(T)^3\xi_\epsilon^U] \\
&+ \epsilon^8\mathbb{E}\left[\frac{B_n(T)^4}{4}\xi_\epsilon^U\right]
\end{aligned}
$$

Combining the above, we examine the terms (w.r.t. to ϵ) one by one, beginning with the highest order.

High order terms

All terms of order 5 and above, when divided by ϵ^4, converge to zero as $\epsilon \to 0+$. This is a direct consequence of the boundedness of ξ_ϵ^U, $\|B_n(T)\|_{\mathbb{L}_\infty} < \infty$, and the fact that $\mathbb{E}[\overline{N}(T)^4] < \infty$.

Fourth-order terms

We have that:

$$
\begin{aligned}
&\lim_{\epsilon\to 0+}\mathbb{E}[\xi_\epsilon^U N^{\tau_\epsilon}(T)^4] = \mathbb{E}[N(T)^4], \\
&\lim_{\epsilon\to 0+}\mathbb{E}[N^{\tau_\epsilon}(T)^2B_n(T)] = \mathbb{E}[N(T)^2B_n(T)],
\end{aligned}
$$

which all hold by dominated convergence, (3.3.1), $0 < \xi_\epsilon^U \leq 2^4$ and $\lim_{\epsilon\to 0+}\xi_\epsilon^U = 1$ a.s.

Third-order terms

Holder's inequality gives:

$$
\frac{1}{\epsilon}\mathbb{E}[|N^{\tau_\epsilon}(T)^3 - N(T)^3|] \leq 2\mathbb{E}[\overline{N}(T)^4]^{\frac{3}{4}}\left(\frac{1}{\epsilon^4}\mathbb{P}(\tau_\epsilon < T)\right)^{\frac{1}{4}},
$$

which goes to zero by (B.1). Similarly, we have:

$$
\frac{1}{\epsilon}\mathbb{E}[|B_n(T)(N(T) - N^{\tau_\epsilon}(T))|] \leq 2\|B_n(T)\|_{\mathbb{L}_\infty}\mathbb{E}[\overline{N}(T)^{4/3}]^{\frac{3}{4}}\left(\frac{1}{\epsilon^4}\mathbb{P}(\tau_\epsilon < T)\right)^{\frac{1}{4}},
$$

which once more goes to zero as $\epsilon \to 0+$.

Second-order term

We have:

$$\frac{1}{\epsilon^2}\mathbb{E}[|N^{\tau_\epsilon}(T)^2 - N(T)^2|] \leq 2\mathbb{E}[(\overline{N}(T))^4]^{\frac{1}{2}} \left(\frac{1}{\epsilon^4}\mathbb{P}(\tau_\epsilon < T)\right)^{\frac{1}{2}},$$

which goes to zero, similarly to the third-order case for the upper bound.

Combining the previous, we get:

$$\limsup_{\epsilon\to 0+} \frac{1}{\epsilon^4}\Big(u(\epsilon) + \epsilon\mathbb{E}[F(T)] + \frac{\epsilon^2}{2}\mathbb{E}[N(T)^2] + \frac{\epsilon^3}{3}\mathbb{E}[N(T)^3] \\ - \epsilon^4\Big(\frac{B_n(T)}{2} - N(T)^2 B_n(T) + \frac{N(T)^4}{4}\Big)\Big) \leq 0,$$

in turn, adding and subtracting $\mathbb{E}[N(T)^4/4]$, we obtain:

$$\limsup_{\epsilon\to 0+} \frac{1}{\epsilon^4}\Big(u(\epsilon) + \epsilon\mathbb{E}[F(T)] + \frac{\epsilon^2}{2}\mathbb{E}[N(T)^2] + \frac{\epsilon^3}{3}\mathbb{E}[N(T)^3] + \frac{\epsilon^4}{4}\mathbb{E}[N(T)^4]\Big) \leq \\ \frac{1}{2}\mathbb{E}[(B_n(T) - N(T)^2)^2],$$

Letting $n \to \infty$, and after some algebra we have:

$$\limsup_{\epsilon\to 0+} \frac{1}{\epsilon^4}\Big(u(\epsilon) + \epsilon\mathbb{E}[F(T)] + \frac{\epsilon^2}{2}\mathbb{E}[N(T)^2] + \frac{\epsilon^3}{3}\mathbb{E}[N(T)^3] + \frac{\epsilon^4}{4}\mathbb{E}[N(T)^4] \\ + \epsilon^4\mathbb{E}[\Gamma(T)^2/2 - N(T)^2\Gamma(T)]\Big) \leq 0. \tag{B.3}$$

Lastly, noting that (B.2) and (B.3) also hold for limit inferior as $\epsilon \to 0+$ concludes the first part of the proof.

To complete the proof, we now deal with (3.3.3). Similarly to how [KS06, Theorem 1] treated the first-order asymptotics of the optimal wealth, we take any sequence $\epsilon_n \in (0, \epsilon^L)$ s.t. $\lim_{n\to\infty} \epsilon_n = 0$ [1]. In turn, recalling the process we were considering for the fourth-order lower bound of the value function, i.e. $Q^{\epsilon_n} := \Delta^{\tau_{\epsilon_n}} + M - M^{\tau_{\epsilon_n}}$ and Taylor expanding

[1]This should come without loss of generality for our purposes since any positive sequence ϵ_n, that tends to zero, eventually lies in $(0, \epsilon^L)$.

$U(1+\epsilon_n Q^{\epsilon_n}(T)+\epsilon_n^2\Gamma^{\tau_{\epsilon_n}}(T)-\epsilon_n F(T))$ around $X^{\epsilon_n}(T)-\epsilon_n F(T)$ we get:

$$\begin{aligned}&\log(1+\epsilon_n Q^{\epsilon_n}(T)+\epsilon_n^2\Gamma^{\tau_{\epsilon_n}}(T)-\epsilon_n F(T))-\log(1+X^{\epsilon_n}(T)-\epsilon_n F(T))=\\&\frac{1}{1+X^{\epsilon_n}(T)-\epsilon_n F(T)}\left(\left(1+\epsilon_n Q^{\epsilon_n}(T)+\epsilon_n^2\Gamma^{\tau_{\epsilon_n}}(T)-\epsilon_n F(T)\right)-(1+X^{\epsilon_n}(T)-\epsilon_n F(T))\right)\\&-\frac{1}{2(1+\xi_n)^2}\left(X^{\epsilon_n}(T)-\epsilon_n Q^{\epsilon_n}(T)-\epsilon_n^2\Gamma^{\tau_{\epsilon_n}}(T)\right)^2,\end{aligned} \tag{B.4}$$

where ξ_n is a r.v. between $X^{\epsilon_n}(T)-\epsilon_n F(T)$ and $\epsilon_n Q^{\epsilon_n}(T)+\epsilon_n^2\Gamma^{\tau_{\epsilon_n}}(T)-\epsilon_n F(T)$ s.t. $\lim_{n\to\infty}\xi_n=0$ in probability. This holds by the fact that $\lim_{n\to\infty}X^{\epsilon_n}(T)=0$ in probability (see [KS99, Lemma 3.6], [KS06, Theorem 1] for similar results that cover the case of log-utility). Passing to any subsequence of ϵ_n we can always find a further subsequence, denoted by ϵ_m, s.t. $\lim_{m\to\infty}X^{\epsilon_m}(T)=0$ a.s., which in turn implies that the r.v. $\sup_m X^{\epsilon_m}(T)$ is well-defined a.s.

Now note that:

$$0<1+\xi_m\le\left(2+\sup_m X^{\epsilon_m}(T)+|F(T)|\right)\vee\left(2+\overline{Q}^{\epsilon}(T)+\overline{\Gamma}(T)+|F(T)|\right)=:B,$$

which implies that $-1/2(1+\xi_m)^2\le -1/2B^2$ and in particular we have $1/B\le 1$. Hence (B.4) becomes:

$$\begin{aligned}&\log(1+\epsilon_m Q^{\epsilon_m}(T)+\epsilon_m^2\Gamma^{\tau_{\epsilon_m}}(T)-\epsilon_m F(T))-\log(1+X^{\epsilon_m}(T)-\epsilon_m F(T))\le\\&\frac{1}{1+X^{\epsilon_m}(T)-\epsilon_m F(T)}\left(\left(1+\epsilon_m Q^{\epsilon_m}(T)+\epsilon_m^2\Gamma^{\tau_{\epsilon_m}}(T)-\epsilon_m F(T)\right)-(1+X^{\epsilon_m}(T)-\epsilon_m F(T))\right)\\&-\frac{1}{2B^2}\left(X^{\epsilon_m}(T)-\epsilon_m Q^{\epsilon_m}(T)-\epsilon_m^2\Gamma^{\tau_{\epsilon_m}}(T)\right)^2.\end{aligned}$$

Now we claim that:

$$\mathbb{E}\left[\frac{1+\epsilon_m Q^{\epsilon_m}(T)+\epsilon_m^2\Gamma^{\tau_{\epsilon_m}}(T)-\epsilon_m F(T)}{1+X^{\epsilon_m}(T)-\epsilon_m F(T)}\right]\le 1.$$

This holds since $1+\epsilon_m Q^{\epsilon_m}+\epsilon_m^2\Gamma^{\tau_{\epsilon_m}}\in\mathcal{X}(1,\epsilon)$. Hence:

$$\mathbb{E}\left[\log\left(\frac{1+\epsilon_m Q^{\epsilon_m}(T)+\epsilon_m^2\Gamma^{\tau_{\epsilon_m}}(T)-\epsilon_m F(T)}{1+X^{\epsilon_m}(T)-\epsilon_m F(T)}\right)\right]\le 0.$$

Using Jensen's inequality, the claim follows. In turn we have:

$$\frac{u(\epsilon_m) - \mathbb{E}[\log(1 + \epsilon_m Q^{\epsilon_m}(T) + \epsilon_m^2 \Gamma(\tau_{\epsilon_m}) - \epsilon_m F(T))]}{\epsilon_m^4} \geq \frac{\mathbb{E}[(X^{\epsilon_m}(T) - \epsilon_m Q^{\epsilon_m}(T) - \epsilon_m^2 \Gamma(\tau_{\epsilon_m}))^2/2B^2]}{\epsilon_m^4}, \tag{B.5}$$

where by the fourth-order expansion of the value function we have that the left-hand side of (B.5) tends to zero as $m \to \infty$. Furthermore, we have:

$$\begin{aligned}
\frac{\mathbb{E}[(X^{\epsilon_m}(T) - \epsilon_m \Delta(T) - \epsilon_m^2 \Gamma(T))^2/2B^2]}{\epsilon_m^4} &\leq \frac{2\mathbb{E}[(X^{\epsilon_m}(T) - \epsilon_m Q^{\epsilon_m}(T) - \epsilon_m^2 \Gamma^{\tau_{\epsilon_m}}(T))^2/2B^2]}{\epsilon_m^4} \\
&+ \frac{2\mathbb{E}[(\epsilon_m Q^{\epsilon_m}(T) + \epsilon_m^2 \Gamma^{\tau_{\epsilon_m}}(T) - \epsilon_m \Delta(T) - \epsilon_m^2 \Gamma(T))^2]}{\epsilon_m^4} \\
&\leq \frac{2\mathbb{E}[(X^{\epsilon_m}(T) - \epsilon_m Q^{\epsilon_m}(T) - \epsilon_m^2 \Gamma^{\tau_{\epsilon_m}}(T))^2/2B^2]}{\epsilon_m^4} \\
&+ 4\mathbb{E}[(\Gamma^{\tau_{\epsilon_m}}(T) - \Gamma(T))^4] \\
&+ C\left(\frac{\mathbb{P}(\tau_{\epsilon_m} < T)}{\epsilon_m^4}\right)^{\frac{1}{2}},
\end{aligned}$$

for $C > 0$. Now the above tends to zero as $m \to \infty$ by (B.5), dominated convergence, (3.3.1) and (B.1). Lastly, Markov's inequality applied on the increasing, non-negative function $x \mapsto x^2/2B^2$ yields:

$$X^{\epsilon_m}(T) - \epsilon_m \Delta(T) - \epsilon_m^2 \Gamma(T) = o_{\mathbb{P}}(\epsilon_m^2),$$

as $m \to \infty$. In fact note that if the above holds for the subsequence ϵ_m, then it also holds for the original (arbitrary) sequence ϵ_n by the double subsequence trick [2], giving the desired result. □

Lemma 4.2.1. Let ζ be a non-negative random variable, and assume that:

$$\mathbb{E}[\zeta^{p_1}] < \infty,$$

for some $p_1 > 0$. Then, for every $p_2 > 0$:

$$\lim_{z \to \infty} \frac{1}{z^{p_2}} \mathbb{E}[\zeta^{p_1 + p_2} \mathbf{1}_{\zeta \leq z}] = 0. \tag{B.6}$$

[2]Note that this refers to the following (general) result: let y_n be a sequence of elements of a topological space. If every subsequence $y_{n(m)}$ has a further subsequence $y_{n(m_k)}$ that converges to y, then $y_n \to y$.

Proof. Fix $p_2 > 0$ and observe that:

$$\frac{1}{z^{p_2}}\mathbb{E}[\zeta^{p_1+p_2}\mathbf{1}_{\zeta\le z}] = \mathbb{E}\left[\zeta^{p_1}\left(\frac{\zeta}{z}\right)^{p_2}\mathbf{1}_{\zeta\le z}\right].$$

Moreover:

$$0 \le \zeta^{p_1}\left(\frac{\zeta}{z}\right)^{p_2}\mathbf{1}_{\zeta\le z} \le \zeta^{p_1}, \qquad z > 0.$$

Hence, using the fact that $\mathbb{E}[\zeta^{p_1}] < \infty$ (also implying finiteness a.s.) and applying dominated convergence we get the desired result since:

$$\lim_{z\to\infty}\left(\frac{\zeta}{z}\right)^{p_2}\mathbf{1}_{\zeta\le z} = 0, \text{ a.s.}$$

□

B.2 Additional calculations for Example III.3

Denote the martingales closed by $F(\infty)$, $F_n(\infty)$ as $\breve{M}$ and $\breve{M}_n$ respectively and recall that we have the following Kunita-Watanabe decompositions, as in §3:

$$\begin{aligned}
\breve{M} &= \Delta + N,\\
\breve{Q} &:= \mathbb{E}[N(\infty)^2|\mathcal{F}(\cdot)] = \Gamma + P,\\
\breve{M}_n &= \Delta_n + N_n,\\
\breve{Q}_n &:= \mathbb{E}[N_n(\infty)^2|\mathcal{F}(\cdot)] = \Gamma_n + P_n,
\end{aligned}$$

where $\Delta(0) = \Gamma(0) = \Delta_n(0) = \Gamma_n(0) = 0$. For sufficiently big r, we get:

$$\|[\breve{M}_n - \breve{M}, \breve{M}_n - \breve{M}](\infty)\|_{\mathbb{L}_1} \to 0,$$

as $n \to \infty$.

Markov's inequality gives that the above convergence also holds in probability. Now, using $[\breve{M}_n-\breve{M}, \breve{M}_n-\breve{M}] = [\Delta_n-\Delta, \Delta_n-\Delta]+[N_n-N, N_n-N]$ implies that $[\Delta_n-\Delta, \Delta_n-\Delta](\infty) \to 0$ in probability as $n \to \infty$. In turn, we get:

$$\sup_{t\ge 0}|\Delta_n(t) - \Delta(t)| \to 0,$$

in probability as $n \to \infty$. For the case of Γ_n, Γ note that we have for any $a > 0$:

$$\mathbb{P}\left(\sup_{t\geq 0} |\breve{Q}_n(t) - \breve{Q}(t)| > a\right) \leq \frac{1}{a}\|N_n(\infty)^2 - N(\infty)^2\|_{\mathbb{L}_1},$$

which holds by Doob's maximal inequality, the fact that $|\mathbb{E}[N_n(\infty)^2 - N(\infty)^2|\mathcal{F}(\cdot)]| \leq \mathbb{E}[|N_n(\infty)^2 - N(\infty)^2||\mathcal{F}(\cdot)]$ and $\cup_{K>0}\{\omega : \sup_{0\leq t\leq K} |\breve{Q}_n(t) - \breve{Q}(t)| > a\} = \{\omega : \sup_{t\geq 0} |\breve{Q}_n(t) - \breve{Q}(t)| > a\}$. In turn, we get:

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\geq 0} |\breve{Q}_n(t) - \breve{Q}(t)| > a\right) &\leq \frac{1}{a}\left(\|N_n(\infty) - N(\infty)\|_{\mathbb{L}_2}^2 + 2\|N(\infty)(N_n(\infty) - N(\infty))\|_{\mathbb{L}_1}\right) \\
&\leq \frac{4}{a}\left(\|F_n(\infty) - F(\infty)\|_{\mathbb{L}_2}^2 + 2\|F(\infty)\|_{\mathbb{L}_2}\|F_n(\infty) - F(\infty)\|_{\mathbb{L}_2}\right) \\
&\leq \frac{4}{a}\left(\|F_n(\infty) - F(\infty)\|_{\mathbb{L}_2}^2 + 2\|F(\infty)\|_{\mathbb{L}_2}\|F_n(\infty) - F(\infty)\|_{\mathbb{L}_2}\right),
\end{aligned}$$

where we've also used that $\|F(\infty)\|_{\mathbb{L}_2}$ is bounded by $\mathbb{E}[(\int_0^\infty e^{-\widetilde{r}t} S_0(t)|Z(t)|\mathrm{d}t)^2]^{1/2} < \infty$ (for sufficiently big $\widetilde{r}$). The above converges to zero as $n \to \infty$, as for sufficiently big $\widetilde{r}$ we have $\|F_n(\infty) - F(\infty)\|_{\mathbb{L}_2} \to 0$. In fact for $\widehat{Q} := \breve{Q} - \breve{Q}(0)$, $\widehat{Q}_n := \breve{Q}_n - \breve{Q}_n(0)$ we have by the above that $\sup_{t\geq 0} |\widehat{Q}_n(t) - \widehat{Q}(t)| \to 0$ in probability. This yields $[\widehat{Q}_n - \widehat{Q}, \widehat{Q}_n - \widehat{Q}](\infty) \to 0$ in probability, implying $[\Gamma_n - \Gamma, \Gamma_n - \Gamma](\infty) \to 0$ in probability; which lastly gives:

$$\sup_{t\geq 0} |\Gamma_n(t) - \Gamma(t)| \to 0,$$

in probability as $n \to \infty$.

B.3 Example III.4 for non-unit-linked liabilities

Borrowing the setup from Example III.4 we can also generalise our results for any polynomial payoff, i.e. for $m \in \mathbb{N}_{>0}$, that is non-unit-linked:

$$\begin{aligned}
\widetilde{\Lambda}(\cdot; m) &= \int_0^{\cdot} e^{-\widetilde{r}t} Z(t)^m \mathrm{d}t, \\
F(\cdot; m) &= \int_0^{\cdot} e^{-\widetilde{r}t} S_0(t) Z(t)^m \mathrm{d}t.
\end{aligned}$$

Now, recall that under $\mathbb{Q}^\nu$ we have:

$$dZ(t) = (\eta - kZ(t))\mathrm{d}t + b\mathrm{d}W^{\mathbb{Q}^\nu}(t), \qquad \eta := k\theta - b\widetilde{\sigma}\nu\rho.$$

In turn, following similar steps as for (3.4.1), we have:

$$\mathbb{E}[F(\infty;m)|\mathcal{F}(t)] = \int_0^t e^{-\widetilde{r}s} S_0(s)Z(s)^m \mathrm{d}s + e^{-\widetilde{r}t} S_0(t)\psi_m(Z(t)),$$

where ψ_m solves:

$$(\widetilde{r} - \mathcal{L}^{\mathbb{Q}^\nu})\psi_m(z) = z^m, \qquad \mathcal{L}^{\mathbb{Q}^\nu} f(z) := (\eta - kz)\partial_z f(z) + \frac{b^2}{2}\partial_{zz} f(z).$$

In fact, by Feynman-Kac (under the implied transversality condition) we have:

$$\psi_m(z) = \mathbb{E}_z^{\mathbb{Q}^\nu}\left[\int_0^\infty e^{-\widetilde{r}t} Z(t)^m \mathrm{d}t\right].$$

Furthermore, note that:

$$(\widetilde{r} - \mathcal{L}^{\mathbb{Q}^\nu})z^m = (\widetilde{r} + mk)z^m - m\eta z^{m-1} - \frac{b^2}{2}m(m-1)z^{m-2}.$$

Using $(\widetilde{r} - \mathcal{L}^{\mathbb{Q}^\nu})\psi_{m-1}(z) = z^{m-1}$, $(\widetilde{r} - \mathcal{L}^{\mathbb{Q}^\nu})\psi_{m-2}(z) = z^{m-2}$ and rearranging the above, we get:

$$(\widetilde{r} - \mathcal{L}^{\mathbb{Q}^\nu})\left(\frac{z^m}{\widetilde{r} + mk} + \frac{m\eta}{\widetilde{r} + mk}\psi_{m-1}(z) + \frac{b^2 m(m-1)}{2(\widetilde{r} + mk)}\psi_{m-2}(z)\right) = z^m,$$

which implies that ψ_m satisfies the following recursive relationship:

$$\begin{aligned}
&\psi_0(z) = \frac{1}{\widetilde{r}}; \qquad \psi_1(z) = \frac{z + \eta/\widetilde{r}}{\widetilde{r} + k},\\
&\psi_m(z) = \frac{z^m}{\widetilde{r} + mk} + \frac{m\eta}{\widetilde{r} + mk}\psi_{m-1}(z) + \frac{b^2 m(m-1)}{2(\widetilde{r} + mk)}\psi_{m-2}(z), \qquad m \geq 2.
\end{aligned}$$

In particular, denoting the martingale closed by $F(\infty;m)$ as $\breve{M}_m$, we have:

$$\begin{aligned}
\mathrm{d}\breve{M}_m(t) = e^{-\widetilde{r}t} S_0(t)\Big(b\rho\partial_z\psi_m(Z(t)) - \nu\widetilde{\sigma}\psi_m(Z(t))\Big)\mathrm{d}W^1(t) +&\\
e^{-\widetilde{r}t} S_0(t) b\sqrt{1-\rho^2}\partial_z\psi_m(Z(t))\mathrm{d}W^2(t)&.
\end{aligned}$$

So, by the Kunita-Watanabe decomposition, we get:

$$\begin{aligned}
\Delta_m &= \int_0^\cdot e^{-\widetilde{r}t} S_0(t)\Big(b\rho\partial_z\psi_m(Z(t)) - \nu\widetilde{\sigma}\psi_m(Z(t))\Big)\mathrm{d}W^1(t),\\
\mathrm{d}N_m(t) &= e^{-\widetilde{r}t} S_0(t) b\sqrt{1-\rho^2}\partial_z\psi_m(Z(t))\mathrm{d}W^2(t),
\end{aligned}$$

where Δ_m, N_m are the respective forms of Δ, N from Example III.3. In fact, for $m = 1$ this reproduces θ^{Δ}, as $\psi_1(z) = (z + \eta/\widetilde{r})/(\widetilde{r} + k)$ and $\partial_z \psi_1(z) = 1/(\widetilde{r} + k)$.

Now, denote the measure induced by $\mathcal{E}(2R_0)$ and the martingale closed by $[N_m, N_m](\infty)$ as $\mathbb{Q}^{2\nu}, \breve{K}_m$ respectively. Then:

$$dZ(t) = (\breve{\eta} - kZ(t))\mathrm{d}t + b\mathrm{d}W^{\mathbb{Q}^{2\nu}}(t), \qquad \breve{\eta} := k\theta - 2b\widetilde{\sigma}\nu\rho,$$

and:

$$\breve{K}_m(t) = \int_0^t b^2(1-\rho^2)e^{-2\widetilde{r}s}S_0(s)^2 \partial_z\psi_m(Z(s))^2 \mathrm{d}s + e^{-\breve{r}t}\mathcal{E}(2R_0)(t)\chi_m(Z(t)),$$

where for $\breve{r} := 2\widetilde{r} - (\nu\widetilde{\sigma})^2 > 0$, χ_m solves:

$$(\breve{r} - \mathcal{L}^{\mathbb{Q}^{2\nu}})\chi_m(z) = b^2(1-\rho^2)\partial_z\psi_m(z)^2, \qquad \mathcal{L}^{\mathbb{Q}^{2\nu}} f(z) := (\breve{\eta} - kz)\partial_z f(z) + \frac{b^2}{2}\partial_{zz} f(z).$$

In fact, by Feynman-Kac (under the implied transversality condition) we have:

$$\chi_m(z) = \mathbb{E}_z^{\mathbb{Q}^{2\nu}}\left[\int_0^\infty e^{-\breve{r}t} b^2(1-\rho^2)\partial_z\psi_m(Z(t))^2 \mathrm{d}t\right].$$

We need to pin down $\partial_z\psi_m^2$. Differentiating the recursion for ψ_m, we get:

$$\partial_z\psi_m(z) = \frac{m}{\widetilde{r} + mk}z^{m-1} + \frac{m\eta}{\widetilde{r} + mk}\partial_z\psi_{m-1}(z) + \frac{b^2 m(m-1)}{2(\widetilde{r} + mk)}\partial_z\psi_{m-2}(z), \qquad m \geq 2,$$

where $\partial_z\psi_0(z) = 0$ and $\partial_z\psi_1(z) = 1/(\widetilde{r} + k)$. In particular, $\partial_z\psi_m$ is of the form: $\partial_z\psi_m(z) = \sum_{j=0}^{m-1} d_{m,j}z^j$. Hence:

$$\sum_{j=0}^{m-1} d_{m,j}z^j = \frac{m}{\widetilde{r} + mk}z^{m-1} + \frac{m\eta}{\widetilde{r} + mk}\sum_{j=0}^{m-2} d_{m-1,j}z^j + \frac{b^2 m(m-1)}{2(\widetilde{r} + mk)}\sum_{j=0}^{m-3} d_{m-2,j}z^j.$$

By the above, and after setting $d_{m,j} := 0$, whenever $j > m-1$ implies the following recursion for $m \geq 2$:

$$d_{m,j} = \begin{cases} \frac{m}{\widetilde{r}+mk}, & j = m-1, \\ \frac{m\eta}{\widetilde{r}+mk}d_{m-1,j}, & j = m-2, \\ \frac{m\eta}{\widetilde{r}+mk}d_{m-1,j} + \frac{b^2 m(m-1)}{2(\widetilde{r}+mk)}d_{m-2,j}, & j = 0, ..., m-3, \end{cases}$$

and $d_{1,0} = 1/(\widetilde{r}+k)$. In turn, $\partial_z \psi_m^2$ is of the form: $\partial_z \psi_m(z)^2 = \sum_{j=0}^{2m-2} q_{m,j} z^j$, where $q_{m,j} = \sum_{k=\max(0,j-(m-1))}^{\min(j,m-1)} d_{m,k} d_{m,j-k}$. Lastly, in order to obtain χ_m, consider:

$$(\breve{r} - \mathcal{L}^{\mathbb{Q}^{2\nu}})\phi_m(z) = z^m.$$

Following a similar process as ψ_m, we get that:

$$\begin{aligned}
\phi_m(z) &= \mathbb{E}_z^{\mathbb{Q}^{2\nu}}\left[\int_0^\infty e^{-\breve{r}t} Z(t)^m \mathrm{d}t\right],\\
\phi_0(z) &= \frac{1}{\breve{r}}; \qquad \phi_1(z) = \frac{z+\breve{\eta}/\breve{r}}{\breve{r}+k},\\
\phi_m(z) &= \frac{z^m}{\breve{r}+mk} + \frac{m\breve{\eta}}{\breve{r}+mk}\phi_{m-1}(z) + \frac{b^2 m(m-1)}{2(\breve{r}+mk)}\phi_{m-2}(z).
\end{aligned}$$

Lastly, by linearity we get:

$$\chi_m(z) = b^2(1-\rho^2)\sum_{j=0}^{2m-2} q_{m,j}\phi_j(z).$$

C Supplementary material for Chapter IV

C.1 A primer on arbitrage and market completeness

In an effort to make Chapter IV as self-contained as possible, we introduce several key notions related to arbitrage and market completeness. These concepts, in turn, provide the foundation for the utility-based pricing framework developed in that chapter. This section follows the notes of Schachermayer in [**?**].

We consider a security market consisting of $1+d$ assets. We write $\widetilde{S} \equiv (\widetilde{S}_i; 0 \le i \le d)$ for their price process, and assume that the price of the asset $\widetilde{S}_0$, referred to as the "bond", "cash account" or "riskless asset", is constant, i.e., $\widetilde{S}_0 \equiv 1$. This assumption is without loss of generality, and is mainly done for notational convenience, since we may always choose the bond as numeraire and thus express the values of the remaining assets w.r.t. it.

The process $\widetilde{S}$ is assumed to be a semimartingale, based on and adapted to a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathcal{F}(\cdot)$ satisfying the usual conditions, s.t. $\mathcal{F}(0)$ is trivial a.s. As usual in mathematical finance, we consider a finite horizon T, but we remark that our results can also be generalised for the case of an infinite horizon.

For the remainder of this section—in order to fix ideas—we shall be mainly working with the case of finite Ω and discrete time, i.e. $\Omega = \{\omega_1, \omega_2, ..., \omega_N\}$; $\mathcal{F}(T) = \mathcal{F}$ is the power set of Ω; $\mathbb{P}(\omega_1), \mathbb{P}(\omega_2), ..., \mathbb{P}(\omega_N) > 0$; and $(\widetilde{S}(t); 0 \leq t \leq T) = (\widetilde{S}(0), \widetilde{S}(1), ..., \widetilde{S}(T))$, for some $N, T \in \mathbb{N}_{>0}$. This assumption implies that all the differences among the spaces $\mathbb{L}_\infty$, $\mathbb{L}_1$, and $\mathbb{L}_0$ disappear, as they are simply isomorphic to $\mathbb{R}^N$. Hence all the functional analysis reduces to simple linear algebra in the setting of the present section.

Passing momentarily to a general semimartingale market, a "portfolio" or "strategy" is a $(1 + d)$-dimensional, predictable process $(\theta_i; 0 \leq i \leq d)$ that is $\widetilde{S}$-integrable . We think of θ_i as the number of assets of type i in the portfolio at each time. The requirement that θ is predictable essentially means that the portfolio is constructed using only the information available up to that time. The "value process", V, associated with a portfolio θ is given by:

$$V(\cdot) = \sum_{i=0}^{d} \theta_i(\cdot)\widetilde{S}_i(\cdot).$$

In order to reconcile the above with the discrete time, finite sample space setting, we focus on portfolios that satisfy two key requirements: (i) the self-financing condition, and (ii) the condition of admissibility. The former essentially entails portfolios that do not involve injections or withdrawals of money after the initial time. In particular, for such strategies, we have:

$$V(\cdot) = V(0) + \int_0^{\cdot} \theta(t) d\widetilde{S}(t),$$

where the above equals $V(0) + \sum_{t=1}^{\cdot} \theta(t)\Delta\widetilde{S}(t)$ in the discrete time, finite Ω setting [3]. To see this, consider such a portfolio with value $V(t-1)$. Just after that time the portfolio is rebalanced, and the new value equals:

$$\sum_{i=0}^{d} \theta_i(t)\widetilde{S}_i(t-1) = V(t) - \sum_{i=0}^{d} \theta_i(t)\left(\widetilde{S}_i(t) - \widetilde{S}_i(t-1)\right).$$

[3] Here $\theta(s)\Delta\widetilde{S}(s)$ denotes the inner product of the vectors $\theta(s)$ and $\Delta\widetilde{S}(s) := \widetilde{S}(s) - \widetilde{S}(s-1)$.

If no money is injected or withdrawn this should equal $V(t-1)$, hence:

$$V(t) - V(t-1) = \sum_{i=1}^{d} \theta_i(t) \left(\widetilde{S}_i(t) - \widetilde{S}_i(t-1) \right), \qquad t \geq 1,$$

using that $\widetilde{S}_0 \equiv 1$. Note that the self-financing condition does not come with any severe loss of generality, since one could create a self-financing portfolio by investing appropriately on the riskless asset.

Regarding the admissibility condition for a general semimartingale model, in order to rule out doubling strategies and similar schemes generating arbitrage profits (by going deeply into the red), we assume that $\exists C \in \mathbb{R}_{\geq 0}$ s.t.:

$$\int_0^{\cdot} \theta(t) d\widetilde{S}(t) \geq -C.$$

Of course, the above holds for all self-financing strategies θ if the underlying probability space Ω is finite.

Returning to our simplified model, denoting the set of self-financing strategies by Θ, we introduce the following definition:

Definition 4.3.1. We call the subspace $\mathcal{K}$ of $\mathbb{L}_0$, defined by:

$$\mathcal{R} = \left\{ \int_0^T \theta(t) d\widetilde{S}(t) : \theta \in \Theta \right\},$$

the set of contingent claims attainable at price 0.

In words, the random variables $f = \int_0^T \theta(t) d\widetilde{S}(t)$, for some $\theta \in \Theta$, represent exactly the contingent claims that can be generated from zero initial capital by following a self-financing strategy θ. Equivalently, they are the $\mathcal{F}_T$-measurable payoffs at time T that an agent can replicate without any initial investment. For $x \in \mathbb{R}$, one can also define the collection of contingent claims attainable at price x as the affine space $\mathcal{R}_x$, obtained by translating $\mathcal{R}$ by the constant function $x\mathbf{1}$.

Definition 4.3.2. A financial market $\widetilde{S}$ satisfies the no-arbitrage condition (NA) if:

$$\mathcal{R} \cap \mathbb{L}_0^{\geq 0} = \{0\},$$

where $\mathbb{L}_0^{\geq 0}$ denotes the space of non-negative measurable functions and 0 denotes the function identically equal to zero.

In this way, the notion of an arbitrage opportunity is made precise: it means that there exists a trading strategy θ such that, starting from zero initial capital, the terminal payoff $\int_0^T \theta(t) d\widetilde{S}(t)$ is almost surely non-negative and not identically zero.

Definition 4.3.3. A probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ is called an equivalent martingale measure for $\widetilde{S}$, if $\mathbb{Q} \sim \mathbb{P}$ and $\widetilde{S}$ is a martingale under $\mathbb{Q}$.

The set of equivalent martingale measures is denoted by $\mathcal{M}^e(\widetilde{S})$. In fact, for financial market $\widetilde{S}$ modelled on a finite stochastic base, the following are equivalent:

1. $\widetilde{S}$ satisfies (NA),
2. $\mathcal{M}^e(\widetilde{S}) \neq \varnothing$.

The above result is usually called the Fundamental Theorem of Asset Pricing (cf. [**?** , Theorem 2.8]). Now, for a market that satisfies (NA) a contingent claim $f \in \mathbb{L}_\infty$ is called "attainable" or "replicable" if:

$$f = x + \int_0^T \theta(t) d\widetilde{S}(t),$$

for some $x \in \mathbb{R}$, $\theta \in \Theta$. In fact, for such a market, the constant x and the process $\int_0^\cdot \theta(t) d\widetilde{S}(t)$ are uniquely determined by the above, and satisfy for every $\mathbb{Q} \in \mathcal{M}^e(\widetilde{S})$ (cf. [**?** , Corollary 2.9]):

$$x = \mathbb{E}^{\mathbb{Q}}[f]; \qquad x + \int_0^\cdot \theta(t) d\widetilde{S}(t) = \mathbb{E}^{\mathbb{Q}}[f|\mathcal{F}(\cdot)].$$

In particular, one can formalise pricing via no-arbitrage by using the following result (cf. [**?** , Theorem 2.11]):

Assume that $\widetilde{S}$ satisfies (NA) and let $f \in \mathbb{L}_\infty$. Define:

$$\underline{\pi}(f) = \inf\left\{\mathbb{E}^{\mathbb{Q}}[f] : \mathbb{Q} \in \mathcal{M}^e(\widetilde{S})\right\},$$
$$\overline{\pi}(f) = \sup\left\{\mathbb{E}^{\mathbb{Q}}[f] : \mathbb{Q} \in \mathcal{M}^e(\widetilde{S})\right\}.$$

Either $\underline{\pi}(f) = \overline{\pi}(f)$, in which case f is replicable at price $\pi(f) := \underline{\pi}(f) = \overline{\pi}(f)$—called its unique arbitrage-free price,

or $\underline{\pi}(f) < \overline{\pi}(f)$, in which case $\left\{\mathbb{E}^{\mathbb{Q}}[f] : \mathbb{Q} \in \mathcal{M}^e(\widetilde{S})\right\}$ is open. In turn, $(\underline{\pi}(f), \overline{\pi}(f))$ equals the set of arbitrage-free prices for the contingent claim f.

A financial market where each $f \in \mathbb{L}_\infty$ is replicable is called "complete". This is formalised by the following result (cf. [**?** , Corollary 2.12]):

For a financial market $\widetilde{S}$ satisfying (NA), the following are equivalent:

1. $\mathcal{M}^e(\widetilde{S})$ consists of a single element $\mathbb{Q}$.

2. Each $f \in \mathbb{L}_\infty$ may be represented as:

$$f = x + \int_0^T \theta(t) d\widetilde{S}(t),$$

   for some $x \in \mathbb{R}$, $\theta \in \Theta$. In this case, $x = \mathbb{E}^{\mathbb{Q}}[f]$, the stochastic integral $\int_0^{\cdot} \theta(t) d\widetilde{S}(t)$ is unique, and we have:

$$\mathbb{E}^{\mathbb{Q}}[f|\mathcal{F}(\cdot)] = x + \int_0^{\cdot} \theta(t) d\widetilde{S}(t).$$

C.2 Proof(s)

Proof of Theorem 4.3.1. Using (3.3.2) write [4]:

$$u(\epsilon) = u_1\epsilon + u_2\epsilon^2 + u_3\epsilon^3 + u_4\epsilon^4 + o(\epsilon^4),$$

[4]In §2 we've shown that $u(\epsilon) < \infty$. In fact, recalling (3.2.7) it also holds that $u(\epsilon) > -\infty$, by potentially restricting ϵ to a sub-interval of $(0, \epsilon^L)$ s.t. $\epsilon x < 1$. One way to see this is to take the positive wealth process in S: $X^x := (1 - \epsilon x) + \epsilon X$ with $X^x(0) = 1$, where X is the process implied by (3.2.7), starting at x. In particular, at maturity $X^x(T) - \epsilon F(T)$ is bounded below by $1 - \epsilon x > 0$.

where:

$$u_1 := -A_1, \;\; u_2 := -\frac{A_2}{2}, \;\; u_3 := -\frac{A_3}{3}, \;\; u_4 := -\left(\frac{A_4}{4} + G\right).$$

Using (4.3.4) and expanding:

$$c(\epsilon) = u(\epsilon) + \frac{u(\epsilon)^2}{2} + \frac{u(\epsilon)^3}{6} + \frac{u(\epsilon)^4}{24} + o(\epsilon^4).$$

Collecting terms, for order ϵ we only have a contribution from u_1; for order ϵ^2 we have contributions from u_2, and $u_1^2/2$; for order ϵ^3 we have contributions from u_3, u_1u_2, and $u_1^3/6$; finally for order ϵ^4 we have contributions from u_4, u_1u_3, $u_2^2/2$, $u_1^2u_2/2$, and $u_1^4/24$. Bringing all these together, we get the desired result. $\square$

Bibliography

[AB20] M. Anthropelos and T.J. Boonen, *Nash Equilibria in Optimal Reinsurance Bargaining*, Insurance: Mathematics and Economics **93** (2020), 196–205.

[AB23] M. Anthropelos and E. Blontzou, *On Valuation and Investments of Pension Plans in Discrete Incomplete Markets*, Risks **11** (2023), no. 6, 103, In the Special Issue: Frontiers in Quantitative Finance and Risk Management.

[AC00] R. Almgren and N. Chriss, *Optimal execution of portfolio transactions*, Journal of Risk **3** (2000), 5–39.

[AK17] M. Anthropelos and C. Kardaras, *Equilibrium in risk-sharing games*, Finance and Stochastics (2017), no. 21 (3), 815–865.

[AK24] ———, *Price impact under heterogeneous beliefs and restricted participation*, Journal of Economic Theory **215** (2024), 105774.

[AKV20] M. Anthropelos, C. Kardaras, and G. Vichos, *The Effective Risk Aversion in Thin Risk-Sharing Markets*, Mathematical Finance **30** (2020), 1565–1590.

[Ant17] M. Anthropelos, *The effect of market power on risk-sharing*, Mathematics and Financial Economics (2017), no. 11, 323–368.

[AR] M. Anthropelos and S. Robertson, *Strategic Informed Trading and the Value of Private Information*, Mathematics and Financial Economics, To appear.

[ARS21] M. Anthropelos, S. Robertson, and K. Spiliopoulos, *Optimal Investment, Derivative Demand and Arbitrage under Price Impact*, Mathematical Finance **31** (2021), 3–35.

[AS24] M. Anthropelos and P. Schneider, *Optimal Investment and Equilibrium Pricing under Ambiguity*, Review of Finance **28** (2024), no. 6, 1759–1805.

[ATHL05] R. Almgren, C. Thum, E. Hauptmann, and H. Li, *Equity market impact*, Risk (2005), 57–62.

[AŽ10a] M. Anthropelos and G. Žitković, *On Agents' Agreement and Partial Equilibrium Pricing in Incomplete Markets*, Mathematical Finance **20** (2010), no. 3, 411–446.

[AŽ10b] ______, *Partial Equilibria with Convex Capital Requirements: Existence, Uniqueness and Stability*, Annals of Finance **6** (2010), 107–135.

[Bas00] S. Basak, *A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk*, Journal of Economic Dynamics and Control **24** (2000), 63–95.

[BD19] A. Buss and B. Dumas, *The dynamic properties of financial-market equilibrium with trading fees*, The Journal of Finance **74** (2019), no. 2, 795–844.

[BFHMK18] B. Bouchard, M. Fukasawa, M. Herdegen, and J. Muhle-Karbe, *Equilibrium returns with transaction costs*, Finance and Stochastics (2018), no. 22 (3), 569–601.

[BSM17] P. Bank, M. Soner, and VoßM., *Hedging with termporary price impact*, Math. Fin. Econ. (2017), no. 11(2), 215–239.

[CC98] D. Cuoco and J. Cvitanic, *Optimal consumption choices for a 'large' investor*, Journal of Economic Dynamics and Control **22** (1998), no. 3, 401–436.

[CCFM17] H Chau, A. Cosso, C Fontana, and O. Mostovyi, *Optimal investment with intermediate consumption under no unbounded profit with bounded risk*, Journal of Applied Probability (2017), no. 54, 710–719.

[CDJ20] Á. Cartea, R. Donelly, and S. Jaimungal, *Hedging non-tradable risks with transaction costs and price impact*, Mathematical Finance (2020), no. 30 (3), 833–868.

[CH89] J.C. Cox and C.F. Huang, *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, Journal of Mathematical Economics **49** (1989), 33–83.

[CH91] ______, *A variational problem arising in financial economics*, Journal of Mathematical Economics **20** (1991), 465–487.

[CL24] F. Cordoni and F. Lillo, *Instabilities in multi-asset and multi-agent market impact games*, Annals of Operations Research **336** (2024), 505–539.

[CLS21] J.H. Choi, K. Larsen, and D.J. Seppi, *Equilibrium effects of intraday order-splitting benchmarks*, Mathematics and Financial Economics (2021), no. 15, 315–352.

[CSW01] J. Cvitaniĉ, W. Schachermayer, and H. Wang, *Utility maximization in incomplete markets with random endowment*, Finance and Stochastics **5** (2001), 259–272.

[Dav06] M.H.A. Davis, *Optimal hedging with basis risk*, pp. 169–187, Springer Berlin Heidelberg, 2006.

[DGP05] D. Duffie, N. Gârleanu, and L.H. Pedersen, *Over-the-counter markets*, Econometrica **73** (2005), no. 6, 1815–1847.

[DS97] F. Delbaen and W. Schachermayer, *The banach space of workable contingent claims in arbitrage theory*, Annales de l'Institut Henri Poincare (B) Probability and Statistics **33** (1997), 113–144.

[ET99] I. Ekeland and R. Temam, *Convex analysis and variational problems*, SIAM, Philadelphia, 1999.

[FIM18] A. Frazzini, R. Israel, and T. Moskowitz, *Transaction costs*, Working paper, available at SSRN: https://ssrn.com/abstract=3229719, 2018.

[FS08] C. Frei and M. Schweizer, *Exponential utility indifference valuation in two brownian settings with stochastic correlation*, Advances in Applied Probability **40** (2008), 401–423.

[FS10] ______, *Exponential utility indifference valuation in a general semimartingale model*, pp. 49–86, Springer Berlin Heidelberg, 2010.

[GH07] M. Grasselli and T.R. Hurd, *Indifference pricing and hedging for volatility derivatives*, Applied Mathematical Finance **14** (2007), 303–317.

[GP16] N. Gârleanu and L. Pedersen, *Dynamic portfolio choice with frictions*, J. Economic Theory (2016), no. 165, 487–516.

[Hen02] V. Henderson, *Valuation of claims on nontraded assets using utility maximization*, Mathematical Finance **12** (2002), 351–373.

[HH02] V. Henderson and D.G. Hobson, *Real options with constant relative risk aversion*, Journal of Economic Dynamics & Control **27** (2002), 329–355.

[HH07] V. Henderson and D. Hobson, *Utility indifference pricing: an overview*, Princeton University Press, 2007.

[HHLT21] H. Hau, P. Hoffmann, S. Langfield, and Y. Timmer, *Discriminatory pricing of over-the-counter derivatives*, Management Science **67** (2021), no. 11, 6660–6677.

[Hig08] N. J. Higham, *Functions of matrices*, SIAM, Philadelphia, PA, 2008.

[HK79] J.M. Harrison and D. Kreps, *Martingales and arbitrage in multiperiod security markets*, Journal of Economic Theory **20** (1979), 381–408.

[HK04] J. Hugonnier and D. Kramkov, *Optimal investment with random endowments in incomplete markets*, The Annals of Applied Probability **14** (2004), 845–864.

[HL16] V. Henderson and G. Liang, *A multidimensional exponential utility indifference pricing model with applications to counterparty risk*, SIAM Journal on Control and Optimization **54** (2016), 690–717.

[HMKP21] M. Herdegen, J. Muhle-Karbe, and D. Possamaï, *Equilibrium asset pricing with transaction costs*, Finance and Stochastics (2021), no. 25, 1–45.

[HP81] J.M. Harrison and S.R. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic Processes and their Applications **11** (1981), 215–260.

[HP91a] H. He and N.D. Pearson, *Consumption and portfolio policies with incomplete markets and short-sale constraints: The finite-dimensional case*, Mathematical Finance **1** (1991), 1–10.

[HP91b] ______, *Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite-dimensional case*, Journal of Economic Theory **54** (1991), 259–304.

[HW04] G. Huberman and S. Werner, *Price manipulation and quasi-arbitrage*, Econometrica **72** (2004), no. 4, 1247–1275.

[Kal02] J. Kallsen, *Derivative pricing based on local utility maximization*, Finance and Stochastics **6** (2002), 115–140.

[KK21] I. Karatzas and C. Kardaras, *Portfolio theory and arbitrage: A course in mathematical finance*, Graduate Studies in Mathematics, American Mathematical Society, 2021.

[KLS87] I. Karatzas, J.P. Lehoczky, and S.E. Shreve, *Optimal portfolio and consumption decisions for a "small investor" on a finite horizon*, SIAM Journal of Control and Optimization **25** (1987), 1557–1586.

[KLSX91] I. Karatzas, J.P. Lehoczky, S.E. Shreve, and G.L. Xu, *Martingale and duality methods for utility maximization in an incomplete market*, SIAM Journal of Control and Optimization **29** (1991), 702–730.

[KMKV14] J. Kallsen, J. Muhle-Karbe, and R. Vierthauer, *Asymptotic power utility-based pricing and hedging*, Mathematics and Financial Economics **8** (2014), 1–28.

[KR11] J. Kallsen and T. Rheinländer, *Asymptotic utility-based pricing and hedging for exponential utility*, Statistics & Decisions **28** (2011), 17–36.

[KS99] D. Kramkov and W. Schachermayer, *The asymptotic elasticity of utility functions and optimal investment in incomplete markets*, The Annals of Applied Probability **9** (1999), 904–950.

[KS03] ______, *Necessary and sufficient conditions in the problem of optimal investment in incomplete markets*, The Annals of Applied Probability **13** (2003), 1504–1516.

[KS06] D. Kramkov and M. Sîrbu, *Sensitivity analysis of utility-based prices and risk-tolerance wealth processes*, Annals of Applied Probability **16** (2006), 2140–2194.

[KS07] ______, *Asymptotic analysis of utility-based hedging strategies for small number of contingent claims*, Stochastic Processes and their Applications **117** (2007), 1606–1620.

[KY19] R. S. J. Koijen and M. Yogo, *A demand system approach to asset pricing*, Journal of Political Economy **127** (2019), no. 4, 1475–1515.

[Kyl85] A.S. Kyle, *Continuous auctions and insider trading*, Econometrica **53** (1985), no. 6, 1315–1335.

[KŽ03] I. Karatzas and G. Žitković, *Optimal consumption from investment and random endowment in incomplete semimartingale markets*, The Annals of Applied Probability **31** (2003), 1821–1858.

[KZ05] A. Kurdila and M. Zabarankin, *Convex functional analysis*, Birkhauser Verlag, 2005.

[LL12] T. Leung and M. Ludkovski, *Accounting for risk aversion in derivatives purchase timing*, Mathematics and Financial Economics **6** (2012), 363–386.

[LS19] X. Luo and A. Schied, *Nash equilibrium for risk-averse investors in a market impact game with transient price impact*, Market Microstructure and Liquidity **05** (2019), 2050001.

[Mer69] R.C. Merton, *Lifetime portfolio selection under uncertainty: The continuous-time case*, Rev. Econom. Statist. **51** (1969), 247–257.

[Mer71] ______, *Optimum consumption and portfolio rules in a continuous-time model*, Journal of Economic Theory **3** (1971), 373–413.

[Mon04] M. Monoyios, *Performance of utility-based strategies for hedging basis risk*, Quantitative Finance **4** (2004), 245–255.

[Mon07] ______, *Optimal hedging and parameter uncertainty*, IMA Journal of Management Mathematics **18** (2007), 331–351.

[Mon10] ______, *Utility-based valuation and hedging of basis risk with partial information*, Applied Mathematical Finance **17** (2010), 519–551.

[Mon13] ______, *Malliavin calculus method for asymptotic expansion of dual control problems*, SIAM Journal of Financial Mathematics **4** (2013), 884–915.

[Mos15] O Mostovyi, *Necessary and sufficient conditions in the problem of optimal investment with intermediate consumption*, Finance and Stochastics (2015), no. 19, 135–159.

[Mos17] ______, *Optimal investment with intermediate consumption and random endowment*, Mathematical Finance (2017), no. 27, 96–114.

[Mos20] ______, *Asymptotic analysis of the expected utility maximization problem with respect to perturbations of the numeraire*, Stochastic Processes and their Applications (2020), no. 130, 4444–4469.

[MR17] S. Malamud and M. Rostek, *Decentralized exchange*, American Economic Review **107** (2017), no. 11, 3320–62.

[MS19] O Mostovyi and M. Sirbu, *Sensitivity analysis of the utility maximization problem with respect to model perturbations*, Finance and Stochastics (2019), no. 23, 595–640.

[MS24] ______, *Quadratic expansions in optimal investment with respect to perturbations of the semimartingale model*, Finance and Stochastics (2024), no. 28, 553–613.

[MT16] C. Ménassé and P. Tankov, *Asymptotic indifference pricing in exponential lévy models*, Applied Mathematical Finance **23** (2016), 197–235.

[MY07] J. Ma and J. Yong, *Forward-backward stochastic differential equations and their applications*, Springer Verlag, 2007.

[MZ04] M. Musiela and T. Zariphopoulou, *An example of indifference prices under exponential preferences*, Finance and Stochastics **8** (2004), 229–239.

[NW22] E. Noh and K. Weston, *Price impact equilibrium with transaction costs and twap trading*, Mathematics and Financial Economics (2022), no. 16, 187–204.

[OŽ09] M. Owen and G. Žitković, *Optimal investment with an unbounded random endowment and utility-based pricing*, Mathematical Finance **19** (2009), 129–159.

[Pli86] S.R. Pliska, *A stochastic calculus model of continuous trading: Optimal portfolio*, Mathematics of Operational Research **11** (1986), 371–382.

[Ros76] S.A. Ross, *The arbitrage theory of capital asset pricing*, Journal of Economic Theory **13** (1976), 341–360.

[RW15] M. Rostek and M. Weretka, *Dynamic thin markets*, Review of Financial Studies **28** (2015), 2946–2992.

[RY22] M. Rostek and J. H. Yoon, *Equilibrium theory of financial markets: Recent developments*, Prepared for The Journal of Economic Literature, 2022.

[SZ05] R. Sircar and T. Zariphopoulou, *Bounds and asymptotic approximations for utility prices when volatility is random*, SIAM Journal on Control and Optimization **43** (2005), 1328–1353.

[SZ19] A. Schied and T. Zhang, *A market impact game under transient price impact*, Mathematics of Operations Research **44** (2019), no. 1, 102–121.

[Vay98] D. Vayanos, *Transaction costs and asset prices: A dynamic equilibrium model*, The Review of Financial Studies **11** (1998), no. 1, 1–58.

[Vay99] ______, *Strategic trading and welfare in a dynamic market*, Review of Economic Studies **66** (1999), no. 2, 219–54.

[Viv11] X. Vives, *Strategic supply function competition with private information*, Econometrica **79** (2011), 1919–1966.

[VV99] D. Vayanos and JL Vila, *Equilibrium interest rate and liquidity premium with transaction costs*, Economic Theory **13** (1999), 517–539.

[Web23] K. Webster, *Handbook of price impact modeling*, CRC Press, Boca Raton, FL, 2023.

[Wes18] K. Weston, *Existence of a radner equilibrium in a model with transaction costs*, Mathematics and Financial Economics **12** (2018), 517–539.

[Zar01] T. Zariphopoulou, *A solution approach to valuation with unhedgeable risks*, Finance and Stochastics **5** (2001), 61–82.