Finance Stochast. 8, ♣–♣ (2004)
DOI: 10.1007/s00780-003-0123-2

Finance and Stochastics

Liquidity risk and arbitrage pricing theory

Umut Çetin[1], **Robert A. Jarrow**[2], **Philip Protter**[3]

[1] Technische Universität Wien, Institut für Finanz- und Versicherungsmathematik, Wiedner Hauptstr. 8–10, 1040 Vienna, Austria
[2] Johnson Graduate School of Management, Cornell University, Ithaca, NY, 14853, USA
[3] School of Operations Research, Cornell University, Ithaca, NY, 14853-3801, USA

Abstract. Classical theories of financial markets assume an infinitely liquid market and that all traders act as price takers. This theory is a good approximation for highly liquid stocks, although even there it does not apply well for large traders or for modelling transaction costs. We extend the classical approach by formulating a new model that takes into account illiquidities. Our approach hypothesizes a stochastic supply curve for a security's price as a function of trade size. This leads to a new definition of a self-financing trading strategy, additional restrictions on hedging strategies, and some interesting mathematical issues.

Key words: ♣ please insert 3–4 keywords ♣

JEL Classification: ♣ please insert 2–3 classifications ♣

Mathematics Subject Classification (1991): ♣ please insert 2–3 classifications ♣

1 Introduction

Classical arbitrage pricing theory is formulated under two primary assumptions: frictionless and competitive markets. A frictionless market is one that has no transaction costs (including taxes) and no restrictions on trade (e.g. short sale constraints). A competitive market is one where any trader can buy or sell unlimited quantities of

This work was performed while Dr. Çetin was at the Center for Applied Mathematics, Cornell University
Supported in part by NSF grant DMS-0202958 and NSA grant MDA-904-03-1-0092
The authors wish to thank M. Warachka and Kiseop Lee for helpful comments, as well as the anonymous referee and Associate Editor for numerous helpful suggestions, which have made this a much improved paper.

Manuscript received: September 2002; final version received: November 2003

the relevant security without changing the security's price. There is a large literature studying arbitrage pricing theory with transaction costs or trading restrictions (see Barles and Soner [3]; Constantinides and Zariphopoulou [6]; Cvitanic and Karatzas [7]; Cvitanic et al. [9]; Jouini [23]; Jouini and Kallal [24]; Jouini et al. [25]; Soner et al. [29]). In general, the standard theory applies but instead of a single arbitrage-free price for a security there exists an interval of arbitrage-free prices. Hedging strategies are modified, but the basic intuition remains unchanged. In contrast, the literature studying the relaxation of the competitive market hypothesis is much less extensive (see Back [1]; Cvitanic and Ma [8]; Jarrow [19],[20]; Bank and Baum [2]). When the competitive market assumption is relaxed, however, the standard theory can completely change. Market manipulation may become an issue and option pricing becomes trader and market structure dependent.

The relaxation of the frictionless and competitive market hypotheses introduces the notion of liquidity risk. Roughly speaking, liquidity risk is the additional risk due to the timing and size of a trade. From a financial engineering perspective, the need is paramount for a simple yet robust method that incorporates liquidity risk into arbitrage pricing theory. The market microstructure literature (see Kyle [26]; Glosten and Milgrom [15]; Grossman and Miller [16]), although conceptually useful, is lacking in this regard. As a first solution to this problem, liquidity risk has recently been incorporated into arbitrage pricing theory as a convenience yield (see Jarrow and Turnbull [22]; Jarrow [21]). Convenience yields have a long history in the context of commodity pricing. This solution to the problem successfully captures that component of liquidity risk due to inventory considerations. And, more importantly, it retains the price taking condition so that classical arbitrage pricing theory can still be applied. Nonetheless, this convenience yield approach to the inclusion of liquidity risk has an important omission. It doesn't explicitly capture the impact of different trade sizes on the price. Consequently, there is no notion of a bid/ask spread for the traded securities in this model structure. This is a significant omission because all markets experience price inelasticities (quantity impacts) and bid/ask spreads. A simple yet robust method for including a quantity impact on the traded security prices in arbitrage pricing theory is currently unknown.

The purpose of this paper is to develop a model for the inclusion of liquidity risk into arbitrage pricing theory that incorporates the impact of differing trade sizes on the price. As such, our approach is consistent with price inelasticities. We do this by hypothesizing the existence of a stochastic supply curve for a security's price as a function of trade size, for which traders act as price takers. This condition implies that the investor's trading strategy has no lasting impact on the price process. Following Heath et al. [18], we study conditions on the supply curve such that there are no arbitrage opportunities in the economy (appropriately defined). Given an arbitrage free evolution, we then discuss complete markets and the pricing of derivatives. This structure can also be viewed as an extension of the model in Jouini [23] where the traded securities have two distinct price processes - a selling price (the bid) and a buying price (the ask). Instead of two prices as in Jouini [23], we have a continuum of stochastic processes indexed by trade size.

This paper provides a new perspective on classical arbitrage pricing theory based on the insight that after its initiation, and before its liquidation, there is no

unique value for a portfolio. Indeed, any price on the supply curve is a plausible price to be used in valuing a portfolio. Among these, at least three economic meaningful values can be identified: the marked-to-market value, the accumulated cost of the portfolio, and the liquidation value (for a related discussion, see Jarrow [19]). This non-uniqueness is important because it changes the mathematics used in tracking a portfolio's performance. The existing portfolio theory literature and portfolio performance evaluation literature need to be modified accordingly. This is a subject for future research.

The first fundamental theorem of asset pricing appropriately generalized, holds in our new setting, while the second fundamental theorem fails. Herein, a martingale measure (appropriately defined) can be unique, and markets still be incomplete. However, a weakening of the second fundamental theorem holds. Markets will be approximately complete in our setting if a martingale measure is unique. Here, approximately complete is roughly defined to be as follows: given any random variable (appropriately integrable), there exists a self-financing trading strategy whose liquidation value is arbitrarily close (in an L^2 sense) to the given random variable. In an approximately complete market, we show that by using trading strategies that are continuous and of finite variation, all liquidity costs can be avoided. This occurs because trading strategies induce no path dependency in the evolution of the price process. Hence, a large trade can be subdivided into infinitely many infinitesimal trades to avoid liquidity costs. Consequently, the arbitrage-free price of any derivative is shown to be equal to the expected value of its payoff under the risk neutral measure. This is the same price as in the classical economy with no liquidity costs. However, in a world with liquidities, the classical hedge will not be the hedge used to replicate the option. Both of these observations are consistent with industry usage of the classical arbitrage free pricing methodology.

Our approach to liquidity risk can be implemented in practice and the stochastic supply curve estimated from market data that contains the price of a trade, the size of the trade, and whether it is a buy or a sell. Our approach to liquidity risk, appropriately generalized, can also be used to study transaction costs. Transaction costs effectively make the traded security's price dependent on the trade size. However, transaction costs induce a supply curve for the security's price distinct from that assumed herein. In particular, our supply curve is C^2 in the quantity purchased. Transactions costs violate this condition at the origin. As such, our solution to the problem is quite different from that used earlier in the literature to handle transaction costs, and it provides a new perspective on the existing results. We do not pursue this line of study in this paper, but instead refer the interested reader to Çetin [4].

Independently generated, the mathematics and theorems contained in Bank and Baum [2] are remarkably similar to those contained herein.[1] Bank and Baum study a continuous time economy with a single large trader whose trades affect the price, generalizing Jarrow's [19,20] discrete time model. In Bank and Baum, the large

[1] One difference is that in Bank and Baum a manipulative-free economy requires an equivalent measure such that every point on the supply curve evolves as a martingale. In our model, an arbitrage free economy requires only an equivalent measure such that the supply curve evaluated at the zero net trade evolves as a martingale.

trader faces a price process represented by a family of continuous semimartingales. The family is indexed by the aggregate position of the large investor's trading strategy. They concentrate on the large trader's real wealth, and characterize its evolution using the Itô-Wentzell formula. The economic issue studied is the large trader's ability or lack thereof to manipulate the market, and in a manipulative free economy, the valuation of contingent claims. Akin to our setting, continuous and finite variation trading strategies play a key role in the analysis because they enable the large trader to avoid market impact costs.

As just illustrated via the Bank and Baum paper, our approach is related to the market manipulation literature. The difference is that under market manipulation, the security's price process can depend not only on the investor's current trade, but also on the entire history of their past trades. The elimination of this path dependency condition in our structure excludes market manipulation and allows us to use classical arbitrage pricing theory, appropriately modified (see Cvitanic and Ma [8]; Jarrow [19]; Bank and Baum [2]). The extension of our model to the study of market manipulation is a subject for future research.

An outline for this paper is as follows. Section 2 describes the basic economy. Sections 3 and 4 study the first and second fundamental theorems of asset pricing, respectively. Section 5 provides an example - the extended Black-Scholes economy. Section 6 extends the analysis to discontinuous sample path evolutions for the supply curve, and Sect. 7 concludes the paper.

2 The model

This section presents the model. We are given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ satisfying the usual conditions where T is a fixed time. $\mathbb{P}$ represents the statistical or empirical probability measure. We also assume that $\mathcal{F}_0$ is trivial, i.e. $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

We consider a market for a security that we will call a stock, although the subsequent model applies equally well to bonds, commodities, foreign currencies, etc. We will initially assume that ownership of the security has no cash flows associated with it. Also traded is a money market account that accumulates value at the spot rate of interest. Without loss of generality, we assume that the spot rate of interest is zero, so that the money market account has unit value for all times.[2]

2.1 Supply curve

We consider an arbitrary trader who acts as a price taker with respect to an exogenously given supply curve for shares bought or sold of this stock within the trading interval. More formally, let $S(t, x, \omega)$ represent the stock price, *per share*, at time $t \in [0, T]$ that the trader pays/receives for an order of size $x \in R$ given the state $\omega \in \Omega$. A positive order $(x > 0)$ represents a buy, a negative order $(x < 0)$ represents a sale, and the order zero $(x = 0)$ corresponds to the marginal trade.

[2] A numéraire invariance theorem is proved in the Appendix for the subsequent economy.

By construction, rather than the trader facing a horizontal supply curve as in the classical theory (the same price for any order size), the trader now faces a supply curve that depends on his order size.[3] Note that the supply curve is otherwise independent of the trader's past actions, endowments, risk aversion, or beliefs. This implies that an investor's trading strategy has no lasting impact on the price process. This restriction distinguishes our economy from the situation where the supply curve also depends on the entire history of the trader's trades.

We now impose some structure on the supply curve.

Assumption 1 *(Supply Curve)*

1. $S(t,x,\cdot)$ *is* $\mathcal{F}_t - measurable$ *and non-negative.*
2. $x \mapsto S(t,x,\omega)$ *is a.e.* t *non-decreasing in* x*, a.s. (i.e.* $x \leq y$ *implies* $S(t,x,\omega) \leq S(t,y,\omega)$ *a.s.* $\mathbb{P}$*, a.e.* t*).*
3. S *is* C^2 *in its second argument,* $\partial S(t,x)/\partial x$ *is continuous in* t*, and* $\partial^2 S(t,x)/\partial x^2$ *is continuous in* t*.*
4. $S(\cdot,0)$ *is a semi-martingale.*
5. $S(\cdot,x)$ *has continuous sample paths (including time* 0*) for all* x*.*

Except for the second condition, these restrictions are self-explanatory. Condition 2 is the situation where the larger the purchase (or sale), the larger the price impact that occurs on the share price. This is the usual situation faced in asset pricing markets, where the quantity impact on the price is due to either information effects or supply/demand imbalances (see Kyle [26]; Glosten and Milgrom [15]; Grossman and Miller [16]). This excludes the more familiar situation in consumer products where there are quantity discounts for large orders. It includes, as a special case, horizontal supply curves. This structure can also be viewed as a generalization of the model in Jouini [23] where the traded securities have distinct selling and buying prices following separate stochastic processes. Here, instead of two prices as in Jouini [23], we have a continuum of stochastic processes indexed by trade size.

Example 2.1 (Supply curve). To present a concrete example of a supply curve, let $S(t,x) \equiv f(t,D_t,x)$ where D_t is an $n-$dimensional, $\mathcal{F}_t$-measurable semimartingale, and $f : R^{n+2} \to R^+$ is Borel measurable, C^1 in t, and C^2 in all its other arguments. This non-negative function f can be viewed as a reduced form supply curve generated by a market equilibrium process in a complex and dynamic economy. Under this interpretation, the vector stochastic process D_t represents the state variables generating the uncertainty in the economy, often assumed to be diffusion processes or at least Markov processes (e.g. a solution to a stochastic differential equation driven by a Levy process). This supply curve is time non-stationary. From a practical perspective, the stationary supply curve formulation, $f(D_t,x)$, may be more useful for empirical estimation.

If the supply curve in this example assumes the form

$$S(t,x) = e^{g(t,D_t,x)} S(t,0)$$

[3] In contrast, the trader is assumed to have no quantity impact due to his trades in the money market account.

where $g : R^{n+2} \to R$ is Borel measurable, C^1 in t, C^2 in all its other arguments, non-decreasing in x, with $g(t, D_t, 0) = 0$, $g(t, D_t, x) < 0$ for $x < 0$, and $g(t, D_t, x) > 0$ for $x > 0$, then it is similar to the liquidity premium model used in Jarrow and Turnbull [22]; Jarrow [21]. In Jarrow [21], the liquidity premium is due to temporary restrictions on trading strategies imposed during extreme market conditions, in particular, purchases/short sales are not allowed when the market has a shortage/surplus. In the absence of these extreme conditions, there is no liquidity premium and the supply curve is identically $S(t, 0)$. More formally, the function $g(t, D_t, x)$ is piecewise constant in x: $g(t, D_t) > 0$ for $x < 0$ when there are shortages, $g(t, D_t) < 0$ for $x > 0$ when there are surpluses, and $g(t, D_t) = 0$ otherwise. This setting does not lead to an arbitrage opportunity, however, because one cannot short or purchase the stock at the times when these extreme conditions arise. In Jarrow's [21] model, the liquidity premium has the interpretation of being a convenience yield.

2.2 Trading strategies

We start by defining the investor's trading strategy.

Definition 2.1 *A trading strategy is a triplet* $((X_t, Y_t : t \in [0, T]), \tau)$ *where* X_t *represents the trader's aggregate stock holding at time t (units of the stock),* Y_t *represents the trader's aggregate money market account position at time t (units of the money market account), and* τ *represents the liquidation time of the stock position, subject to the following restrictions: (a)* X_t *and* Y_t *are predictable and optional processes, respectively, with* $X_{0-} \equiv Y_{0-} \equiv 0$*, and (b)* $X_T = 0$ *and* τ *is a predictable* $(\mathcal{F}_t : 0 \le t \le T)$ *stopping time with* $\tau \le T$ *and* $X = H1_{[0,\tau)}$ *for some predictable process* $H(t, \omega)$.

We are interested in a particular type of trading strategy - those that are self-financing. By construction, a self-financing trading strategy generates no cash flows for all times $t \in [0, T)$. That is, purchase/sales of the stock must be obtained via borrowing/investing in the money market account. This implies that Y_t is uniquely determined by (X_t, τ). The goal is to define this self-financing condition for Y_t given an arbitrary stock holding (X_t, τ).

Definition 2.2 *A self-financing trading strategy (s.f.t.s.) is a trading strategy* $((X_t, Y_t : t \in [0, T]), \tau)$ *where (a)* X_t *is càdlàg if* $\partial S(t, 0)/\partial x \equiv 0$ *for all t, and* X_t *is càdlàg with finite quadratic variation (*$[X, X]_T < \infty$*) otherwise, (b)* $Y_0 = -X_0 S(0, X_0)$*, and (c) for* $0 < t \le T$*,*

$$Y_t = Y_0 + X_0 S(0, X_0) + \int_0^t X_{u-} dS(u, 0) - X_t S(t, 0) - \sum_{0 \le u \le t} \Delta X_u [S(u, \Delta X_u) - S(u, 0)] - \int_0^t \frac{\partial S}{\partial x}(u, 0) d[X, X]^c_u. \quad (2.1)$$

Condition (a) imposes restrictions on the class of acceptable trading strategies. Under the hypotheses that X_t is càdlàg and of finite quadratic variation, the right side of expression (2.1) is always well-defined although the last two terms (always being non-positive) may be negative infinity. The classical theory, under frictionless and competitive markets, does not need these restrictions. An example of a trading strategy that is allowed in the classical theory, but disallowed here, is $X_t = 1_{\{S(t,0)>K\}}$ for some constant $K > 0$ where $S(t,0)$ follows a Brownian motion. Under the Brownian motion hypothesis this is a discontinuous trading strategy that jumps infinitely often immediately after $S(t,0) = K$ (the jumps are not square summable), and hence Y_t is undefined.

Condition (b) implies the strategy requires zero initial investment at time 0. When studying complete markets in a subsequent section, condition (b) of the s.f.t.s. is removed so that $Y_0 + X_0 S(0, X_0) \neq 0$.

Condition (c) is the self-financing condition at time t. The money market account equals its value at time 0, plus the accumulated trading gains (evaluated at the marginal trade), less the cost of attaining this position, less the price impact costs of discrete changes in share holdings, and less the price impact costs of continuous changes in the share holdings. This expression is an extension of the classical self-financing condition when the supply curve is horizontal. To see this note that using condition (b) with expression (2.1) yields the following simplified form of the self-financing condition:

$$\begin{aligned} Y_t + X_t S(t,0) = &\int_0^t X_{u-} dS(u,0) \\ &- \sum_{0\leq u\leq t} \Delta X_u [S(u, \Delta X_u) - S(u,0)] \\ &- \int_0^t \frac{\partial S}{\partial x}(u,0) d[X,X]_u^c \quad \text{for } 0 \leq t \leq T. \end{aligned} \tag{2.2}$$

The left side of expression (2.2) represents the classical "value" of the portfolio at time 0. The right side gives its decomposition into various components. The first term on the right side is the classical "accumulated gains/losses" to the portfolio's value. The last two terms on the right side capture the impact of illiquidity, both entering with a negative sign.

To understand the origin of the liquidity terms, consider the following heuristic derivation. Assume that X is a semi-martingale. Intuitively, the self-financing condition is

$$\begin{aligned} dY_t &= -S(t+dt, dX_t) dX_t \\ &= -S(t,0) dX_t - [S(t+dt, dX_t) - S(t+dt,0)] dX_t \\ &\quad -[S(t+dt,0) - S(t,0)] dX_t. \end{aligned}$$

Because S is continuous, we can rewrite this last expression as

$$dY_t = -S(t,0) dX_t - [S(t+dt, dX_t) - S(t+dt,0)] dX_t - d[X^c, S]_t.$$

Decomposing X into its continuous and discontinuous parts, we rewrite this expression one more time:

$$dY_t = -S(t,0)dX_t - [S(t+dt, dX_t^c) - S(t+dt,0)]dX_t^c \\ -[S(t,\Delta X_t) - S(t,0)]\Delta X_t - d[X^c, S]_t.$$

Because $\partial S(\cdot, x)/\partial x$ is continuous and ignoring the higher order terms,

$$-[S(t+dt, dX_t^c) - S(t+dt,0)]dX_t^c = -\frac{\partial S}{\partial x}(t,0)d[X,X]_t^c.$$

Integration by parts gives

$$S(t,0)dX_t = -X_{t-}dS(t,0) - d[X^c, S]_t.$$

Using these two expressions yields the desired result:

$$dY_t = X_{t-}dS(t,0) - \Delta X_t[S(t,\Delta X_t) - S(t,0)] - \frac{\partial S}{\partial x}(t,0)d[X,X]_t^c.$$

This heuristic derivation can be justified by a limiting argument based on simple trading strategies (see the Appendix).

2.3 The marked-to-market value of a s.f.t.s. and its liquidity cost

This section defines the marked-to-market value of a trading strategy and its liquidity cost. At any time prior to liquidation, there is no unique value of a trading strategy or portfolio. Indeed, any price on the supply curve is a plausible price to be used in valuing the portfolio. At least three economically meaningful possibilities can be identified: (i) the immediate liquidation value (assuming that $X_t > 0$ gives $Y_t + X_t S(t, -X_t)$), (ii) the accumulated cost of forming the portfolio (Y_t), and (iii) the portfolio evaluated at the marginal trade ($Y_t + X_t S(t,0)$).[4] This last possibility is defined to be the *marked-to-market value* of the self-financing trading strategy (X, Y, τ). It represents the value of the portfolio under the classical price taking condition.

Motivated by expression (2.2), we define the liquidity cost to be the difference between the accumulated gains/losses to the portfolio, computed as if all trades are executed at the marginal trade price $S(t,0)$, and the marked-to-market value of the portfolio.

Definition 2.3 *The liquidity cost of a s.f.t.s.* (X, Y, τ) *is*

$$L_t \equiv \int_0^t X_{u-}dS(u,0) - [Y_t + X_t S(t,0)].$$

The following lemma follows from the preceding definition.

[4] These three valuations are (in general) distinct except at one date, the liquidation date. At the liquidation time τ, the value of the portfolio under each of these three cases are equal because $X_\tau = 0$.

Lemma 2.1 *(Equivalent Characterization of the Liquidity Costs).*

$$L_t = \sum_{0\le u\le t} \Delta X_u[S(u,\Delta X_u) - S(u,0)] + \int_0^t \frac{\partial S}{\partial x}(u,0)d[X,X]^c_u \ \ge 0$$

where $L_{0-} = 0,\ L_0 = X_0[S(0,X_0) - S(0,0)]$ *and* L_t *is non-decreasing in* t.

Proof The first equality follows directly from the definitions. The second inequality and the subsequent observation follow from the fact that $S(u,x)$ is increasing in x. □

We see here that the liquidity cost is non-negative and non-decreasing in t. It consists of two components. The first is due to discontinuous changes in the share holdings. The second is due to the continuous component. This expression is quite intuitive. Note that because $X_{0-} = Y_{0-} = 0$, $\Delta L_0 = L_0 - L_{0-} = L_0 > 0$ is possible.

Remark 2.1

1. When X is of bounded variation, $[X,X]^c_u$ is zero.
2. When X is continuous (that is, $\Delta X_u = 0$ for $u > 0$), then the first term in the liquidity costs equals its value at zero, i.e. L_0.
3. When X is both continuous and of bounded variation, the liquidity costs of the trading strategy equal L_0.

3 The extended first fundamental theorem

This section studies the characterization of an arbitrage free market and generalizes the first fundamental theorem of asset pricing to an economy with liquidity risk.

To evaluate a self-financing trading strategy, it is essential to consider its value after liquidation. This is equivalent to studying the portfolio's real wealth, as contrasted with its marked-to-market value or paper wealth, see Jarrow [19]. Using this insight, an arbitrage opportunity can now be defined.

Definition 3.1 *An arbitrage opportunity is a s.f.t.s.* (X,Y,τ) *such that* $\mathbb{P}\{Y_T \ge 0\} = 1$ *and* $\mathbb{P}\{Y_T > 0\} > 0$.

Before we begin, let us clarify the behavior of the stochastic integral at time $t = 0$. If S is a semimartingale with decomposition $S_t = S_0 + M_t + A_t$, $M_0 = A_0 = 0$, where X is a càdlàg and adapted process, then $\int_0^t X_{s-}dS_s = \int_0^t X_{s-}dM_s + \int_0^t X_{s-}dA_s$. The initial term is $X_{0-}\Delta S_0$, which of course is 0, since $X_{0-} = 0$. On the other hand if X itself is predictable then $\int_0^t X_s dS_s = X_0\Delta S_0 + \int_0^t X_s dM_s + \int_0^t X_s dA_s$. Some authors write $\int_{0+}^t X_s dS_s$ to denote the stochastic integral without its jump at 0. With this interpretation, $\int_0^t X_s dS_s = X_0\Delta S_0 + \int_{0+}^t X_s dS_s$. Note that we assume $S(\cdot,0)$ is continuous at $t = 0$, thus $\int_0^0 X_u dS(u,0) = 0$ for all predictable X that are $S(\cdot,0)$-integrable.

We first need to define some mathematical objects. Let $s_t \equiv S(t,0)$, $(X_- \cdot s)_t \equiv \int_0^t X_{u-} dS(u,0)$, and for $\alpha \geq 0$, let $\Theta_\alpha \equiv \{\text{s.f.t.s } (X,Y,\tau) \mid (X_- \cdot s)_t \geq -\alpha \text{ for all } t \text{ almost surely } \}$.

Definition 3.2 *Given an $\alpha \geq 0$, a s.f.t.s. (X,Y,τ) is said to be α-admissible if $(X,Y,\tau) \in \Theta_\alpha$. A s.f.t.s. is admissible if it is α-admissible for some α.*

Lemma 3.1 *($Y_t + X_t S(t,0)$ is a supermartingale). If there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that $S(\cdot,0)$ is a $\mathbb{Q}$-local martingale, and if $(X,Y,\tau) \in \Theta_\alpha$ for some α, then $Y_t + X_t S(t,0)$ is a $\mathbb{Q}$-supermartingale.*

Proof From Definition 2.3 we have that $Y_t + X_t S(t,0) = (X_- \cdot s)_t - L_t$. Under the $\mathbb{Q}$ measure, $(X_- \cdot s)_t$ is a local $\mathbb{Q}$-martingale. Since $(X,Y,\tau) \in \Theta_\alpha$ for some α, it is a supermartingale (Duffie [12]). But, by Lemma 2.1, L_t is non-negative and non-decreasing. Therefore, $Y_t + X_t S(t,0)$ is a supermartingale too. □

Theorem 3.1 *(A sufficient condition for no arbitrage) If there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that $S(\cdot,0)$ is a $\mathbb{Q}$-local martingale, then there is no arbitrage for $(X,Y,\tau) \in \Theta_\alpha$ for any α.*

Proof Under this hypothesis, by Lemma 3.1, $Y_t + X_t S(t,0)$ is a supermartingale. Note that $Y_\tau + X_\tau S(\tau,0) = Y_\tau$ by the definition of the liquidation time. Thus, for this s.f.t.s., $\mathbb{E}^{\mathbb{Q}}[Y_\tau] = \mathbb{E}^{\mathbb{Q}}[Y_\tau + X_\tau S(\tau,0)] \leq 0$. But, by the definition of an arbitrage opportunity, $\mathbb{E}^{\mathbb{Q}}[Y_\tau] > 0$. Hence, there exist no arbitrage opportunities in this economy. □

The intuition behind this theorem is straightforward. The marked-to-market portfolio is a hypothetical portfolio that contains zero liquidity costs (see Definition 2.3). If $S(\cdot,0)$ has an equivalent martingale measure, then these hypothetical portfolios admit no arbitrage. But, since the actual portfolios differ from these hypothetical portfolios only by the subtraction of non-negative liquidity costs (Lemma 2.1), the actual portfolios cannot admit arbitrage either.

In order to get a sufficient condition for the existence of an equivalent local martingale measure, we need to define the notion of a free lunch with vanishing risk as in Delbaen and Schachermayer [11]. This will require a preliminary definition.

Definition 3.3 *A free lunch with vanishing risk (FLVR) is either: (i) an admissible s.f.t.s. that is an arbitrage opportunity or (ii) a sequence of ϵ_n-admissible s.f.t.s. $(X^n,Y^n,\tau^n)_{n\geq 1}$ and a non-negative F_T-measurable random variable, f_0, not identically 0 such that $\epsilon_n \to 0$ and $Y_T^n \to f_0$ in probability.*[5]

To state the theorem, we need to introduce a related, but fictitious economy. Consider the economy introduced previously, but suppose instead that $S(t,x) \equiv S(t,0)$. When there is no confusion, we denote $S(t,0)$ by the simpler notation s_t. In this fictitious economy, a s.f.t.s. $\left(X,Y^0,\tau\right)$ satisfies the classical condition with $X_0 = 0$, the value of the portfolio is given by $Z_t^0 \equiv Y_t^0 + X_t s_t$ with $Y_t^0 =$

[5] Delbaen and Schachermayer [11] Proposition 3.6 page 477 shows that this definition is equivalent to FLVR in the classical economy.

$(X \cdot s)_t - X_t s_t$ for all $0 \leq t \leq T$, and X is allowed to be a general $S(\cdot, 0) \equiv s$ integrable predictable process (see the remark following expression (2.1)). So, in this fictitious economy, our definitions of an arbitrage opportunity, an admissible trading strategy, and a NFLVR collapse to those in [11].

Theorem 3.2 *(First fundamental theorem) Suppose there are no arbitrage opportunities in the fictitious economy. Then, there is no free lunch with vanishing risk (NFLVR) if and only if there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that $S(\cdot, 0) \equiv s$ is a $\mathbb{Q}-$local martingale.*

The proof is in the Appendix.

4 The extended second fundamental theorem

This section studies the meaning and characterization of a complete market and generalizes the second fundamental theorem of asset pricing to an economy with liquidity risk. For this section we assume that there exists an equivalent local martingale measure $\mathbb{Q}$ so that the economy is arbitrage free and there is no free lunch with vanishing risk (NFLVR).

Also for this section, we generalize the definition of a s.f.t.s (X, Y, τ) slightly to allow for non-zero investments at time 0. In particular, a s.f.t.s. (X, Y, τ) in this section will satisfy Definition 2.2 with the exception that condition (b) is removed. That is, a s.f.t.s. need not have zero initial value ($Y_0 + X_0 S(0, X_0) \neq 0$).[6]

To proceed, we need to define the space $\mathcal{H}^2_Q$ of semimartingales with respect to the equivalent local martingale measure $\mathbb{Q}$. Let Z be a special semimartingale with canonical decomposition $Z = N + A$, where N is a local martingale under $\mathbb{Q}$ and A is a predictable finite variation process. The $\mathcal{H}^2$ norm of Z is defined to be

$$\|Z\|_{\mathcal{H}^2} = \left\| [N, N]^{1/2}_\infty \right\|_{L^2} + \left\| \int_0^\infty |dA_s| \right\|_{L^2}$$

where the L^2- norms are with respect to the equivalent local martingale measure $\mathbb{Q}$.

Throughout this section we make the assumption that $s(\cdot) = S(\cdot, 0) \in \mathcal{H}^2_Q$. Since we're assuming $s \in \mathcal{H}^2_Q$, it is no longer necessary to require that $X \cdot s$ is uniformly bounded from below.

Definition 4.1 *A contingent claim is any $\mathcal{F}_T-$ measurable random variable C with $\mathbb{E}^{\mathbb{Q}}(C^2) < \infty$.*

Note that the contingent claim is considered at a time T, prior to which the trader's stock position is liquidated. If the contingent claim's payoff depends on

[6] In this section we could also relax condition (b) of a trading strategy, Definition 2.1, to remove the requirement that $X_T = 0$. However, as seen below, it is always possible to approximate any random variable with such a trading strategy. Consequently, this restriction is without loss of generality in the context of our model. This condition was imposed in the previous section to make the definition of an arbitrage opportunity meaningful in a world with liquidity costs.

the stock price at time T, then the dependence of the contingent claim's payoff on the shares purchased/sold at time T must be made explicit. Otherwise, the contingent claim's payoff is not well-defined. An example helps to clarify this necessity.

Consider a European call option on the stock with a strike price[7] of K and maturity $T_0 \leq T$.[8] To write the modified boundary condition for this option incorporating the supply curve for the stock, we must consider two cases: cash delivery and physical delivery.

1. If the option has cash delivery, the long position in the option receives cash at maturity if the option ends up in-the-money. To match the cash settlement, the synthetic option position must be liquidated prior to time T_0. When the synthetic option position is liquidated, the underlying stock position is also liquidated. The position in the stock at time T_0 is, thus, zero.
   If we sell the stock at time T_0 to achieve this position, then the boundary condition is $C \equiv \max[S(T_0, -1) - K, 0]$ where $\Delta X_{T_0} = -1$ since the option is for one share of the stock. However, as we show below, one could also liquidate this stock position just prior to time T_0 using a continuous and finite variation process, so that $\Delta X_{T_0} = 0$. This alternative liquidation strategy might be chosen in an attempt to avoid liquidity costs at time T_0. In this case, the boundary condition is $C \equiv \max[S(T_0, 0) - K, 0]$. Note that using this latter liquidation strategy, the option's payoff is only approximately obtained (to a given level of accuracy) because liquidation occurs just before T_0.
2. If the option has physical delivery, then the synthetic option position should match the physical delivery of the stock in the option contract. With physical delivery, the option contract obligates the short position to deliver the stock shares. To match the physical delivery, the stock position in the synthetic option is not sold. Unfortunately, our model requires the stock position to be liquidated at time T_0. Formally, physical delivery is not possible in our construct. However, to approximate physical delivery in our setting, we can impose the boundary condition $C \equiv \max[S(T_0, 0) - K, 0]$ where $\Delta X_{T_0} = 0$. This boundary condition is consistent with no liquidity costs being incurred at time T_0, which would be the case with physical delivery of the stock.[9]

Definition 4.2 *The market is complete if given any contingent claim C, there exists a s.f.t.s. (X, Y, τ) with $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T X_u^2 d[s,s]_u\right) < \infty$ such that $Y_T = C$.*

To understand the issues involved in replicating contingent claims, let us momentarily consider a contingent claim C in $L^2(d\mathbb{Q})$ where there exists a s.f.t.s.

[7] To be consistent with the previous construct, one should interpret K as the strike price normalized by the value of the money market account at time T_0.

[8] Recall that interest rates are zero, so that the value of the liquidated position at time T_0 is the same as the position's value at time T.

[9] We are studying an economy with trading only in the stock and money market account. Expanding this economy to include trading in an option expands the liquidation possibilities prior to time T. Included in this set of expanded possibilities is the delivery of the physical stock to offset the position in an option, thereby avoiding any liquidity costs at time T. Case 2 is the method for capturing no liquidity costs in our restricted setting.

(X, Y, τ) such that $C = c + \int_0^T X_u ds_u$ where $c \in \mathbb{R}$ and $\mathbb{E}^{\mathbb{Q}} \left\{ \int_0^T X_u^2 d[s,s]_u \right\} < \infty$. Note that $\mathbb{E}^{\mathbb{Q}}(C) = c$ since $\int_0^0 X_u ds_u = X_0 \Delta s_0 = 0$ by the continuity of s at time 0. This is the situation where, in the absence of liquidity costs, a long position in the contingent claim C is redundant. In this case, Y_0 is chosen so that $Y_0 + X_0 s_0 = c$. But, the liquidity costs in trading this stock position are (by Lemma 2.1):

$$L_t = \sum_{0 \leq u \leq t} \Delta X_u [S(u, \Delta X_u) - S(u,0)] + \int_0^t \frac{\partial S}{\partial x}(u,0) d[X,X]_u^c \geq 0.$$

We have from Definition 2.2 that

$$Y_T = Y_0 + X_0 s_0 + \int_0^T X_{u-} ds_u - X_T s_T - L_T + L_0$$

and[10] $\int_0^T X_{u-} ds_u = \int_0^T X_u ds_u$ so that

$$Y_T = C - X_T s_T - L_T + L_0.$$

By assumption, we have liquidated by time T, giving $X_T = 0$. Thus, we have

$$Y_T = C - (L_T - L_0) \leq C.$$

That is, considering liquidity costs, this trading strategy sub-replicates a long position in this contingent claim's payoffs. Analogously, if we use $-X$ to hedge a short position in this contingent claim, the payoff is generated by

$$\overline{Y}_T = -C - (\overline{L}_T - \overline{L}_0) \leq -C$$

where $\overline{Y}$ is the value in the money market account and $\overline{L}$ is the liquidity cost associated with $-X$. The liquidation value of the trading strategies (long and short) provide a lower and upper bound on attaining the contingent claim's payoffs.

Remark 4.1

1. If $\frac{\partial S}{\partial x}(\cdot, 0) \equiv 0$, then $L_\cdot = L_0$ if X is a continuous trading strategy. So, under this hypothesis, all claims C where there exists a s.f.t.s. (X, Y, τ) such that $C = c + \int_0^T X_u ds_u$ with X continuous can be replicated. For example, if $S(\cdot, 0)$ is a geometric Brownian motion (an extended Black-Scholes economy), a call option can be replicated since the Black-Scholes hedge is a continuous s.f.t.s.
2. If $\frac{\partial S}{\partial x}(\cdot, 0) \geq 0$ (the general case), then $L = L_0$ if X is a finite variation and continuous trading strategy. So, under this hypothesis, all claims C where there exists a s.f.t.s. (X, Y, τ) such that $C = c + \int_0^T X_u ds_u$ with X of finite variation and continuous can be replicated.

[10] $\int_0^T X_u ds_u = \int_0^T X_{u-} ds_u + \sum_{0 \leq u \leq T} \Delta X_u \Delta s_u$ and $\Delta X_u \Delta s_u = 0$ for all u since $\Delta s_u = 0$ for all u by the continuity of s.

The remark above shows that if we can approximate X using a finite variation and continuous trading strategy, in a limiting sense, we may be able to avoid all the liquidity costs in the replication strategy. In this regard, the following lemma is relevant.

Lemma 4.1 *(Approximating continuous and finite variation s.f.t.s.) Let* $C \in L^2(d\mathbb{Q})$. *Suppose there exists a predictable* X *with* $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T X_u^2 d[s,s]_u\right) < \infty$ *so that* $C = c + \int_0^T X_u ds_u$ *for some* $c \in \mathbb{R}$. *Then, there exists a sequence of s.f.t.s.* $(X^n, Y^n, \tau^n)_{n\geq 1}$ *with* X^n *bounded, continuous and of finite variation such that* $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T (X_u^n)^2 d[s,s]_u\right) < \infty$, $X_0^n = 0$, $X_T^n = 0$, $Y_0^n = \mathbb{E}^{\mathbb{Q}}(C)$ *for all* n *and*

$$Y_T^n = Y_0^n + X_0^n S(0, X_0^n) + \int_0^T X_{u-}^n ds_u - X_T^n S(T,0)$$
$$-L_T^n \to c + \int_0^T X_u ds_u = C \tag{4.1}$$

in $L^2(d\mathbb{Q})$.

Proof Note that for any predictable X that is integrable with respect to s, $\int_0^T X_u ds_u = \int_0^T X_u 1_{(0,T]}(u) ds_u$ since $\int_0^T 1_{(0,T]} X_u ds_u = \int_0^T X_u ds_u - X_0 \Delta s_0$ and $\Delta s_0 = 0$. Therefore, we can without loss of generality assume that $X_0 = 0$.

Given any $H \in \mathbb{L}$ (the set of adapted processes that have left continuous paths with right limits a.s.) with $H_0 = 0$, we define, H^n, by the following:

$$H_t^n(\omega) = n \int_{t-\frac{1}{n}}^t H_u(\omega) du,$$

for all $t \geq 0$, letting H_u equal 0 for $u < 0$. Then H is the a.s. pointwise limit of the sequence of adapted processes H^n that are continuous and of finite variation. Note that $H_0^n = 0$ for all n. Theorem 2 in Chapter IV of Protter [27] remains valid if $\mathbf{b}\mathbb{L}$ is replaced by the set of bounded, continuous processes with paths of finite variation on compact time sets. Let X with $X_0 = 0$ be predictable and $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T X_u^2 d[s,s]_u\right) < \infty$. Since $X \cdot s$ is defined to be the $\lim_{k\to\infty} \overline{X}^k \cdot s$, where the convergence is in $\mathcal{H}^2$ and $\overline{X}^k = X 1_{\{|X|\leq k\}}$, and using the above observation, there exists a sequence of continuous and bounded processes of finite variation, $(X^n)_{n\geq 1}$, such that $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T (X_u^n)^2 d[s,s]_u\right) < \infty$, $X_0^n = 0$ for all n and

$$\int_0^T X_u^n ds_u \to \int_0^T X_u ds_u,$$

in $L^2(d\mathbb{Q})$ (see Theorems 2, 4, 5 and 14 in Chapter IV of Protter [27] in this respect.)

Furthermore, Theorem A.4 and Corollary A.2 in the Appendix allow us to choose $X_T^n = 0$ for all n. Now, choose $Y^n = \mathbb{E}^{\mathbb{Q}}(C)$ for all n and define Y_t^n for $t > 0$ by (2.1). Let $\tau^n = T$ for all n. Then, the sequence $(X^n, Y^n, \tau^n)_{n\geq 1}$ will satisfy (4.1). Note that $L^n \equiv 0$ for all n and $\int_0^T X_{u-}^n ds_u = \int_0^T X_u^n ds_u$. □

This lemma motivates the following definition and extension of the second fundamental theorem of asset pricing.

Definition 4.3 *The market is approximately complete if given any contingent claim C, there exists a sequence of s.f.t.s. (X^n, Y^n, τ^n) with $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T (X_u^n)^2 d[s,s]_u\right) < \infty$ for all n such that $Y_T^n \to C$ as $n \to \infty$ in $L^2(d\mathbb{Q})$.*

Theorem 4.1 *(Second fundamental theorem) Suppose there exists a* **unique** *probability measure $\mathbb{Q} \sim \mathbb{P}$ such that $S(\cdot,0) = s$ is a $\mathbb{Q}$-local martingale. Then, the market is approximately complete.*

Proof The proof proceeds in two steps. Step 1 shows that the hypothesis guarantees that a fictitious economy with no liquidity costs is complete. Step 2 shows that this result implies approximate completeness for an economy with liquidity costs.

Step 1 Consider the economy introduced in this paper, but suppose that $S(\cdot,x) \equiv S(\cdot,0)$. In this fictitious economy, a s.f.t.s. $\left(X, Y^0, \tau\right)$ satisfies the classical condition with $Y_t^0 \equiv Y_0 + X_0 S(0,0) + \int_0^t X_{u-} ds_u - X_t s_t$. The classical second fundamental theorem (see Harrison and Pliska [17]) applies: the fictitious market is complete if and only if $\mathbb{Q}$ is unique.

Step 2 By Step 1, given $\mathbb{Q}$ is unique, the fictitious economy is complete and, moreover, s has the martingale representation property. Hence, there exists a predictable X such that $C = c + \int_0^T X_u ds_u$ with $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T X_u^2 d[s,s]_u\right) < \infty$ (see Sect. 3 of Chapt. IV of [27] in this respect). Then, by applying the lemma above, the market is approximately complete. □

Suppose the martingale measure is unique. Then, by the theorem we know that given any contingent claim C, there exists a sequence of s.f.t.s. $(X^n, Y^n, \tau^n)_{n\geq 1}$ with $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T (X_u^n)^2 d[s,s]_u\right) < \infty$ for all n so that $Y_T^n = Y_0^n + X_0^n S(0, X_0^n) - L_T^n + \int_0^T X_{u-}^n dS(u,0) \to C$ in $L^2(d\mathbb{Q})$. We call any such sequence of s.f.t.s., $(X^n, Y^n, \tau^n)_{n\geq 1}$ an *approximating sequence* for C.

Definition 4.4 *Let C be a contingent claim and Ψ^C be the set of approximating sequences for C. The time 0 value of the contingent claim C is given by*

$$\inf\left\{\liminf_{n\to\infty} Y_0^n + X_0^n S(0, X_0^n) : (X^n, Y^n, \tau^n)_{n\geq 1} \in \Psi^C\right\}.$$

Corollary 4.1 *(Contingent claim valuation) Suppose there exists a unique probability measure $\mathbb{Q} \sim \mathbb{P}$ such that $S(\cdot,0) = s$ is a $\mathbb{Q}$-local martingale. Then, the time 0 value of any contingent claim C is equal to $\mathbb{E}^{\mathbb{Q}}(C)$.*

Proof Let $(X^n, Y^n, \tau^n)_{n\geq 1}$ be an approximating sequence for C. Then, $\mathbb{E}^{\mathbb{Q}}(Y_T^n - C)^2 \to 0$, and thus, $\mathbb{E}^{\mathbb{Q}}(Y_T^n - C) \to 0$. However, since $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T (X_u^n)^2 d[s,s]_u\right) < \infty$ for all n, $\int_0^{\cdot} X_{u-}^n ds_u$ is a $\mathbb{Q}$-martingale for each n. This yields $\mathbb{E}^{\mathbb{Q}}(Y_T^n) = Y_0^n + X_0^n S(0, X_0) - \mathbb{E}^{\mathbb{Q}}(L_T^n)$. Combining this with the fact that $L^n \geq 0$ for each n and $\mathbb{E}^{\mathbb{Q}}(Y_T^n - C) \to 0$ gives $\liminf_{n\to\infty} Y_0^n +$

$X_0^n S(0, X_0) \geq \mathbb{E}^{\mathbb{Q}}(C)$ for all approximating sequences. However, as proven in Lemma 4.1, there exists some approximating sequence $(\overline{X}^n, \overline{Y}^n, \overline{\tau}^n)_{n\geq 1}$ with $\overline{L}^n = 0$ for all n. For this sequence, $\liminf_{n\to\infty} \overline{Y}_0^n + \overline{X}_0^n S(0, X_0) = \mathbb{E}^{\mathbb{Q}}(C)$. □

Remark 4.2

1. The above value is consistent with no arbitrage. Indeed, suppose the contingent claim is sold at price $p > \mathbb{E}^{\mathbb{Q}}(C)$. Then, one can short the contingent claim at p and construct a sequence of continuous and of finite variation s.f.t.s., $(X^n, Y^n, \tau^n)_{n\geq 1}$, with $Y_0^n = \mathbb{E}^{\mathbb{Q}}(C)$, $X_0^n = 0$ and $\lim_{n\to\infty} Y_T^n = C$ in L^2, hence, in probability, creating a FLVR. However, this is not allowed since $\mathbb{Q}$ is an equivalent martingale measure for s. Similarly, one can show that the price of the contingent claim cannot be less than $\mathbb{E}^{\mathbb{Q}}(C)$.
2. Given our supply curve formulation, this corollary implies that continuous trading strategies of finite variation can be constructed to both (i) approximately replicate any contingent claim, and (ii) avoid all liquidity costs. This elucidates the special nature of *continuous* trading strategies in a continuous time setting.

5 Example (extended Black-Scholes economy)

To illustrate the previous theory, we consider an extension of the Black-Scholes economy that incorporates liquidity risk. A detailed discussion of this example along with some empirical evidence regarding the pricing of traded options in the extended Black-Scholes economy can be found in Çetin et al. [5].

5.1 The economy

Let

$$S(t, x) = e^{\alpha x} S(t, 0) \ \ with\ \alpha > 0 \tag{5.1}$$

$$S(t, 0) \equiv \frac{s_0 e^{\mu t + \sigma W_t}}{e^{rt}} \tag{5.2}$$

where μ, σ are constants and W is a standard Brownian motion initialized at zero.

For this section, let the spot rate of interest be constant and equal to r per unit time. The marginal stock price follows a geometric Brownian motion. The normalization by the money market account's value is made explicit in expression (5.2). Expressions (5.1) and (5.2) characterize an extended Black-Scholes economy. It is easy to check that this supply curve satisfies Assumption 1 in Sect. 2.

Under these assumptions, there exists a unique martingale measure for $S(\cdot, 0) = s$, see Duffie [12]. Hence, we know that the market is arbitrage-free and approximately complete.

5.2 Call option valuation

Consider a European call option with strike price K and maturity date T on this stock with cash delivery. Given cash delivery, in order to avoid liquidity costs at time T, the payoff[11] to the option at time T is selected to be $C_T = \max[S(T,0) - K, 0]$.

Under this structure, by the corollary to the second fundamental theorem of asset pricing, the value of a long position in the option is:

$$C_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}(\max[S(T,0) - K, 0]).$$

It is well-known that the expectation in this expression gives the Black-Scholes-Merton formula:

$$s_0 N(h(0)) - Ke^{-rT}N(h(0) - \sigma\sqrt{T})$$

where $N(\cdot)$ is the standard cumulative normal distribution function and

$$h(t) \equiv \frac{\log s_t - \log K + r(T-t)}{\sigma\sqrt{T-t}} + \frac{\sigma}{2}\sqrt{T-t}.$$

Applying Itô's formula, the classical replicating strategy, $X = (X_t)_{t\in[0,T]}$, implied by the classical Black-Scholes-Merton formula is given by

$$X_t = N(h(t)). \tag{5.3}$$

This hedging strategy is continuous, but not of finite variation.

In this economy, we have that $(\frac{\partial S}{\partial x}(t,0) = \alpha e^0 s_t = \alpha s_t)$. Hence, although the call's value is the Black-Scholes formula, the standard hedging strategy will not attain this value. Indeed, using this strategy, it can be shown that the classical Black-Scholes hedge leads to the following non-zero liquidity costs (from expression (2.1)):[12]

$$L_T = X_0(S(0,X_0) - S(0,0)) + \int_0^T \frac{\alpha\left(N'\left(h(u)\right)\right)^2 s_u}{T-u}du. \tag{5.4}$$

In contrast, an approximate hedging strategy that is continuous and of finite variation having zero liquidity costs is the sequence of s.f.t.s. $(X^n, Y^n, \tau^n)_{n\geq 1}$ with

$$X_t^n = 1_{[\frac{1}{n}, T-\frac{1}{n})}(t) n \int_{(t-\frac{1}{n})^+}^t N(h(u))du, \text{ if } 0 \leq t \leq T - \frac{1}{n}, \tag{5.5}$$

$$X_t^n = (nTX^n_{(T-\frac{1}{n})} - nX^n_{(T-\frac{1}{n})}t), \text{ if } T - \frac{1}{n} \leq t \leq T,$$

and $Y_0^n = \mathbb{E}^{\mathbb{Q}}(C_T)$. In the limit, this trading strategy attains the call's time T value, i.e. $Y_T^n = Y_0^n + \int_0^T X^n_{u-}ds_u \to C_T = \max[S(T,0) - K, 0]$ in $L^2(d\mathbb{Q})$.

[11] The strike price needs to be normalized by the value of the money market account.

[12] Note that both L_T and Y_T^n are normalized by the value of the money market account.

6 Discontinuous supply curve evolutions

This section extends the previous economy to allow stock price supply curves with discontinuous sample paths.

6.1 The supply curve and s.f.t.s.'s

For this extension we replace Assumption 1 with

Assumption 2 *(Supply curve)*

1. $S(t,x,\cdot)$ *is* $\mathcal{F}_t - measurable$ *and non-negative.*
2. $x \mapsto S(t,x,\omega)$ *is a.e.* t *non-decreasing in* x*, a.s. (i.e.* $x \leq y$ *implies* $S(t,x,\omega) \leq S(t,y,\omega)$ *a.s.* $\mathbb{P}$*, a.e.* t*).*
3. S *is* C^2 *in its second argument.*
4. $S(\cdot,0)$ *is a semi-martingale and* $\partial S(\cdot,0)/\partial x$ *has finite quadratic variation.*

Remark 6.1 Sample path continuity of $S(t,0)$ is replaced by $\partial S(\cdot,0)/\partial x$ has finite quadratic variation. In order for $[\partial S(\cdot,0)/\partial x, [X,X]]$ to be well defined we need this restriction. It will be satisfied, for example, if $S(t,x) = f(D_t,x)$ where f is C^3 in its first argument, C^2 in its second argument, and where D is a semimartingale.

The definition of a trading strategy remains unchanged under Assumption 2. The definition of a self-financing trading strategy is changed only marginally to include possible jumps in the first partial derivative of the supply curve, i.e.

Definition 6.1 *A self-financing trading strategy (s.f.t.s.) is a trading strategy* $((X_t, Y_t : t \in [0,T]), \tau)$ *where (a)* X *is càdlàg if* $\partial S(t,0)/\partial x \equiv 0$ *for all* t *, and* X *is càdlàg with finite quadratic variation (*$[X,X]_T < \infty$*) otherwise, (b)* $Y_0 = -X_0 S(0,X_0)$*, and (c) for* $0 \leq t \leq T$*,*

$$Y_t = Y_0 + X_0 S(0,X_0) + \int_0^t X_{u-} dS(u,0) - X_t S(t,0) - \sum_{0 \leq u \leq t} \Delta X_u [S(u,\Delta X_u) - S(u,0)] - \int_0^t \frac{\partial S}{\partial x}(u-,0) d[X,X]^c_u.$$

The justification of this definition, based on a limiting argument, is contained in the Appendix.

6.2 The extended first fundamental theorem

Under Assumption 2, Theorem 3.1 extends. A similar proof applies.

Theorem 6.1 *(A sufficient condition for no arbitrage) Given Assumption 2, if there exists a probability measure* $\mathbb{Q} \sim \mathbb{P}$ *such that* $S(\cdot,0)$ *is a* $\mathbb{Q}$*-local martingale, then there is no arbitrage for* $(X,Y,\tau) \in \Theta_\alpha$ *for any* α*.*

6.3 The extended second fundamental theorem

Suppose there exists an equivalent local martingale measure $\mathbb{Q}$ for $s \in \mathcal{H}^2_Q$ so that the market is arbitrage free. Under Assumption 2, the definitions of a contingent claim, a complete market, and an approximately complete market remain unchanged. The extended second fundamental theorem of asset pricing still holds. The same proof applies as in the original theorem. The original proof did not use sample path continuity, but only the condition that the intercept of the supply curve has totally inaccessible jumps.

Theorem 6.2 *(Second fundamental theorem) Assume Assumption 2. Let $S(\cdot, 0) = s$ have totally inaccessible jumps. Suppose there exists a* unique *probability measure $\mathbb{Q} \sim \mathbb{P}$ such that $S(\cdot, 0) = s$ is a $\mathbb{Q}$-local martingale. Then, the market is approximately complete.*

7 Conclusion

This paper extends classical arbitrage pricing theory to include liquidity risk by studying an economy where the security's price depends on the trade size. Extended first and second fundamental theorems of asset pricing are investigated. The economy is shown to be arbitrage free if and only if the stochastic process for the price of a marginal trade has an equivalent martingale probability measure. The second fundamental theory of asset pricing fails to hold in our setting. In our setting, a martingale measure (appropriately defined) can be unique, and markets still be incomplete. However, a weakening of the second fundamental theorem holds. Markets will be approximately complete in our setting if the martingale measure is unique. In an approximately complete market, derivative prices are shown to be equal to the classical arbitrage free price of the derivative.

Appendix

A.1 Proof of the first fundamental theorem

This theorem uses Assumption 1, sample path continuity of $S(t, x)$. The proof proceeds in two steps. Step 1 constructs a fictitious economy where all trades are executed at the marginal stock price. The theorem is true in this fictitious economy by the classical result. Step 2 then shows the theorem in this fictitious economy is sufficient to obtain the result in our economy.

Prior to this proof, we need to make the following observation in order to utilize the classical theory. The classical theory (see [11] or alternatively [28] for an expository version) has trading strategies starting with $X_0 = 0$, while we have trading strategies with $X_{0-} = 0$ but not $X_0 = 0$. Without loss of generality, in the subsequent proof, we can restrict ourselves to predictable processes with $X_0 = 0$. Here is the argument. Recall $s_u = S(u, 0)$. In our setup, choose Y^0 so that $X_0 S(0,0) + Y^0_0 = 0$ and $X_t S(t,0) + Y^0_t = X_0 S(0,0) + Y^0_0 + \int_{0+}^{T} X_u ds_u =$

$\int_{0+}^{T} X_u ds_u$. Define $\widehat{X} = 1_{(0,T]}X$. $\widehat{X}$ is predictable, $\widehat{X}_0 = 0$, and $\int_{0+}^{T} X_u ds_u = \int_0^T \widehat{X}_u ds_u$. The analysis can be done for $\widehat{X}$.

A.1.1 **Step 1.** The fictitious economy

Consider the fictitious economy introduced in Sect. 3. Delbaen and Schachermayer prove the following in Sect. 4 of [11]:

Theorem A.1 *Given Assumption 1 and no arbitrage, there is NFLVR in the fictitious economy if and only if there exists a measure $\mathbb{Q} \sim \mathbb{P}$ such that $S(\cdot, 0)$ is a $\mathbb{Q}$-local martingale.*

Since the stochastic integral of a predictable process can be written as a limit (uniformly on compacts in probability) of stochastic integrals with continuous and finite variation integrands (see Appendix A.3 below), we have the following corollary.[13]

Corollary A.1 *Suppose there is no arbitrage opportunity in the fictitious economy. Given Assumption 1, if there's an (FLVR) in the fictitious economy, there exists a sequence of ϵ_n-admissible trading strategies X^n, continuous and of FV, and a nonnegative F_T-measurable random variable f_0, not identically zero, such that $\epsilon_n \to 0$ and $(X^n \cdot S)_T \to f_0$ in probability.*

The proof of this corollary is straightforward, and hence we omit it.

A.1.2 **Step 2**. The illiquid economy

In the economy with liquidity risk, restricting consideration to s.f.t.s. (X, Y, τ) with X finite variation and continuous processes, by Lemma 2.1, we have that $Y_t = (X \cdot s)_t - X_t S(t, 0)$. At time T, we have $Y_T = (X \cdot s)_T$. This is the value of the same s.f.t.s. in the fictitious economy. We use this insight below.

Lemma A.1 *Given Assumption 1, let X be an α-admissible trading strategy which is continuous and of FV in the fictitious economy. Then there exists a sequence of $(\alpha + \epsilon_n)$-admissible trading strategies, in the illiquid economy, $(H^n, Y^n, \tau^n)_{n\geq 1}$ of FV and continuous on $[0, \tau^n)$, such that Y_T^n tends to $(X \cdot S)_T$, in probability, and $\epsilon_n \to 0$.*

Proof Let $T_n = T - \frac{1}{n}$. Define

$$f_n(t) = 1_{[T_n \leq t \leq T_{n+1}]} \frac{X_{T_n}}{T_n - T_{n+1}} (t - T_{n+1}) \tag{A.1}$$

so that $f_n(T_n) = X_{T_n}$ and $f_n(T_{n+1}) = 0$. Note that $f_n(t) \to 0, a.s., \forall t$. Define

$$X_t^n = X_t 1_{[t<T_n]} + f_n(t). \tag{A.2}$$

[13] In the original paper [11], there is a missing hypothesis in the statement of their theorem related to this corollary. We include here and in other results as needed the missing hypothesis of no arbitrage. We are grateful to Professor Delbaen for providing us with a counterexample that shows one does in fact need this hypothesis [10].

By this definition, X^n is continuous and of FV. Note that T is a fixed time and not a stopping time, so X^n is predictable. Moreover,

$$(X^n \cdot S)_t = (X \cdot S)_{t\wedge T_n} + \int_0^t f_n(s)dS(s,0). \tag{A.3}$$

Notice that $|f_n(\omega)| \leq \sup_t |X_t(\omega)| \equiv K(\omega) \in R$ since X is continuous on $[0,T]$. Thus, f_n is bounded by an $S(\cdot,0)$-integrable function. Therefore, by dominated convergence theorem for stochastic integrals (see [27], p.145) $\int f_n(s)dS(s,0)$ tends to 0 in u.c.p. on the compact time interval $[0,T]$, and therefore $X^n \cdot S \to X \cdot S$ in u.c.p. on $[0,T]$.[14]

Now, let $(\epsilon_n)_{n\geq 1}$ be a sequence of positive real numbers converging to 0 such that $\sum_n \epsilon_n < \infty$. Define $\tau^n = \inf\{t > 0 : (X^n \cdot S)_t < -\alpha - \epsilon_n\} \wedge T$. τ^n is a predictable stopping time by the continuity of $S(\cdot, x)$. Due to u.c.p convergence of $X^n \cdot S$ to $X \cdot S$, passing to a subsequence if necessary, we have the following:

$$\mathbb{P}\left(\sup_{0\leq t\leq T} |(X^n \cdot S)_t - (X \cdot S)_t| \geq \epsilon_n\right) \leq \epsilon_n. \tag{A.4}$$

Notice that $\mathbb{P}(\tau^n < T) \leq \epsilon_n$, i.e. $\tau^n \to T$ in probability. Moreover, $\tau^n \geq T_n$ because $X^n = X$ up to time T_n. Choose $H^n = X^n 1_{[0,\tau^n)}$. Consider the sequence of trading strategies $(H^n, \tau^n)_{n\geq 1}$. Note that $(H^n \cdot S)_t \geq -\alpha - \varepsilon_n$ for all $t \in [0,\tau^n]$ since $H^n_{\tau^n} = 0$ for all n. Therefore, $(H^n, \tau^n)_{n\geq 1}$ is a sequence of $(\alpha + \epsilon_n)$-admissible trading strategies. The value of the portfolio at liquidation for each trading strategy is given by

$$Y^n_{\tau^n} = X^n(\tau^n)\left[S\left(\tau^n, -X^n(\tau^n)\right) - S(\tau^n, 0)\right] + (X^n \cdot S)_{\tau^n} \tag{A.5}$$

since H^n is of FV and jumps only at τ^n for each n by the continuity of X^n. Therefore, it remains to show $X^n(\tau^n) \to 0$ in probability since this, together with $\tau^n \to T$ in probability, will prove the theorem. Indeed, $\sum_n \mathbb{P}(\tau^n < T) \leq \sum_n \epsilon_n < \infty$. Therefore, by by the first Borel-Cantelli lemma, $\mathbb{P}[\tau^n < T \text{ i.o.}] = 0$, which implies $X^n(\tau^n) = X^n(T) = 0$, with probability 1, for all but at most finitely many n. □

Lemma A.2 *Suppose there is no arbitrage opportunity in the fictitious economy. Given Assumption 1, there is NFLVR in the fictitious economy if and only if there is NFLVR in the illiquid economy.*

Proof Suppose there is NFLVR in the fictitious economy. Since, given any s.f.t.s. (X, Y, τ) in the illiquid economy, $Y_\tau \leq (X \cdot S)_\tau$, it follows there exists NFLVR in the illiquid economy. Conversely, suppose there is FLVR in the fictitious economy. In view of Corollary A.1, there's a sequence, $(X^n)_{n\geq 1}$, with each X^n continuous, of FV, and ϵ_n-admissible trading strategies such that $(X^n \cdot S)_T \to f_0$ in probability where f_0 is as before and $\epsilon_n \to 0$. However, by the previous lemma, there exists a sequence of α_n -admissible trading strategies, $(H^n, Y^n, \tau^n)_{n\geq 1}$, where $\alpha_n \to 0$, in the illiquid economy such that $Y^n_{\tau^n} \to f_0$ in probability, which gives an FLVR in the illiquid economy. □

[14] One can also show this using integration by parts.

Theorem A.2 *(First fundamental theorem) Suppose there is no arbitrage opportunity in the fictitious economy. Given Assumption 1, there is no free lunch with vanishing risk (NFLVR) in the illiquid economy if and only if there exists a measure* $\mathbb{Q} \sim \mathbb{P}$ *such that* $S(\cdot, 0)$ *is a* $\mathbb{Q}$*-local martingale.*

Proof By the previous lemma, (NFLVR) in the illiquid economy is equivalent to (NFLVR) in the fictitious economy, which is equivalent to existence of a martingale measure by Theorem A.1. □

A.2 Construction of the self-financing condition for a class of trading strategies

The purpose of this section is to provide justification for Definition 2.2 and Definition 6.1 in the text. This proof uses only the weaker hypotheses of Assumption 2.

Let t be a fixed time and let (σ_n) be a sequence of random partitions of $[0, t]$ tending to identity in the following form:

$$\sigma_n : 0 = T_0^n \leq T_1^n \leq \ldots \leq T_{k_n}^n = t$$

where T_k^n's are stopping times. For successive trading times, t_1 and t_2, the self-financing condition can be written as

$$Y_{t_2} - Y_{t_1} = -\left(X_{t_2} - X_{t_1}\right)\left[S(t_2, X_{t_2} - X_{t_1})\right].$$

Note that $Y_t = Y_0 + \sum_{k\geq 1}(Y_{T_k^n} - Y_{T_{k-1}^n})$ for all n. Therefore, we'll define Y_t to be the following limit whenever it exists:

$$Y_0 - \lim_{n\to\infty} \sum_{k\geq 1} (X_{T_k^n} - X_{T_{k-1}^n}) S(T_k^n, X_{T_k^n} - X_{T_{k-1}^n}). \tag{A.6}$$

Example A.1 In the classical case, $S(t, x) = S(t, 0)$ for all $x \in \mathbb{R}$. Thus, self-financing condition becomes

$$Y_{t_2} - Y_{t_1} = -[X_{t_2} - X_{t_1}]S(t_2, 0)$$

and initial trades must satisfy $Y(0) = -X(0)S(0, 0)$ instead. Therefore,

$$\begin{aligned}
Y_t &= Y_0 - \lim_{n\to\infty} \sum_{k\geq 1} (X_{T_k^n} - X_{T_{k-1}^n}) S(T_k^n, 0) \\
&= Y(0) - \lim_{n\to\infty} \left[\sum_{k\geq 1} X_{T_k^n} S(T_k^n, 0) - \sum_{k\geq 1} X_{T_{k-1}^n} S(T_k^n, 0) \right] \\
&= Y(0) - \lim_{n\to\infty} \Bigg[\sum_{k\geq 1} X_{T_k^n} S(T_k^n, 0) \\
&\qquad - \sum_{k\geq 1} X_{T_{k-1}^n} (S(T_k^n, 0) - S(T_{k-1}^n, 0)) - \sum_{k\geq 1} X_{T_{k-1}^n} S(T_{k-1}^n, 0) \Bigg]
\end{aligned}$$

$$
\begin{aligned}
&= Y_0 - X_t S(t,0) + X_0 S(0,0) \\
&\quad + \lim_{n\to\infty} \sum_{k\geq 1} X_{T^n_{k-1}} (S(T^n_k, 0) - S(T^n_{k-1}, 0)) \\
&= -X_t S(t,0) + \int_0^t X_{u-} dS(u,0).
\end{aligned}
$$

Notice that the limit agrees with the value of $Y(t)$ in classical case. So, we have a framework that contains the existing theory.

Theorem A.3 *For X càdlàg and has finite quadratic variation (QV), the value in the money market account is given by*

$$
\begin{aligned}
Y_t = &-X_t S(t,0) + \int_0^t X_{u-} dS(u,0) - \int_0^t S^{(1)}_x(u-,0) d[X,X]^c_u \\
&- \sum_{0\leq u\leq t} \left[S(u, \Delta X_u) - S(u,0)\right] \Delta X_u.
\end{aligned} \tag{A.7}
$$

where $S^{(n)}_x$ is the n-th partial derivative of S with respect to x.

Proof Expression (A.6) is

$$
\begin{aligned}
Y_t &= Y_0 \\
&\quad - \lim_{n\to\infty} \sum_{k\geq 1} \left(X_{T^n_k} - X_{T^n_{k-1}}\right) S(T^n_k, \left(X_{T^n_k} - X_{T^n_{k-1}}\right)) \\
&= -X(0) S(0, X_0) \\
&\quad - \lim_{n\to\infty} \sum_{k\geq 1} \left(X_{T^n_k} - X_{T^n_{k-1}}\right) \left[S(T^n_k, \left(X_{T^n_k} - X_{T^n_{k-1}}\right)) - S(T^n_k, 0)\right] \\
&\quad - \lim_{n\to\infty} \sum_{k\geq 1} \left(X_{T^n_k} - X_{T^n_{k-1}}\right) S(T^n_k, 0)).
\end{aligned}
$$

We know from Example A.1 that the last sum converges to $-X_0 S(0,0) + X_t S(t,0) - \int_0^t X_{u-} dS(u,0)$. Let $A = A(\epsilon, t)$ be a set of jumps of X that has a.s. a finite number of times s, and let $B = B(\epsilon, t)$ be such that $\sum_{s\in B} (\Delta X_s)^2 \leq \epsilon^2$, where A and B are disjoint and $A \cup B$ exhaust the jumps of X on $(0,t]$, see proof of Itô's formula in [27]. Thus,

$$
\begin{aligned}
&\lim_{n\to\infty} \sum_{k\geq 1} \left(X_{T^n_k} - X_{T^n_{k-1}}\right) \left[S(T^n_k, \left(X_{T^n_k} - X_{T^n_{k-1}}\right)) - S(T^n_k, 0)\right] \\
&= \lim_{n\to\infty} \sum_{k,A} \left(X_{T^n_k} - X_{T^n_{k-1}}\right) \left(S(T^n_k, \left(X_{T^n_k} - X_{T^n_{k-1}}\right)) - S(T^n_k, 0)\right) \\
&\quad + \lim_{n\to\infty} \sum_{k,B} \left(X_{T^n_k} - X_{T^n_{k-1}}\right) \left(S(T^n_k, \left(X_{T^n_k} - X_{T^n_{k-1}}\right)) - S(T^n_k, 0)\right)
\end{aligned}
$$

where $\sum_{k,A}$ denotes $\sum_{k\geq 1} 1_{[A\cap(T^n_{k-1},T^n_k]\neq\emptyset]}$, and $\sum_{k,B}$ denotes $\sum_{k\geq 1} 1_{[B\cap(T^n_{k-1},T^n_k]=\emptyset]}$ Since A has only finitely many elements, ω by ω, the first limit equals

$$\sum_{u\in A} \left[S(u,\Delta X_u) - S(u,0)\right]\Delta X_u. \tag{A.8}$$

Applying Taylor's formula to each $S(T^n_k,\cdot)$, the second limit becomes

$$\begin{aligned}
&\lim_{n\to\infty}\sum_{k,B} S^{(1)}_x(T^n_k,0)\left(X_{T^n_k}-X_{T^n_{k-1}}\right)^2\\
&+\lim_{n\to\infty}\sum_{k,B}\left(X_{T^n_k}-X_{T^n_{k-1}}\right)R\left(T^n_k,\left|X_{T^n_k}-X_{T^n_{k-1}}\right|\right)\\
=&\lim_{n\to\infty}\sum_{k\geq 1} S^{(1)}_x(T^n_k,0)\left(X_{T^n_k}-X_{T^n_{k-1}}\right)^2\\
&-\lim_{n\to\infty}\sum_{k,A} S^{(1)}_x(T^n_k,0)\left(X_{T^n_k}-X_{T^n_{k-1}}\right)^2\\
&+\lim_{n\to\infty}\sum_{k,B}\left(X_{T^n_k}-X_{T^n_{k-1}}\right)R\left(T^n_k,\left|X_{T^n_k}-X_{T^n_{k-1}}\right|\right)\\
=&\lim_{n\to\infty}\sum_{k\geq 1} S^{(1)}_x(T^n_{k-1},0)\left(X_{T^n_k}-X_{T^n_{k-1}}\right)^2\\
&+\lim_{n\to\infty}\sum_{k\geq 1}\left[S^{(1)}_x(T^n_k,0)-S^{(1)}_x(T^n_{k-1},0)\right]\left(X_{T^n_k}-X_{T^n_{k-1}}\right)^2\\
&-\lim_{n\to\infty}\sum_{k,A} S^{(1)}_x(T^n_k,0)\left(X_{T^n_k}-X_{T^n_{k-1}}\right)^2\\
&+\lim_{n\to\infty}\sum_{k,B}\left(X_{T^n_k}-X_{T^n_{k-1}}\right)R\left(T^n_k,\left|X_{T^n_k}-X_{T^n_{k-1}}\right|\right),
\end{aligned}\tag{A.9}$$

where R is the remainder term in Taylor's formula. The sum of the first three limits converges to[15]

$$\begin{aligned}
&\int_0^t S^{(1)}_x(u-,0)d[X,X]_u+\left[S^{(1)}_x(\cdot,0),[X,X]\right]_t-\sum_{u\in A} S^{(1)}_x(u,0)(\Delta X_u)^2\\
=&\int_0^t S^{(1)}_x(u-,0)d[X,X]_u+\sum_{0<u\leq t}\Delta S^{(1)}_x(u,0)(\Delta X_u)^2\\
&-\sum_{u\in A} S^{(1)}_x(u,0)(\Delta X_u)^2.
\end{aligned}\tag{A.10}$$

[15] Note that the assumption that $S^{(1)}_x(\cdot,0)$ has a finite QV is not needed when $S^{(1)}_x(\cdot,0)$ is continuous. In this case, the second limit is zero. This follows from the fact that X has a finite QV and $S^{(1)}_x(\cdot,0)$ is uniformly continuous, ω by ω, over the compact domain $[0,T]$.

Now we will show as ϵ tends to 0, the last term in (A.9) vanishes. Assuming temporarily that $|S_x^{(2)}| < K < \infty$ uniformly in x and t,

$$\begin{aligned}
&\left|R\left(T_k^n, \left|X_{T_k^n} - X_{T_{k-1}^n}\right|\right)\right| \\
\leq &\sup_{0\leq|x|\leq\left|X_{T_k^n}-X_{T_{k-1}^n}\right|} \left|S_x^{(1)}(T_k^n, x) - S_x^{(1)}(T_k^n, 0)\right| \left|\left(X_{T_k^n} - X_{T_{k-1}^n}\right)\right| \\
\leq &\sup_{0\leq|y|\leq|x|\leq\left|X_{T_k^n}-X_{T_{k-1}^n}\right|} \left|S_x^{(2)}(T_k^n, y)x\left(X_{T_k^n} - X_{T_{k-1}^n}\right)\right| \\
\leq &K\left(X_{T_k^n} - X_{T_{k-1}^n}\right)\left(X_{T_k^n} - X_{T_{k-1}^n}\right),
\end{aligned}$$

where the second inequality follows from the Mean Value Theorem. Therefore, the last sum in (A.9) is less than or equal to, in absolute value,

$$\begin{aligned}
&K \lim_{n\to\infty} \sum_{k,B} \left(\left|X_{T_k^n} - X_{T_{k-1}^n}\right|\right)^3 \\
< &K \lim_{n\to\infty} \sup_{k,B} |X_{T_k^n} - X_{T_{k-1}^n}| \sum_k \left(\left|X_{T_k^n} - X_{T_{k-1}^n}\right|\right)^2 \\
\leq &K\epsilon[X,X]_t.
\end{aligned}$$

Note that ϵ can be made arbitrarily small and X has a finite QV. Furthermore, since all summands are positive, as $\epsilon \to 0$, (A.8) converges to

$$\sum_{0<u\leq t} \left[S(u, \Delta X_u) - S(u, 0)\right] \Delta X_u$$

and (A.10) converges to

$$\begin{aligned}
&\int_0^t S_x^{(1)}(u-,0)d[X,X]_u + \sum_{0<u\leq t} \Delta S_x^{(1)}(u,0)(\Delta X_u)^2 \\
&- \sum_{0<u\leq t} S_x^{(1)}(u,0)(\Delta X_u)^2 \\
= &\int_0^t S_x^{(1)}(u-,0)d[X,X]_u - \sum_{0<u\leq t} S_x^{(1)}(u-,0)(\Delta X_u)^2 \\
= &\int_0^t S_x^{(1)}(u-,0)d[X,X]_u^c.
\end{aligned}$$

For the general case, let $V_k^x = \inf\{t > 0 : S^{(2)}(t,x) > k\}$. Define $\tilde{S}(t,x) := S(t,x)1_{[0,V_k^x)}$.Therefore, (A.7) holds for $\tilde{S}$, for each k. Now, a standard argument using set unions, as in the proof of Itô's formula in [27], establishes (A.7) for S. □

A.3 Approximating stochastic integrals with continuous and of FV integrands

The next lemma (Lemma A.3) is well known and can be found in ♣[?]♣.

Lemma A.3 *Let X be a special semimartingale with the canonical decomposition $X = N + A$, where N is a local martingale and A is predictable. Suppose S has totally inaccessible jumps. Then A is continuous.*

We make the following assumption. (Note that this assumption is satisfied in all classical market models studies, since [for example] a Lévy process has only totally inaccessible jumps, and indeed by a classic theorem of P. A. Meyer, all "reasonable" strong Markov processes have only totally inaccessible jumps.)

Assumption 3 *$S(\cdot, 0)$ has only totally inaccessible jumps.*

We recall a few definitions.

Definition A.1 *Let X be a special semimartingale with canonical decomposition $X = \bar{N} + \bar{A}$. The $\mathcal{H}^2$ norm of X is defined to be*

$$\|X\|_{\mathcal{H}^2} = \left\| \left[\bar{N}, \bar{N}\right]_\infty^{1/2} \right\|_{L^2} + \left\| \int_0^\infty \left| d\bar{A}_u \right| \right\|_{L^2}.$$

The space $\mathcal{H}^2$ of semimartingales consists of all special semimartingales with finite $\mathcal{H}^2$ norm.

Definition A.2 *The predictable σ-algebra $\mathcal{P}$ on $R_+ \times \Omega$ is the smallest σ-algebra making all processes in $\mathbb{L}$ measurable where $\mathbb{L}$ is the set of processes that have paths that are left continuous with right limits. We let $\mathbf{b}\mathcal{P}$ denote bounded processes that are $\mathcal{P}$-measurable.*

Definition A.3 *Let $X \in \mathcal{H}^2$ with $X = \bar{N} + \bar{A}$ its canonical decomposition, and let $H, J \in \mathbf{b}\mathcal{P}$. We define $d_X(H, J)$ by*

$$d_X(H,J) \equiv \left\| \left(\int_0^T (H_u - J_u)^2 \, d\left[\bar{N}, \bar{N}\right]_u \right)^{1/2} \right\|_{L^2} + \left\| \int_0^T |H_u - J_u| \left| d\bar{A}_u \right| \right\|_{L^2}$$

From here on, we suppose $s \in \mathcal{H}^2$ with the canonical decomposition $s = \bar{N} + \bar{A}$.

Theorem A.4 *Let $\epsilon > 0$. For any H bounded, continuous and of FV, there exists H^ϵ, bounded, continuous and of FV, with $H_T^\epsilon = 0$ such that $d_s\left(H, H^\epsilon\right) < \epsilon$.*

Proof Define

$$H_t^m = H_t 1_{[0,T_m]} + H_{T_m} \frac{T - t}{T - T_m} 1_{(T_m, T]}$$

where $T_m = T - \frac{1}{m}$. We'll first show $d_s\left(H, H 1_{[0,T_m]}\right) \to 0$ as $m \to \infty$.

To show $\left\| \left(\int_0^T \left(H_u(\omega) - H_u(\omega) 1_{[0,T_m]} \right)^2 d \left[\bar{N}, \bar{N} \right]_u (\omega) \right)^{1/2} \right\|_{L^2} \to 0$, first observe that $[\bar{N}, \bar{N}] = \langle \bar{N}, \bar{N} \rangle + M$, where $\langle \bar{N}, \bar{N} \rangle$ is the compensator, hence predictable, of $[\bar{N}, \bar{N}]$ and M is a local martingale. Since M is a local martingale, there exists a sequence $(T_n)_{n \geq 1}$ of stopping times increasing to ∞ such that M^{T_n} is a martingale for each n. Thus, given a bounded G, $G \cdot M^{T_n}$ is a martingale implying $\mathbb{E}[(G \cdot M^{T_n})_t] = 0$ for all t. Moreover,

$$\begin{aligned} |(G \cdot M^{T_n})_t| &\leq |G| \cdot [\bar{N}, \bar{N}]_t^{T_n} + |G| \cdot \langle \bar{N}, \bar{N} \rangle_t^{T_n} \\ &\leq |G| \cdot [\bar{N}, \bar{N}]_t + |G| \cdot \langle \bar{N}, \bar{N} \rangle_t \end{aligned} \tag{A.11}$$

where the first equality is the triangle inequality and the second follows from $[\bar{N}, \bar{N}]$ and $\langle \bar{N}, \bar{N} \rangle$ being increasing. Furthermore, $G 1_{[0,T_n]}$ converges to G hence by Dominated Convergence Theorem for stochastic integrals, $G \cdot M^{T_n}$ converges to $G \cdot M$ in ucp. Moreover, by (A.11), since G is bounded and $[\bar{N}, \bar{N}]$ and $\langle \bar{N}, \bar{N} \rangle$ are integrable, $\mathbb{E}[(G \cdot M^{T_n})_t]$ converges to $\mathbb{E}[(G \cdot M)_t]$ by ordinary Dominated Convergence Theorem. Therefore, $\mathbb{E}[(G \cdot M)_t] = 0$ for all t. Hence, we have

$$\mathbb{E}\left[G \cdot [\bar{N}, \bar{N}]_t \right] = \mathbb{E}\left[G \cdot \langle \bar{N}, \bar{N} \rangle_t \right].$$

Jump times of $[\bar{N}, \bar{N}]$ are those of $\bar{N}$, which are totally inaccessible as a corollary to the previous lemma. Therefore, by the same lemma, $\langle \bar{N}, \bar{N} \rangle$ is continuous. Now,

$$\int_0^T \left(H_u(\omega) - H_u(\omega) 1_{[0,T_m]} \right)^2 d \langle \bar{N}, \bar{N} \rangle_u (\omega) \leq \int_0^T \left(H_u(\omega) \right)^2 d \langle \bar{N}, \bar{N} \rangle_u (\omega) < \infty,$$

for all m, for almost all ω. Thus, by Lebesgue's Dominated Convergence Theorem

$$\int_0^T \left(H_u(\omega) - H_u(\omega) 1_{[0,T_m]} \right)^2 d \langle \bar{N}, \bar{N} \rangle_u (\omega) \to 0, \text{ a.s..}$$

since $\langle \bar{N}, \bar{N} \rangle$ is continuous. Moreover,

$$\left\| \left(\left(H - H 1_{[0,T_m]} \right)^2 \cdot file \left\langle \bar{N}, \bar{N} \right\rangle \right)^{1/2} \right\|_{L^2} \leq \left\| \left(H^2 \cdot \left\langle \bar{N}, \bar{N} \right\rangle \right)^{1/2} \right\|_{L^2} < \infty$$

since $H \cdot s \in \mathcal{H}^2$. A second application of Dominated Convergence Theorem yields

$$\left\| \left(\int_0^T \left(H_u(\omega) - H_u(\omega) 1_{[0,T_m]} \right)^2 d \left\langle \bar{N}, \bar{N} \right\rangle_u (\omega) \right)^{1/2} \right\|_{L^2} \to 0.$$

Since, for any bounded $|G|$, $\mathbb{E}\left[G \cdot [\bar{N}, \bar{N}]_t \right] = \mathbb{E}\left[G \cdot \langle \bar{N}, \bar{N} \rangle_t \right]$, for all t,

$$\left\| \left(\int_0^T \left(H_u(\omega) - H_u(\omega) 1_{[0,T_m]} \right)^2 d \left[\bar{N}, \bar{N} \right]_u (\omega) \right)^{1/2} \right\|_{L^2} \to 0,$$

too. By the previous lemma, $\bar{A}$ is continuous as well, so $\left\| \int_0^T \left| H_u - H_u 1_{[0,T_m]} \right| \left| d\bar{A}_u \right| \right\|_{L^2} \to 0$ by a similar argument. Hence, $d_s\left(H, H1_{[0,T_m]}\right) \to 0$ as $m \to \infty$.

It remains to show $d_s\left(H_{T_m} \frac{T-t}{T-T_m} 1_{(T_m,T]}, 0\right) \to 0$, as $m \to \infty$. First note that

$$\begin{aligned} &\int_0^T H_{T_m}^2(\omega) \left(\frac{T-u}{T-T_m}\right)^2 1_{(T_m,T]} d\left\langle \bar{N}, \bar{N} \right\rangle_u (\omega) \\ &\leq \int_0^T K d\left\langle \bar{N}, \bar{N} \right\rangle_u (\omega) < \infty \end{aligned}$$

where $K = \left\| \max_{0 \leq t \leq T} H_t^2(\omega) \right\|_\infty < \infty$ since H is bounded. Thus, by the Dominated Convergence Theorem,

$$\int_0^T H_{T_m}^2(\omega) \left(\frac{T-u}{T-T_m}\right)^2 1_{(T_m,T]} d\left\langle \bar{N}, \bar{N} \right\rangle_u (\omega) \to 0, \text{ a.s..}$$

Moreover, another application of the Dominated Convergence Theorem yields

$$\lim_{m \to \infty} \mathbb{E}\left[\int_0^T H_{T_m}^2 \left(\frac{T-u}{T-T_m}\right)^2 1_{(T_m,T]} d\left\langle \bar{N}, \bar{N} \right\rangle_u \right] = 0.$$

A similar argument shows

$$\left\| \int_0^T \left| H_{T_m} \left(\frac{T-u}{T-T_m}\right) \right| 1_{(T_m,T]} \left| dA_u \right| \right\|_{L^2} \to 0$$

which completes the proof. □

Corollary A.2 *Let $\epsilon > 0$. For any H, bounded, continuous and of FV, there exists H^ϵ, bounded, continuous and of FV, with $H_T^\epsilon = 0$ such that $\|H \cdot s - H^\epsilon \cdot s\|_{L^2} < \epsilon$.*

Proof This follows from a combination of Theorem A.4 and Theorem 5 of Chapter IV in [27]. □

A.4 A Numéraire invariance theorem

This proof only uses the weaker supply curve Assumption 2. Let $(S(\cdot,x)_{x \in \mathbb{R}}, B)$ represent an economy where $S(\cdot,x)$ is the stock price process per share for an order of size x and B is a strictly positive semimartingale representing bond price. We define the discounted stock price process $\bar{S}(\cdot,x) = \frac{S(\cdot,x)}{B}$. We have shown that for a trading strategy $(\bar{X}, \bar{Y}, \bar{\tau})$ in the economy $(\bar{S}(\cdot,x)_{x \in \mathbb{R}}, 1)$, the self-financing condition is

$$\bar{Y}_t = (\bar{X}_- \cdot \bar{s})_t - \bar{X}_t \bar{s}_t - \bar{L}_t^1 - \bar{L}_t^2, \qquad \text{for } 0 \leq t \leq \bar{\tau} \tag{A.12}$$

where $\bar{s} := \bar{S}(\cdot,0)$, $\bar{L}^1_t = \sum_{0\le u\le t} \Delta X_u[\bar{S}(u,\Delta X_u) - \bar{S}(u,0)]$ and $\bar{L}^2_t = \int_0^t \frac{\partial \bar{S}}{\partial x}(u-,0)d[X,X]^c_u$. Using similar arguments, one can show that for a trading strategy, (X,Y,τ) in the economy $(S(\cdot,x)_{x\in\mathbb{R}},B)$, the self-financing condition is

$$B_tY_t = (X_-\cdot s)_t - X_ts_t - L^1_t - L^2_t + (Y_-\cdot B)_t, \qquad \text{for } 0\le t\le\tau, \quad \text{(A.13)}$$

where s, L^1 and L^2 are defined similarly. L^1 and $\bar{L}^1$ are submartingales being increasing processes, thus their càdlàg versions exist. We'll use these càdlàg versions and therefore they are semimartingales of finite variation (FV) and, thus, Y being càdlàg is justified. The proof below is similar to the theorem on numéraire invariance for the classical case, which can be found in [14], or alternatively in [28] for the most general classical setting.

Theorem A.5 *Let X be a predictable càdlàg process with finite QV and τ be a predictable stopping time. (X,Y,τ) is a s.f.t.s. in the economy $(S(\cdot,x)_{x\in\mathbb{R}},B)$ if and only if it is a s.f.t.s. in $(\bar{S}(\cdot,x)_{x\in\mathbb{R}},1)$.*

Proof Given X and τ, (X,Y,τ) is a s.f.t.s. in $(S(\cdot,x)_{x\in\mathbb{R}},B)$ if Y satisfies (A.13). Similarly, $(X,\bar{Y},\tau)$ is a s.f.t.s. in $(\bar{S}(\cdot,x)_{x\in\mathbb{R}},1)$ if $\bar{Y}$ satisfies (A.12). Moreover, Y and $\bar{Y}$ are uniquely determined given X and τ. Therefore, it remains to show $Y=\bar{Y}$. Note that L^1 and L^2 are of FV. Thus,

$$\begin{aligned}
\frac{1}{B}L^1 &= L^1_-\cdot\frac{1}{B} + \frac{1}{B}\cdot L^1\\
&= L^1_-\cdot\frac{1}{B} + \sum_{0\le u\le t}\Delta X_u[S(u,\Delta X_u)-S(u,0)]\frac{1}{B_u}\\
&= L^1_-\cdot\frac{1}{B} + \bar{L}^1\\
\frac{1}{B}L^2 &= L^2\cdot\frac{1}{B} + \frac{1}{B}\cdot L^2 = L^2\cdot\frac{1}{B} + \frac{\partial\bar{S}}{\partial x}(\cdot,0)\cdot[X,X]^c\\
&= L^2\cdot\frac{1}{B} + \bar{L}^2\\
\frac{1}{B}(Y_-\cdot B) &= \frac{1}{B_-}\cdot(Y_-\cdot B) + (Y_-\cdot B)_-\cdot\frac{1}{B} + \left[\frac{1}{B},Y_-\cdot B\right]\\
&= \frac{1}{B_-}Y_-\cdot B + (Y_-\cdot B)_-\cdot\frac{1}{B} + \left[\frac{1}{B},Y_-\cdot B\right]\\
\frac{1}{B}(X_-\cdot s) &= \frac{1}{B_-}X_-\cdot s + (X_-\cdot s)_-\cdot\frac{1}{B} + \left[\frac{1}{B},X_-\cdot s\right]\\
&= X_-\cdot(\frac{1}{B_-}\cdot s) + (X_-\cdot s)_-\cdot\frac{1}{B} + \left[\frac{1}{B},X_-\cdot s\right]\\
&= X_-\cdot\bar{s} - X_-s_-\cdot\frac{1}{B} - X_-\cdot\left[\frac{1}{B},s\right]\\
&\quad +(X_-\cdot s)_-\cdot\frac{1}{B} + \left[\frac{1}{B},X_-\cdot s\right]
\end{aligned}$$

$$
\begin{aligned}
&= X_- \cdot \bar{s} - X_- s_- \cdot \frac{1}{B} - \left[\frac{1}{B}, X_- \cdot s\right] \\
&\quad + (X_- \cdot s)_- \cdot \frac{1}{B} + \left[\frac{1}{B}, X_- \cdot s\right] \\
&= X_- \cdot \bar{s} - X_- s_- \cdot \frac{1}{B} + (X_- \cdot s)_- \cdot \frac{1}{B}
\end{aligned}
$$

Dividing both sides of expression (A.13) by B and by using the above expressions and rearranging the terms yield

$$
\begin{aligned}
Y &= X_- \cdot \bar{s} - X\bar{s} - \bar{L}^1 - \bar{L}^2 \\
&\quad + \left((X_- \cdot s)_- - X_- s_- - L^1_- - L^2 + (Y_- \cdot B)_-\right) \cdot \frac{1}{B} \\
&\quad + \frac{1}{B_-} Y_- \cdot B + \left[\frac{1}{B}, Y_- \cdot B\right] \\
&= \bar{Y} + B_- Y_- \cdot \frac{1}{B} + \frac{1}{B_-} Y_- \cdot B + \left[\frac{1}{B}, Y_- \cdot B\right] \\
&= \bar{Y} + Y_- \cdot \left(B_- \cdot \frac{1}{B} + \frac{1}{B_-} \cdot B + \left[\frac{1}{B}, B\right]\right) \\
&= \bar{Y} + Y_- \cdot (1) \\
&= \bar{Y}
\end{aligned}
$$

□

References

1. Back, K.: Asymmetric information and options. Rev. Financial Stud. **6**(3), 435–472 (1993)
2. Bank, P., Baum, D.: Hedging and portfolio optimization in illiquid financial markets with a large trader. Mathe. Finance (forthcoming) (2002)
3. Barles, G., Soner, H.: Option pricing with transaction costs and a nonlinear Black-Scholes equation. Finance Stochast. **2**, 369–397 (1998)
4. Çetin, U.: Default and liquidity risk modeling. Ph.D. thesis, Cornell University 2003
5. Çetin, U., Jarrow, R., Protter, P., Warachka, M.: Option pricing with liquidity risk. Working paper, Cornell University 2002
6. Constantinides, G., Zariphopoulou, T.: Bounds on prices of contingent claims in an intertemporal economy with proportional transaction costs and general preferences. Finance Stochast. **3**, 345–369 (1999)
7. Cvitanic, J., Karatzas, I.: Hedging and portfolio optimization under transaction costs: a martingale approach. Math. Finance **6**, 133–165 (1996)
8. Cvitanic, J, Ma, J.: Hedging options for a large investor and forward-backward SDEs. Ann. Appl. Probab. **6**, 370–398 (1996)
9. Cvitanic, J., Pham, H., Touze, N.: A closed-form solution to the problem of super-replication under transaction costs. Finance Stochast. **3**, 35–54 (1999)
10. Delbaen, F.: Personal communication 2003
11. Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Math. Ann. **300**, 463–520 (1994)
12. Duffie, D.: Dynamic asset pricing theory, 2nd edn. New Jersey: Princeton University Press 1996
13. El Karoui, N.: Sur les Montées des semi-martingales. Astérisque **52–53**, 63–72 (1978)

14. Geman, H., El Karoui, N., Rochet, J.-C.: Changes of Numéraire, change of probability measure and option pricing. J. Appl. Probab. **32**, 443–458 (1995)
15. Glosten, L., Milgrom, P.: Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. J. Financial Econ. 14(March), 71–100 (1985)
16. Grossman, S., Miller, M.: Liquidity and market structure. J. Finance **43**(3), 617–637 (1988)
17. Harrison, J.M., Pliska, S.: Martingales and stochastic integrals in the theory of continuous trading. Stochast. Proc. Appl. **11**, 215–260 (1981)
18. Heath, D., Jarrow, R., Morton, A.: Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. Econometrica **60**(1), 77–105 (1992)
19. Jarrow, R.: Market manipulation, bubbles, corners and short squeezes. J. Financial Q. Analy.(September), 311–336 (1992)
20. Jarrow, R.: Derivative security markets, market manipulation and option pricing. J. Financial Q. Analy. **29**(2), 241–261 (1994)
21. Jarrow, R.: Default parameter estimation using market prices. Financial Analy. September/October, 75–92 (2001)
22. Jarrow, R., Turnbull, S.: An integrated approach to the hedging and pricing of eurodollar derivatives. J. Risk Insurance **64**(2), 271–299 (1997)
23. Jouini, E.: Price functionals with bid-ask spreads: an axiomatic approach. J. Math. Econ. **34**(4), 547–558 (2000)
24. Jouini, E., Kallal, H.: Martingales and arbitrage in securities markets with transaction costs. J. Econ. Theory **66**(1), 178–197 (1995)
25. Jouini, E., Kallal, H., Napp, C.: Arbitrage in financial markets with fixed costs. J. Math. Econ. **35**(2), 197–221 (2001)
26. Kyle, A.: Continuous auctions and insider trading. Econometrica **53**, 1315–1335 (1985)
27. Protter, P.: Stochastic integration and differential equations, 2nd edn. Berlin Heidelberg New York: Springer 2004
28. Protter, P.: A partial introduction to financial asset pricing theory. Stochast. Proc. Appl. **91**, 169–203 (2001)
29. Soner, H.M., Shreve, S., Cvitanic, J.: There is no nontrivial hedging portfolio for option pricing with transaction costs. Ann. Appl. Probab. **5**, 327–355 (1995)