Liquidity Risk and Option Pricing Theory

Robert A. Jarrow,[*] and Philip Protter [†]

February 2, 2005

Abstract

This paper summarizes the recent advances of Çetin [6], Çetin, Jarrow and Protter [7], Çetin, Jarrow, Protter and Warachka [8], Blais [4], and Blais and Protter [5] on the inclusion of liquidity risk into option pricing theory. This research provides new insights into the relevance of the classical techniques used in continuous time finance for practical risk management.

1 Introduction

Classical asset pricing theory assumes that traders act as price takers, that is, the theory assumes that investors' trades have no impact on the prices paid or received. The relaxation of this price taking assumption and its impact on realized returns in asset pricing models is called *liquidity risk*. Liquidity risk has been extensively studied in the market microstructure literature, but not in the asset pricing literature. In the market microstructure literature, it is well-known that a quantity impact on prices can be due to asymmetric information or differential risk tolerances (see Kyle [29], Glosten and Milgrom [20], or Grossman and Miller [21] in this regard). In an extreme form, liquidity risk has also been studied in the market manipulation literature (see Cvitanic and Ma [12], Jarrow [24], and Bank and Baum [2]). And, as argued in Çetin [6], liquidity risk is related to the transaction costs literature because

[*]Johnson Graduate School of Management, Cornell University, Ithaca, NY, 14853.

[†]Supported in part by NSF grant DMS-0202958 and NSA grant MDA-904-03-1-0092; ORIE – 219 Rhodes Hall, Cornell University, Ithaca, NY 14853-3801 USA

transaction costs induce a quantity impact on prices paid/received (see also Barles and Soner [3], Constantinides and Zariphopoulou [10], Cvitanic and Karatzas [11], Cvitanic, Pham, Touzi [13], Jouini [26], Jouini and Kallal [27], Jouini, Kallal, Napp [28], Soner, Shreve and Cvitanic [32] in this regard).

The purpose of this paper is to review the recent research of Çetin [6], Çetin, Jarrow and Protter [7], and Çetin, Jarrow, Protter and Warachka [8] on the inclusion of liquidity risk into option pricing theory, and the recent of results of Blais and Protter [4] and [5] where these results are interpreted through an analysis of book data. This approach embeds liquidity risk into the classical theory by having investors act as price takers with respect to a C^2 supply curve for the shares. In essence, instead of a single price for all shares traded, investors face a twice continuously differentiable price/quantity schedule. In this framework, it is assumed that the quantity impact on the price transacted is temporary.[1] Given this extension, Çetin, Jarrow and Protter [7] show that appropriate generalizations of the first and second fundamental theorems of asset pricing hold. Briefly stated, in this model, markets are arbitrage free if and only if there exists an equivalent martingale measure. In addition, markets will be approximately complete (in the L^2 sense), if the martingale measure is unique. The converse of this last implication does not hold.

The first and second fundamental theorems extend in this model due to the fact that trading strategies that are both continuous and of finite variation can approximate (in the L^2 sense) arbitrary predictable trading strategies. And, these continuous and finite variation trading strategies can be shown to avoid all 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 used to replicate the option. Instead, a continuous and finite variation approximation will be used. Both of these observations are consistent with industry usage of the classical arbitrage free pricing methodology. But, they have another set of strong implications for practice.

If one is interested in understanding the quantity impact of trades on prices in options markets as well, then this theory does not readily apply. Indeed, under the C^2 supply curve with continuous trading strategies, all

[1]Permanent quantity impacts on prices relates to the previously cited market manipulation literature and it is not studied here.

liquidity costs can be avoided when trading in the underlying shares. Although liquidity costs exist, they are non-binding. Consequently, there can be no quantity impact on option prices in such an economy,[2] otherwise arbitrage opportunities exist. To accommodate upward sloping supply curves for options, either the supply curve for the stock must have a discontinuity at 0 (it must violate the C^2 hypothesis) or continuous trading strategies must be excluded. Both extensions are possible. The first extension relates to the transaction cost literature (see Çetin [6]). The second extension is investigated in Çetin, Jarrow, Protter and Warachka [8]. This second extension is important because continuous trading strategies are not feasible in practice, and only approximating simple trading strategies can be applied. Yet, for simple trading strategies, liquidity costs are binding. This liquidity cost impact implies that the markets are no longer complete, and exact replication is not possible, implying an upward sloping supply curve for options can exist. Çetin, Jarrow, Protter and Warachka [8] show, in this context, how to super-replication options with minimum liquidity costs. The cost of the super-replication strategy provides an upper bound on the supply curve for the option market.

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 investigates a model with supply curves for options, Section 7 relates the supply curve formulation to transaction costs, Section 8 discusses examples inspired by an analysis of data, and Section 9 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 with no dividends. Also traded is a money market account that accumulates value at the spot rate of interest. Without loss of generality, we assume that the

[2]Recall that the previous theory implies that there is a *unique* price for an option (long or short), independent of the quantity of shares in the underlying traded.

spot rate of interest is zero, so that the money market account has unit value for all times.[3]

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.

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.[4] 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.

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

[3]A numéraire invariance theorem is proved in Çetin, Jarrow and Protter [7].

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

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 [29], Glosten and Milgrom [20], Grossman and Miller [21]). It includes, as a special case, horizontal supply curves.[5]

Example 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* $f : R^{n+2} \rightarrow R^+$ *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).*

2.2 Trading Strategies

We start by defining the investor's trading strategy.

Definition 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* $(\mathcal{F}_t : 0 \leq t \leq T)$ *stopping time with* $\tau \leq T$ *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).

[5] This structure can also be viewed as a generalization of the model in Jouini [26] where the traded securities have distinct selling and buying prices following separate stochastic processes.

Definition 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* $0 < t \leq T$*,*

$$\begin{aligned} Y_t &= Y_0 + X_0 S(0, X_0) + \int_0^t X_{u-} dS(u,0) - X_t S(t,0) \\ &\quad - \sum_{0 \leq u \leq t} \Delta X_u [S(u, \Delta X_u) - S(u,0)] - \int_0^t \frac{\partial S}{\partial x}(u,0) d[X,X]_u^c. \end{aligned} \tag{1}$$

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 (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 (1) yields the following simplified form of the self-financing condition:

$$\begin{aligned} Y_t + X_t S(t,0) &= \int_0^t X_{u-} dS(u,0) \\ &\quad - \sum_{0 \le u \le t} \Delta X_u [S(u, \Delta X_u) - S(u,0)] \\ &\quad - \int_0^t \frac{\partial S}{\partial x}(u,0) d[X,X]^c_u \quad \text{for } 0 \le t \le T. \end{aligned} \tag{2}$$

The left side of expression (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.

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 ($Y_t + X_t S(t,0)$).[6] 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), 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.

[6] 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 $X_\tau = 0$.

Definition 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.

Lemma 1 *(Equivalent Characterization of the Liquidity Costs).*

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

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.

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 [24]. Using this insight, an arbitrage opportunity can now be defined.

Definition 4 *An arbitrage opportunity is a s.f.t.s.* (X, Y, τ) *such that* $\mathbb{P}\{Y_T \ge 0\} = 1$ *and* $\mathbb{P}\{Y_T > 0\} > 0$.

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 $\alpha \ge 0$, let $\Theta_\alpha \equiv \{$s.f.t.s $(X, Y, \tau) \mid (X_- \cdot s)_t \ge -\alpha$ for all t almost surely $\}$.

Definition 5 *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 2 *($Y_t + X_t S(t, 0)$ 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 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 [16]). But, by Lemma 1, L_t is non-negative and non-decreasing. Therefore, $Y_t + X_t S(t, 0)$ is a supermartingale too.

Theorem 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 2, $Y_t + X_t S(t, 0)$ 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 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 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 [15]. This will require a preliminary definition.

Definition 6 *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.*[7]

[7] Delbaen and Schachermayer [15] Proposition 3.6 page 477 shows that this definition is equivalent to FLVR in the classical economy.

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. (X, Y^0, τ) 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 $Y_t^0 = (X \cdot s)_t - X_t s_t$ 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 (1)). So, in this fictitious economy, our definitions of an arbitrage opportunity, an admissible trading strategy, and a NFLVR collapse to those in [15].

Theorem 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 with the exception that condition (b) is removed. That is, a s.f.t.s. need not have zero initial value $(Y_0 + X_0 S(0, X_0) \neq 0)$.[8]

[8] 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.

To proceed, we need to define the space $\mathcal{H}_Q^2$ 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

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

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

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

Definition 7 *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 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[9] of K and maturity $T_0 \leq T$.[10] 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.

[9]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.

[10]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.

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.[11]

Definition 8 *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. (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

[11] 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.

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 1):

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

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[12] $\int_0^T X_{u-} ds_u = \int_0^T X_u ds_u$ 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

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

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

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

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 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.*

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

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.*

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 3 *(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*

$$Y_T^n = Y_0^n + X_0^n S(0, X_0^n) + \int_0^T X_{u-}^n ds_u - X_T^n S(T, 0) - L_T^n \to c + \int_0^T X_u ds_u = C \tag{3}$$

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 [30] 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 ^{2} and $\overline{X}^k = X1_{\{|X|\leq k\}}$, 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 [30] in this respect.)

Furthermore, Theorem 12 and Corollary 3 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 (1). Let $\tau^n = T$ for all n. Then, the sequence $(X^n, Y^n, \tau^n)_{n\geq 1}$ will satisfy (3). 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 9 *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 3 *(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. (X, Y^0, τ) 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 [22]) 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 Section 3 of Chapter IV of [30] 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 10 *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 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 3, 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 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. This example along with some empirical evidence regarding the pricing of traded options in the extended Black-Scholes economy can be found in Çetin, Jarrow, Protter, Warachka [8].

5.1 The Economy

Let

$$S(t,x) = e^{\alpha x} S(t,0) \quad \textit{with } \alpha > 0 \tag{4}$$

$$S(t,0) \equiv \frac{s_0 e^{\mu t + \sigma W_t}}{e^{rt}} \tag{5}$$

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). Expressions (4) and (5) characterize an extended Black-Scholes economy. It is easy to check that this supply curve satisfies Assumption 1 in Section 2.

Under these assumptions, there exists a unique martingale measure for $S(\cdot,0) = s$, see Duffie [16]. 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[13] to the option at time T is selected to be $C_T = \max[S(T,0) - K, 0]$.

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

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

$$X_t = N(h(t)). \tag{6}$$

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)):[14]

$$L_T = X_0(S(0,X_0) - S(0,0)) + \int_0^T \frac{\alpha\left(N'\left(h(u)\right)\right)^2 s_u}{T-u}du. \tag{7}$$

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

$$X_t^n = 1_{[\frac{1}{n}, T-\frac{1}{n})}(t) n \int_{(t-\frac{1}{n})^+}^t N(h(u))du, \text{ if } 0 \leq t \leq T - \frac{1}{n}, \tag{8}$$

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

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. $Y_T^n = Y_0^n + \int_0^T X_{u-}^n ds_u \to C_T = \max[S(T,0) - K, 0]$ in $L^2(d\mathbb{Q})$.

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

6 Economies with Supply Curves for Derivatives

Extended first and second fundamental theorems hold in the above economy, with a C^2 supply curve for the stock and allowing continuous trading strategies, consequently, there is a unique price for any option on the stock. This implies that the supply curve for trading an option is horizontal, exhibiting no quantity impact on the price. Otherwise, there would exist arbitrage opportunities (given trading in options and the stock). This is inconsistent with practice. And, it seems that any model analyzing liquidity risk, should imply supply curves for both stocks and options.

The reason they exist for stocks, but not options in the above model, is that continuous trading strategies of finite variation enable the investor to avoid all liquidity costs in the stock. Hence, although liquidity costs exist, they are non-binding, and the classical theory still applies (albeit in a modified and approximate manner). To make liquidity costs binding (as they are in practice), one must either remove the C^2 condition or disallow continuous trading strategies. The removal of the C^2 condition has been studied in the transaction cost literature (see Çetin [6], Barles and Soner [3], Constantinides and Zariphopoulou [10], Cvitanic and Karatzas [11], Cvitanic, Pham, Touzi [13], Jouini [26], Jouini and Kallal [27], Jouini, Kallal, Napp [28], Soner, Shreve and Cvitanic [32]) and will be discussed here directly in section 7, and in the context of estimating supply curves when summarizing Marcel Blais [4] and M. Blais and P. Protter [5] in section 8 below. The exclusion of continuous trading strategies, but still retaining the C^2 condition, has been studied by Çetin, Jarrow, Protter and Warachka [8]. This exclusion is consistent with practice because continuous trading strategies are impossible to utilize, except as approximations via simple trading strategies. But, with simple trading strategies, liquidity costs are binding. We discuss this extension next.

We modify the previous theory by considering only the class of the *discrete trading strategies* defined as any simple s.f.t.s. X_t where

$$X_t \in \left\{ x_{\tau_0} 1_{\{\tau_0\}} + \sum_{j=1}^{N} x_{\tau_j} 1_{(\tau_{j-1},\tau_j]} \;\middle|\; \begin{array}{l} 1.\ \tau_j \text{ are } \mathbb{F} \text{ stopping times for each } j, \\ 2.\ x_{\tau_j} \text{ is in } \mathcal{F}_{\tau_{j-1}} \text{ for each } j \text{ (predictable)}, \\ 3.\ \tau_0 \equiv 0 \text{ and } \tau_j > \tau_{j-1} + \delta \text{ for a fixed } \delta > 0\,. \end{array} \right\}$$

These trading strategies are discontinuous because once a trade is executed, the subsequent trade is separated by a minimum of $\delta > 0$ time units, as in Cheridito [9]. For the remainder of the paper, lower case values x and y denote discrete trading strategies.

By restricting the class of trading strategies in this manner, we retain an arbitrage-free environment (the extended first fundamental theorem still applies), although the minimum distance δ between trades prevents the market from being approximately complete. In an incomplete (not approximately complete market), the cost of replicating an option depends on the chosen trading strategy. Hence, the extended second fundamental theorem fails. This failure implies that there can be a quantity impact on the price of an option, i.e. the supply curve for an option need not be horizontal.

To investigate no arbitrage constraints on this supply curve, we can study the super-replication of options via the use of discrete trading strategies. For any discrete trading strategy, the liquidity cost equals

$$L_T = \sum_{j=0}^{N} \left[x_{\tau_{j+1}} - x_{\tau_j}\right] \left[S(\tau_j, x_{\tau_{j+1}} - x_{\tau_j}) - S(\tau_j, 0)\right] . \tag{10}$$

For a discrete trading strategy with $x_T = 0$, the hedging error is given by

$$C_T - Y_T = C_T - \left[y_0 + x_0 S(0,0) + \sum_{j=0}^{N-1} x_{\tau_{j+1}} \left[S(\tau_{j+1}, 0) - S(\tau_j, 0)\right] \right] + L_T \, .$$

Thus, there are two components to this hedging error. The first quantity,

$$\left[y_0 + x_0 S(0,0) + \sum_{j=0}^{N-1} x_{\tau_{j+1}} \left[S(\tau_{j+1}, 0) - S(\tau_j, 0)\right] \right] - C_T \, , \tag{11}$$

is the error in replicating the option's payoff C_T. The second component is the *liquidity cost* L_T defined in equation (10).

To provide an upper bound on the price a particular quantity of options, one can investigate the minimum cost of super-replication. For a *single* call option on the stock, this cost can be obtained as follows. Define $Z_t = X_t S(t,0) + Y_t$ as the time t marked-to-market value of the replicating portfolio. The optimization problem is:

$$\min_{(X,Y)} Z_0 \quad \text{s.t.} \quad Z_T \geq C_T = \max\{S(T,0) - Ke^{-rT}, 0\} \tag{12}$$

where

$$Z_T = y_0 + x_0 S(0,0) + \sum_{j=0}^{N-1} x_{\tau_{j+1}} [S(\tau_{j+1}, 0) - S(\tau_j, 0)] - L_T \,.$$

The solution to this problem requires a numerical approximation. One such numerical procedure involving the binomial approximation is discussed in Çetin, Jarrow, Protter and Warachka [8].[15]

Since liquidity costs in the underlying stock are quantity dependent, the cost of super-replication will also be quantity dependent. The cost of super-replicating a number of shares of the option then provides an upper bound on the entire supply curve for the option (as a function of the quantity constructed).

In Çetin, Jarrow, Protter and Warachka [8], for various traded options, some empirical evidence is provided showing that the difference between the classical price and the super-replication cost to an option is economically significant.

7 Transaction Costs

As previously stated, transaction costs can be viewed as a special case of our liquidity risk formulation where the $\mathcal{C}^2$ hypothesis on the supply curve is violated. This section provides the justification for this statement. We discuss three kinds of transaction costs, where all costs are *per share* unless otherwise stated. The three kinds are *fixed transaction costs, proportionate transaction costs*, and *mixed fixed and proportionate transaction costs*. We are now ignoring liquidity issues, and we use the mathematics of the supply curve framework to study only transaction costs. To emphasize this distinction, we now call the supply curve the *transaction curve*. Our goal is to see when continuous trading is feasible.[16] This section largely follows Umut

[15]For multiple call options on the stock, the right side of the equation is premultiplied by the number of option shares.

[16]Obviously in practice continuous trading is not truly feasible, since one cannot physically trade continuously. However, there remains the issue of whether one would desire to approximate a continuous trading strategy arbitrarily closely with discrete trading strategies. If continuous trading strategies have infinite transaction costs, then any approximating sequence would have unboundedly large transaction costs, and be undesirable to utilize. It is the desirability of using approximating sequences that is really being investigated below.

Çetin's thesis [6].

Definition 11 *We define three kinds of transaction costs:*

1. ***Fixed transaction costs*** *are defined by a transaction curve giving the (per share stock price) by*

$$S(t,x) = S(t,0) + \frac{a}{x}.$$

2. ***Proportionate transaction costs*** *depend proportionately on the dollar value of the trade, and are given by*

$$S(t,x) = S(t,0)(1 + \beta sign(x))$$

where $\beta > 0$ is the proportionate transaction cost per unit value.

3. ***Combined fixed and proportionate transaction costs*** *vary with the specific application. Two examples are:*[17]

   - *(Fidelity)*

$$S(t,x) = S(t,0) + \frac{\beta}{x} + sign(x)\gamma 1_{\{|x|>\delta\}}$$

   where β, γ and δ are positive constants;

   - *(Vanguard)*

$$S(t,x) = S(t,0) + \frac{\max\{\alpha, |x|c\}}{x}$$

   where α and c are positive constants.

Our first result implies that in the presence of fixed transaction costs, only piecewise constant trading strategies need be considered for modeling purposes.

Theorem 4 *Continuous trading in the presence of fixed transaction costs creates infinite costs in finite time.*

[17] These characterizations were obtained from information provided on both the Fidelty and Vanguard web sites in 2004.

Proof: We assume the structure given in Definition 11, part 1. First assume that our trading strategy X is of the form $X = \sum_{i=0}^{n-1} X_i 1_{[T_i, T_{i+1})}$, for a sequence of trading times $0 = T_0 \leq T_1 \leq \cdots \leq T_n = T$. Then the cumulative trading costs are $\sum_{i=0}^{n-1} a 1_{\{X_i \neq X_{i-1}\}}$, and if we further assume that always $X_i \neq X_{i-1}$, then it is equal simply to na.

Now suppose our trading strategy X has continuous paths. Let $TC(X)$ denote the transaction costs of following the strategy X. We have

$$TC(X) = \limsup_{n\to\infty} \sum_{T_i^n \in \Pi_k} a 1_{\{X_{T_i^n} \neq X_{T_{i+1}^n}\}} = \limsup_{n\to\infty} a N_{\Pi_n}(X),$$

where Π_n is a sequence of random partitions tending to the identity[18] and $N_{\Pi_n}(X)$ is the number of times that $X_{T_i^n} \neq X_{T_{i-1}^n}$ for the random stopping times of Π_n. Note that $\limsup_{n\to\infty} N_{\Pi_n}(X) = \infty$ unless X is a.s. piecewise constant. Thus continuous trading strategies incur infinite transaction costs. Finally, if our trading strategy has both jumps and continuous parts to it, the transaction costs will exceed the costs for each part, hence will also be infinite.

The situation for proportional transaction costs is more complicated. Here it is possible to trade continuously, provided one follows a trading strategy with paths of finite variation (which is not the case, for example, with the standard Black-Scholes hedge of a European call or put option).

Theorem 5 *Continuous trading in the presence of proportional transaction costs is infinite if the trading strategy has paths of infinite variation. If the strategy X has paths of finite variation on $[0,T]$ for a subset Λ of Ω, then the cumulative transaction costs are $b\int_0^T S(s,0)|dX_s|$ a.s. on Λ, where $|dX_s|$ denotes the total variation Stieltjes path by path integral, and they are infinite on Λ^c. (b is the constant in Definition 11, part 2.)*

Proof: Let Π_n be a sequence of random partitions tending to the identity on $[0,T]$. Let X be a continuous trading strategy. Then the cumulative transaction costs for proportional costs can be written as

$$TC(X) = \limsup_{n\to\infty} \sum_{T_i^n \in \Pi_k} S(T_k^n, 0)|\Delta X_{n,k}|b$$

[18]This is terminology from [30]; it means that each Π_n is a finite increasing sequence of stopping times covering the interval $[0,T]$, and the mesh of Π_n tends to 0 as $n \to \infty$.

where $\Delta X_{n,k} = X_{T^n_k} - X_{T^n_{k-1}}$. When X has paths of finite variation, this converges to the path by path Stieltjes integral $b\int_0^T S(s,0)|dX_s|$, and when X paths of infinite variation is diverges to ∞. Since it is a path by path result, we deduce the theorem.

Theorem 6 *Continuous trading in the presence of combined fixed and proportional transaction costs creates infinite costs in finite time.*

Proof: Suppose our trading strategy X has continuous paths. Let $TC(X)$ denote the transaction costs of following the strategy X. We have

$$TC(X) \geq \limsup_{n\to\infty} \sum_{T^n_i \in \Pi_k} \delta 1_{\{X_{T^n_i} \neq X_{T^n_{i+1}}\}} \geq \limsup_{n\to\infty} \delta N_{\Pi_n}(X),$$

for some constant δ, and where Π_n is a sequence of random partitions tending to the identity, and $N_{\Pi_n}(X)$ is the number of times that $X_{T^n_i} \neq X_{T^n_{i-1}}$ for the random stopping times of Π_n. This leads to infinite costs as in the proof of Theorem 4.

8 Examples of Supply Curves

We now discuss recent results of Marcel Blais [4] and M. Blais and P. Protter [5]. These results are inspired by an analysis of a trading book, provided to Blais and Protter by Morgan Stanley, via the good office of Robert Ferstenberg (see [18]). See [4],[5] for a detailed description of the data and its more profound implications. Note that the classical theory, with unlimited liquidity, is embedded in the structure previously discussed, using a standard price process $S_t = S(t,0)$. In this case the supply curve

$$x \to S(t,x) \text{ reduces to } x \to S(t,0):$$

that is, it is a line with slope 0 and vertical axis intercept $S(t,0)$. If one supposes that the supply curve is linear, that is of the form

$$x \to S(t,x) = M_t x + b_t,$$

then if the classical theory is accurate one must have $M_t = 0$. Taking this as the null hypothesis, Blais [4] has shown that it can be rejected at the 0.9999 significance level. From this we conclude that the supply curve exists and is non-trivial.

Using linear regression, Blais [4] has further shown that for *liquid stocks*[19] the supply curve is linear, with time varying slope and intercept; thus for liquid stocks the supply curve can be written

$$x \to S(t,x) = M_t x + b_t,$$

where $b_t = S(t,0)$. Moreover it is reasonable to assume that $(M_t)_{t\geq 0}$ is itself a stochastic process with continuous paths. We have then the following theorem for this case, which is a special case of Theorem 11 of Section 10.

Theorem 7 *For a liquid stock with linear supply curve of the form*

$$x \to S(t,x) := M_t x + b_t,$$

and a càdlàg trading strategy X with finite quadratic variation, the value in the money market account for a self financing trading strategy is given by

$$Y_t = -X_t S(t,0) + \int_0^t X_{u-} dS(u,0) - \int_0^t M_u d[X,X]_u.$$

Note that in this theorem, the quadratic variation differential term can have jumps.

The case for *non-liquid stocks* presents a new problem, and the previously established theory breaks down at one particular point, because the standing hypothesis that the supply curve $x \to S(t,x)$ is $\mathcal{C}^2$ no longer holds. Indeed, in this case the data shows that the supply curve is jump linear, with one jump, which can be thought of as the bid-ask spread. Fortunately the only place Çetin et al [7] use the $\mathcal{C}^2$ hypothesis is in the derivation of the self financing strategy, and in the jump linear case the simple structure allows Blais and Protter to eliminate this hypothesis.[20] Since we no longer assume the supply curve is continuous, we can let $S(t,0-)$ denote the marginal ask, and $S(t,0)$ will denote the marginal bid, whence we can let $\gamma(t) = S(t,0) - S(t,0-)$

[19] We deliberately leave the definition of a "liquid stock" vague. For a precise definition, see [5]. Examples of liquid stocks include BP, ATT, IBM.

[20] The $\mathcal{C}^2$ hypothesis of the supply curve in the space variable is of course also not needed for the linear supply curve case, and thus in practice perhaps it is not needed at all.

denote the bid/ask spread. Let $\Lambda = \{(s,\omega) : \Delta X_s(\omega) < 0\}$, and assume that the supply curve has a jump linear form given by

$$S(t,x) = \begin{cases} \beta(t)x + S(t,0)\,(x \geq 0) \\ \alpha(t)x + S(t,0-)\,(x < 0). \end{cases} \tag{14}$$

We have in this case the following result([5]):

Theorem 8 *For an illiquid stock with jump linear supply curve of the form given in equation (14) and a càdlàg trading strategy X with finite quadratic variation, the value in the money market account for a self financing trading strategy is given by*

$$\begin{aligned} Y_t &= -X_t S(t,0) + \int_0^t X_{u-} dS(u,0) - \int_0^t \beta(s)1_{\Lambda^c}(s) + \alpha(s)1_{\Lambda}(s) d[X,X]_s \\ &\quad - \int_0^t 1_{\Lambda}(s) d[\gamma, X]_s, \end{aligned}$$

where $\gamma(t) = S(t,0) - S(t,0-)$ denotes the bid/ask spread.

Proof: It is clear that the money market process Y should satisfy

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

We know from Example 2 that the last sum converges to $-X_0S(0,0) + X_tS(t,0) - \int_0^t X_{u-}dS(u,0)$. Let us thus focus on the term

$$- \lim_{n\to\infty} \sum_{k\geq 1} \left(X_{T_k^n} - X_{T_{k-1}^n}\right) \left[S(T_k^n, \left(X_{T_k^n} - X_{T_{k-1}^n}\right)) - S(T_k^n,0)\right]. \tag{15}$$

Due to our jump linear hypothesis on the structure of the supply curve, we can re-write the sum in expression (15) as:

$$
\begin{aligned}
& \sum_{k\geq 1}\left(X_{T_k^n}-X_{T_{k-1}^n}\right)\left[S(T_k^n,\left(X_{T_k^n}-X_{T_{k-1}^n}\right))-S(T_k^n,0)\right] \\
= & \sum_{k\geq 1}\Delta X_{n,k}1_{\{\Delta X_{n,k}\geq 0\}}\left[\beta(t)\Delta X_{n,k}+b^+(t)-b^+(t)\right] \\
& +\sum_{k\geq 1}\Delta X_{n,k}1_{\{\Delta X_{n,k}<0\}}\left[\alpha(t)\Delta X_{n,k}+b^-(t)-b^+(t)\right] \\
= & \sum_{k\geq 1}(\Delta X_{n,k})^2 1_{\{\Delta X_{n,k}\geq 0\}}\beta(t)+\sum_{k\geq 1}(\Delta X_{n,k})^2 1_{\{\Delta X_{n,k}<0\}}\alpha(t) \\
& +\gamma(t)\Delta X_{n,k}1_{\{\Delta X_{n,k}<0\}},
\end{aligned}
$$

where we have written $\Delta X_{n,k}$ as a shorthand for $X_{T_k^n}-X_{T_{k-1}^n}$, and where $b^+(t)=S(t,0)$, $b^-(t)=S(t,0-)$, and also $\Lambda=\{(s,\omega):\Delta X_s(\omega)<0\}$.

Next we take the limits as indicated in expression (15) and using standard theorems from stochastic calculus (see, e.g., [30]) we get convergence uniformly on compact time sets in probability to the expression

$$
-\int_0^t \beta(s)1_{\Lambda^c}(s)+\alpha(s)1_\Lambda(s)d[X,X]_s-\int_0^t 1_\Lambda(s)d[\gamma,X]_s,
$$

and the result follows.

In contrast to the fixed transaction cost situation of the previous section, Theorem 7 demonstrates that bid-ask spreads (as characterized in expression (14)), do not necessarily generate infinite liquidity costs and, therefore, bid-ask spreads do not preclude the use of continuous trading strategies.

9 Conclusion

This paper reviews the work of Çetin [6], Çetin, Jarrow and Protter [7], Çetin, Jarrow, Protter and Warachka [8], Blais [4], and Blais and Protter [5] which extends classical arbitrage pricing theory to include liquidity risk. This is accomplished by studying an economy where the security's price depends on the trade size. Extended first and second fundamental theorems of asset pricing are shown to hold. For the first theorem, 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 also approximately holds: markets will be approximately complete if the martingale measure is unique. In an approximately complete market, derivative prices are shown to equal the classical arbitrage free price of the derivative security. This implies horizontal supply curves for a derivative on the stock. To obtain upward sloping supply curves for derivatives, continuous trading strategies need to be excluded. This extension implies an incomplete market. Minimal cost super-replicating trading strategies are discussed in this regard. Last, an analysis of the theory and how it applies to data in both the liquid and illiquid cases is reviewed.

References

[1] Back, K., 1993, "Asymmetric Information and Options," *The Review of Financial Studies*, 6 (3), 435 - 472.

[2] P. Bank and D. Baum, 2002, "Hedging and Portfolio Optimization in Illiquid Financial Markets with a Large Trader," forthcoming, *Mathematical Finance*.

[3] Barles, G. and H. Soner, 1998, "Option Pricing with Transaction Costs and a Nonlinear Black-Scholes Equation," *Finance and Stochastics*, 2, 369 - 397.

[4] Blais, M., 2005, *Liquidity and Data,* Ph.D. thesis, Cornell University.

[5] M. Blais and P. Protter, 2005, "An Analysis of the Supply Curve for Liquidity Risk through Book Data," in preparation.

[6] Çetin, U., 2003, *Default and Liquidity Risk Modeling,* Ph.D. thesis, Cornell University.

[7] Çetin, U., R. Jarrow, P. Protter, 2004, "Liquidity Risk and Arbitrage Pricing Theory," *Finance and Stochastics*, **8**, 311-341.

[8] Çetin, U., R. Jarrow, P. Protter, M. Warachka, 2004, Pricing Options in an Extended Black Scholes Economy with Illiquidity: Theory and Empirical Evidence, working paper, Cornell University.

[9] Cheridito, P., 2003. Arbitrage in Fractional Brownian Motion Models. Finance and Stochastics 7, 417-554.

[10] Constantinides, G. and T. Zariphopoulou, 1999, "Bounds on Prices of Contingent Claiims in an Intertemporal Economy with Proportional Transaction Costs and General Preferences," *Finance and Stochastics*, 3, 345 - 369.

[11] Cvitanic, J. and I. Karatzas, 1996, "Hedging and Portfolio Optimization under Transaction Costs: a Martingale Approach," *Mathematical Finance*, 6 , 133 - 165.

[12] Cvitanic, J and J. Ma, 1996, "Hedging Options for a Large Investor and Forward-Backward SDEs," *Annals of Applied Probability,* 6, 370 - 398.

[13] Cvitanic, J., H. Pham, N. Touze, 1999, "A Closed-form Solution to the Problem of Super-replication under Transaction Costs," *Finance and Stochastics*, 3, 35 - 54.

[14] Delbaen, F., 2003, Personal communication.

[15] Delbaen, F. and Schachermayer, W., 1994, "A General Version of the Fundamental Theorem of Asset Pricing", *Mathematische Annalen*, 300 , 463 - 520.

[16] Duffie, D., 1996, *Dynamic Asset Pricing Theory*, 2nd edition, Princeton University Press: New Jersey.

[17] El Karoui, N., 1978, "Sur les Mont´ees des Semi-martingales," *Astérisque*, 52-53, 63-72.

[18] Ferstenberg, R., 2004, Private communication.

[19] Geman, H., El Karoui, N., and J.-C. Rochet, 1995, "Changes of Numéraire, change of probability measure and option pricing," *J. Appl. Probability*, 32, 443 - 458.

[20] Glosten, L. and P. Milgrom, 1985, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14 (March), 71 - 100.

[21] Grossman, S. and M. Miller, 1988, "Liquidity and Market Structure," *Journal of Finance*, 43 (3), 617 - 637.

[22] Harrison, J.M and S. Pliska, 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications*, 11, 215 - 260.

[23] Heath, D.,R. Jarrow and A. Morton, 1992, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica*, 60 (1), 77 - 105.

[24] Jarrow, R., 1992, "Market Manipulation, Bubbles, Corners and Short Squeezes," *Journal of Financial and Quantitative Analysis*, September, 311 - 336.

[25] Jarrow, R., 1994, "Derivative Security Markets, Market Manipulation and Option Pricing," *Journal of Financial and Quantitative Analysis*, 29 (2), 241 - 261.

[26] Jouini, E., 2000, "Price Functionals with Bid-Ask spreads: An Axiomatic Approach," *Journal of Mathematical Economics*, 34 (4), 547 - 558.

[27] Jouini, E. and H. Kallal, 1995, "Martingales and Arbitrage in Securities Markets with Transaction Costs," *Journal of Economic Theory*, 66 (1), 178 - 197.

[28] Jouini, E., H. Kallal and C. Napp, 2001, "Arbitrage in Financial Markets with Fixed costs," *Journal of Mathematical Economics*, 35 (2), 197 - 221.

[29] Kyle, A., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315-1335.

[30] Protter, P., 2004, *Stochastic Integration and Differential Equations, Second Edition*, Springer-Verlag: Heidelberg.

[31] Protter, P., 2001, "A Partial Introduction to Financial Asset Pricing Theory,"*Stochastic Processes and their Applications*, 91, 169 - 203.

[32] Soner, H.M., S. Shreve and J. Cvitanic, 1995, "There is no Nontrivial Hedging Portfolio for Option Pricing with Transaction Costs," *Annals of Applied Probability*, 5, 327 - 355.

10 Appendix

10.1 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 [15] or alternatively [31] 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 $X_t S(t,0) + Y_t^0 = X_0 S(0,0) + Y_0^0 + \int_{0+}^T X_u ds_u = \int_{0+}^T X_u ds_u$. 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 $\widehat{X}$.

10.1.1 Step 1. The Fictitious Economy

Consider the fictitious economy introduced in Section 3. Delbaen and Schachermayer prove the following in Section 4 of [15]:

Theorem 9 *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.[21]

[21] In the original paper [15], 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 [14].

Corollary 2 *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.

10.1.2 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 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 4 *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^n_T tends to $(X \cdot S)_T$, in probability, and $\epsilon_n \to 0$.*

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

$$f_n(t) = 1_{[T_n \leq t \leq T_{n+1}]} \frac{X_{T_n}}{T_n - T_{n+1}} (t - T_{n+1}) \tag{16}$$

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

$$X^n_t = X_t 1_{[t<T_n]} + f_n(t). \tag{17}$$

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,

$$(X^n \cdot S)_t = (X \cdot S)_{t\wedge T_n} + \int_0^t f_n(s) dS(s, 0). \tag{18}$$

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 [30], 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 $[0,T]$.[22]

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:

$$\mathbb{P}\left(\sup_{0\leq t\leq T} |(X^n \cdot S)_t - (X \cdot S)_t| \geq \epsilon_n\right) \leq \epsilon_n. \tag{19}$$

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

$$Y^n_{\tau^n} = X^n(\tau^n)\left[S\left(\tau^n, -X^n(\tau^n)\right) - S(\tau^n, 0)\right] + (X^n \cdot S)_{\tau^n} \tag{20}$$

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, $\mathbb{P}[\tau^n < T$ 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 5 *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 2, 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,

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

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.

Theorem 10 *(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 9.

10.2 Construction of the self-financing condition for a class of trading strategies

The purpose of this section is to provide justification for Definition 2 in the text. This proof uses only the weaker hypotheses of Assumption 2.[23]

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:

$$\sigma_n : 0 = T^n_0 \leq T^n_1 \leq \ldots \leq T^n_{k_n} = t$$

where T^n_k'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 $Y_t = Y_0 + \sum_{k\geq 1}(Y_{T^n_k} - Y_{T^n_{k-1}})$ for all n. Therefore, we'll define Y_t to be the following limit whenever it exists:

$$Y_0 - \lim_{n\to\infty} \sum_{k\geq 1}(X_{T^n_k} - X_{T^n_{k-1}})S(T^n_k, X_{T^n_k} - X_{T^n_{k-1}}). \tag{21}$$

Example 2 *In the classical case, $S(t, x) = S(t, 0)$ for all $x \in \mathbb{R}$. Thus, self-financing condition becomes*

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

[23]Note that we have already justified the notion of a self financing strategy in the jump linear illiquid case of section 8

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

$$\begin{aligned}
Y_t &= Y_0 - \lim_{n\to\infty} \sum_{k\geq 1} (X_{T_k^n} - X_{T_{k-1}^n}) S(T_k^n, 0) \\
&= Y(0) - \lim_{n\to\infty} \left[\sum_{k\geq 1} X_{T_k^n} S(T_k^n, 0) - \sum_{k\geq 1} X_{T_{k-1}^n} S(T_k^n, 0) \right] \\
&= Y(0) - \lim_{n\to\infty} \left[\sum_{k\geq 1} X_{T_k^n} S(T_k^n, 0) \right. \\
&\quad \left. - \sum_{k\geq 1} X_{T_{k-1}^n} (S(T_k^n, 0) - S(T_{k-1}^n, 0)) - \sum_{k\geq 1} X_{T_{k-1}^n} S(T_{k-1}^n, 0) \right] \\
&= Y_0 - X_t S(t,0) + X_0 S(0,0) \\
&\quad + \lim_{n\to\infty} \sum_{k\geq 1} X_{T_{k-1}^n} (S(T_k^n, 0) - S(T_{k-1}^n, 0)) \\
&= -X_t S(t,0) + \int_0^t X_{u-} dS(u,0).
\end{aligned}$$

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

$$\begin{aligned}
Y_t &= -X_t S(t,0) + \int_0^t X_{u-} dS(u,0) - \int_0^t S_x^{(1)}(u-,0) d[X,X]_u^c \\
&\quad - \sum_{0\leq u\leq t} \left[S(u, \Delta X_u) - S(u,0) \right] \Delta X_u. \qquad (22)
\end{aligned}$$

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

Proof: The proof of this theorem is reminiscent of the proof of Theorem 8.

Expression (21) is

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

We know from Example 2 that the last sum converges to $-X_0S(0,0) + X_tS(t,0) - \int_0^t X_{u-}dS(u,0)$. 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 [30]. Thus,

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

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

$$\sum_{u\ \in A}\left[S(u, \Delta X_u) - S(u, 0)\right]\Delta X_u. \tag{23}$$

Applying Taylor's formula to each $S(T_k^n,\cdot)$, the second limit becomes

$$\begin{aligned}
& \lim_{n\to\infty}\sum_{k,B} S_x^{(1)}(T_k^n,0)\left(X_{T_k^n}-X_{T_{k-1}^n}\right)^2 \\
& + \lim_{n\to\infty}\sum_{k,B}\left(X_{T_k^n}-X_{T_{k-1}^n}\right) R\left(T_k^n,\left|X_{T_k^n}-X_{T_{k-1}^n}\right|\right) \\
= \;& \lim_{n\to\infty}\sum_{k\geq 1} S_x^{(1)}(T_k^n,0)\left(X_{T_k^n}-X_{T_{k-1}^n}\right)^2 \\
& - \lim_{n\to\infty}\sum_{k,A} S_x^{(1)}(T_k^n,0)\left(X_{T_k^n}-X_{T_{k-1}^n}\right)^2 \\
& + \lim_{n\to\infty}\sum_{k,B}\left(X_{T_k^n}-X_{T_{k-1}^n}\right) R\left(T_k^n,\left|X_{T_k^n}-X_{T_{k-1}^n}\right|\right) \\
= \;& \lim_{n\to\infty}\sum_{k\geq 1} S_x^{(1)}(T_{k-1}^n,0)\left(X_{T_k^n}-X_{T_{k-1}^n}\right)^2 \\
& + \lim_{n\to\infty}\sum_{k\geq 1}\left[S_x^{(1)}(T_k^n,0)-S_x^{(1)}(T_{k-1}^n,0)\right]\left(X_{T_k^n}-X_{T_{k-1}^n}\right)^2 \qquad (24) \\
& - \lim_{n\to\infty}\sum_{k,A} S_x^{(1)}(T_k^n,0)\left(X_{T_k^n}-X_{T_{k-1}^n}\right)^2 \\
& + \lim_{n\to\infty}\sum_{k,B}\left(X_{T_k^n}-X_{T_{k-1}^n}\right) R\left(T_k^n,\left|X_{T_k^n}-X_{T_{k-1}^n}\right|\right),
\end{aligned}$$

where R is the remainder term in Taylor's formula. The sum of the first three limits converges to[24]

$$\begin{aligned}
& \int_0^t S_x^{(1)}(u-,0)d[X,X]_u + \left[S_x^{(1)}(\cdot,0),[X,X]\right]_t - \sum_{u\in A} S_x^{(1)}(u,0)(\Delta X_u)^2 \\
= \;& \int_0^t S_x^{(1)}(u-,0)d[X,X]_u + \sum_{0<u\leq t}\Delta S_x^{(1)}(u,0)(\Delta X_u)^2 \\
& - \sum_{u\in A} S_x^{(1)}(u,0)(\Delta X_u)^2. \qquad (25)
\end{aligned}$$

[24]Note that the assumption that $S_x^{(1)}(\cdot,0)$ has a finite QV is not needed when $S_x^{(1)}(\cdot,0)$ is continuous. In this case, the second limit is zero. This follows from the fact that X has a finite QV and $S_x^{(1)}(\cdot,0)$ 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 (24) 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 (24) 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 $\epsilon \to 0$, (23) converges to

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

and (25) 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, (22) holds for $\tilde{S}$, for each k. Now, a standard argument using set unions, as in the proof of Itô's formula in [30], establishes (22) for S.

10.3 Approximating stochastic integrals with continuous and of FV integrands

The next lemma (Lemma 6) is well known and can be found in [30].

Lemma 6 *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 2 *$S(\cdot, 0)$ has only totally inaccessible jumps.*

We recall a few definitions.

Definition 12 *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 13 *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 14 *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 12 *Let $\epsilon > 0$. For any H bounded, continuous and of FV, there exists H^ϵ, bounded, continuous and of FV, with $H^\epsilon_T = 0$ such that $d_s(H, H^\epsilon) < \epsilon$.*

Proof: Define

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

where $T_m = T - \frac{1}{m}$. We'll first show $d_s\left(H, H1_{[0,T_m]}\right) \to 0$ 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,

$$\begin{aligned} |(G \cdot M^{T_n})_t| &\leq |G| \cdot [\bar{N}, \bar{N}]^{T_n}_t + |G| \cdot \langle \bar{N}, \bar{N} \rangle^{T_n}_t \\ &\leq |G| \cdot [\bar{N}, \bar{N}]_t + |G| \cdot \langle \bar{N}, \bar{N} \rangle_t \end{aligned} \tag{26}$$

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, $G1_{[0,T_n]}$ 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 (26), 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,

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

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,

$$\left\|\left(\left(H - H1_{[0,T_m]}\right)^2\cdot\left\langle\bar{N},\bar{N}\right\rangle\right)^{1/2}\right\|_{L^2} \leq \left\|\left(H^2\cdot\left\langle\bar{N},\bar{N}\right\rangle\right)^{1/2}\right\|_{L^2} < \infty$$

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 $\left\|\int_0^T \left|H_u - H_u1_{[0,T_m]}\right|\left|d\bar{A}_u\right|\right\|_{L^2} \to 0$ 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 $K = \|\max_{0\leq t\leq T} H_t^2(\omega)\|_\infty < \infty$ 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 3 *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 12 and Theorem 5 of Chapter IV in [30].