# OPTION PRICING AND HEDGING WITH EXECUTION COSTS AND MARKET IMPACT

OLIVIER GUÉANT

*Université Paris-Diderot*

JIANG PU

*Institut Europlace de Finance*

This paper considers the pricing and hedging of a call option when liquidity matters, that is, either for a large nominal or for an illiquid underlying asset. In practice, as opposed to the classical assumptions of a price-taking agent in a frictionless market, traders cannot be perfectly hedged because of execution costs and market impact. They indeed face a trade-off between hedging errors and costs that can be solved by using stochastic optimal control. Our modeling framework, which is inspired by the recent literature on optimal execution, makes it possible to account for both execution costs and the lasting market impact of trades. Prices are obtained through the indifference pricing approach. Numerical examples are provided, along with comparisons to standard methods.

KEY WORDS: option pricing, option hedging, illiquid markets, optimal execution, stochastic optimal control.

## 1. INTRODUCTION

Classical option pricing theory is based on the hypothesis of a frictionless market in which agents are price takers: there are no transaction costs and traders have no impact—neither temporary nor permanent—on prices. These assumptions are not realistic, but the resulting option pricing models (for instance, the Black–Scholes model or the Heston model) are widely used and provide useful results as long as the underlying asset is liquid and the nominal is not too large. However, for options on illiquid assets or for options with a large nominal compared to the volume commonly traded on the market of the underlying asset, execution costs and market impact cannot be ignored.

Several improvements to the Black–Scholes model have been made to account for transaction costs. The basic idea is that high-frequency hedging costs are prohibitive due to transaction fees, whereas low-frequency hedging leads to large tracking errors. Leland (1985) proposed one of the first models to deal with transaction costs in the context of

option pricing. Other models of frictional markets, with either fixed transaction costs or transaction costs proportional to the traded volume, include Barles and Soner (1998), Cvitanić and Karatzas (1996), and Cvitanić, Pham, and Touzi (1999).

Two other routes have been considered to account for market imperfections in option pricing models.

The first route is usually referred to as the "supply curve" approach. In this approach, introduced by Çetin, Jarrow, and Protter (2004) (see also Bank and Baum 2004; Çetin et al. 2006; Çetin and Rogers 2007), traders are not price takers, and the price they pay depends on the quantity they trade. Although appealing, this framework leads to prices identical to those in the Black–Scholes model. Çetin, Soner, and Touzi (2010) consider the same approach but they restrict the set of admissible strategies (see also Longstaff 2001) to obtain positive liquidity costs and prices that eventually depart from those in the Black–Scholes model. We model execution costs (liquidity costs) differently because our framework is inspired by the literature on optimal execution (see Almgren and Chriss 2001; Schied, Schöneborn, and Tehranchi 2010; Guéant 2015).

The second route has to do with the impact of $\Delta$-hedging on the dynamics of the underlying asset,[1] and the resulting feedback effect on the price of the option. An important amount of literature exists on this topic, therefore we refer to Platen and Schweizer (1998), Schönbucher and Wilmott (2000), and Sircar and Papanicolaou (1998) for the different modeling approaches. To take account of this effect, we use the same linear form of permanent market impact as in most papers on optimal execution.

In addition to these two routes, a new approach has recently emerged, in relation with the literature on optimal execution. Rogers and Singh (2010) and Li and Almgren (2013) consider approaches inspired by this literature, and similar to ours. In their settings, the authors consider execution costs that are not linear in the volume executed but instead are convex to account for liquidity effects.

Rogers and Singh consider an objective function that penalizes both execution costs and the mean-squared hedging error at maturity. They obtain, in this close-to-mean-variance framework, a closed-form approximation for the optimal hedging strategy when illiquidity costs are small.

Li and Almgren, motivated by the swings observed in US stock prices in July 2012, consider a model with both permanent and temporary impacts, whereas Rogers and Singh (2010) do not examine permanent market impact. They use a mean–variance optimization criterion where the hedging error is the main variable. They consider the case of quadratic execution costs and use a constant-$\Gamma$ approximation in order to obtain a closed-form expression for the hedging strategy.

Instead of focusing on a special case that leads to closed-form expressions, our goal is to consider a general model. We use a general form for the execution costs and we examine the influence of permanent market impact. The optimization criterion we consider is an expected utility applied to final wealth. Therefore, we characterize the optimal strategy with a partial differential equation (PDE), and we rely on numerical methods to approximate the solution of the PDE and the optimal hedging strategy. Another difference is that we account for the interest rate. Furthermore, neither Li and Almgren (2013) nor Rogers and Singh (2010) distinguish physical delivery from cash settlement. In this paper, we show how hedging strategies and option prices are impacted by the type of settlement.

---

[1]This impact was observed in July 2012 through sawtooth patterns on the prices of five major US stocks (see Lehalle et al. 2012, 2013; Li and Almgren 2013).

In terms of (partial) hedging strategies, both Li and Almgren (2013) and Rogers and Singh (2010) obtain optimal strategies that are mean reverting around the classical $\Delta$. This is not the case in our expected utility framework: our optimal strategy does not oscillate around a solution without execution costs and market impact. The optimal strategies in our model are smoother than classical $\Delta$-hedging strategies because the trader seeks to avoid round trips on the stock, which entail execution costs upon purchase and sale of shares. In the case of a physical settlement, smooth strategies are also linked to the fact that the trader is averse both to price risk and to the binary risk of having to deliver versus not delivering. By comparing our strategies with classical $\Delta$-hedging strategies for different frequencies of rebalancing, we show that our approach makes it possible to reach very low levels of variance while mitigating execution costs. In particular, we observe that for the frequency of rebalancing that leads to the same level of execution costs, the variance of the profit and loss (PnL) associated with our strategy is lower than the one obtained with a classical $\Delta$-hedging strategy.

In addition to optimal hedging strategies, our expected CARA utility framework provides prices by using the indifference pricing approach. We compute the amount a client needs to pay to compensate, in utility terms, the payoff of the option when the trader uses the optimal hedging strategy. We find that the price of a call is higher in the presence of execution costs than in the classical model, and that this price is an increasing function of the illiquidity of the underlying asset and the nominal of the option.

Although we concentrate on the case of a call option throughout the paper, the same approach can be used for other types of options. In particular, a similar approach is used to price and hedge Accelerated Share Repurchase contracts (see Jaimungal, Kinzebulatov, and Rubisov 2013; Guéant, Pu, and Royer 2015). These contracts are Asian-type options with Bermudan-style exercise dates and a physical delivery.

The remainder of the text is organized as follows. In Section 2, we present the basic hypotheses of our model and we introduce the Hamilton–Jacobi–Bellman equation associated with the problem. In Section 3, we solve the control problem without permanent market impact, and we show that the price of the option satisfies a nonlinear PDE. In Section 4, we then show how our solution can be extended to the case where there is a permanent market impact. In Section 5, we discuss numerical methods to solve the problem. In Section 6, we present the outcomes from several examples, and we compare our model with the Bachelier model.

## 2. SETUP OF THE MODEL

### 2.1. Notations

We consider a filtered probability space $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ that corresponds to the available information on the market, namely, the market price of a stock up to the observation time. For $0 \leq s < t \leq T$, we denote $\mathcal{P}(s, t)$ the set of $\mathbb{R}$-valued progressively measurable processes on $[s, t]$.

The problem we consider is a bank (or a trader) selling a call option on a stock to a client.[2] The call option has a nominal $N$ (in shares), a strike $K$, and a maturity $T$.

---

[2]The reasoning is the same for a put option or if the bank is buying the option. We consider the specific case of a call option to highlight the difference between physical delivery and cash settlement.

*2.1.1. Execution process.* Because of the execution costs, the bank is not able to replicate the option. However, it buys and sells shares progressively to (partially) hedge its risky position. To model the execution process, we first introduce the market volume process $(V_t)_t$, that is assumed to be deterministic, nonnegative, and bounded. The trading is constrained to not go too fast, relative to the market volume, by imposing a maximum participation rate $\rho_m$.

The number of shares in the hedging portfolio is therefore modeled as:[3]

$$q_t = q_0 + \int_0^t v_s ds,$$

where the stochastic process $v$ belongs to the set of admissible strategies $\mathcal{A}$ defined by

$$\mathcal{A} := \{v \in \mathcal{P}(0, T), |v_t| \leq \rho_m V_t, \text{ a.e. in } (0, T) \times \Omega\}.$$

REMARK 2.1. The market volume process can be used to model overnight risk by assuming $V_t = 0$ when the market is closed.

*2.1.2. Price process.* The price process of the underlying asset is defined under the historical probability as an Ito process of the form:[4]

$$dS_t = \mu dt + \sigma dW_t + kv_t dt,$$

where $k \geq 0$ models permanent market impact and where $\mu$ is typically a view on the future trend of the underlying asset.

We consider a linear form for the permanent market impact to avoid dynamic arbitrage (see the analysis of Gatheral 2010). The more general framework proposed in Guéant (2013) could be another possibility but we believe it is more suited to intraday problems.

REMARK 2.2. We consider a drifted Bachelier dynamics for the price instead of the classical Black–Scholes framework. This is also the case in Li and Almgren (2013). The underlying reason for this choice is that we consider a CARA utility function which is, a priori, incompatible with a geometric Brownian motion. The prices we obtain are therefore subject to criticisms when the option maturity is long (see, e.g., Schachermayer and Teichmann 2008, for a comparison between the Bachelier and the Black-Scholes option pricing models).

*2.1.3. Cash account and execution costs.* The cash account of the bank follows a dynamics linked to the hedging strategy. It is, in particular, impacted by execution costs. These execution costs are modeled through the introduction of a function $L \in C(\mathbb{R}, \mathbb{R}_+)$ that verifies the following conditions:

- $L(0) = 0,$

---

[3]$q_0$ is the number of shares in the portfolio at inception. In illiquid markets, especially for corporate deals including options, the buyer of the call may provide an initial number of shares (see the discussion in Section 3). We shall consider below the case where $q_0 = 0$ and the case where $q_0$ is set to the initial Bachelier $\Delta$ thanks to an initial trade with the buyer.

[4]There is no reason to consider a risk-neutral probability in our framework, as one cannot replicate the payoff of a call option because of the execution costs.

- $L$ is an even function,
- $L$ is increasing on $\mathbb{R}_+$,
- $L$ is strictly convex,
- $L$ is asymptotically superlinear, that is:

$$\lim_{\rho \to +\infty} \frac{L(\rho)}{\rho} = +\infty.$$

For any $v \in \mathcal{A}$, the cash account $X$ evolves as

$$dX_t = r\, X_t dt - v_t S_t dt - V_t L\left(\frac{v_t}{V_t}\right) dt,$$

where $r$ is the risk-free rate.

REMARK 2.3. In applications, $L$ is often a power function of the form $L(\rho) = \eta\, |\rho|^{1+\phi}$ with $\phi > 0$, or a function of the form $L(\rho) = \eta\, |\rho|^{1+\phi} + \psi |\rho|$ with $\phi, \psi > 0$ where $\psi$ takes into account proportional costs such as the bid-ask spread or a stamp duty. In particular, the initial Almgren–Chriss framework corresponds to $L(\rho) = \eta \rho^2$ ($\phi = 1$, $\psi = 0$).

*2.1.4. Payoff of the option.* At time $T$, we consider either a physical settlement or a cash settlement.

Let us consider the case of a physical settlement. If the option is exercised, then the bank receives $KN$ and needs to deliver $N$ shares. Because the hedging portfolio contains $q_T$ shares at time $T$, the bank has to buy $N - q_T$ shares to be able to deliver. Thus, if the option is exercised, then the payoff of the bank is

$$X_T + KN - (N - q_T)S_T - \mathcal{L}(q_T, N) = X_T + q_T S_T + N(K - S_T) - \mathcal{L}(q_T, N),$$

where $\mathcal{L}(q, q')$ models the additional cost over the Mark to Market (MtM) price to go from a portfolio with $q$ shares to a portfolio with $q'$ shares.

In the case where the option is not exercised, the payoff is

$$X_T + q_T S_T - \mathcal{L}(q_T, 0),$$

because the trader needs to liquidate the portfolio. The term $\mathcal{L}(q_T, 0)$ is the discount incurred to liquidate the remaining shares.

If we assume that the option is exercised if and only if the stock price is above $K$,[5] then the total payoff in the case of physical settlement is

$$X_T + q_T S_T - N(S_T - K)_+ - 1_{S_T \geq K} \mathcal{L}(q_T, N) - 1_{S_T < K} \mathcal{L}(q_T, 0).$$

In the case of a cash settlement, the only difference is when the option is exercised. In that case, the bank pays $N(S_T - K)$ and liquidates its portfolio (with usually a lot of shares). The liquidation leads to the following payoff for the bank:

$$X_T + q_T S_T - N(S_T - K)_+ - \mathcal{L}(q_T, 0).$$

---

[5]The threshold might be less than the strike $K$ in an illiquid market.

In general, for both the cash and the physical settlements, the payoff is therefore of the form

$$X_T + q_T S_T - \Pi(q_T, S_T),$$

where $\Pi(q, S) \geq N(S - K)_+$.

### 2.1.5. Optimization.

The stochastic optimal control problem we consider is

$$\sup_{v \in \mathcal{A}} \mathbb{E} \left[ -\exp \left( -\gamma \left( X_T + q_T S_T - \Pi(q_T, S_T) \right) \right) \right],$$

where $\gamma$ is the absolute risk aversion parameter of the bank.

REMARK 2.4. The penalty function $\mathcal{L}$ needs to be specified. When there is no permanent market impact, a natural choice is $\mathcal{L}(q, q') = \ell(|q - q'|)$ where $\ell$ is an increasing and convex function.[6] One candidate for $\ell$ is the risk-liquidity premium of a block trade as in Guéant (2015). Another candidate is the risk-liquidity premium associated with liquidation at some constant participation rate $\rho$ (for instance, $\rho_{\mathrm{m}}$):

$$\ell(q) = \int_T^{T'} L(\rho) V_t dt + \frac{\gamma}{2} \sigma^2 \int_T^{T'} q_t^2 dt = \frac{L(\rho)}{\rho} |q| + \frac{\gamma}{2} \sigma^2 \int_T^{T'} \left( |q| - \rho \int_T^t V_s ds \right)^2 dt,$$

where $T'$ is the first time such that $\int_T^{T'} \rho V_t dt = |q|$.

## 2.2. The Value Function and the Hamilton–Jacobi–Bellman (HJB) Equation

To solve the above stochastic optimal control problem, we define the value function $u$ by

$$u(t, x, q, S) = \sup_{v \in \mathcal{A}_t} \mathbb{E} \left[ -\exp \left( -\gamma \left( X_T^{t,x,v} + q_T^{t,q,v} S_T^{t,S,v} - \Pi(q_T^{t,q,v}, S_T^{t,S}) \right) \right) \right],$$

where

$$\mathcal{A}_t := \{ v \in \mathcal{P}(t, T), |v_s| \leq \rho_{\mathrm{m}} V_s, \text{ a.e. in } (t, T) \times \Omega \},$$

and where

$$X_{t'}^{t,x,v} = x + \int_t^{t'} \left( r X_s^{t,x,v} - v_s S_s^{t,S,v} - V_s L \left( \frac{v_s}{V_s} \right) \right) ds$$

$$q_{t'}^{t,q,v} = q + \int_t^{t'} v_s ds$$

$$S_{t'}^{t,S,v} = S + \mu(t' - t) + \int_t^{t'} \sigma dW_s + k(q_{t'}^{t,q,v} - q).$$

---

[6]When permanent market impact is taken into account, $\mathcal{L}$ must have a specific form to avoid dynamic arbitrages (see Section 4).

The HJB equation associated with this problem is the following:

$$-\partial_t u - \mu \partial_S u - \frac{1}{2}\sigma^2 \partial^2_{SS} u - \sup_{|v| \leq \rho_m V_t} \left\{ v \partial_q u + \left( rx - vS - L\left(\frac{v}{V_t}\right) V_t \right) \partial_x u + kv \partial_S u \right\} = 0,$$

with the terminal condition:

$$u(T, x, q, S) = -\exp\left(-\gamma \left(x + qS - \Pi(q, S)\right)\right).$$

REMARK 2.5.  In the case of a physical settlement, this terminal condition is not continuous.

## 3.  CHARACTERIZATION OF THE SOLUTION

We first consider the case without permanent market impact ($k = 0$). In that case, we consider a function $\mathcal{L}$ of the form $\mathcal{L}(q, q') = \ell(|q' - q|)$, where $\ell$ is a convex and even function that increases on $\mathbb{R}_+$ (as exemplified in Remark 2.4).

The following lemma states that we can factor out the compounded MtM value of the current portfolio (we omit the superscripts to improve readability):

LEMMA 3.1.

$$X_T + q_T S_T = e^{r(T-t)}(x + qS)$$

$$+ e^{r(T-t)} \left( \int_t^T e^{-r(s-t)} q_s(\mu - r S_s)ds + \int_t^T e^{-r(s-t)} q_s \sigma\, dW_s - \int_t^T e^{-r(s-t)} V_s L\left(\frac{v_s}{V_s}\right) ds \right).$$

This lemma shows that the value function $u(t, x, q, S)$ is of the form

$$u(t, x, q, S) = -\exp\left(-\gamma e^{r(T-t)}(x + qS)\right) \inf_{v \in \mathcal{A}_t} J_t(q, S, v),$$

where

$$J_t : \mathbb{R} \times \mathbb{R} \times \mathcal{A}_t \to \mathbb{R}$$

$$(q, S, v) \mapsto J_t(q, S, v)$$

is defined as

$$J_t(q, S, v) = \mathbb{E}\left[ \exp\left(-\gamma \left( e^{r(T-t)} \left( \int_t^T e^{-r(s-t)} q_s(\mu - r S_s)ds + \int_t^T e^{-r(s-t)} q_s \sigma\, dW_s \right.\right.\right.\right.$$

$$\left.\left.\left.\left. - \int_t^T e^{-r(s-t)} V_s L\left(\frac{v_s}{V_s}\right) ds \right) - \Pi(q_T, S_T) \right)\right)\right].$$

We also define

$$\theta(t, q, S) = \inf_{v \in \mathcal{A}_t} \frac{e^{-r(T-t)}}{\gamma} \log(J_t(q, S, v)).$$

The following proposition states that $\theta$ is well defined, and gives a lower bound for $\theta$:

PROPOSITION 3.2.  $\forall (t, q, S) \in [0, T] \times \mathbb{R} \times \mathbb{R}$, $\theta(t, q, S)$ is finite.

*Moreover, if $\mu = r = 0$ then*

$$\theta(t, q, S) \geq N\mathbb{E}\left[(S_T - K)_+\right].$$

The function $\theta$ has a natural interpretation. Let us consider a call option deal between the bank and a client, where at time 0:

- the bank writes the call option with either a physical or cash settlement and the client pays a price $P$, and
- the client gives $q_0$ shares to the bank and receives $q_0 S_0$ in cash from the bank.

In utility terms, the bank gives the following value to this deal:

$$u(0, X_0 - q_0 S_0 + P, q_0, S_0) = -\exp\left(-\gamma e^{rT} (X_0 + P - \theta(0, q_0, S_0))\right).$$

As a consequence, if $P = \theta(0, q_0, S_0)$, and if we assume that the cash is invested at rate $r$, then the bank is indifferent to making the deal or not making the deal. Therefore, $\theta(0, q_0, S_0)$ is the indifference price of the call option deal.

This definition of a price for the call option depends on $q_0$: the initial number of stocks in the portfolio. This echoes the fact that, in practice, building the initial $\Delta$ position (as computed in a classical model) is usually costly for options with a large nominal.

This interpretation of $\theta$ also enables us to see the inequality of Proposition 3.2 in a different manner. When $\mu = r = 0$, then the price in our setting is always greater than the price when there is no execution cost (Bachelier model).[7]

Our focus now is on the function $\theta$ because it is the price of the call option. Our first result on $\theta$ states that it is a convex function of $q$.

PROPOSITION 3.3. *For* $(t, S) \in [0, T] \times \mathbb{R}$, $\quad q \in \mathbb{R} \mapsto \theta(t, q, S)$ *is a convex function.*

REMARK 3.4. We cannot expect the same result for $S$ because the final payoff is not continuous in $S$ in general (see the physical delivery case).

The main property for $\theta$ is the following PDE characterization:

PROPOSITION 3.5. *Let us introduce*

$$H(p) = \sup_{|\rho| \leq \rho_m} p\rho - L(\rho).$$

$\theta$ *is a viscosity solution of the following equation:*

$$-\partial_t \theta + r\theta + (\mu - rS)q - \mu \partial_S \theta - \frac{1}{2}\sigma^2 \partial_{SS}^2 \theta - \frac{1}{2}\gamma\sigma^2 e^{r(T-t)}(\partial_S \theta - q)^2 + V_t H(\partial_q \theta) = 0,$$

*with the terminal condition* $\theta(T, q, S) = \Pi(q, S)$ *in the classical sense.*

The PDE satisfied by $\theta$ is a nonlinear equation and, in particular, the price of the call option is not proportional to the nominal. To go from a nominal equal to $N$ to a nominal equal to 1, we introduce the function $\tilde{\theta}$ defined by

$$\tilde{\theta}(t, \tilde{q}, S) = \frac{1}{N}\theta(t, N\tilde{q}, S).$$

---

[7]The price of a call (with a unitary nominal and when $r = 0$) in the Bachelier model is given by $\mathbb{E}[(S_T^{t,S} - K)_+] = (S - K)\Phi(\frac{S-K}{\sigma\sqrt{T-t}}) + \sigma\sqrt{T-t}\varphi(\frac{S-K}{\sigma\sqrt{T-t}})$, where $\varphi$ and $\Phi$ are, respectively, the probability density function and the cumulative distribution function of a standard normal variable.

Then, $\tilde{\theta}$ satisfies the following equation in the viscosity sense:

$$-\partial_t\tilde{\theta} + r\tilde{\theta} + (\mu - rS)\tilde{q} - \mu\partial_S\tilde{\theta} - \frac{1}{2}\sigma^2\partial^2_{SS}\tilde{\theta} - \frac{1}{2}\gamma N\sigma^2 e^{r(T-t)}(\partial_S\tilde{\theta} - \tilde{q})^2 + \frac{V_t}{N}H(\partial_{\tilde{q}}\tilde{\theta}) = 0,$$

with the terminal condition

$$\tilde{\theta}(T, \tilde{q}, S) = \frac{1}{N}\Pi(N\tilde{q}, S).$$

In other words, we need to rescale the risk aversion parameter $\gamma$, the market volume process $(V_t)_t$, and the liquidation penalty function $\mathcal{L}$ in order to go from a call of nominal $N$ to a call of nominal 1.

Each term in the PDE

$$\partial_t\theta = \underbrace{r\theta}_{\text{(I)}} + \underbrace{(\mu - rS)q}_{\text{(II)}} \underbrace{-\mu\partial_S\theta - \frac{1}{2}\sigma^2\partial^2_{SS}\theta}_{\text{(III)}} \underbrace{-\frac{1}{2}\gamma\sigma^2 e^{r(T-t)}(\partial_S\theta - q)^2}_{\text{(IV)}} + \underbrace{V_t H(\partial_q\theta)}_{\text{(V)}}$$

has a specific interpretation:

- The term (I) is the classical term linked to discounting at the risk-free rate $r$.
- The term (II) corresponds to the premium linked to holding shares instead of cash. If indeed one holds $q$ shares, on average the MtM wealth is increased by $\mu q$ per unit of time, whereas the amount of cash equivalent to $q$ shares (that is $qS$) increases the MtM wealth by $rqS$ per unit of time.
- The term (III) is linked to the dynamics of the stock price.
- The interdependence between the number of shares $q$ in the hedging portfolio and the dynamics of the price occurs through (IV), and more precisely through the term $(\partial_S\theta - q)^2$. Although there is no $\Delta$ in this model because the market is incomplete, this term measures the difference between the first derivative of the option price with respect to the price of the underlying asset and the number of shares in the hedging portfolio: it looks therefore like the measure of a mis-hedge.
- And, (V) is the classical term in the literature on optimal execution. It models the execution costs and the participation limit $\rho_m$. In particular, the optimal participation rate at time $t$ is $\rho^*(t, q_t, S_t) = \frac{v^*(t, q_t, S_t)}{V_t} = H'(\partial_q\theta(t, q_t, S_t))$.

REMARK 3.6. If we replace $\mu$ with $rS$ and $\sigma$ with $\sigma S$, then the terms (I), (II), and (III) are exactly the same as those in the Black–Scholes PDE.

Furthermore, the PDE satisfied by $\theta$ is (surprisingly) not derived from a control problem because

$$(p_q, p_S) \mapsto -\frac{1}{2}\gamma\sigma^2 e^{r(T-t)}(p_S - q)^2 + V_t H(p_q)$$

is neither convex, nor concave.

In fact, it derives from a zero-sum game (see the appendix of Bardi and Capuzzo-Dolcetta 2008) in which the first player controls $q$ through

$$dq_t = v_t dt,$$

and player 2 controls the drift of the price

$$dS_t = (\mu + \alpha_t)dt + \sigma dW_t.$$

The payoff of the zero-sum game associated with the above equation is

$$\mathbb{E}\left[\int_0^T e^{r(T-t)}\left(L\left(\frac{v_t}{V_t}\right)V_t - \frac{e^{-r(T-t)}}{2\gamma\sigma^2}(\alpha_t + \gamma e^{r(T-t)}\sigma^2 q_t)^2\right.\right.$$
$$\left.\left. -\frac{1}{2}\gamma e^{r(T-t)}\sigma^2 q_t^2 + (rS_t - \mu)q_t\right)dt + \Pi(q_T, S_T)\right],$$

where player 1 minimizes and player 2 maximizes.

## 4. THE PROBLEM WITH PERMANENT MARKET IMPACT

We now turn to the case where there is a permanent market impact. To stay in the framework of Gatheral's paper on permanent market impact without dynamic arbitrage (Gatheral 2010), we consider in this section that $\mu = r = 0$, and we use the linear form of permanent market impact introduced in Section 2. We will show that, up to a change of variables, the problem is—from a mathematical point of view—the same as in the absence of permanent market impact.

To avoid dynamic arbitrage, we have to specify $\mathcal{L}$. At time $T$, if one wants to go from a portfolio with $q$ shares to a portfolio with $q'$ shares, one must pay the liquidity costs related to the volume transacted. This is modeled by $\ell(q' - q)$, as in the previous case without permanent market impact. However, we must also take permanent market impact into account. The amount paid to go from a portfolio with $q$ shares at time $T$ to a portfolio with $q'$ at time $T'$ is (on average and ignoring temporary market impact):

$$\mathbb{E}\left[\int_T^{T'} dq_t S_t \,\middle|\, S_T\right] = (q' - q)S_T + \int_T^{T'} k(q' - q_t)dq_t$$
$$= (q' - q)S_T + \frac{1}{2}k(q' - q)^2.$$

Hence, we define $\mathcal{L}$ by

$$\mathcal{L}(q, q') = \ell(|q' - q|) + \frac{1}{2}k(q' - q)^2.$$

Let us now come to the change of variables. In the previous section, we showed that $u(t, x, q, S)$, in the absence of permanent market impact, can be written as

$$u(t, x, q, S) = -\exp\left(-\gamma\left(x + qS - \theta(t, q, S)\right)\right).$$

Using the same method as in Section 3, we show that, with permanent market impact, $u$ can be written as

$$u(t, x, q, S) = -\exp\left(-\gamma\left(x + qS - \frac{1}{2}kq^2 + \frac{1}{2}kq_0^2 - \theta(t, q, S - k(q - q_0))\right)\right).$$

In other words, we introduce the function

$$\theta(t, q, \tilde{S}) = x + q(\tilde{S} + k(q - q_0)) - \frac{1}{2}kq^2 + \frac{1}{2}kq_0^2 + \frac{1}{\gamma}\log\left(-u(t, x, q, \tilde{S} + k(q - q_0))\right).$$

As in the previous case, $\theta(0, q_0, S_0)$ is the price of the call at time 0 when the deal starts with an exchange of $q_0$ shares against $q_0 S_0$ in cash.

REMARK 4.1. The new variable $\tilde{S}_t = S_t - k(q_t - q_0)$ is the price from which we remove the influence of the permanent market impact.

Now, using the same techniques as in Section 3, we prove the following proposition:

PROPOSITION 4.2. *Let us assume that $\mu = r = 0$. Then, $\theta$ is a viscosity solution of the following equation:*

$$-\partial_t \theta - \frac{1}{2}\sigma^2 \partial^2_{\tilde{S}\tilde{S}}\theta - \frac{1}{2}\gamma\sigma^2(\partial_{\tilde{S}}\theta - q)^2 + V_t H(\partial_q \theta) = 0,$$

*with the terminal condition*

$$\theta(T, q, \tilde{S}) = \Pi(q, \tilde{S} + k(q - q_0)) - \frac{1}{2}kq^2 + \frac{1}{2}kq_0^2.$$

The introduction of a permanent market impact only changes the terminal condition of the PDE. In the case of a cash settlement, the terminal condition is

$$\theta(T, q, \tilde{S}) = N(\tilde{S} + k(q - q_0) - K)_+ + \ell(q) + \frac{1}{2}kq_0^2.$$

In the case of a physical settlement, the terminal condition is

$$\theta(T, q, \tilde{S}) = N(\tilde{S} + k(q - q_0) - K)_+$$

$$+ 1_{\tilde{S}+k(q-q_0)\geq K}\left(\ell(N - q) + \frac{1}{2}kN(N - 2q)\right) + 1_{\tilde{S}+k(q-q_0)< K}\ell(q) + \frac{1}{2}kq_0^2.$$

## 5. NUMERICAL METHODS

We now present two numerical methods to approximate the solution of our hedging and pricing problem.

### 5.1. Numerical Method for the PDE

Propositions 3.5 and 4.2 show that the hedging and pricing problem boils down to solving a PDE in dimension 3. Factoring out the nominal of the call, the PDE is

$$\partial_t \tilde{\theta} = \underbrace{r\tilde{\theta} + (\mu - r S)\tilde{q} - \mu\partial_S\tilde{\theta} - \frac{1}{2}\sigma^2\partial^2_{SS}\tilde{\theta}}_{(A)} - \underbrace{\frac{1}{2}\gamma N\sigma^2 e^{r(T-t)}(\partial_S\tilde{\theta} - \tilde{q})^2}_{(B)} + \underbrace{\frac{V_t}{N}H(\partial_{\tilde{q}}\tilde{\theta})}_{(C)},$$

with a final condition that depends on the nature of the settlement and on whether or not we consider permanent market impact (if the permanent market impact is taken into account, then we assume $\mu = r = 0$).

To approximate a solution of this PDE with the terminal condition corresponding to our problem, we first split the equation into three parts to consider a numerical scheme based on operator splitting. For (A), we consider an implicit finite difference scheme. We always start with this step that smoothes the terminal condition because there is a singularity at time $T$. For (B), we use a monotonic explicit scheme *à la* Godunov, except at the boundaries (see below). For (C), we use a semi-Lagrangian method because it provides the optimal control directly (see Falcone and Ferretti 2002, for more details on

the classical numerical methods for HJB equations). Regarding the boundary conditions, we can use a grid for $q$ that is sufficiently large to search for an optimum inside the domain. However, this is not the case as far as the finite differences in $S$ are concerned (this is related to the fact that $\theta$ describes a zero-sum game). Therefore, we need to specify boundary conditions for the minimum and maximum values of $S$ ($S_{\min}$ is assumed to be far below the strike $K$, and $S_{\max}$ is assumed to be far greater than $K$). Because $\theta$ is the price of the call option, a natural condition is $\partial_{SS}\theta = 0$ at $S_{\min}$ and $S_{\max}$. However, this condition leads to a globally nonmonotone scheme (see the seminal paper by Crandall and Lions 1984, on monotone schemes). In practice, this scheme provides good results (see below). However, because it requires setting boundary conditions that are not exactly in line with the underlying financial problem, we also develop an alternative method.

## 5.2. Tree-Based Approach

The above numerical method is based on a finite difference scheme and therefore requires artificially setting the boundary conditions. A way to avoid setting the boundary conditions is to use a tree-based approach. The underlying idea is to discretize the problem and to use the same change of variables as in the continuous model to obtain a way to approximate $\theta$.

We consider the subdivision $t_0 = 0, \ldots, t_j = j\Delta t, \ldots, t_J = J\Delta t = T$. We also consider the sequences $(X_j)_j$, $(q_j)_j$, $(S_j)_j$ defined by the following equations:

- $S_{j+1} = S_j + \mu\Delta t + \sigma\sqrt{\Delta t}\epsilon_{j+1}$, where the $\epsilon_j$s are i.i.d. with $\mathbb{E}[\epsilon_j] = 0$ and $\mathbb{V}[\epsilon_j] = 1$,
- $q_{j+1} = q_j + v_j\Delta t$,
- $X_{j+1} = e^{r\Delta t}X_j - v_j S_j\Delta t - L(\frac{v_j}{V_{j+1}})V_{j+1}\Delta t$.

Our goal is to maximize over $\{(v_j)_{0\leq j<J}, |v_j| \leq \rho_m V_{j+1}\}$, the following expression:

$$\mathbb{E}\left[-\exp\left(-\gamma\left(X_J + q_J S_J - \Pi(q_J, S_J)\right)\right)\right].$$

For that purpose, we introduce the value functions:

$$u_j(x, q, S) = \mathbb{E}\left[-\exp\left(-\gamma\left(X_J + q_J S_J - \Pi(q_J, S_J)\right)\right) \mid X_j = x, q_j = q, S_j = S\right].$$

The Bellman equation associated with the problem is

$$\forall j \in \{0, \ldots, J-1\}, \qquad u_j(x, q, S) =$$

$$\sup_{|v|\leq\rho_m V_{j+1}} \mathbb{E}\left[u_{j+1}\left(e^{r\Delta t}x - vS\Delta t - L\left(\frac{v}{V_{j+1}}\right)V_{j+1}\Delta t, q + v\Delta t, S + \mu\Delta t + \sigma\sqrt{\Delta t}\epsilon_{j+1}\right)\right],$$

and

$$u_J(x, q, S) = -\exp\left(-\gamma\left(x + qS - \Pi(q, S)\right)\right).$$

If we write $u_j(x, q, S) = -\exp(-\gamma e^{r(J-j)\Delta t}(x + qS - \theta_j(q, S)))$, then the Bellman equation becomes

$$\theta_j(q, S) = \frac{e^{-r(J-j)\Delta t}}{\gamma}\inf_{|v|\leq\rho_m V_{j+1}}\log\mathbb{E}\left[\exp\left(\gamma e^{r(J-(j+1))\Delta t}\left(qS(e^{r\Delta t}-1) + L\left(\frac{v}{V_{j+1}}\right)V_{j+1}\Delta t\right.\right.\right.$$

$$-(q + v\Delta t)(\mu\Delta t + \sigma\sqrt{\Delta t}\epsilon) + \theta_{j+1}(q + v\Delta t, S + \mu\Delta t + \sigma\sqrt{\Delta t}\epsilon)\bigg)\bigg)\bigg], \quad 0 \le j < J,$$

and

$$\theta_J(q, S) = \Pi(q, S).$$

We now consider a trinomial tree that corresponds to

$$\epsilon = \begin{cases} \alpha & \text{with probability } \frac{1}{2\alpha^2} \\ 0 & \text{with probability } 1 - \frac{1}{\alpha^2} \\ -\alpha & \text{with probability } \frac{1}{2\alpha^2}, \end{cases}$$

where $\alpha > 1$.[8]

Each node of the tree represents a pair $(j, S_j)$ where $S_j \in \{S_0 + \mu j\Delta t + \sigma\sqrt{\Delta t}p\alpha, -j \le p \le j\}$, and the tree is naturally recombining because the drift is constant and the noise symmetrical. At a given node $(j, S_j)$, we compute the value of $\theta_j(q, S_j)$ for $q$ on a specified grid (the natural boundaries when $\mu = r = 0$ are $q_{\min} = 0$ and $q_{\max} = N$ for a call option) by using the above recursive equations. In particular, if the market volume is assumed to be constant (equal to $V$), then the step $\Delta q$ of the grid in $q$ should be such that $\rho_{\mathrm{m}}V$ is a multiple of $\Delta q$.

Recursively, by backward induction, we end up at node $(0, S_0)$ with the price of the call for any $q_0$. Also, we get the optimal strategy at each node as a function of $q$.

This method is preferable over the first one because there is no issue with respect to the boundaries in $S$. However, like all tree methods, it ignores the risk of an important price move over a short period of time.

## 6. NUMERICAL EXAMPLES AND COMPARISON WITH THE BACHELIER MODEL

### 6.1. Examples without Permanent Market Impact

To exemplify the use of our model and the effectiveness of our numerical methods, we consider the following reference scenario with no permanent market impact. This reference scenario corresponds to rounded values for the stock Total SA (the most important component of the CAC 40 Index):

- $S_0 = 45$,
- $\sigma = 0.6 \, \text{€/day}^{1/2}$ or $.\text{day}^{-1/2}$—it corresponds to an annual volatility approximately equal to 21%,
- $T = 63$ days,
- $V = 4{,}000{,}000$ shares/day,
- $N = 20{,}000{,}000$ shares,
- $L(\rho) = \eta|\rho|^{1+\phi}$ with $\eta = 0.1/\text{stock/day}$, and $\phi = 0.75$.

For our reference case, 6.1 we consider $\mu = r = 0$. Our choice for the risk aversion

---

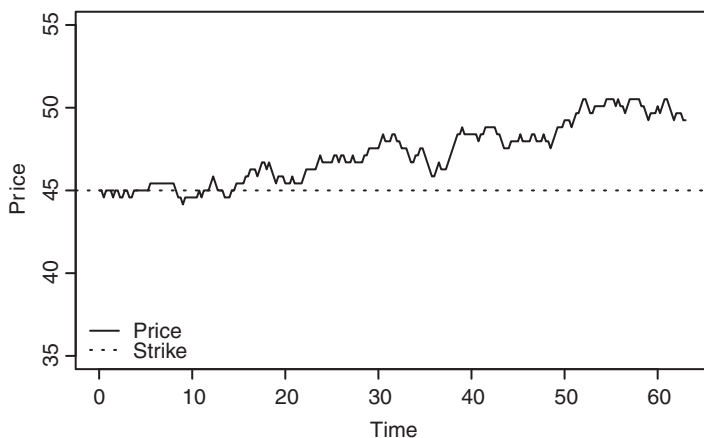[8]In examples, we consider $\alpha = \sqrt{2}$.
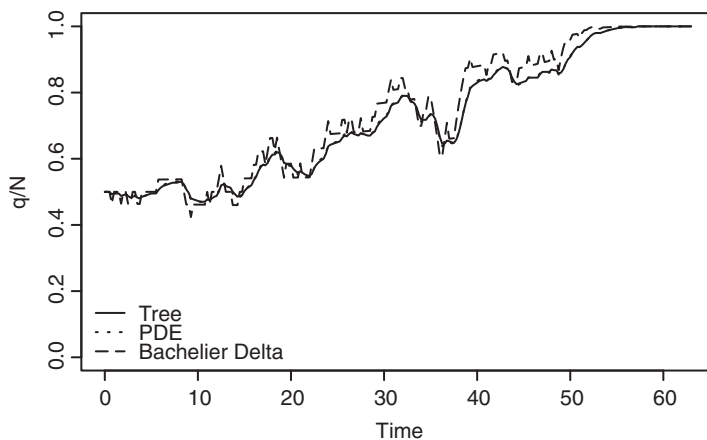
FIGURE 6.1. Trajectory of the stock price.



FIGURE 6.2. Results with the tree-based method and the finite difference scheme (PDE).

parameter is $\gamma = 2 \cdot 10^{-7}/€$. We consider a call option with strike $K = 45$ (at-the-money call option).

Also, we consider (by default) the case where $\rho_m = 500\%$ so that, in practice, there is no constraint on the participation rate. For the terminal cost, we use the form presented in Remark 2.4 with a participation rate equal to $\rho_m$.

Figure 6.2 presents the outcomes when $q_0 = 0.5N$ (initial Bachelier $\Delta$) in the case of a physical delivery for the trajectory of the stock price (compatible with the structure of the tree with four levels of nodes per day) represented in Figure 6.1. This trajectory corresponds to an exercise of the option at time $T$. Similar results could be obtained in the case of $S_T < K$.

We use the two numerical methods presented above to illustrate the optimal strategy, and we also plot the Bachelier $\Delta$ as a benchmark. Figure 6.2 shows that the finite difference scheme and the tree-based method provide almost identical results as far as the optimal strategy is concerned. Moreover, this optimal strategy is different from the

TABLE 6.1
Prices of the Call Option for the Two Numerical Methods

| Model/Method | Bachelier | Tree-based approach | PDE approach |
|---|---|---|---|
| Price | 1.900 | 2.060 | 2.067 |

hedging strategy in a Bachelier model. As opposed to other papers in the literature, the optimal strategy in our model does not oscillate around the Bachelier $\Delta$ but instead is conservative. Our strategy is smoother because the trader avoids buying too many shares to avoid selling them afterward, due to the execution costs.

In terms of prices, the results obtained with the PDE method and with the tree-based approach are very close (Table 6.1). However, the difference between the Bachelier price and the price in our approach is significant.[9]

### 6.2. Influence of the Parameters

We now illustrate the main drivers of the difference between our approach and the Bachelier model.

*6.2.1. Execution costs.* First, because there are execution costs in the model, we illustrate the role of $\eta$. We consider the previous scenario with a physical delivery, but with $\eta \in \{0.01, 0.05, 0.1, 0.2\}$. The results obtained with the tree-based approach are shown in Figure 6.3.
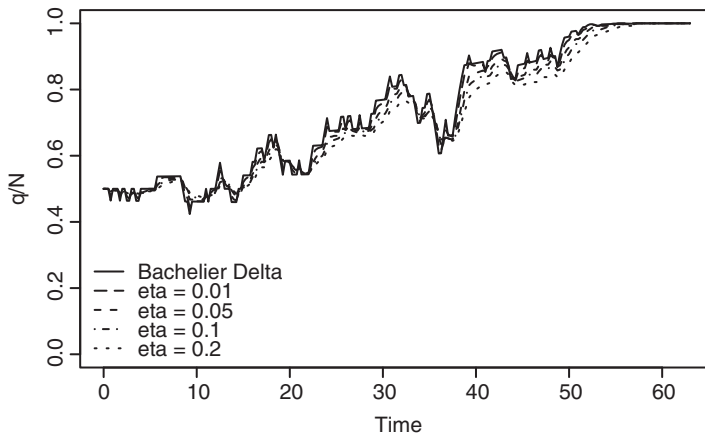


FIGURE 6.3. Optimal strategies for different values of $\eta$ (tree-based approach).

The effect of execution costs is clear: the higher the execution costs, the smoother the optimal strategy. The trader wants to avoid costly erratic changes in his or her portfolio because of execution costs. Furthermore, the optimal portfolio gets closer to $0.5N$ when

---

[9]We divide the prices by $N$ to obtain meaningful figures.

TABLE 6.2

Prices of the Call Option for Different Levels of Liquidity (Tree-Based Approach)

| $\eta$ | 0.2 | 0.1 | 0.05 | 0.01 | 0 (Bachelier) |
|---|---|---|---|---|---|
| Price of the call | 2.144 | 2.060 | 2.007 | 1.943 | 1.900 |

the execution costs increase. This is the same idea as above: because the trader does not know whether he or she will eventually have to deliver $N$ shares or 0, he or she wants to avoid round trips. Therefore, the trader stays closer to $0.5N$ when the liquidity of the underlying asset decreases.

Table 6.2 shows that the price of the call increases with $\eta$, as expected.

*6.2.2. Initial position.* Another parameter linked to liquidity is $q_0$. To understand the role of the initial number of shares, we show in Figure 6.4 the optimal strategies for $q_0 = 0$ and $q_0 = 0.5N$. To be even more realistic, we add a participation constraint $\rho_m = 50\%$ in Figure 6.5.

The associated prices are given in Table 6.3.

Table 6.3 shows that there is a substantial difference between the price of the call option when $q_0 = 0$ and when $\frac{q_0}{N} = 0.5$, especially when a participation constraint is imposed. The rationale for this difference is the cost of building a position consistent with the risk
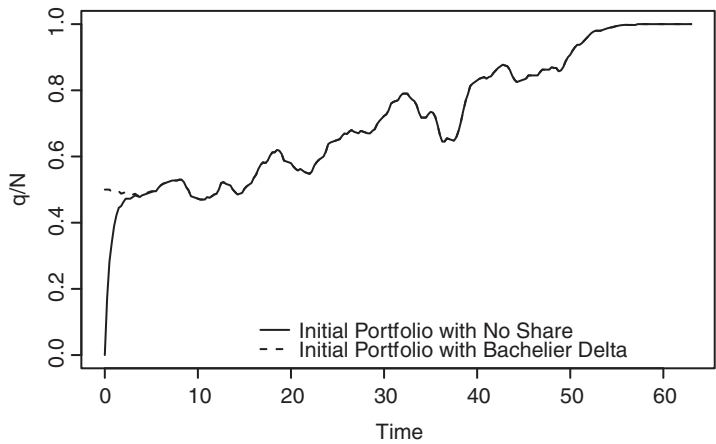


FIGURE 6.4. Optimal portfolio for different values of $q_0$ (tree-based approach).

TABLE 6.3

Prices of the Call Option for Different Values of the Initial Portfolio and Different Participation Constraints (Tree-Based Approach)

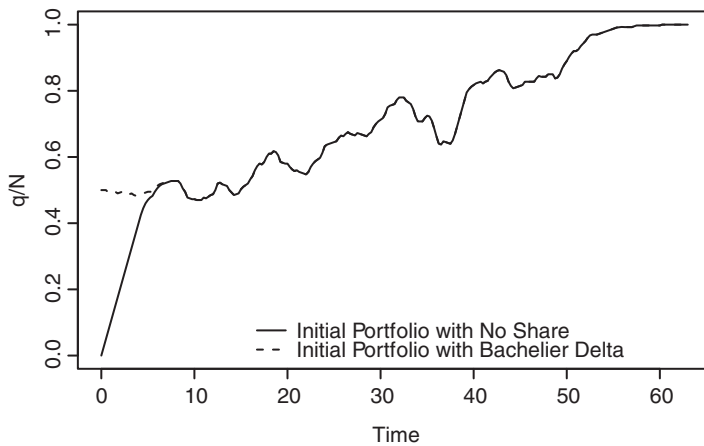| Values of the parameters | $q_0 = 0$ | $q_0 = 0, \rho_m = 50\%$ | $q_0 = 0.5N$ | $q_0 = 0.5N, \rho_m = 50\%$ |
|---|---|---|---|---|
| Price of the call | 2.182 | 2.653 | 2.060 | 2.100 |

FIGURE 6.5. Optimal portfolio for different values of $q_0$ when $\rho_m = 50\%$ (tree-based approach).

linked to the option. This is clearly seen in Figures 6.4 and 6.5. The two portfolios are almost the same after a few days. However, the first few days are used by the trader to buy shares in order to obtain a portfolio close to the portfolio he or she would have had, had he started with the $\Delta$ in the Bachelier model.

*6.2.3. Price risk and risk aversion.* One of the main parameters when dealing with options is volatility. Here, the influence of the parameter $\sigma$ is clear. The more volatile the stock, the closer to the Bachelier $\Delta$ the hedging strategy is. Also, the price of a call is an increasing function of $\sigma$. What is interesting when it comes to risk is not $\sigma$ but $\gamma$, the risk aversion parameter, because there are two risks of two different natures:

- The first risk is linked to the optional dimension of the contract: the trader has to deliver either $N$ shares or none. Being averse to this risk encourages the trader to stay close to a neutral portfolio with $q = 0.5N$.
- The second risk is linked to the price at which the shares are bought or sold: the trader knows that, at time $T$, his or her portfolio will consist of either 0 or $N$ shares depending on $S_T$, and the price the trader pays to buy and sell the shares is random. Being averse to price risk encourages the trader to have a portfolio that evolves in the same direction as the price, as is the case in the Bachelier model.

Several values of $\gamma$ are considered in Figures 6.6 and 6.7 to see these two effects. Figure 6.6 shows that the second effect dominates for small values of $\gamma$. When $\gamma$ is really small, the trader is not really interested in hedging and he or she just wants to minimize the cost of delivering the shares (if the option is exercised). Therefore, the hedging strategy is smooth for very small values of $\gamma$. As $\gamma$ increases, the hedging strategy follows the price movement more and more, like the Bachelier $\Delta$: this is the second risk. Now, to see the first effect, we need to consider high values of $\gamma$. Figure 6.7 shows that when $\gamma$ increases above a certain threshold, then the hedging strategy becomes more and more conservative and "close" to $0.5N$: the trader does not want to buy too much because he or she is afraid of being forced to sell afterward.
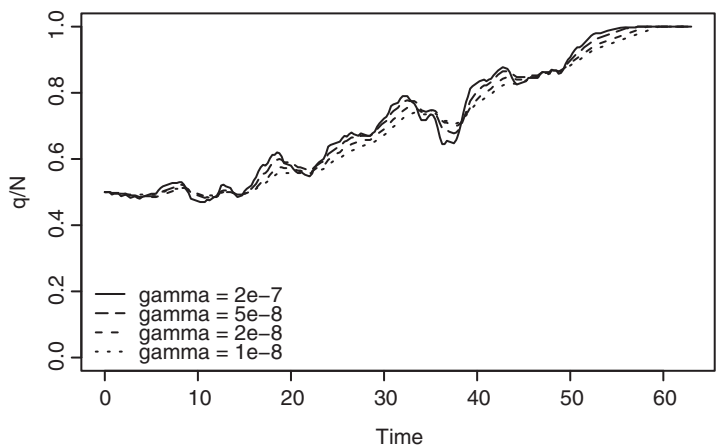
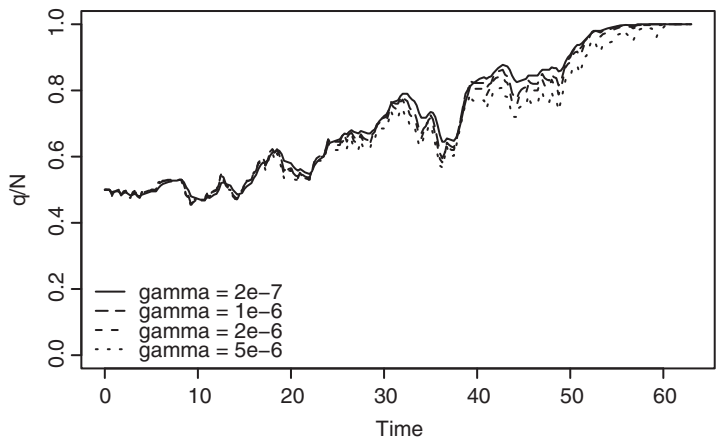FIGURE 6.6. Optimal portfolio for small values of $\gamma$ (tree-based approach).



FIGURE 6.7. Optimal portfolio for large values of $\gamma$ (tree-based approach).

TABLE 6.4
Prices of the Call Option for Different Values of $\gamma$ (Tree-Based Approach)

| $\gamma$ | $1 \cdot 10^{-8}$ | $2 \cdot 10^{-8}$ | $5 \cdot 10^{-8}$ | $2 \cdot 10^{-7}$ | $1 \cdot 10^{-6}$ | $2 \cdot 10^{-6}$ | $5 \cdot 10^{-6}$ |
|---|---|---|---|---|---|---|---|
| Price of the call | 1.955 | 1.968 | 1.994 | 2.060 | 2.207 | 2.308 | 2.521 |

In terms of prices, the effect however is unambiguous. Table 6.4 shows that the more risk averse the trader is, the more he or she charges for the risk.

*6.2.4. Drift and interest rates.* We have discussed the role of the main parameters. To finish this section on comparative statics, we focus on the respective roles of $r$ and $\mu$. In the classical Bachelier (or Black–Scholes) setting, there is no place for the drift of the
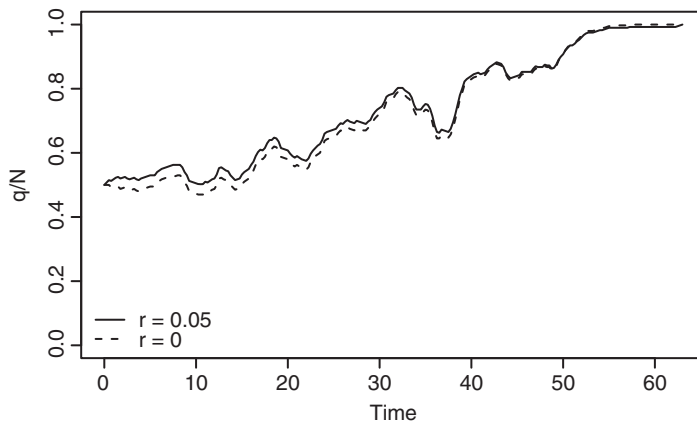
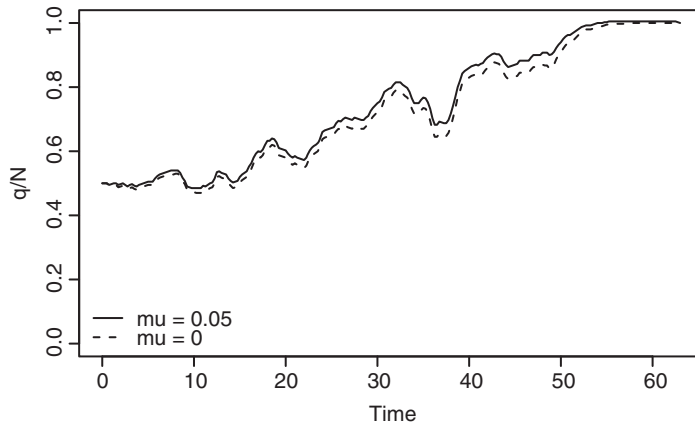FIGURE 6.8. Optimal portfolio for different values of $r$ (tree-based approach).



FIGURE 6.9. Optimal portfolio for different values of $\mu$ (tree-based approach).

underlying asset because the payoff can be replicated. Here, the situation is different. There is indeed, in addition to the hedging problem, another problem of portfolio management in which the trader has to choose the optimal repartition between cash and stock. Figure 6.8, where $\mu = 0$, shows that an increase of $r$ from 0% to 5% leads to two effects. As in the classical theory, an increase in the interest rate leads to more shares in the hedging portfolio. This is what we observe, except at the end of the period. The second effect, explaining the behavior near time $T$, is a pure portfolio management effect. Because a cash position is profitable ($r = 5\%$) compared to a long position in the stock, the trader puts less weights on stocks compared to the situation $r = 0$.

The same portfolio management effect is at play as far as $\mu$ is concerned. When $\mu$ increases, holding shares is more profitable and the hedging portfolio contains more shares (see Figure 6.9).

### 6.3. The Difference between Physical and Cash Settlement

One important difference between our model and most of the models in the literature is that we differentiate between physical and cash settlements. To illustrate this point, we use the reference scenario again but with a participation constraint $\rho_m = 50\%$. Figure 6.10
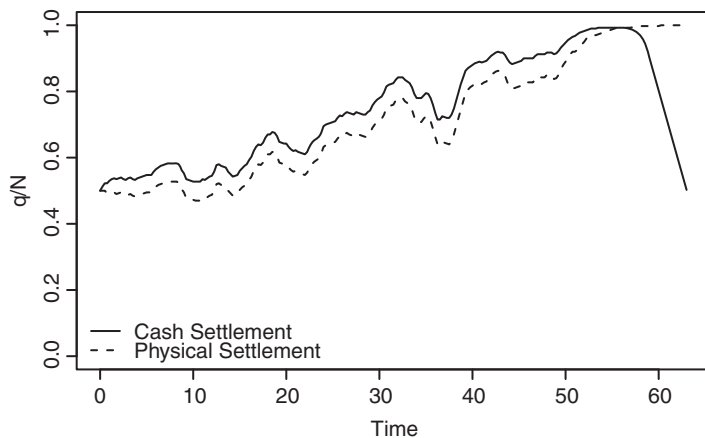


FIGURE 6.10. The difference between a physical settlement and a cash settlement (tree-based approach).

shows that there is an important difference between the two types of settlements when the nominal is large. In both cases, the optimal strategy consists of buying (selling) when the price of the underlying asset is moving up (down) to hedge the position. Hence, when the price $S_t$ is far above $K$ for $t$ close to $T$, the hedging portfolio contains a large number of shares. In the case of a physical delivery, this is fine because the trader has to deliver $N$ shares at expiry if the price stays above $K$. However, in the case of a cash settlement, the trader needs to deliver cash. Figure 6.10 shows that, in order to have cash (and in fact to liquidate the position), the trader progressively sells his or her shares near expiry (given the final cost function we considered, the trader continues to sell with a participation rate to the market equal to 50% after time $T$). To explain what happens far from the expiry date, we rely on a second and more subtle effect. Because the portfolio contains a lot of shares near expiry if $S$ is high, the actual payoff depends on $S$ not only through $(S - K)_+$ but also through $\ell(q)$, which depends implicitly on $S$. Hence, if one applies the classical reasoning, the hedging strategy consists in buying more shares of the underlying stock when the price is increasing. This is the rationale for the difference we observe between the two strategies in Figure 6.10, except near time $T$.

Table 6.5 shows that a cash settlement is more expensive than a physical settlement. The underlying reason is the final liquidation cost when the option expires in the money.

### 6.4. Comparison with the Bachelier Model

Proposition 3.2 argues that the price in the Bachelier model is lower than the price in our model when $\mu = r = k = 0$. This is natural because our model includes additional costs linked to liquidity. An important point then is to understand what happens in practice when one uses the Bachelier model and has to pay the execution costs when

TABLE 6.5
Prices of the Call Option for Cash Delivery and Physical Delivery (Tree-Based Approach)

|  | Cash settlement $\rho_m = 50\%$ | Physical settlement $\rho_m = 50\%$ |
| --- | --- | --- |
| Price of the call | 2.401 | 2.100 |

rebalancing the $\Delta$-hedging portfolio (at discrete points in time). This scenario highlights the fundamental trade-off between a low mis-hedge (when $\Delta$-hedging is done at high frequency) and low execution costs (when $\Delta$-hedging is done at low frequency).

The formula for the $\Delta$ in a Bachelier model (when $r = 0$) is

$$\Delta_t^B = \mathbb{P}[\, S_T \geq K | \, S_t] = \Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right).$$

In order to carry out a fair comparison between our model and the Bachelier model, we consider several frequencies for the $\Delta$-hedging process.

Let $t_0 = 0, \ldots, t_i = i\delta t, \ldots, t_M = M\delta t$ be a subdivision of $[0, T]$. If at time $t_i$ the $\Delta$ of the Bachelier model is $\Delta_{t_i}^B$, then we assume that the difference in $\Delta$ (i.e., $\Delta_{t_i}^B - \Delta_{t_{i-1}}^B$) is executed using a perfect time-weighted average price (TWAP) algorithm over the period $[t_i, t_{i+1}]$. In other words, the execution speed is[10]

$$v_t = \begin{cases} v_{(0)} := \frac{q_{t_1} - q_0}{\delta t} = \frac{\Delta_0^B - q_0}{\delta t} = 0 & \text{if } t < t_1 \\ v_{(i)} := \frac{q_{t_{i+1}} - q_{t_i}}{\delta t} = \frac{\Delta_{t_i}^B - \Delta_{t_{i-1}}^B}{\delta t} & \text{for } t \in [t_i, t_{i+1}),\ 1 \leq i < M. \end{cases}$$

Over each time interval $[t_i, t_{i+1})$ the price obtained by the trader (excluding the execution costs) is the TWAP over the period:

$$\text{TWAP}_{i,i+1} = \frac{1}{\delta t}\int_{t_i}^{t_{i+1}} S_t dt.$$

A classical result on Brownian bridges leads to the fact that $\text{TWAP}_{i,i+1}|\{S_{t_i}, S_{t_{i+1}}\}$ is Gaussian with

$$\mathbb{E}[\text{TWAP}_{i,i+1}|S_{t_i}, S_{t_{i+1}}] = \frac{S_{t_i} + S_{t_{i+1}}}{2} \quad \text{and} \quad \mathbb{V}[\text{TWAP}_{i,i+1}|S_{t_i}, S_{t_{i+1}}] = \frac{\sigma^2 \delta t}{12}.$$

Now, the execution costs can be computed as

$$\int_0^T L\left(\frac{v_t}{V}\right) V dt = \sum_{i=0}^{M-1}\int_{t_i}^{t_{i+1}} L\left(\frac{v_t}{V}\right) V dt = V\delta t \sum_{i=1}^{M-1} L\left(\frac{v_{(i)}}{V}\right).$$

For the terminal condition, we consider the case of a physical settlement, and the terminal condition of Remark 2.4 with the value of $\rho_m$ of the reference scenario.

To obtain statistics on the PnL when using the Bachelier model, we consider a Monte-Carlo algorithm with 10,000 draws. We draw each trajectory $(S_{t_i})_i$ for the price on the

---
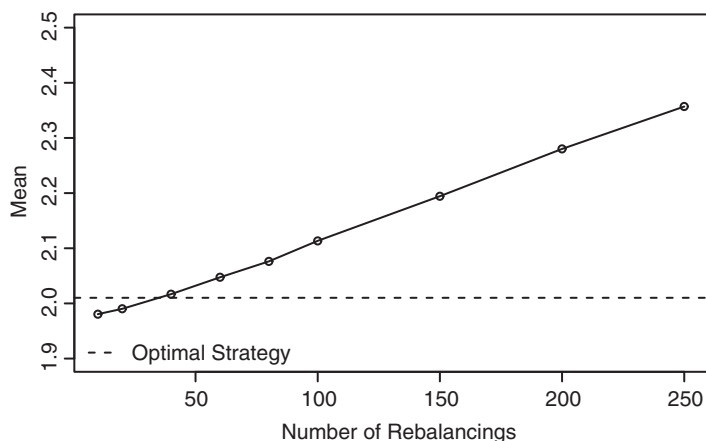
[10]We assume that $q_0 = \Delta_0^B$.

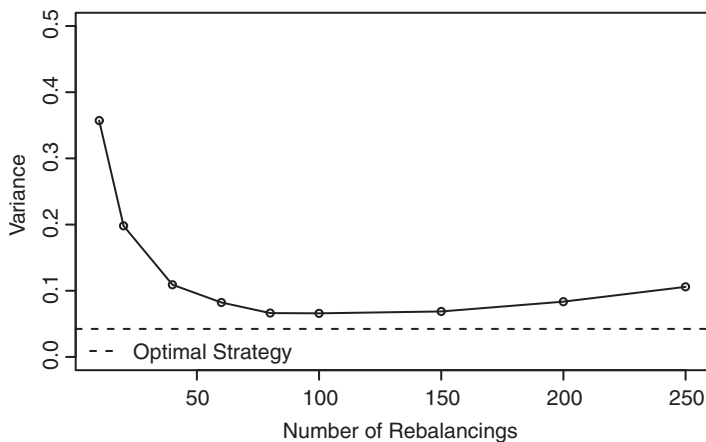FIGURE 6.11. Average cost of the hedging strategy.



FIGURE 6.12. Variance of the cost of the hedging strategy.

time grid $(t_i)_i$. The trajectories are drawn by using Gaussian increments with a standard deviation parameter $\sigma\sqrt{\delta t}$. Then, we draw values for the TWAPs, and we compute a sample PnL associated with our strategy for the sample trajectory $(S_{t_i})_i$. The mean and variance of the PnL (in fact $-$PnL) for several values of $M$ (the number of portfolio rebalancings) are given in Figures 6.11 and 6.12. To compare the outcomes from the Bachelier model to the outcomes from our model, the same Monte-Carlo procedure is used with 253 points in time (four points per day), but the number of shares to be bought or sold over each period $[t_i, t_{i+1})$ is computed using the PDE method to approximate $\theta$.

As expected, Figure 6.11 shows that the cost of $\Delta$-hedging increases with the frequency of the rebalancings. The level of the average costs when one uses our model corresponds approximately to $M = 40$ rebalancings. However, the variance of our strategy is very small compared to the variance of a strategy that consists of $M = 40$ rebalancings with the Bachelier $\Delta$. In fact, the variance of the PnL associated with our strategy is smaller than the variance of the PnL associated with the Bachelier strategy for any values of the number of rebalancings. In the case of $\Delta$-hedging, the variance indeed decreases with $M$

for small values of $M$ but reaches a minimum value (greater than the variance of our strategy) and then increases for large values of $M$ because of the presence of execution costs, which generates variance.

### 6.5. Numerical Examples with Permanent Market Impact

So far, we have only considered the case where there is no permanent market impact. As Proposition 4.2 argues, adding permanent market impact only changes the final condition of the PDE. We use the finite difference scheme to solve the PDE of our reference case when $k = 3 \cdot 10^{-7}$. The optimal strategies are given in Figure 6.13, and the impacted prices are represented in Figure 6.14.

Taking the permanent market impact into account leads to buying more rapidly when the price goes up and selling more rapidly when the price goes down. In fact, there are several effects at play.
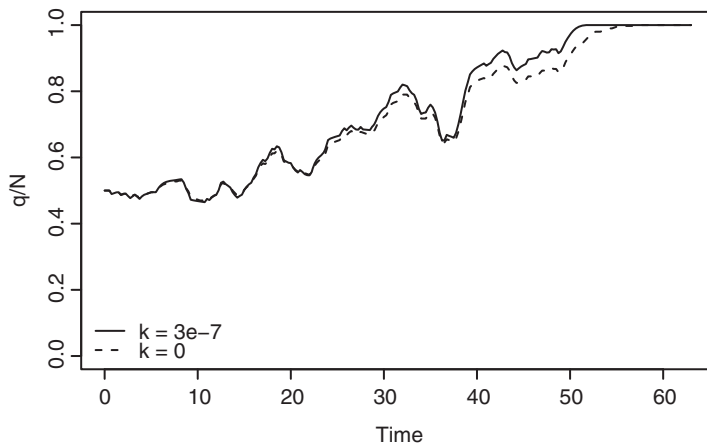


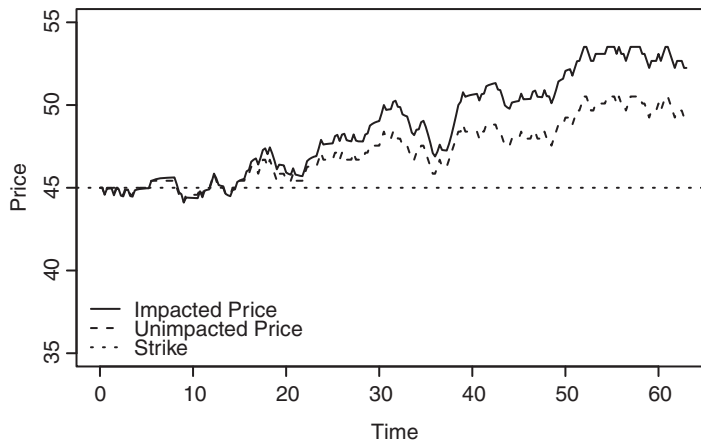FIGURE 6.13. Optimal strategy with and without permanent market impact.



FIGURE 6.14. Impact of the strategy on the price of the underlying asset.

TABLE 6.6
Prices of the Call Option with and without Permanent Market Impact (PDE Approach)

|                  | $k = 0$ | $k = 3 \cdot 10^{-7}$ |
| ---------------- | ------- | --------------------- |
| Price of the call | 2.067   | 2.689                 |

- The first effect is mechanical: when the price of the underlying asset goes up, the position in the underlying asset goes up and it pushes the price of the underlying asset up. Conversely, when the price of the underlying asset goes down, the position in the underlying asset goes down and that pushes the price of the underlying asset down.
- The second effect is strategic: the trader is risk averse and he or she prefers to know whether he or she will have to liquidate at time $T$. Hence, when the trader's position is above a certain threshold, and the price is far above the strike, he or she can buy to push the price up to decrease the level of uncertainty.
- Furthermore, because of the permanent market impact, the trader might be tempted to sell shares to push the price down so that the option expires worthless. We do not observe this effect in Figure 6.13 (in our experiments, it occurs sometimes when the price nears the strike).

In terms of prices, we obtain the results presented in Table 6.6. The table shows that the difference is substantial because the price at which shares are bought is higher and the price at which shares are sold is lower.

## 7. CONCLUSION

In this paper, we present a new model to price and hedge options in the case of an illiquid underlying asset or when the nominal of the option is too large to neglect the execution costs. We show that the price of a call option when the execution costs are taken into account is the solution of a three-variable nonlinear PDE that can be solved using various numerical techniques. The comparisons to classical models show the relevance of our approach. Although our paper focuses on the case of a call option, it can be generalized to other options, with or without physical delivery at maturity. For instance, Guéant et al. (2015) uses a similar framework to price an Accelerated Share Repurchase contract—an execution contract that can be seen as a Bermudan option with an Asian payoff.

## APPENDIX A: PROOFS

*Proof of Lemma 3.1.* By definition,

$$d(e^{-rt} X_t) = e^{-rt} \left( -v_t S_t dt - V_t L \left( \frac{v_t}{V_t} \right) dt \right).$$

Therefore,

$$e^{-rT}X_T - e^{-rt}x = \int_t^T e^{-rs}\left(-v_s S_s ds - V_s L\left(\frac{v_s}{V_s}\right)ds\right)$$
$$= e^{-rt}qS - e^{-rT}q_T S_T + \int_t^T q_s e^{-rs}(-rS_s ds + \mu ds + \sigma dW_s) - \int_t^T e^{-rs}V_s L\left(\frac{v_s}{V_s}\right)ds,$$

where we used an integration by parts. Reorganizing the terms results in

$$X_T + q_T S_T = e^{r(T-t)}(x + qS) \qquad \qquad \Box$$
$$+ e^{r(T-t)}\left(\int_t^T e^{-r(s-t)}q_s(\mu - rS_s)ds + \int_t^T e^{-r(s-t)}q_s\sigma dW_s - \int_t^T e^{-r(s-t)}V_s L\left(\frac{v_s}{V_s}\right)ds\right).$$

*Proof of Proposition 3.2.* We define

$$I(t, q, S, v) = \int_t^T e^{-r(s-t)}q_s(\mu - rS_s)ds + \int_t^T e^{-r(s-t)}q_s\sigma dW_s$$
$$- \int_t^T e^{-r(s-t)}V_s L\left(\frac{v_s}{V_s}\right)ds - e^{-r(T-t)}\Pi(q_T, S_T),$$

and

$$w(t, q, S) = \inf_{v\in\mathcal{A}_t} J_t(q, S, v) = \inf_{v\in\mathcal{A}_t}\mathbb{E}\left[\exp\left(-\gamma e^{r(T-t)}I(t, q, S, v)\right)\right].$$

It is straightforward to verify that $I(t, q, S, 0)$ has a Laplace transform defined on $\mathbb{R}$. Therefore,

$$w(t, q, S) \leq \mathbb{E}\left[\exp\left(-\gamma e^{r(T-t)}I(t, q, S, 0)\right)\right] < +\infty.$$

This proves that $\theta$ is bounded from above.
Coming to the other inequality, Jensen's inequality gives

$$w(t, q, S) \geq \inf_{v\in\mathcal{A}_t}\exp\left(-\gamma e^{r(T-t)}\mathbb{E}[I(t, q, S, v)]\right)$$
$$\geq \exp\left(-\gamma e^{r(T-t)}\mathbb{E}\left[\int_t^T e^{-r(s-t)}q_s(\mu - rS_s)ds - e^{-r(T-t)}N(S_T - K)_+\right]\right)$$

Because $q$ is bounded and because $S_s$ is a Gaussian random variable for all $s$, $w(t, q, s) > -\infty$. Therefore, $\theta$ is also bounded from below.
If $\mu = r = 0$, then

$$\theta(t, q, S) = \frac{1}{\gamma}\log(w(t, q, S)) \geq N\mathbb{E}[(S_T - K)_+]. \qquad \qquad \Box$$

*Proof of Proposition 3.3.* We first prove that $I$ is a concave function of $(v, q)$. Given $t \in [0, T]$ and $S \in \mathbb{R}$, the functions

$$(q, v) \mapsto \int_t^T e^{-r(s-t)}q_s^{t,q,v}(\mu - rS_s)ds = \int_t^T e^{-r(s-t)}\left(q + \int_t^s v_u du\right)(\mu - rS_s)ds$$

and

$$(q, v) \mapsto \int_t^T e^{-r(s-t)}q_s^{t,q,v}\sigma dW_s = \int_t^T e^{-r(s-t)}\left(q + \int_t^s v_u du\right)\sigma dW_s$$

are linear and hence concave.

Because $L$ is convex, the following function is concave:

$$(q, v) \mapsto - \int_t^T e^{-r(s-t)} V_s L\left(\frac{v_s}{V_s}\right) ds.$$

Because $\Pi(q_T, S_T)$ is a convex function of $q_T = q + \int_t^T v_s ds$,

$$(q, v) \mapsto I(t, q, S, v)$$

is a concave function (sum of four concave functions).

To go on, we need a lemma (which is a consequence of Hölder inequality):    □

LEMMA A.1.    *Let $X$ and $Y$ be two random variables. Let $\lambda \in [0, 1]$. The following inequality holds:*

$$\log \mathbb{E}\left[\exp\left(\lambda X + (1 - \lambda) Y\right)\right] \leq \lambda \log \mathbb{E}\left[\exp\left(X\right)\right] + (1 - \lambda) \log \mathbb{E}\left[\exp\left(Y\right)\right].$$

*Let us now recall that*

$$\theta(t, q, S) = \frac{e^{-r(T-t)}}{\gamma} \inf_{v \in \mathcal{A}_t} \log \mathbb{E}\left[\exp\left(-\gamma e^{r(T-t)} I(t, q, S, v)\right)\right].$$

*If $t \in [0, T]$, $S \in \mathbb{R}$, $\hat{q}, \check{q} \in \mathbb{R}$ and $\lambda \in [0, 1]$, then for $\hat{v}, \check{v} \in \mathcal{A}_t$, we have the following inequality:*

$$\theta(t, \lambda \hat{q} + (1 - \lambda)\check{q}, S) \leq \frac{e^{-r(T-t)}}{\gamma} \log \mathbb{E}\Big[ \exp\Big( -\gamma e^{r(T-t)} I\Big(t, \lambda \hat{q}$$
$$+ (1 - \lambda)\check{q}, \lambda \hat{v} + (1 - \lambda)\check{v}, S\Big)\Big)\Big].$$

*Using Lemma A.1 and the concavity of $I$, we have*

$$\theta(t, \lambda \hat{q} + (1 - \lambda)\check{q}, S) \leq \lambda \frac{e^{-r(T-t)}}{\gamma} \log \mathbb{E}\left[\exp\left(-\gamma e^{r(T-t)} I(t, \hat{q}, \hat{v}, S)\right)\right]$$
$$+ (1 - \lambda) \frac{e^{-r(T-t)}}{\gamma} \log \mathbb{E}\left[\exp\left(-\gamma e^{r(T-t)} I(t, \check{q}, \check{v}, S)\right)\right].$$

*Because this inequality holds for all $\hat{v}, \check{v} \in \mathcal{A}_t$, we can take the infima over them on the right-hand side:*

$$\theta(t, \lambda \hat{q} + (1 - \lambda)\check{q}, S) \leq \lambda \frac{e^{-r(T-t)}}{\gamma} \inf_{\hat{v} \in \mathcal{A}_t} \log \mathbb{E}\left[\exp\left(-\gamma e^{r(T-t)} I(t, \hat{q}, \hat{v}, S)\right)\right]$$
$$+ (1 - \lambda) \frac{e^{-r(T-t)}}{\gamma} \inf_{\check{v} \in \mathcal{A}_t} \log \mathbb{E}\left[\exp\left(-\gamma e^{r(T-t)} I(t, \check{q}, \check{v}, S)\right)\right]$$
$$\leq \lambda \theta(t, \hat{q}, S) + (1 - \lambda)\theta(t, \check{q}, S),$$

*which proves the Proposition.*

*Proof of Proposition 3.5.* Given the hypotheses, it is classical to prove that $u$ is a viscosity solution of

$$-\partial_t u - \mu \partial_S u - \frac{1}{2}\sigma^2 \partial_{SS}^2 u - \sup_{|v| \leq \rho_m V_t} \left\{ v \partial_q u + \left( r X - v S - L\left(\frac{v}{V_t}\right) V_t \right) \partial_x u \right\} = 0.$$

Let us consider $\varphi \in C^{1,1,2}((0, T) \times \mathbb{R} \times \mathbb{R})$ and $(t^*, q^*, S^*)$ such that:

- $\theta^* - \varphi$ has a local maximum at $(t^*, q^*, S^*)$,
- $\theta^*(t^*, q^*, S^*) = \varphi(t^*, q^*, S^*)$.

Define $\quad \psi(t, x, q, S) = -\exp[-\gamma e^{r(T-t)}(x + qS - \varphi(t, q, S))] \in C^{1,1,1,2}((0, T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R})$.

Because $\theta(t, q, S) = x + qS + \frac{e^{-r(T-t)}}{\gamma} \log(-u(t, x, q, S))$, $\forall x^* \in \mathbb{R}$, $(t^*, x^*, q^*, S^*)$ is such that $u_* - \psi$ has a local minimum at $(t^*, x^*, q^*, S^*)$.

Using the supersolution property of $u$, we obtain

$$\partial_t \psi(t^*, x^*, q^*, S^*) + \mu \partial_S \psi(t^*, x^*, q^*, S^*) + \frac{1}{2}\sigma^2 \partial_{SS}^2 \psi(t^*, x^*, q^*, S^*)$$
$$+ \sup_{|v| \leq \rho_m V_t} \left\{ v \partial_q \psi(t^*, x^*, q^*, S^*) + \left( rx^* - vS^* - L\left(\frac{v}{V_t}\right) V_t \right) \partial_x \psi(t^*, x^*, q^*, S^*) \right\} \leq 0.$$

As $\psi(t, x, q, S) = -\exp[-\gamma e^{r(T-t)}(x + qS - \varphi(t, q, S))]$, we have

- $\partial_t \psi(t, x, q, S) = \gamma e^{r(T-t)} \psi(t, x, q, S) \partial_t \varphi(t, q, S) + \gamma r e^{r(T-t)} \psi(t, x, q, S)(x + qS - \varphi(t, q, S))$
- $\partial_x \psi(t, x, q, S) = -\gamma e^{r(T-t)} \psi(t, x, q, S)$
- $\partial_q \psi(t, x, q, S) = -\gamma e^{r(T-t)} \psi(t, x, q, S)(S - \partial_q \varphi(t, q, S))$
- $\partial_S \psi(t, x, q, S) = -\gamma e^{r(T-t)} \psi(t, x, q, S)(q - \partial_S \varphi(t, q, S))$
- $\partial_{SS}^2 \psi(t, x, q, S) = \gamma^2 e^{2r(T-t)} \psi(t, x, q, S)(q - \partial_S \varphi(t, q, S))^2 + \gamma e^{r(T-t)} \psi(t, x, q, S) \partial_{SS}^2 \varphi(t, q, S)$.

Hence,

$$0 \geq -\gamma e^{r(T-t^*)} \psi(t^*, x^*, q^*, S^*) \Bigg( -\partial_t \varphi(t^*, q^*, S^*) - r(q^* S^* - \varphi(t^*, q^*, S^*))$$

$$+ \mu(q^* - \partial_S \varphi(t^*, q^*, S^*)) - \frac{1}{2}\sigma^2 \partial_{SS}^2 \varphi(t^*, q^*, S^*) - \frac{1}{2}\gamma \sigma^2 e^{r(T-t^*)}(q^* - \partial_S \varphi(t^*, q^*, S^*))^2$$

$$+ \inf_{|v| \leq \rho_m V_t} \left\{ -v \partial_q \varphi(t^*, q^*, S^*) - L\left(\frac{v}{V_t}\right) V_t \right\} \Bigg).$$

Therefore,

$$0 \geq -\partial_t \varphi(t^*, q^*, S^*) + r \varphi(t^*, q^*, S^*) + q^*(\mu - r S^*) - \mu \partial_S \varphi(t^*, q^*, S^*)$$

$$-\frac{1}{2}\sigma^2 \partial_{SS}^2 \varphi(t^*, q^*, S^*) - \frac{1}{2}\gamma \sigma^2 e^{r(T-t^*)}(q^* - \partial_S \varphi(t^*, q^*, S^*))^2 + V_t H(\partial_q \varphi(t^*, q^*, S^*)).$$

This proves that $\theta$ is a subsolution of the equation.

The same reasoning applies to the supersolution property and this proves the result. $\square$

REFERENCES

ALMGREN, R., and N. CHRISS (2001): Optimal Execution of Portfolio Transactions, *J. Risk* 3, 5–40.

BANK, P., and D. BAUM (2004): Hedging and Portfolio Optimization in Financial Markets with a Large Trader, *Math. Finance* 14, 1–18.

BARDI, M., and I. CAPUZZO-DOLCETTA (2008): *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations Systems & Control: Foundations & Applications*. Boston: Birkhäuser.

BARLES, G., and H. M. SONER (1998): Option Pricing with Transaction Costs and a Nonlinear Black-Scholes Equation, *Finance Stoch.* 2, 369–397.

ÇETIN, U., and L. C. ROGERS (2007): Modelling Liquidity Effects in Discrete Time, *Math. Finance* 17, 15–29.

ÇETIN, U., R. JARROW, and P. PROTTER (2004): Liquidity Risk and Arbitrage Pricing Theory, *Finance Stoch.* 8, 311–341.

ÇETIN, U., R. JARROW, P. PROTTER, and M. WARACHKA (2006): Pricing Options in an Extended Black-Scholes Economy with Illiquidity: Theory and Empirical Evidence, *Rev. Financ. Stud.* 19, 493–529.

ÇETIN, U., H. M. SONER, and N. TOUZI (2010): Option Hedging for Small Investors under Liquidity Costs, *Finance Stoch.* 14, 317–341.

CRANDALL, M. G., and P.-L. LIONS (1984): Two Approximations of Solutions of Hamilton-Jacobi Equations, *Math. Comp.* 43, 1–19.

CVITANIĆ, J., and I. KARATZAS (1996): Hedging and Portfolio Optimization under Transaction Costs: A Martingale Approach, *Math. Finance* 6, 133–165.

CVITANIĆ, J., H. PHAM, and N. TOUZI (1999): A Closed-Form Solution to the Problem of Super-Replication under Transaction Costs, *Finance Stoch.* 3(1), 35–54.

FALCONE, M., and R. FERRETTI (2002): Semi-Lagrangian Schemes for Hamilton–Jacobi Equations, Discrete Representation Formulae and Godunov Methods. *J. Comput. Phys.* 175(2), 559–575.

GATHERAL, J. (2010): No-Dynamic-Arbitrage and Market Impact, *Quant. Finance* 10(7), 749–759.

GUÉANT, O. (2013): Permanent Market Impact Can Be Nonlinear. Working Paper.

GUÉANT, O. (2015): Optimal Execution and Block Trade Pricing: A General Framework. *Appl. Math. Finance* 22(4), 336–65.

GUÉANT, O., J. PU, and G. ROYER (2015): Accelerated Share Repurchase: Pricing and Execution Strategy. *Int. J. Theor. Appl. Finance* 18(3), 1550019.

JAIMUNGAL, S., D. KINZEBULATOV, and D. RUBISOV (2013): Optimal Accelerated Share Repurchase. Working Paper.

LEHALLE, C.-A., M. LASNIER, P. BESSON, H. HARTI, W. HUANG, N. JOSEPH, and L. MASSOULARD (2012): What Does the Saw-Tooth Pattern on US Markets on 19 July 2012 Tell Us about the Price Formation Process, Crédit Agricole Cheuvreux Quant Note.

LEHALLE, C.-A., and S. LARUELLE (2013): *Market Microstructure in Practice*. Singapore: World Scientific Publishing.

LELAND, H. E. (1985): Option Pricing and Replication with Transactions Costs, *J. Finance* 40(5), 1283–1301.

LI, T. M., and R. ALMGREN (2013): A Fully-Dynamic Closed-Form Solution for Δ-Hedging with Market Impact. Working Paper.

Longstaff, F. A. (2001): Optimal Portfolio Choice and the Valuation of Illiquid Securities, *Rev. Financ. Stud.* 14, 407–431.

Platen, E., and M. Schweizer (1998): On Feedback Effects from Hedging Derivatives, *Math. Finance* 8, 67–84.

Rogers, L. C., and S. Singh (2010): The Cost of Illiquidity and Its Effects on Hedging, *Math. Finance* 20(4), 597–615.

Schachermayer, W., and J. Teichmann (2008): How Close Are the Option Pricing Formulas of Bachelier and Black–Merton–Scholes? *Math. Finance* 18(1), 155–170.

Schied, A., T. Schöneborn, and M. Tehranchi (2010): Optimal Basket Liquidation for CARA Investors Is Deterministic, *Appl, Math. Finance* 17(6), 471–489.

Schönbucher, P. J., and P. Wilmott (2000): The Feedback Effects of Hedging in iIlliquid Markets, *SIAM J. Appl. Math.* 61, 232–272.

Sircar, R., and G. Papanicolaou (1998): Generalized Black-Scholes Models Accounting for Increased Market Volatility from Hedging Strategies, *Appl. Math. Finance* 5(1), 45–82.