Hedging Contingent Claims in Markets with Jumps

by

J. Shannon Kennedy

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2007

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Contrary to the Black–Scholes paradigm, an option-pricing model which incorporates the possibility of jumps more accurately reflects the evolution of stocks in the real world. However, hedging a contingent claim in such a model is a non-trivial issue: in many cases, an infinite number of hedging instruments are required to eliminate the risk of an option position. This thesis develops practical techniques for hedging contingent claims in markets with jumps.

A regime-switching model accommodates jumps in (i) the parameters of the stochastic process that drives the underlying asset; and (ii) the price path of the underlying asset itself. We develop numerical techniques for solving the system of partial differential equations that yields the option values in a regime-switching model. When the possible jump sizes of the asset price are drawn from a finite set, all sources of instantaneous risk from an option position can be eliminated by adding a finite number of hedging instruments. We explore a variety of dynamic hedging strategies for a market governed by such a regime-switching process, including techniques that eliminate just some, or all, of the instantaneous risk. The pricing and hedging methodologies are adapted to swing options, a path-dependent derivative traded in the energy markets.

A more realistic representation of jumps in the price path of the underlying asset is made by allowing the amplitudes to be drawn from a continuum. In this case, an infinite number of hedging instruments are required to eliminate the instantaneous risk of an option position, implying that perfect hedging is impossible, even with continuous rebalancing. We demonstrate in a jump-diffusion market that, by imposing delta neutrality and suitably bounding the jump risk and transaction costs at each instant of a continuously rebalanced hedge, the terminal hedging error can be made arbitrarily small. This theoretical treatment motivates a discretely rebalanced dynamic hedging strategy. Hedging examples are considered for options with both European and American-style exercise rights, in a

jump-diffusion market with and without transaction costs. We also investigate semi-static hedging, a buy-and-hold strategy that attempts to replicate the value of a target option at some future time.

Lévy processes constitute a broad class of stochastic processes that exhibit jumps—the jump-diffusion process is a representative member of this group. However, some Lévy processes can generate an infinite number of small jumps over any time period. We demonstrate how our dynamic strategy for hedging under jump diffusion can be used to hedge under any Lévy process.

Acknowledgements

I would first like to sincerely thank my supervisors, Peter Forsyth and George Labahn, for their support and guidance during my years at Waterloo. I would also like to thank my committee members for the time they have dedicated to reading and reviewing my thesis: Matt Davison, Yuying Li, Zoran Miskovic, and Ken Vetzal. In addition, I am grateful to Phelim Boyle for giving me the opportunity to work on some interesting research projects.

I have no doubt that I will look back fondly on my time at Waterloo, due in no small part to the people I have met here. I would like to thank my colleagues in the Scientific Computing Lab, both past and present; in particular, Amélie, Dave, Heath, Omar, Ron, and Simon. Special mention must be made of Yann: I have learned much from him over the years, and without his previous work, my thesis would not have been possible. I will miss the weekly basketball games with the Scicom campus rec team, as well as the trips to Swiss Chalet with Dana and Stacy.

I would like to thank my parents, Gerald and Jean Kennedy. My older sister, Geri-Lynn, and brother-in-law, Pius, have helped me immensely since I moved to Ontario in 2000, and have my eternal thanks. I should also thank my younger sister, Stephanie, or I'll never hear the end of it! My friends back home, in particular Rhonda, have made my time away from Newfoundland much easier by staying in touch.

Thank you
Shannon

Contents

List of Tables

List of Figures

List of Algorithms

Chapter 1

Introduction

1.1 Overview

The Black–Scholes model [10] serves as the standard approach to option pricing and hedging. It furnishes a partial differential equation (PDE) whose solution yields the unique no-arbitrage value of a contingent claim, and also provides a way to perfectly hedge an option position using just the underlying asset and a bond. However, the Black–Scholes model is based on a number of assumptions that do not reflect the true behaviour of stocks in the real world [22]. For one, the asset price paths generated by the governing stochastic process are continuous—they will never experience a jump. Furthermore, the parameters that describe the stochastic process, namely the instantaneous expected return and volatility, are constant.

A more realistic option-valuation model can be achieved by employing a stochastic process that incorporates jumps. With a regime-switching model, the parameters that describe the stochastic process will jump between discrete values. In addition, the price path of the underlying asset can experience a discontinuity when the regime changes, with the possible jump sizes typically drawn from a finite set. If the financial parameters remain

fixed with the jump amplitudes of the asset drawn from a continuum, a jump-diffusion model is obtained. Empirical evidence suggests that both the regime-switching and jump-diffusion models are an improvement over the Black–Scholes paradigm [39, 4, 23, 1]. However, option valuation is more challenging: a system of PDEs must be solved to price contingent claims in a regime-switching model, while the pricing equation for the jump-diffusion model is a partial integro-differential equation (PIDE). More importantly, contingent claims can no longer be perfectly hedged using just the underlying asset and a bond—the markets are said to be incomplete.

A contingent claim in the regime-switching model is subject to two sources of risk: the diffusion risk resulting from the Brownian component of the stochastic process, and the regime-switching risk due to the changes in state that may occur. A hedge portfolio consisting of only the underlying asset and bond is not capable of eliminating all sources of instantaneous risk. However, when the jump sizes of the asset are drawn from a finite set, the instantaneous risk may be removed by adding a finite number of hedging instruments. The regime-switching model thus provides an ideal setting to investigate hedging strategies in an incomplete market, including techniques that remove all instantaneous risk (i.e. a perfect hedge), and those that treat only some of the risk. For example, a delta hedge eliminates the instantaneous diffusion risk, but leaves the regime-switching risk unhedged.

One-factor mean-reverting stochastic processes are typically used to represent the spot price in valuation models for energy derivatives, including those for natural gas [44] and oil [28]. However, this process is not capable of capturing the long-term behaviour of the forward curve [44]. On the other hand, a two-state regime-switching process provides a better fit to the market [21], and solving the resulting system of PDEs is less computationally intensive than treating the equations associated with a multi-factor stochastic process. Swing options written on natural gas are a type of path-dependent derivative that provide flexibility in the timing and amount of delivery [63]. The pricing methodology yields the

optimal exercise strategy the holder of the option should execute. On the opposite side of the trade, the seller of the swing option must protect their short position by hedging.

The jump behaviour of a stock's price path is more accurately modelled by allowing the amplitudes to be drawn from a continuum—a jump-diffusion process captures this feature. A contingent claim in the jump-diffusion model is subject to two sources of risk: the diffusion risk resulting from the Brownian component of the stochastic process, and the jump risk. It is straightforward to remove the instantaneous diffusion risk by imposing delta neutrality. However, the jump risk can only be eliminated by adding an infinite number of hedging instruments, implying that perfect hedging is impossible in practice. Moreover, in general a market participant only has an estimate of the risk-adjusted price dynamics, available through calibration, with no information about the real-world properties (e.g. arrival rate) of the jumps. Nonetheless, it is possible to treat the instantaneous jump risk in a systematic way using a finite and practical number of hedging instruments. Such dynamic hedging strategies are applicable to contingent claims with both European and American-style exercise rights. In addition, since any Lévy process can be approximated as a jump diffusion, the hedging strategies developed for hedging under jump diffusion can be adapted to hedge under a Lévy processes.

Dynamic hedging necessitates frequent rebalancing. Since several instruments are required to hedge under jump diffusion, transaction costs may become prohibitive if they are not accounted for within the dynamic hedging strategy. It is therefore important to incorporate the effect of transaction costs when choosing the hedge portfolio weights that mitigate jump risk. Another way to lessen transaction costs is through semi-static hedging, a buy-and-hold strategy that attempts to replicate the value of a target option at some future time.

1.2 Contributions

The main contributions of this thesis are as follows:

- Numerical techniques are developed for solving the system of partial differential equations that yields the option prices in a regime-switching model.

- An explicit representation is derived for the instantaneous jump risk in a jump-diffusion model. We demonstrate in a jump-diffusion market that, by imposing delta neutrality and suitably bounding the jump risk and transaction costs at each instant of a continuously rebalanced hedge, the terminal hedging error can be made arbitrarily small. This theoretical treatment motivates a discretely rebalanced dynamic hedging strategy.

- The efficacy of the dynamic hedging strategy is demonstrated for European and American options in a jump-diffusion market without transaction costs. The experiments indicate that, using a pricing measure obtained from calibration and selecting the hedge weights in a manner that uses no knowledge about the real-world behaviour of the jumps, results in hedging error that is greatly reduced as compared to a delta hedge.

- The dynamic hedging strategy is demonstrated for European and American options in a jump-diffusion market where transaction costs are present in the form of a relative bid-ask spread. The examples show that the hedge portfolio can be constructed to keep the overall transaction costs manageable while still providing adequate protection from jump risk.

Chapter 2

Option Pricing and Hedging

The Black–Scholes model offers a tractable framework for pricing and hedging options. However, it is based on a number of assumptions that do not conform with what is observed in the real world. In Section 2.1 we introduce the Black–Scholes model and discuss its shortcomings. Some of the deficiencies with Black–Scholes can be overcome by using a stochastic process that more accurately reflects the true behaviour of stocks in the market. We consider two such improvements: Section 2.2 deals with regime-switching processes, while Section 2.3 discusses jump-diffusion processes. Hedging is more involved under these two processes than in the Black–Scholes model. Section 2.4 motivates the importance of hedging, and introduces the general hedging framework.

2.1 The Black–Scholes Model

A contingent claim (or derivative security) is a contract whose value depends on an underlying asset or factor. For example, a call option gives its holder the right, but not the obligation, to buy some financial asset in the future for a fixed price K. The party that purchases the option is said to hold a long position while the seller of the option is short

the contract.

The seminal result in the theory of option pricing is the Black–Scholes partial differential equation (PDE), introduced in 1973 [10]. The stochastic process governing the underlying risky asset in the Black–Scholes model is geometric Brownian motion with drift (GBM), represented by the stochastic differential equation (SDE)

$$dS \;=\; \alpha S\,dt \;+\; \sigma S\,dZ\,, \tag{2.1}$$

where α is the instantaneous expected rate of return, σ is the volatility, and dZ represents the increment of a standard Wiener process. The choice of GBM has several important consequences: asset price paths are continuous and non-negative; asset returns (i.e. $\log S$) are normally distributed over any time span; and asset returns over disjoint time intervals are independent, such that returns over non-overlapping periods are uncorrelated [23].

An arbitrage opportunity is the possibility of making an instantaneous profit without exposing oneself to any risk. In a market free of arbitrage opportunities, any riskless portfolio must earn the risk-free rate r. This fundamental hypothesis, together with other standard assumptions (e.g. continuous trading, no transaction costs, etc.), are used in conjunction with Itô's lemma to derive the Black–Scholes equation: the no-arbitrage value $V(S,t)$ of a European[1] option written on the underlying asset is found by solving

$$\frac{\partial V}{\partial t} \;+\; \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \;+\; rS\frac{\partial V}{\partial S} \;-\; rV \;=\; 0\,, \tag{2.2}$$

with the value at expiry $t = T$ dictated by the option payoff.

The initial value $V(S_0, 0)$ of an option found by solving the PDE in (2.2) has a practical significance apart from being the fair selling price. At time zero the option writer—who

[1] A European option may only be exercised at expiry, whereas an American-style contract may be exercised at any time prior to maturity. To price an American option, a free boundary problem is solved via its linear complementarity formulation [36].

is short the contract—establishes a hedge portfolio, at a cost of $V(S_0, 0)$, consisting of the stock and bond. This portfolio is continuously rebalanced throughout the life of the option such that the amount of stock held at any instant is given by the delta $\frac{\partial V}{\partial S}$. This choice totally eliminates the risk of the overall position and, furthermore, no infusion or withdrawal of money is required during rebalancing—the *delta hedge* is self financing. At expiry, the hedge portfolio's value is exactly enough to cover the payoff if the option is exercised. Hence, the Black–Scholes analysis provides both a way to price an option and a strategy for the seller of the contract to perfectly hedge their short position [75].

Due to its tractability and intuitive appeal, the Black–Scholes model is *the* option valuation archetype. Nevertheless, as with any idealization, it should be viewed as a leading-order approximation. Assets in the Black–Scholes universe are driven by GBM, but this stochastic process is not entirely consistent with the behaviour of stocks observed in the real world. For example:

- Asset price paths are not always continuous—they may exhibit jumps;
- The distribution of returns observed in the market displays a higher degree of kurtosis than the normal distribution. That is, the real-world density of returns has fatter tails and is more sharply peaked [34];
- GBM implies that returns over disjoint time intervals are independent, such that there is no autocorrelation of any type. It is indeed true that there is very little linear autocorrelation observed in the market. However, the absolute returns

$$\left| \log\left(\frac{S_t}{S_{t-1}}\right) \right|$$

tend to display positive autocorrelation. This manifests as 'volatility clustering', where large price swings are followed by large price swings and vice versa [22]. A constant volatility is not compatible with this phenomenon.

The failure of the constant volatility assumption is also evident in the structure of implied volatility[2] observed in the real world (i.e. the smile/skew). According to Black–Scholes, the volatility for calls and puts is constant across all possible strike prices K and maturity times T. However, it is a persistent characteristic of the options market that the volatility implied by the prices is not constant but depends on K and T [23]. The inability of the Black–Scholes model to accommodate this feature in a substantial way is perhaps the most obvious inadequacy it possesses.

It is clear there are deficiencies with Black–Scholes that must be rectified in order to achieve a more realistic option-valuation model. This may be accomplished by using a stochastic process that more accurately reflects the true evolution of stocks. We will consider two such improvements, namely regime-switching and jump-diffusion processes. However, option pricing under these two stochastic processes is more involved than in the Black–Scholes model. Furthermore, unlike the Black–Scholes model, a stock and bond are not sufficient to form a *complete market* whereby any contingent claim can be replicated by dynamically trading in these two instruments.

2.2 Option Pricing under Regime-Switching Processes

Perhaps the most natural extension of the Black–Scholes model is made by allowing the volatility to evolve randomly. For the models proposed in [43] and [41], this is achieved by letting the asset price and variance ($v = \sigma^2$) follow a bivariate diffusion

$$\begin{aligned} dS &= \alpha S\,dt \;+\; \sqrt{v}S\,dZ_1 \\ dv &= p(S,v,t)\,dt \;+\; q(S,v,t)\,dZ_2\,, \end{aligned} \tag{2.3}$$

[2]To price a call or put in the Black–Scholes model, five inputs are required: S, K, r, T, σ. When given the first four inputs and the price of a call or put, the implied volatility σ^* can be imputed.

where $p(S, v, t)$ is often chosen so that v is mean reverting. GARCH models constitute another class of *stochastic volatility* models. A simpler way to introduce another source of randomness is for the volatility in a one-factor diffusion to assume values from a discrete set, with shifts between these volatilities controlled by a continuous Markov chain (Appendix A gives an overview of Markov chains). This specification yields a regime-switching model, where each different level of volatility corresponds to a distinct regime. This model turns out to be less computationally intensive than a two-factor model driven by (2.3).

An N-state analogue to Black–Scholes is the regime-switching model driven by the process

$$dS \;=\; \big(\alpha_i - \tilde{\eta}_i\big) S\, dt \;+\; \sigma^i S\, dZ \;+\; \sum_{j=1}^{N} S\big(\eta^{ij} - 1\big)\, dX_{ij}\,. \tag{2.4}$$

The value of the Markov chain $X = i$ $(i = 1, 2, \ldots, N)$ indicates the current regime number, and the i's appended to the financial parameters in (2.4) denote that the instantaneous expected return α_i and volatility σ^i are regime dependent. The term dX_{ij} governs the transition of the Markov chain between the current state i and the other states j, with

$$dX_{ij} \;=\; \begin{cases} 1 & \text{with probability} \quad \lambda_{ij}dt + \delta_{ij} \\ 0 & \text{with probability} \quad 1 - \lambda_{ij}dt - \delta_{ij} \end{cases} \tag{2.5}$$

and where it is understood that there can only be one transition[3] over an instant; i.e. for $l \neq k$, $Prob\,[(dX_{il} = 1) \cap (dX_{ik} = 1)] = 0$. For $i \neq j$, $\lambda_{ij} \geq 0$ is the transition intensity for state i to j, while

$$\lambda_{ii} \;=\; -\sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}\,.$$

[3]If the Markov chain does not change state then $dX_{ii} = 1$, which corresponds to the trivial transition $i \to i$.

When the transition $i \to j$ occurs, the asset price experiences a deterministic jump of size η^{ij}, $S \to S\eta^{ij}$, where

$$\tilde{\eta}_i \;=\; \sum_{\substack{j=1 \\ j\neq i}}^{N} \lambda_{ij}\left(\eta^{ij}-1\right) \;=\; \sum_{j=1}^{N} \lambda_{ij}\eta^{ij}$$

is the instantaneous expected return due to jumps out of state i. Note that $\eta^{ii} = 1$, otherwise a jump in the asset price will occur over any instant there is not a regime change. Jumps in the underlying that occur without a change in regime may be incorporated by introducing a compound Poisson process in (2.4).

In a two-state model where the underlying evolves according to GBM (i.e. process (2.4) with $N = 2$ and $\eta^{12} = \eta^{21} = 1$), the regimes are characterized as either being of high volatility or low volatility: this is known as a two-state regime-switching lognormal model. For this special case there exist closed-form solutions for European option prices, due to the unique properties of GBM and the availability of the occupation-time conditional density for a two-state continuous Markov chain [58]. For any regime-switching market, European options are amenable to pricing by Monte Carlo simulation.

As with any model, a PDE pricing framework is desirable; the coupled system of option-pricing PDEs is developed in [55]. In the literature, these equations are derived using risk-neutral valuation techniques, namely by applying Itô's lemma to the discounted expected payoff (i.e. the option value) and exploiting the fact that the resulting expression must be a martingale [14, 54]. A more intuitive derivation based on a hedging argument is presented in Appendix A. The PDEs can be solved numerically to price American options and a variety of exotic derivatives, including Asian and lookback options [12].

For the stochastic process in (2.4), the system of option-pricing PDEs is

$$\text{for } i = 1, 2, \ldots, N:$$
$$\frac{\partial V^i}{\partial t} + \frac{1}{2}(\sigma^i)^2 S^2 \frac{\partial^2 V^i}{\partial S^2} + \big(r - \tilde{\eta}_i^{\mathbb{Q}}\big) S \frac{\partial V^i}{\partial S} - r V^i + \sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} \big(V^j - V^i\big) = 0\,, \tag{2.6}$$

where $V^k = V^k(S\eta^{ik}, t)$, such that $V^i = V^i(S\eta^{ii}, t) = V^i(S, t)$ (since $\eta^{ii} = 1$) and $V^j = V^j(S\eta^{ij}, t)$.

There are two fundamental probability measures in mathematical finance. The $\mathbb{P}$ measure is the real-world measure that governs the stochastic evolution of the market, while the $\mathbb{Q}$ measure is the risk-adjusted or pricing measure used for no-arbitrage valuation. The $\mathbb{P}$ and $\mathbb{Q}$ measures are *equivalent*, as they have the same null sets [9]; for example, an event with a probability of zero under $\mathbb{P}$ must also have a probability of zero under $\mathbb{Q}$, and vice versa. If there are no arbitrage opportunities in a market, then there exists at least one pricing measure. Furthermore, if the market is complete, this $\mathbb{Q}$ measure is unique. If required, the appropriate superscript (i.e. either $\mathbb{P}$ or $\mathbb{Q}$) is appended to the model parameters to distinguish the measure with which it is associated.

The Black–Scholes market is complete, so the pricing measure is unique and is found by setting $\alpha^{\mathbb{Q}} = r$ and $\sigma^{\mathbb{Q}} = \sigma^{\mathbb{P}}$. However, the base market in the regime-switching model (i.e. underlying asset + bond) is incomplete, with the consequence that many risk adjustments are possible. In practice, the pricing parameters are obtained by calibrating to the option prices in the market.

The regime-switching model governed by (2.4) offers many improvements over Black–Scholes. From an econometric viewpoint, Hardy [39] carries out a maximum likelihood fit to North American stock indices using a two-state regime-switching lognormal process,

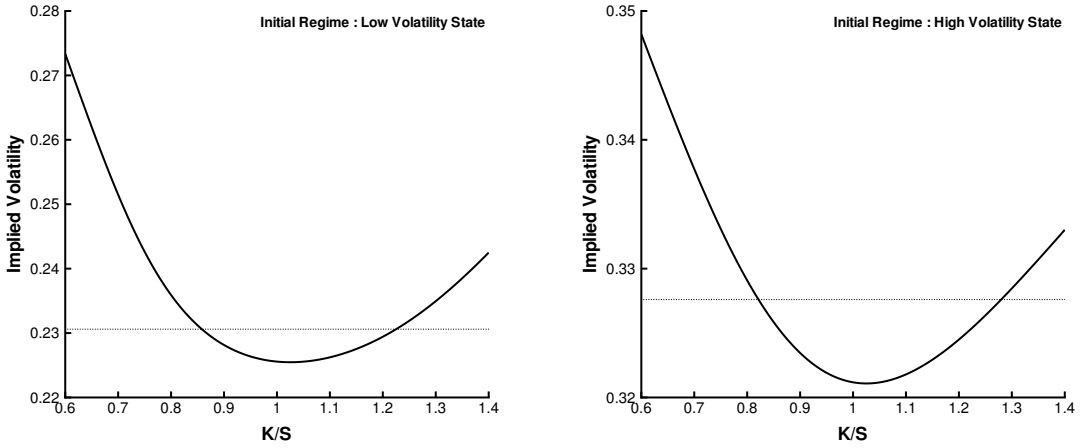


FIGURE 2.1: *Implied volatilities generated by using half-year call option values from a two-state regime-switching lognormal model as the market prices. In the left panel, the calls are priced assuming the current regime is the low-volatility state; the right panel assumes the current regime is the high-volatility state. The volatilities are* $\sigma^L = 0.2$ *and* $\sigma^H = 0.4$ *with intensities* $\lambda^{\mathbb{Q}}_{LH} = \frac{3}{4}$ *and* $\lambda^{\mathbb{Q}}_{HL} = 3$*, and interest rate* $r = 0.05$*. The dotted lines correspond to the square root of the expected variance.*

and finds it performs better than other standard processes. In addition, she demonstrates that the tails of the underlying's transition density are fatter than those generated by GBM. Volatility clustering is an innate characteristic of a regime-switching model. Regime-switching models are also capable of producing the volatility smile/skew observed in the market: by treating the option values from a particular two-state lognormal model as the market prices, the Black–Scholes implied volatilities are imputed, and plotted in Figure 2.1. Furthermore, by including jumps—which introduce a correlation between changes in the volatility and asset price—skews can be generated. The authors in [4] find such a model performs well when calibrating to vanilla and exotic options.

The ideal option valuation model should contain jumps to capture the short-term smile dynamics and include stochastic volatility to encapsulate the behaviour of the smile over the longer term. The regime-switching process (2.4) incorporates these two components in a tractable model.

2.3 Option Pricing under the Jump-Diffusion Process

The model of the previous section may include jumps in the underlying when a regime changes. However, as opposed to being restricted to a few discrete amplitudes as in (2.4), the possible jump sizes in the real world will cover a continuum. By augmenting GBM so that it includes the possibility of random jumps, the jump-diffusion process

$$dS \;=\; \big(\alpha - \kappa\lambda\big)S\,dt \;+\; \sigma S\,dZ \;+\; S\big(J-1\big)\,d\pi \tag{2.7}$$

is obtained,[4] where π is a Poisson process with

$$d\pi \;=\; \begin{cases} 1 & \text{with probability} \quad \lambda dt \\ 0 & \text{with probability} \quad 1-\lambda dt\ . \end{cases} \tag{2.8}$$

In addition, $\lambda > 0$ is the jump intensity and J is the random variable representing the jump sizes, with distribution $g(J)$ and mean $\kappa + 1$. The stochastic mechanism controlling the jumps is termed a compound Poisson process as it admits two sources of uncertainty—the arrival of the jumps is random and when a jump does occur, the amplitude is itself an unpredictable quantity.

Merton introduced the jump-diffusion model in 1976, and considered a normal distribution for the logarithm of the jump sizes [56]. In this case, an analytical pricing formula exists for European options. Kou [50] assumes the logarithm of jump amplitudes follows a double exponential distribution. Again, a variety of closed-form pricing formulae exist.

Following standard arguments [e.g. 23, 1], the value of a European option is found by

[4] The representation (2.7) is shorthand for the more mathematically rigorous

$$dS \;=\; \big(\alpha - \kappa\lambda\big)S\,dt \;+\; \sigma S\,dZ \;+\; d\left(\sum_{i=1}^{\pi} S\big(J_i - 1\big)\right).$$

solving the partial integro-differential equation (PIDE)

$$\mathcal{L}V = 0 \tag{2.9}$$

where

$$\begin{aligned}\mathcal{L}V \equiv & -\frac{\partial V}{\partial t} \\ & - \left(\frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + \left(r - \lambda^{\mathbb{Q}}\kappa^{\mathbb{Q}}\right)S\frac{\partial V}{\partial S} - rV + \lambda^{\mathbb{Q}}\left[\int_0^\infty V(JS,\tau)g^{\mathbb{Q}}(J)\,dJ - V(S,t)\right]\right).\end{aligned}$$

To price an American-style contract, the variational inequality

$$\min(\mathcal{L}V, V - V_e) = 0 \tag{2.10}$$

is solved [5], where V_e denotes the early exercise payoff of the claim. The market in the jump-diffusion model is incomplete. The risk-adjusted parameters can be determined by calibrating to observed option prices, but this calibration is an ill-posed problem, with possibly many parameter sets fitting the option data to a high degree of accuracy [40, 24].

The use of a jump-diffusion model offers many improvements over Black–Scholes. Most obvious is the fact that asset price paths may now contain discontinuities—the possibility of a one-off random shock in the form of a jump is thus incorporated. The probability of a large change in stock price is greater than under GBM, resulting in the distribution of returns having heavier tails than the normal density. The jump-diffusion model is also capable of generating the implied volatility structure observed in the market [23]. To demonstrate this, the option values under a specific jump-diffusion model, with $\log(J) \sim N(\mu^{\mathbb{Q}}, \gamma^{\mathbb{Q}})$, are treated as market prices, and the Black–Scholes volatilities are imputed. The resulting volatility smile and skew are clearly evident in Figure 2.2.

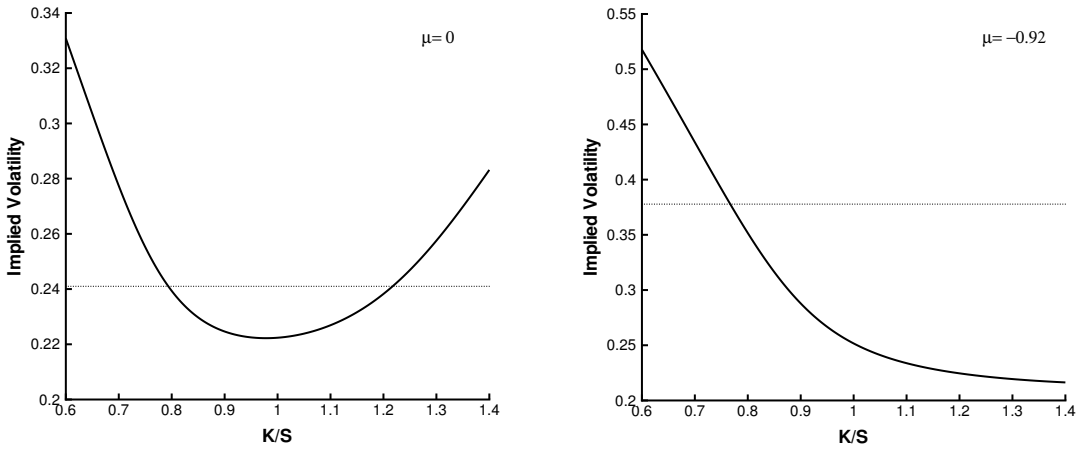


FIGURE 2.2: *Implied volatilities generated by using half-year call option values from a jump-diffusion model as the market prices, where* $\log(J) \sim N(\mu^{\mathbb{Q}}, \gamma^{\mathbb{Q}})$. *The common financial parameters are* $\sigma = 0.2$, $r = 0.05$, $\lambda^{\mathbb{Q}} = 0.1$, *and* $\gamma^{\mathbb{Q}} = 0.425$. *In the left panel, the calls are priced using* $\mu^{\mathbb{Q}} = 0$; *for the right panel,* $\mu^{\mathbb{Q}} = -0.92$. *The dotted lines represent the total volatility* $\sqrt{\sigma^2 + \lambda^{\mathbb{Q}}\left[(\mu^{\mathbb{Q}})^2 + (\gamma^{\mathbb{Q}})^2\right]}$ *[60].*

2.4 Hedging

Consider two counterparties to an options trade: an investor who buys (i.e. is long) a half-year at-the-money[5] European call with strike \$100, and the option writer who sells (i.e. is short) the contract. Using the jump-diffusion parameters from the right panel of Figure 2.2, and $S_0 = 100$, this call will cost \$8.31. Assume the investor obtains a loan to establish her option position while, on the other hand, the option writer takes the \$8.31 and invests it at the risk-free rate. The possible outcomes for the investor and option writer after six months are presented in Figure 2.3. A large increase in the value of the underlying will be ruinous for the option writer. The option writer therefore has to do something more intelligent than simply invest the option premium in a bond.

In general, hedging strategies fall into two categories:

[5]For call options, 'at the money' means $S = K$, 'in the money' means $S > K$, and 'out of the money' means $S < K$. The opposite terminology holds for put options.

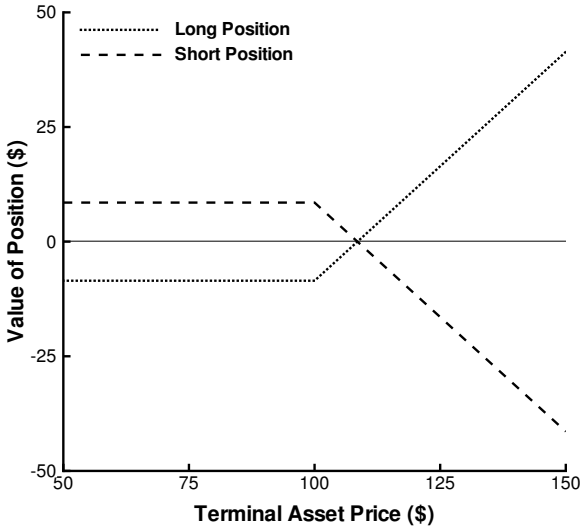


FIGURE 2.3: *Terminal value of long & short positions in a half-year European call option with strike* \$100*, worth \$8.31 at inception. Each position also includes a bond: the holder of the option owes \$8.52 at expiry of the option (i.e.* $t = 0.5$*), while the issuer has \$8.52 in the bank at this time.*

1. **Dynamic Hedging** The hedge portfolio is frequently adjusted to minimize the instantaneous risk. For example, the delta hedge in the Black–Scholes model is continuously rebalanced in order to remove the diffusion risk over every instant.

2. **Static Hedging** This technique attempts to replicate the future value of the short contract with a collection of other instruments. Once established, the hedge is left alone—it is a buy-and-hold strategy. Static hedging is often used in situations where dynamic hedging is impractical. For instance, delta hedging of barrier options in the Black–Scholes framework is problematic since the delta becomes infinite as the spot approaches the barrier; the use of a static hedge is suggested in [32], [15] and [2].

Static hedging is not suitable for many path-dependent contracts, such as options with American-style exercise rights. Furthermore, with dynamic hedging the most recently calibrated pricing parameters can be used at each rebalance time. This thesis will concentrate

on dynamic hedging, although Chapter 10 concerns *semi-static* hedging.

To hedge a target option V, a hedge portfolio is established that contains an amount B in cash, is long e units of the underlying asset S, and long N additional hedging instruments $\vec{I} = [I_1, I_2, ..., I_N]^T$ (written on the underlying) with weights $\vec{\phi} = [\phi_1, \phi_2, ..., \phi_N]^T$. When combined with the short position in the target option, the resulting overall hedged position has value

$$\Pi \;=\; -V \;+\; \underbrace{eS \;+\; \vec{\phi}\cdot\vec{I} \;+\; B}_{\text{Hedge Portfolio } \mathcal{H}}\,.$$

Let transaction costs be present in the form of a relative bid-ask spread, given by BA. If the weight of an instrument is $\rho(t_{n-1})$ before rebalancing and $\rho(t_n)$ after rebalancing, the transaction cost is assumed to be

$$\big|\rho(t_n) - \rho(t_{n-1})\big| \times \frac{\text{BA}}{2} \times \text{ Instrument Value}\,.$$

At time zero the option seller receives the premium $V(S_0, 0)$, and they choose the hedge portfolio weights $e(0)$ and $\vec{\phi}(0)$. These trades must be financed by the bank account, such that

$$\begin{aligned} B(0) \;=\; V(S_0,0) \;-\; e(0)S_0 \;-\; \vec{\phi}(0)\cdot\vec{I}(S_0,0) \\ -\; \big|e(0)\big|\Big(\frac{\text{BA}_S}{2}\Big)S_0 \;-\; \sum_{j=1}^{N}\big|\phi_j(0)\big|\Big(\frac{\text{BA}_j}{2}\Big)I_j(S_0,0)\,, \end{aligned}$$

where BA_S is the relative bid-ask spread for the underlying, and BA_j is the relative bid-ask spread for the j^{th} hedging instrument in $\vec{I}$. In addition, it is assumed the hedging instruments have non-negative value. At each rebalance time t_n the hedge portfolio weights are recalculated. The long position in the underlying asset is updated by purchasing $e(t_n) - e(t_{n-1})$ shares, where $e(t_n)$ is the new computed weight and t_{n-1} denotes the time of the

last rebalancing. The long positions in the hedging instruments are updated similarly by purchasing $\vec{\phi}(t_n) - \vec{\phi}(t_{n-1})$ units. These trades are again financed by the cash account, which after rebalancing contains

$$\begin{aligned} B(t_n) \;=\; & \exp\{r(t_n - t_{n-1})\}B(t_{n-1}) - \Big[e(t_n) - e(t_{n-1})\Big]\left[1 + \operatorname{sgn}\big(e(t_n) - e(t_{n-1})\big)\frac{\mathrm{BA}_S}{2}\right] S_{t_n} \\ & - \sum_{j=1}^{N}\Big[\phi_j(t_n) - \phi_j(t_{n-1})\Big]\left[1 + \operatorname{sgn}\big(\phi_j(t_n) - \phi_j(t_{n-1})\big)\frac{\mathrm{BA}_j}{2}\right] I_j(S_{t_n}, t_n)\,. \end{aligned}$$

The above can be written in the alternative form

$$\begin{aligned} B(t_n) \;=\; & \exp\{r(t_n - t_{n-1})\}B(t_{n-1}) - \Big[e(t_n) - e(t_{n-1})\Big]S_{t_n} - \Big[\vec{\phi}(t_n) - \vec{\phi}(t_{n-1})\Big]\cdot\vec{I}(S_{t_n}, t_n) \\ & - \underbrace{\left[\Big|e(t_n) - e(t_{n-1})\Big|\Big(\frac{\mathrm{BA}_S}{2}\Big)S_{t_n} + \sum_{j=1}^{N}\Big|\phi_j(t_n) - \phi_j(t_{n-1})\Big|\Big(\frac{\mathrm{BA}_j}{2}\Big)I_j(S_{t_n}, t_n)\right]}_{\text{transaction costs}} \end{aligned}$$

to make explicit the cost of transactions due to the bid-ask spread.

An instant after rebalancing, the overall hedged position has value

$$\Pi(t_n) \;=\; -V(S_{t_n}, t_n) \;+\; e(t_n)S_{t_n} \;+\; \vec{\phi}(t_n)\cdot\vec{I}(S_{t_n}, t_n) \;+\; B(t_n)\,.$$

The value of Π at the exercise/expiry time T^* of the target option indicates how well the hedge performed: a perfect hedge will have $\Pi(T^*) = 0$, indicating the hedge portfolio's terminal value exactly covers the payoff from the short position in V.

The markets in the regime-switching and jump-diffusion models are incomplete in the sense that a perfect hedge cannot be constructed using only the stock and bond. For an N-state regime-switching model in which the underlying is tradeable, the introduction of an additional $N-1$ instruments will complete the market. The instantaneous diffusion and regime-switching risk of an option position can be eliminated using these instruments, such

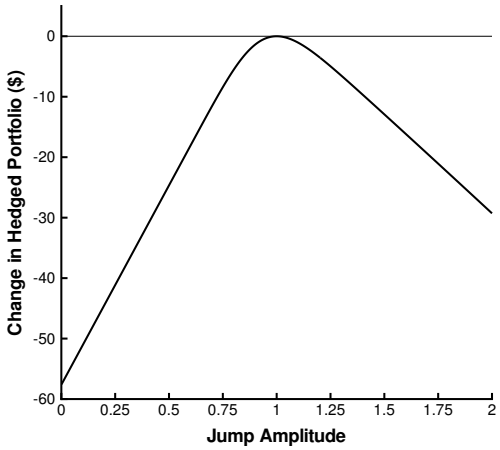


FIGURE 2.4: *Change in the value of a delta-hedged position due to a jump in the price of the underlying asset. The half-year call option ($K = \$100$) being hedged costs \$8.31 at inception, and the jump in the underlying asset is assumed to take place an instant after the option is sold. The jump amplitudes are distributed according to* $\log(J) \sim N(\mu^{\mathbb{Q}}, \gamma^{\mathbb{Q}})$*, with the parameters of the model given by* $\sigma = 0.2$*,* $r = 0.05$*,* $\lambda^{\mathbb{Q}} = 0.1$*,* $\gamma^{\mathbb{Q}} = 0.425$*, and* $\mu^{\mathbb{Q}} = -0.92$*.*

that perfect hedging is possible when rebalancing is done continuously [58]. On the other hand, the incompleteness of the jump-diffusion model is pathological in the sense that it can only be eliminated by introducing an infinite number of instruments [1]. The instantaneous diffusion risk may still be eliminated by imposing delta neutrality,[6] but the best that can be hoped for in terms of the jump risk is reducing it in some way. In this case perfect hedging is impossible, even in the theoretical limit of continuous rebalancing.

Not only is hedging more involved for the incomplete markets we are considering, but ignoring the new sources of risk (i.e. regime-switching and jump risk) may have disastrous consequences. For example, in the jump-diffusion model, delta hedging an option with a convex payoff will result in a loss whenever a jump occurs [59], regardless of its size or direction. Consider a delta hedge that is short an at-the-money, half-year European call with strike \$100, and long the underlying. This hedge eliminates the diffusion risk but

[6] A portfolio Π is said to be delta neutral if its delta $\frac{\partial \Pi}{\partial S}$ is zero.

ignores the jump risk. If a jump occurs, the loss to the position may be substantial, as is evident in the plot contained in Figure 2.4.

Chapter 3

Implementation of the Hedging Strategies

In this chapter, we describe the computational framework for implementing the hedging strategies that will be considered in the sequel. We also discuss the various sources of error that enter into the hedging experiments. The effect of these errors is explored in the context of the most complex example in this thesis, namely hedging an American put under jump diffusion in the presence of transaction costs.

3.1 Steps for Implementing the Hedging Strategies

We use Monte Carlo simulation to measure the effectiveness of a hedging procedure, whereby the strategy is executed for many realizations of the underlying's price path. For each path, the (discounted) relative profit and loss (P&L)

$$\mathrm{e}^{-rT^*}\frac{\Pi(T^*)}{V(S_0,0)} \tag{3.1}$$

is calculated at the exercise/expiry time T^* of the target option V, where $\Pi(T^*)$ is the terminal value of the overall hedged position. Statistical measures of the relative P&L, including the mean and standard deviation, indicate how well the strategy performs (the distribution of profit and loss for a perfect hedge is the Dirac delta function: it has a zero mean and a zero standard deviation).

Regardless of the stochastic process governing the market, the same fundamental steps are carried out for the hedging experiments:

1. The various quantities required for hedging, including the option values and deltas, are precomputed and stored on a series of (S, t) grids;

2. A set of asset paths are simulated;

3. For each path, the hedge portfolio is formed at time zero and rebalanced at those times dictated by the particular strategy; the financial details of forming and rebalancing a hedge portfolio are discussed in Section 2.4. The relative P&L is recorded at the exercise/expiry time of the target option.

Note that the precomputed data is used when calculating the hedge weights, which will usually entail solving a linear system[1] formed through table lookup and interpolation. Since the linear system may be poorly conditioned in certain situations, we use a Truncated Singular Value Decomposition (TSVD) [38] to solve it: small singular values are set to zero—where the cutoff[2] is imposed via a user controlled parameter—and the modified decomposition is used to solve the system in the standard way [65].

We now further describe the above steps and the sources of error (**SoE**) that are introduced.

[1]The linear system typically results from the treatment of a constrained optimization. We also experimented with NAG subroutines to solve the optimization problems directly; in most cases the results agreed. In some cases, however, the NAG solution had a large norm (as compared to a regularized solution).

[2]If $\omega_{\max}$ is the maximum singular value of the matrix, then any singular value less than $\text{CUTOFF} \times \omega_{\max}$ is set to zero.

3.1.1 Precomputing Option Data

The option values are computed by numerically solving the appropriate pricing equation. In the case of the regime-switching model, the system of PDEs (2.6) is solved numerically using methods developed in the next chapter. The numerical treatment of the option-pricing PIDE (2.9) from the jump-diffusion model is based on d'Halluin et al. [34], and is summarized in Appendix B. Once the option values are computed, the Greeks[3] required for hedging are found by numerical differentiation.

SoE.1 Due to truncation, round-off, and localization error, the option quantities found using the numerical solution of the governing equation will not be exact.

The option pricing and Monte Carlo simulation are carried out by separate software. As such, in order to be used in the Monte Carlo simulations, the option data calculated by the pricer must be output onto an (S,t) grid. All times on this output data grid are represented exactly on the mesh used in the option pricer, so no interpolation is necessary in the t-direction. However, it will be computationally efficient to have equally-spaced asset nodes on the data grid used by the hedging simulations, while the asset node spacing of the pricer's mesh is contract dependent (e.g. when pricing a call or put, the asset node spacing is finer near the strike). Therefore, results from the pricer are interpolated (in the S-direction) onto the output data grid using a Lagrange polynomial.

SoE.2 Due to interpolation error, the option data at a point (S,t) on the grid used in the hedging simulations will be different from what would be calculated if (S,t) was represented exactly on the pricer's computational mesh.

For American options, the early exercise region to be used in the hedging simulations is determined from the option pricer. For example, consider an American put, where $V(S,t)$ is the numerical solution and V_e is the payoff: if $V(S_i,t_n) = V_e$ and $V(S_{i+1},t_n) > V_e$, then

[3]This refers to the various partial derivatives of the option value, such as the delta $\frac{\partial V}{\partial S}$ and gamma $\frac{\partial^2 V}{\partial S^2}$.

$[0, S_i]$ is treated as the early exercise region for time t_n.

SoE.3 For time t_n, the true upper boundary for the early exercise region of the numerical solution could be anywhere in (S_i, S_{i+1}). Hence, the possible error in the boundary point depends on the resolution of the asset node spacing of the mesh used by the option pricer.

3.1.2 Simulating Asset Price Paths

The asset price paths are simulated using the Euler–Maruyama method for numerically solving SDEs [48, 77] (this is the stochastic analog to the Euler forward method for ordinary differential equations (ODE)). Algorithm 1 describes the steps for generating a sample path from the regime-switching process (2.4), while Algorithm 2 simulates the jump-diffusion process (2.7). The sample paths for these two processes may in fact be simulated exactly [see, e.g., 23], but Algorithms 1 and 2 are general in the sense they can be used if the drift and diffusion components of the SDE are more elaborate (e.g. time-dependent parameters, mean-reverting processes).

SoE.4 The scheme for numerically solving the SDE has timestepping error, since the Euler–Maruyama method converges strongly with $\mathcal{O}(\sqrt{\Delta t})$ and converges weakly with $\mathcal{O}(\Delta t)$ [49].

The hedging strategy will be executed for all asset paths, and the relative profit and loss (3.1) recorded for each. Statistical measures of the relative P&L will be used to gauge the efficacy of the strategy.

SoE.5 The statistical measures for the relative P&L are subject to Monte Carlo sampling error, which is $\mathcal{O}\left(\frac{1}{\sqrt{M}}\right)$ when M paths are generated.

Algorithm 1 Simulating a path of the N-state regime-switching process (2.4)

Given R_0: Initial regime
S_0: Initial asset price
Δt: Timestep, with $\big(\max_i |\lambda_{ii}|\big)\Delta t \ll 1$
T: Length of path
for $k = 0$ to $\big(\frac{T}{\Delta t} - 1\big)$ **do**
Generate u_k from a $U(0,1)$ distribution
if $u_k \leq |\lambda_{R_k R_k}|\Delta t$ **then** {THE REGIME CHANGES—Determine a new regime}
Generate u_k' from a $U(0,1)$ distribution
for $rnum = 1$ to N **do**
if $rnum \neq R_k$ **then**
if $u_k' \leq \dfrac{1}{|\lambda_{R_k R_k}|} \displaystyle\sum_{\substack{j=1 \\ j \neq R_k}}^{rnum} \lambda_{R_k j}$ **then** {THE REGIME CHANGES TO $rnum$}
$R_{k+1} = rnum$
BREAK FROM INNER FOR LOOP
end if
end if
end for
end if
FIND S_{k+1}
//
Generate ϕ_k from a standard normal distribution
if $R_k \neq R_{k+1}$ **then**
$S_{k+1} = S_k + (\alpha_{R_k} - \tilde{\eta}_{R_k})S_k\Delta t + \sigma^{R_k} S_k \sqrt{\Delta t}\,\phi_k + (\eta^{R_k \to R_{k+1}} - 1)S_k$
else
$S_{k+1} = S_k + (\alpha_{R_k} - \tilde{\eta}_{R_k})S_k\Delta t + \sigma^{R_k} S_k \sqrt{\Delta t}\,\phi_k$
end if
end for

Algorithm 2 Simulating a path of the jump-diffusion process (2.7)

Given S_0: Initial asset price
Δt: Timestep, with $\lambda \Delta t \ll 1$
T: Length of path
for $k = 0$ to $\left(\frac{T}{\Delta t} - 1\right)$ **do**
FIND S_{k+1}
//
Generate ϕ_k from a standard normal distribution
Generate u_k from a $U(0,1)$ distribution
if $u_k \leq \lambda \Delta t$ **then** {THERE IS A JUMP}
Generate J' from a distribution with density $g(J)$
$S_{k+1} = S_k + (\alpha - \kappa\lambda) S_k \Delta t + \sigma S_k \sqrt{\Delta t}\, \phi_k + (J' - 1) S_k$
else
$S_{k+1} = S_k + (\alpha - \kappa\lambda) S_k \Delta t + \sigma S_k \sqrt{\Delta t}\, \phi_k$
end if
end for

3.1.3 Executing a Hedging Simulation

The data on the output grid generated by the option pricer is used during the hedging simulations. For European options, all rebalancing times are known beforehand, so the grid will be constructed such that these times are represented exactly. At each rebalancing time, the hedge weights are chosen using option data that is a function of the prevailing asset price S: if an S value is not represented on the data grid, a Lagrange polynomial is used to interpolate in the S-direction. Due to the early exercise feature of American-style contracts, option data may be required for any time. If neither the asset value or time of (S, t) is represented on the data grid, bilinear interpolation is used.

When hedging under jump diffusion, the hedge portfolio weights are computed by solving a linear system whose entries are integrals. These integrals are precomputed on the data grid using fast Fourier transform (FFT) techniques; the details are provided in Appendix B. During the hedging simulations, the appropriate integral values are determined using the

same interpolation techniques as employed for the other option data.

SoE.6 Option data used by the hedging simulations will be subject to interpolation error. Also, there is numerical integration error when hedging under jump diffusion due to the use of the FFT.

For an American put, the early exercise boundary is output by the option pricer as a set of points $\{(S_n, t_n)\}_{n=1}^{\mathcal{N}}$. Within the hedging simulations, the boundary is represented as piecewise linear function based on this set. If a point (S, t) on the simulated asset path breaches this boundary, then it is deemed optimal to exercise the put.

SoE.7 The early exercise region of an American put—which is not linear—is represented as a piecewise linear function in the hedging simulations.

3.2 The Effect of the Error Sources

Before we examine the sources of error, we demonstrate the simulation algorithms in Algorithms 1 and 2 by employing them for Monte Carlo pricing.[4] We value a half-year, at-the-money European call option with a strike of \$100 in both a regime-switching and jump-diffusion market. The pricing parameters for the three-state regime-switching market come from [4] (and are stated in Section 4.1.4), while the $\mathbb{Q}$ measure for the jump-diffusion market is given in Table 3.1. The results in Table 3.2 demonstrate the Monte Carlo values are in agreement with those prices found by solving the appropriate option-pricing equation.

To examine the effect of the error sources presented in Section 3.1, we consider a specific hedging example. A half-year, at-the-money American put with a strike of \$100 is to be hedged over its lifetime, in a jump-diffusion market with transaction costs, using eight three-month American calls and puts with strikes near $K = \$100$. The $\mathbb{P}$ and $\mathbb{Q}$ measures that

[4] The value of a European option with payoff $f(S_T)$ is given by $\mathrm{e}^{-rT}\mathbb{E}^{\mathbb{Q}}\big[f(S_T)|S_0\big]$, which can be estimated via Monte Carlo simulation [11].

Probability Measure	λ	μ	γ	σ	α
Risk-adjusted ($\mathbb{Q}$)	0.1000	-0.9200	0.4250	0.2000	0.0500
Real-world ($\mathbb{P}$)	0.1000	-0.5588	0.4250	0.2000	0.1779

TABLE 3.1: *The pricing $\mathbb{Q}$ measure and real-world $\mathbb{P}$ measure that characterize the jump-diffusion model used to investigate the effect of the errors introduced into the hedging simulations, where $\log(J) \sim N(\mu, \gamma)$. The dividend yield $q = 0$ and $\alpha^{\mathbb{Q}} = r = 0.05$.*

	Jump Diffusion	Regime Switching		
		Regime 1	Regime 2	Regime 3
Monte Carlo Price	8.301	4.070	8.847	3.928
95% Confidence Interval	$\pm$0.021	$\pm$0.010	$\pm$0.026	$\pm$0.008
Exact Value	8.305	4.064	8.865	3.930

TABLE 3.2: *Price of a half-year European call option, in both a regime-switching and jump-diffusion market, obtained by Monte Carlo simulation. The timestep is 0.000125 years, and 1,000,000 simulations are carried out. The exact value comes from solving the corresponding option-pricing equation. The regime-switching pricing parameters are drawn from [4], while the $\mathbb{Q}$ measure for the jump-diffusion model is given in Table 3.1.*

characterize this market are given in Table 3.1. Details of the hedging strategy, and this example in particular, are discussed in Chapter 8. In short, the hedge portfolio is scheduled to be rebalanced twenty times, at intervals of 0.025 years. For all puts in the portfolio, the early exercise region is monitored—since we assume there are no dividends, it will never be optimal to exercise a call before expiry. If it is deemed optimal for a hedging put to be exercised, it is removed from the hedge portfolio and the hedge is rebalanced. If it is ever optimal to exercise the target put, the hedge is immediately liquidated to cover the short position.

The base simulation set consists of 250,000 asset paths, with a Monte Carlo timestep of $\Delta t = \frac{1}{12000}$ years—the timestep is made small to reduce the Monte Carlo timestepping error. A linear Lagrange polynomial is used for all one-dimensional interpolation. The option pricer employs a pre-determined number of timesteps, and the asset node spacing of its mesh is finer around the strike. The time spacing of the data grid used for the hedging simulations is 0.0125 years. The asset nodes on this grid extend from $S = 0$ to $S = 1000$, with $\Delta S = 0.125$ for $S \in [0, 480]$, $\Delta S = 0.25$ for $S \in [480, 720]$, and $\Delta S = 0.3125$ for $S \in [720, 1000]$. Note that the hedging results for this base case, in terms of statistical measures of the relative P&L (3.1), will be the first entry in all of the forthcoming tables. The 0.2% quantile is a tail measure that gives an indication of the extreme losses the hedger may be exposed to.

The sources of error **SoE.1**–**SoE.3** and **SoE.6**–**SoE.7** are concerned with the precomputation of the option data and its use in the hedging simulations. To gauge the effect of these errors, we refine the computational mesh in the option pricer by a factor of two in both the time and asset price direction, and double $S_{\max}$ (i.e. the value where the mesh is truncated at infinity in the S-direction). The data grid used by the hedging simulations is refined and extended in the same manner. In addition, the time spacing for the early exercise boundary of the American puts is reduced by a factor of two. The original and new

Refinement & Extension	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
Original	-0.0605	0.0215	-0.1403	-0.0144
2×	-0.0604	0.0214	-0.1401	-0.0144

TABLE 3.3: *Effect of mesh/grid refinement and extension on the hedging results, in terms of the relative profit and loss. For both the mesh in the option pricer and the data grid used by the hedging simulations, a new node is inserted between each pair of coarse asset value nodes, and the timestep is halved. In addition, the maximum extent in the S-direction is doubled. For the α% quantile, approximately α% of the 250,000 simulations resulted in a relative P&L less than the reported amount.*

Lagrange Interpolant	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
Linear	-0.0605	0.0215	-0.1403	-0.0144
Cubic	-0.0605	0.0215	-0.1404	-0.0144

TABLE 3.4: *Effect of the interpolation method on the hedging results, in terms of the relative profit and loss.*

results are presented in Table 3.3 for the same 250,000 paths, and are seen to be very close. The results in Table 3.4 correspond to using a cubic Lagrange polynomial for interpolation, with all other settings the same as in the base case. Again, with the same 250,000 asset paths, the results are close. The use of the linear interpolator is preferred, as it is faster and will not introduce spurious maxima or minima.

Both **SoE.4** and **SoE.5** are concerned with the Monte Carlo error. The effect of the timestepping error, **SoE.4**, is examined by cutting the Monte Carlo timestep in half (all other settings are the same as in the base case). The original and new results are presented in Table 3.5, and are in good agreement. The effect of the sampling error, **SoE.5**, is studied by reducing the number of asset paths to 100,000 and increasing the number to 1,000,000. The results are given in Table 3.6, where more digits are used than before to give context to the confidence intervals [78]. We see there is very little difference between using 100,000, 250,000, and 1,000,000 asset paths.

Monte Carlo Timestep	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
$\frac{1}{12000}$ years	-0.0605	0.0215	-0.1403	-0.0144
$\frac{1}{24000}$ years	-0.0604	0.0215	-0.1398	-0.0145

TABLE 3.5: *Effect of the Monte Carlo timestep on the hedging results, in terms of the relative profit and loss.*

Number of Simulations	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
100,000	-0.06051	0.02152	-0.14068	-0.01453
	[-0.06065,-0.06038]	[0.02143,0.02161]	[-0.14285,-0.13880]	[-0.01470,-0.01437]
250,000	-0.06046	0.02146	-0.14028	-0.01440
	[-0.06054,-0.06037]	[0.02140,0.02152]	[-0.14179,-0.13941]	[-0.01453,-0.01428]
1,000,000	-0.06047	0.02150	-0.14066	-0.01451
	[-0.06058,-0.06042]	[0.02147,0.02153]	[-0.14122,-0.13992]	[-0.01458,-0.01445]

TABLE 3.6: *Effect of the number of generated price paths on the hedging results, in terms of the relative profit and loss. The statistical measures are given with the corresponding 95% confidence interval [78]. Each set of paths is generated using a different seed in the random number generator.*

	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
Original	-0.0605	0.0215	-0.1403	-0.0144
Modified	-0.0604	0.0215	-0.1398	-0.0146

TABLE 3.7: *Cumulative effect of the error sources on the hedging results, in terms of the relative profit and loss.*

Finally, the results of Table 3.7 come from augmenting the base case using the modifications captured by Table 3.3 (refinement/extension), Table 3.5 (reduced Monte Carlo timestep), and Table 3.6 (1,000,000 simulated asset paths). The statistical measures are very close.

3.3 Summary

In this chapter, we introduced the general computational framework that will be used for the hedging experiments of this thesis. We presented the various sources of error that enter into the Monte Carlo simulations, and explored these errors in the context of a specific hedging example. This example demonstrates that the typical specifications we use are sufficient to yield accurate hedging results. The statistical measures of the relative profit and loss give an indication of the efficacy of a hedging strategy. The forthcoming tables of relative P&L statistics will use two decimal places which, based on the experiments of this chapter, is a conservative estimate of the accuracy level that can be obtained due to the sources of error.

Chapter 4

Pricing and Hedging under Regime-Switching Processes

In this chapter, we develop techniques for numerically solving the system of option-pricing PDEs from the regime-switching model. This pricing engine is used to investigate hedging under regime-switching processes. The incompleteness in the base regime-switching market (i.e. underlying asset + bond) is mild, in the sense that it can be removed by adding a finite number of instruments. Therefore, the regime-switching model provides an ideal setting to compare perfect[1] hedging—which eliminates all instantaneous risk—with the various forms of imperfect hedging, such as the delta hedge, that can be implemented in an incomplete market. We will investigate various hedging strategies, in both a two-state market without jumps in the underlying and a three-state market with jumps.

[1] In the rest of this thesis, a "perfect hedge" will denote a hedge that is asymptotically perfect as the rebalancing interval tends to zero.

4.1 Numerical Solution of the Option-Pricing PDEs

4.1.1 The System of Option-Pricing PDEs

Consider the N-state regime-switching market introduced in Section 2.2: for regime $X = i$, the evolution of the underlying asset is governed by the stochastic process

$$dS \;=\; \big(\alpha_i^{\mathbb{P}} - \tilde{\eta}_i^{\mathbb{P}}\big) S\, dt \;+\; \sigma^i S\, dZ^{\mathbb{P}} \;+\; \sum_{j=1}^{N} S\big(\eta^{ij} - 1\big)\, dX_{ij}^{\mathbb{P}}\,, \tag{4.1}$$

with $\eta^{ij} \geq 0$ a deterministic constant and $\eta^{ii} = 1$. The system of option-pricing PDEs for $(S, \tau) \in [0, \infty) \times [0, T]$ is

$$\begin{aligned} &\text{for } \; i = 1, 2, \ldots, N : \\ &\frac{\partial V^i}{\partial \tau} \;=\; \frac{1}{2}\big(\sigma^i\big)^2 S^2 \frac{\partial^2 V^i}{\partial S^2} \;+\; \big(r - \tilde{\eta}_i^{\mathbb{Q}}\big) S \frac{\partial V^i}{\partial S} \;-\; \big(r - \lambda_{ii}^{\mathbb{Q}}\big) V^i \;+\; \sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} V^j\,, \end{aligned} \tag{4.2}$$

with $\tau = T - t$ and $V^k = V^k(S\eta^{ik}, \tau)$. In addition, $\lambda_{ij}^{\mathbb{Q}} \geq 0$ is the risk-adjusted intensity for transitions from state i to j $(j \neq i)$, and[2]

$$\lambda_{ii}^{\mathbb{Q}} \;=\; -\sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}^{\mathbb{Q}}\,. \tag{4.3}$$

Furthermore,

$$\tilde{\eta}_i^{\mathbb{Q}} \;=\; \sum_{j=1}^{N} \lambda_{ij}^{\mathbb{Q}} \eta^{ij}\,. \tag{4.4}$$

[2]The $\lambda^{\mathbb{Q}}$'s will typically be given in the form of an $N \times N$ rate matrix $\Lambda^{\mathbb{Q}}$. Condition (4.3) is satisfied, by definition, in these matrices. See Appendix A for details.

The initial condition $V^i(S, \tau = 0)$ for the system of equations (4.2) is given by the option payoff. As $S \to 0$, the system of PDEs reduces to

$$\begin{aligned} &\text{for } \; i = 1, 2, \ldots, N : \\ &\frac{\partial V^i}{\partial \tau} \;=\; -\left(r - \lambda_{ii}^{\mathbb{Q}}\right) V^i \;+\; \sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} V^j \,. \end{aligned} \tag{4.5}$$

As $S \to \infty$, we make the common assumption that, for $i = 1, 2, \ldots, N$,

$$\frac{\partial^2 V^i}{\partial S^2} \simeq 0 \,, \tag{4.6}$$

which implies

$$V^i(S, \tau) \;\simeq\; A^i(\tau) S \;+\; B^i(\tau) \tag{4.7}$$

as $S \to \infty$. When (4.7) is substituted into (4.2), we obtain the systems of ordinary differential equations

$$\frac{d\vec{A}}{d\tau} \;=\; \left[\Lambda^{\mathbb{Q}} - \mathrm{diag}(\tilde{\eta}_1^{\mathbb{Q}}, \tilde{\eta}_2^{\mathbb{Q}}, \ldots, \tilde{\eta}_N^{\mathbb{Q}})\right] \vec{A}(\tau) \tag{4.8}$$

and

$$\frac{d\vec{B}}{d\tau} \;=\; \left[\Lambda^{\mathbb{Q}} - \mathrm{diag}(r, r, \ldots, r)\right] \vec{B}(\tau) \,, \tag{4.9}$$

where $\Lambda^{\mathbb{Q}}$ is the $N \times N$ risk-adjusted rate matrix.

4.1.2 Discretization of the PDEs

The computational mesh is truncated in the S-direction at some large value $S_{\max}$. The existence of relative jump sizes greater than unity leads to a numerical issue, since the system of equations (4.2) requires option values for $S > S_{\max}$. To localize the problem, we use an augmented jump amplitude near $S_{\max}$,

$$\bar{\eta}^{ij}(S) \;=\; \begin{cases} \eta^{ij} & 0 \le S \le \frac{S_{\max}}{\eta^{ij}} \\ \dfrac{S_{\max}}{S} & \frac{S_{\max}}{\eta^{ij}} < S \le S_{\max}\ , \end{cases} \tag{4.10}$$

such that $S\bar{\eta}^{ij}$ is always within $[0, S_{\max}]$. In addition, since the jump size is now a function of S, the quantity $\tilde{\eta}_i^{\mathbb{Q}}$ in (4.4) is no longer constant but is also a function of S:

$$\hat{\eta}_i^{\mathbb{Q}}(S) \;=\; \sum_{j=1}^{N} \lambda_{ij}^{\mathbb{Q}}\, \bar{\eta}^{ij}(S)\,. \tag{4.11}$$

The system (4.2), localized to $[0, S_{\max}] \times [0, T]$ as described above, is discretized using standard finite differences for the partial derivatives and θ-method timestepping. The discrete option values are $V_{m,n}^i$, with

$$\begin{aligned} i &\quad \text{regime number } (1, \ldots, N) \\ m &\quad \text{asset node index } (1, \ldots, \mathcal{M}) \\ n &\quad \text{timestep index } (1, \ldots, \mathcal{N})\,. \end{aligned}$$

The discrete equations for regime i are

$$\begin{aligned}
&\Big[1+\theta\Delta\tau\big(\alpha^i_m+\beta^i_m+r-\lambda^{\mathbb{Q}}_{ii}\big)\Big]V^i_{m,n+1} \;-\; \theta\Delta\tau\alpha^i_m V^i_{m-1,n+1} \;-\; \theta\Delta\tau\beta^i_m V^i_{m+1,n+1} \\
&\quad= \;\Big[1-\big(1-\theta\big)\Delta\tau\big(\alpha^i_m+\beta^i_m+r-\lambda^{\mathbb{Q}}_{ii}\big)\Big]V^i_{m,n} \\
&\qquad\qquad +\;\big(1-\theta\big)\Delta\tau\alpha^i_m V^i_{m-1,n} \;+\; \big(1-\theta\big)\Delta\tau\beta^i_m V^i_{m+1,n} \\
&\qquad\qquad +\;\big(1-\theta\big)\Delta\tau\sum_{\substack{j=1\\ j\neq i}}^{N}\lambda^{\mathbb{Q}}_{ij}\chi\big(S_m\bar\eta^{ij},V^j_{m^*_j,n},V^j_{m^*_j+1,n}\big) \\
&\qquad\qquad\qquad +\;\theta\Delta\tau\sum_{\substack{j=1\\ j\neq i}}^{N}\lambda^{\mathbb{Q}}_{ij}\chi\big(S_m\bar\eta^{ij},V^j_{m^*_j,n+1},V^j_{m^*_j+1,n+1}\big)\,,
\end{aligned} \tag{4.12}$$

where $V^j = V^j(S_m\bar\eta^{ij},\tau_n)$ has been approximated by linearly interpolating between $V^j_{m^*_j,n}$ and $V^j_{m^*_j+1,n}$, represented by

$$\chi\big(S_m\bar\eta^{ij},V^j_{m^*_j,n},V^j_{m^*_j+1,n}\big) \tag{4.13}$$

with

$$S_{m^*_j} \;\le\; S_m\bar\eta^{ij} \;\le\; S_{m^*_j+1}\,. \tag{4.14}$$

The spatial derivatives are discretized so that α^i_m and β^i_m are always positive, as this will preclude the appearance of spurious oscillations in certain situations (e.g. with fully implicit timestepping). Therefore either central, forward or backward differencing is used for $\frac{\partial V^i}{\partial S}$, depending on the resulting sign of α^i_m and β^i_m: central differencing is preferred, as

it is more accurate [33]. Therefore, α_m^i and β_m^i for $m = 2, \ldots, \mathcal{M}-1$ are chosen from

$$\begin{aligned}
\alpha^i_{m,central} &= \frac{(\sigma^i)^2 S_m^2}{(S_m - S_{m-1})(S_{m+1} - S_{m-1})} - \frac{\big(r - \hat{\eta}_i^{\mathbb{Q}}(S_m)\big) S_m}{S_{m+1} - S_{m-1}} \\
\beta^i_{m,central} &= \frac{(\sigma^i)^2 S_m^2}{(S_{m+1} - S_m)(S_{m+1} - S_{m-1})} + \frac{\big(r - \hat{\eta}_i^{\mathbb{Q}}(S_m)\big) S_m}{S_{m+1} - S_{m-1}} \\
\\
\alpha^i_{m,forward} &= \frac{(\sigma^i)^2 S_m^2}{(S_m - S_{m-1})(S_{m+1} - S_{m-1})} \\
\beta^i_{m,forward} &= \frac{(\sigma^i)^2 S_m^2}{(S_{m+1} - S_m)(S_{m+1} - S_{m-1})} + \frac{\big(r - \hat{\eta}_i^{\mathbb{Q}}(S_m)\big) S_m}{S_{m+1} - S_m} \\
\\
\alpha^i_{m,backward} &= \frac{(\sigma^i)^2 S_m^2}{(S_m - S_{m-1})(S_{m+1} - S_{m-1})} - \frac{\big(r - \hat{\eta}_i^{\mathbb{Q}}(S_m)\big) S_m}{S_m - S_{m-1}} \\
\beta^i_{m,backward} &= \frac{(\sigma^i)^2 S_m^2}{(S_{m+1} - S_m)(S_{m+1} - S_{m-1})} \,.
\end{aligned} \tag{4.15}$$

For $S = S_1 = 0$, the discrete equations (4.12) handle the reduced system (4.5), with $\alpha_1^i = \beta_1^i = 0$. For $S = S_{\mathcal{M}} = S_{\max}$, the assumption in (4.6) for $S \to \infty$ leads to

$$V^i_{\mathcal{M},n} \;=\; A^i(\tau_n) S_{\max} + B^i(\tau_n) \,. \tag{4.16}$$

The ODE systems (4.8) and (4.9) can be solved to find $\vec{A}(\tau_n)$ and $\vec{B}(\tau_n)$, respectively, with the initial conditions determined from the option payoff. Therefore, $V^i_{\mathcal{M},n}$ from (4.16) is imposed as a Dirichlet condition at $S = S_{\max}$ for each timestep. Note that, as an alternative to imposing a Dirichlet condition at $S_{\max}$, assumption (4.6) also leads to a discretization of (4.2) at $S = S_{\mathcal{M}} = S_{\max}$ using backward differencing, with

$$\alpha^i_{\mathcal{M}} = -\,\frac{\big(r - \hat{\eta}^{\mathbb{Q}}_{\mathcal{M}}(S_{\mathcal{M}})\big) S_{\mathcal{M}}}{S_{\mathcal{M}} - S_{\mathcal{M}-1}} \qquad \text{and} \qquad \beta^i_{\mathcal{M}} = 0 \,. \tag{4.17}$$

However, since $\hat{\eta}^{\mathbb{Q}}(S_{\mathcal{M}}) \leq 0$, the scheme can no longer be positive coefficient, although this doesn't tend to cause problems in practice [76].

Both the penalty method for pricing American options [36] and the computational procedure for pricing under a jump process with a continuum of possible amplitudes (within a regime) [34] are easily incorporated into the numerical solution.

4.1.3 Stability and Convergence

Recall that α_m^i and β_m^i are selected so they are non-negative, while $\lambda_{ij}^{\mathbb{Q}} \geq 0$ $(i \neq j)$, $-\lambda_{ii}^{\mathbb{Q}} \geq 0$, and $r \geq 0$. Furthermore, linear interpolation guarantees that

$$\min\left(V^j_{m_j^*,n+1}, V^j_{m_j^*+1,n+1}\right) \;\leq\; \chi\big(S_m\bar{\eta}^{ij}, V^j_{m_j^*,n+1}, V^j_{m_j^*+1,n+1}\big) \;\leq\; \max\left(V^j_{m_j^*,n+1}, V^j_{m_j^*+1,n+1}\right).$$

When a Dirichlet boundary condition is imposed at $S = S_{\max}$, a straightforward analysis similar to that in [34] shows that, for fully implicit timestepping ($\theta = 1$), the discretization (4.12) is unconditionally l_∞ stable. In addition, we expect first-order convergence.

We now consider a fixed point iteration for solving the equations in (4.12) for $0 \leq \theta \leq 1$. Let

$$V_n^i \;=\; \begin{bmatrix} V^i_{1,n} \\ V^i_{2,n} \\ \vdots \\ V^i_{\mathcal{M},n} \end{bmatrix}$$

be the $\mathcal{M}$-vector of option values in regime i at time τ_n. The discrete equations (4.12) for regime i may be written in matrix form as

$$\left[I - \theta M^i\right] V^i_{n+1} \;=\; \left[I + (1-\theta)M^i\right] V^i_n \;+\; (1-\theta)\Delta\tau\Xi(V^i_n) \;+\; \theta\Delta\tau\Xi(V^i_{n+1}) \tag{4.18}$$

where

$$\left[M^i V_n^i\right]_{row\ m} \;=\; \Delta\tau\left[\alpha_m^i V_{m-1,n}^i \;-\; \left(\alpha_m^i + \beta_m^i + r - \lambda_{ii}^{\mathbb{Q}}\right) V_{m,n}^i \;+\; \beta_m^i V_{m+1,n}^i\right]$$

and

$$\Xi\left(V_n^i\right) \;=\; \begin{bmatrix} \tilde{V}_n^1 & \tilde{V}_n^2 & \cdots & \tilde{V}_n^i & \cdots & \tilde{V}_n^N \\ \downarrow & \downarrow & & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_{i1}^{\mathbb{Q}} \\ \lambda_{i2}^{\mathbb{Q}} \\ \vdots \\ 0 \;\leftarrow\; i^{\text{th}}\ \text{Element} \\ \vdots \\ \lambda_{iN}^{\mathbb{Q}} \end{bmatrix}$$

with the m^{th} element of the $\mathcal{M}$-vector $\tilde{V}_n^j$ for $j = 1, \ldots, N$ given by

$$\chi\left(S_m \bar{\eta}^{ij}, V_{m_j^*,n}^j, V_{m_j^*+1,n}^j\right),$$

defined in (4.13) and (4.14). When a Dirichlet condition is imposed at $S = S_{\max}$, the last row of (4.18) is modified to

$$\begin{bmatrix} \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} V_{n+1}^i \;=\; \vec{u}_{RHS}\,,$$

where the last element in $\vec{u}_{RHS}$ is the imposed value.

At each timestep, the option values for every regime can be found by employing the fixed point iteration of Algorithm 3. The fixed point iteration will converge under the same restrictions that guarantee the l_∞ stability of the fully implicit scheme.

Theorem 4.1 (Convergence of the Fixed Point Iteration). *Assume the following*

Algorithm 3 Solve for $i = 1, \ldots, N$

$$\left[I - \theta M^i\right] V^i_{n+1} \;=\; \left[I + (1-\theta)M^i\right] V^i_n \;+\; (1-\theta)\Delta\tau\Xi(V^i_n) \;+\; \theta\Delta\tau\Xi(V^i_{n+1})$$

for $i = 1$ to N **do**
  $V^{i,0}_{n+1} = V^i_n$ {Initial guess are the option values from the previous timestep, except for any Dirichlet nodes}
end for
$k = 0$
while $ERROR > TOL$ **do**
  for $i = 1$ to N **do**
    SOLVE
    $\left[I - \theta M^i\right] V^{i,k+1}_{n+1} = \left[I + (1-\theta)M^i\right] V^i_n \;+\; (1-\theta)\Delta\tau\Xi(V^i_n) \;+\; \theta\Delta\tau\Xi(V^{i,k}_{n+1})$
  end for
  $$ERROR = \max_i \left\{ \max_m \left(\frac{|V^{i,k+1}_{m,n+1} - V^{i,k}_{m,n+1}|}{\max\left(1, |V^{i,k}_{m,n+1}|, |V^i_{m,n}|\right)} \right) \right\}$$
  $k = k + 1$
end while
TERMINAL ITERATE GIVES V^i_{n+1} FOR $i = 1, \ldots, N$.

hold:

- $\alpha^i_m \geq 0, \quad \beta^i_m \geq 0$;
- $\lambda^{\mathbb{Q}}_{ij} \geq 0 \quad (i \neq j)$;
- $\lambda^{\mathbb{Q}}_{ii} = -\sum_{\substack{j=1 \\ j\neq i}}^{N} \lambda^{\mathbb{Q}}_{ij}$ $\quad$ (*i.e. Condition* (4.3));
- $r \geq 0$;
- *A Dirichlet boundary condition is imposed at* $S = S_{\mathcal{M}} = S_{\max}$.

Then, the fixed point iteration of Algorithm 3 is globally convergent. In addition, with $\lambda^{\mathbb{Q}}_{\max} \;=\; \max_i |\lambda^{\mathbb{Q}}_{ii}|$,

$$||e^{k+1}||_\infty \leq \left(\theta\lambda^{\mathbb{Q}}_{\max}\Delta\tau + \mathcal{O}\big((\lambda^{\mathbb{Q}}_{\max}\Delta\tau)^2\big)\right) ||e^k||_\infty \,. \tag{4.19}$$

Proof of Theorem 4.1. For the k^{th} iteration, define the error as

$$e_m^{i,k} = \underbrace{V_{m,n+1}^i}_{\substack{\text{Exact solution to} \\ \text{discrete equations}}} - \underbrace{V_{m,n+1}^{i,k}}_{k^{th} \text{ guess}}, \tag{4.20}$$

where $e_{\mathcal{M}}^{i,k} = 0$ due to the Dirichlet condition at $S = S_{\max}$. Substituting (4.20) into Algorithm 3 yields

$$\begin{aligned}\Big[1 + \theta\Delta\tau\big(\alpha_m^i + \beta_m^i + r - \lambda_{ii}^{\mathbb{Q}}\big)\Big] e_m^{i,k+1} &= \theta\Delta\tau\alpha_m^i e_{m-1}^{i,k+1} + \theta\Delta\tau\beta_m^i e_{m+1}^{i,k+1} \\ &\quad + \theta\Delta\tau\sum_{\substack{j=1 \\ j\neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} \chi\big(S_m\bar{\eta}^{ij}, e_{m_j^*}^{j,k}, e_{m_j^*+1}^{j,k}\big)\end{aligned} \tag{4.21}$$

for $m < \mathcal{M}$. Now, let $||e^{i,k}||_\infty = \max_m |e_m^{i,k}|$. Since (4.21) uses linear interpolation, and all coefficients of the right-hand side are non-negative, we get for $m < \mathcal{M}$

$$\begin{aligned}\Big[1 + \theta\Delta\tau\big(\alpha_m^i + \beta_m^i + r + |\lambda_{ii}^{\mathbb{Q}}|\big)\Big] |e_m^{i,k+1}| &\leq \theta\Delta\tau\alpha_m^i ||e^{i,k+1}||_\infty \\ &\quad + \theta\Delta\tau\beta_m^i ||e^{i,k+1}||_\infty \\ &\quad + \theta\Delta\tau\sum_{\substack{j=1 \\ j\neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} ||e^{j,k}||_\infty ,\end{aligned}$$

where we have also used the fact that $-\lambda_{ii}^{\mathbb{Q}} = |\lambda_{ii}^{\mathbb{Q}}|$. In a similar fashion, defining $||e^k||_\infty =$

$\max_i ||e^{i,k}||_\infty = \max_i \left\{ \max_m |e_m^{i,k}| \right\}$ gives

$$\begin{aligned}\Big[1 + \theta\Delta\tau\big(\alpha_m^i + \beta_m^i + r + |\lambda_{ii}^{\mathbb{Q}}|\big)\Big]|e_m^{i,k+1}| \;&\leq\; \theta\Delta\tau\alpha_m^i||e^{k+1}||_\infty \\ &\quad+\; \theta\Delta\tau\beta_m^i||e^{k+1}||_\infty \\ &\quad+\; \theta\Delta\tau|\lambda_{ii}^{\mathbb{Q}}| \cdot ||e^k||_\infty\,,\end{aligned}$$

where condition (4.3) has been used. The above must be true for all regimes and asset nodes, so in particular it must be satisfied for the regime i^* and asset node m^* such that $|e_{m^*}^{i^*,k+1}| = ||e^{k+1}||_\infty$. Therefore

$$\begin{aligned}\Big[1 + \theta\Delta\tau\big(\alpha_{m^*}^{i^*} + \beta_{m^*}^{i^*} + r + |\lambda_{i^*i^*}^{\mathbb{Q}}|\big)\Big]||e^{k+1}||_\infty \;&\leq\; \theta\Delta\tau\alpha_{m^*}^{i^*}||e^{k+1}||_\infty \\ &\quad+\; \theta\Delta\tau\beta_{m^*}^{i^*}||e^{k+1}||_\infty \\ &\quad+\; \theta\Delta\tau|\lambda_{i^*i^*}^{\mathbb{Q}}| \cdot ||e^k||_\infty\,.\end{aligned} \tag{4.22}$$

The expression in (4.22) implies

$$\begin{aligned}||e^{k+1}||_\infty \;&\leq\; \frac{\theta|\lambda_{i^*i^*}^{\mathbb{Q}}|\Delta\tau}{1 + \theta(r + |\lambda_{i^*i^*}^{\mathbb{Q}}|)\Delta\tau}||e^k||_\infty \\ &\leq\; \frac{\theta\lambda_{\max}^{\mathbb{Q}}\Delta\tau}{1 + \theta(r + \lambda_{\max}^{\mathbb{Q}})\Delta\tau}||e^k||_\infty \\ &<\; ||e^k||_\infty\,,\end{aligned}$$

where we have defined $\lambda_{\max}^{\mathbb{Q}} = \max_i |\lambda_{ii}^{\mathbb{Q}}|$. Finally, the expression in (4.19) comes from a Taylor expansion of the above.

□

It is possible to use the most up-to-date option values for regimes $j < i$ within the inner

loop of Algorithm 3 to compute $\Xi(V_{n+1}^{i,k})$—the fixed point iteration is again convergent, with the proof following along the same lines as above.

Remark 4.1 (Solution of the Discrete Equations). *It is possible, of course, to assemble the linear system* (4.18) *for each regime into a large sparse matrix and carry out a direct solve—using modern ordering methods [68], this should not be very expensive. However, since in practice* $\lambda_{\max}^{\mathbb{Q}}\Delta\tau$ *is small, only 2–3 iterations of Algorithm 3 are required to reduce the error by* 10^{-6}*. Direct methods are therefore unlikely to produce a significant improvement. In any event, some type of iteration is required to price American options.*

4.1.4 A Numerical Example

As a specific example, we price an option in a three-state regime-switching model where the underlying jumps during a change in regime. The market is from [4], and was obtained by calibrating to vanilla options written on the S&P 500. The interest rate[3] for this market is $r = 0.02$, with diffusion volatilities $\vec{\sigma} = [0.0955, 0.0644, 0.0241]$. The rate matrix of risk-adjusted intensities is

$$\Lambda^{\mathbb{Q}} = \begin{bmatrix} -3.5613 & 0.2405 & 3.3208 \\ 1.1279 & -1.2008 & 0.0729 \\ 2.9882 & 0.2025 & -3.1907 \end{bmatrix} \tag{4.23}$$

with jump amplitudes

$$\eta = \begin{bmatrix} 1 & 0.9095 & 1.0279 \\ 1.2502 & 1 & 1.6512 \\ 0.9693 & 0.7732 & 1 \end{bmatrix}, \tag{4.24}$$

[3]We use a constant interest rate for the regime-switching markets considered in this thesis. However, a regime-dependent interest rate is easily handled.

where the entry in row i and column j corresponds to the transition from state i to j.

Consider a European call option with $T = 0.5$ and a strike of \$100. The system of PDEs (4.2) is solved numerically using Algorithm 3 with fully implicit timestepping ($\theta = 1$), $\Delta\tau$ constant, and the Dirichlet condition (4.16) imposed at $S = S_{\max}$. The asset node spacing is relatively finer around the strike, and the mesh extends from $S_1 = 0$ to $S_{\max} = 1500$. The option values and deltas at inception for each regime are plotted in Figure 4.1, and the results of a convergence study are given in Table 4.1. For each refinement in the convergence study, new nodes are inserted between each pair of coarse mesh nodes and the timestep size $\Delta\tau$ is halved. The results indicate the asymptotic rate of convergence is linear, as expected.

When Crank-Nicolson timestepping ($\theta = \frac{1}{2}$) is used with Rannacher smoothing[4], quadratic convergence is observed. Although the discrete equations (4.12) with $\theta = \frac{1}{2}$ are no longer guaranteed to represent a positive-coefficient scheme, the use of Rannacher smoothing tends to damp out the spurious oscillations that can arise due to non-smoothness [64].

To determine if $S_{\max} = 1500$ is large enough to avoid localization error w.r.t. the option values at inception, we consider the region of interest as $S \in [0, 500]$. When $S_{\max}$ is increased from $S_{\max} = 1500$ to $S_{\max} = 5000$, the absolute change in option values is less than 10^{-16} for all asset nodes in $[0, 500]$. Therefore, $S_{\max} = 1500$ is clearly sufficient. In general, the upper limit of the computational mesh should be chosen to ensure the augmented jump amplitude (4.10) has no discernible influence on the option values in the region of interest.

[4]Rannacher smoothing refers to the use of a small number of implicit timesteps before employing Crank-Nicolson [67].

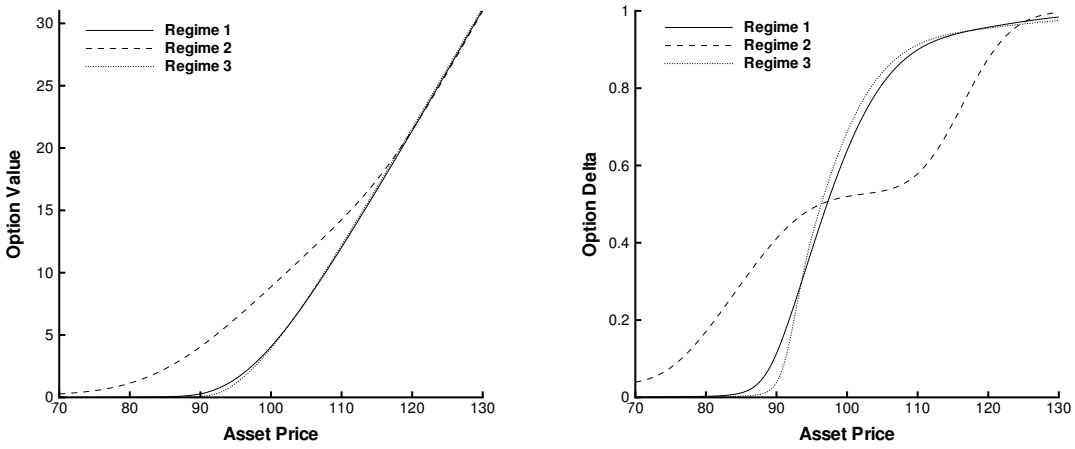


FIGURE 4.1: *Option value (left panel) and delta (right panel) for a half-year European call option in the three-state regime-switching market from [4]. The strike price is* \$100.

Number of Nodes	Number of Timesteps	Regime 1		Regime 2		Regime 3	
		Value	Ratio	Value	Ratio	Value	Ratio
213	100	4.115331	—	8.861139	—	4.020679	—
425	200	4.089879	—	8.860966	—	3.975688	—
849	400	4.063196	0.95	8.858187	0.06	3.929001	0.96
1697	800	4.063566	-72.17	8.861186	-0.93	3.929624	-74.88
3393	1600	4.063792	1.64	8.862876	1.77	3.929936	2.00
6785	3200	4.063908	1.94	8.863749	1.94	3.930088	2.05
13569	6400	4.063966	1.98	8.864193	1.97	3.930163	2.03
27137	12800	4.063996	1.99	8.864417	1.98	3.930200	2.02
54273	25600	4.064010	2.00	8.864529	1.99	3.930219	2.01

TABLE 4.1: *Value of a half-year European call option, in each regime of the three-state market from [4], for successive refinements of the computational mesh. Ratio is the ratio of changes between one level of refinement and the next. Fully implicit timestepping is used. The option is priced at* $S = 100$, $t = 0$ *and has a strike price of* \$100.

4.2 Hedging in a Market with Shifts in Volatility

4.2.1 The Risk in a Regime-Switching Market

Consider the N-state regime-switching market (4.1). To hedge a short position in a target option V, the bond B and e units of the underlying are used in conjunction with an additional M instruments, represented by $\vec{F}$, to form the overall hedged position

$$\Pi \;=\; -V \;+\; \underbrace{eS \;+\; \vec{\phi}\cdot\vec{F} \;+\; B}_{\text{Hedge Portfolio } \mathcal{H}},$$

where the weights of the instruments in $\vec{F}$ are given by $\vec{\phi}$. Assume that regime i is the current state. By applying Itô's lemma, the change in Π over an instant dt is

$$\begin{aligned}
d\Pi \;&=\; -dV \;+\; edS \;+\; \vec{\phi}\cdot d\vec{F} \;+\; dB \\
&=\; \left[-\left(\frac{\partial V^i}{\partial t} + \left(\alpha_i^{\mathbb{P}} - \tilde{\eta}_i^{\mathbb{P}}\right)S\frac{\partial V^i}{\partial S} + \frac{1}{2}\left(\sigma^i\right)^2 S^2 \frac{\partial^2 V^i}{\partial S^2}\right) \;+\; \left(\alpha_i^{\mathbb{P}} - \tilde{\eta}_i^{\mathbb{P}}\right)eS \right. \\
&\qquad\qquad \left. +\; \vec{\phi}\cdot\left(\frac{\partial \vec{F}^i}{\partial t} + \left(\alpha_i^{\mathbb{P}} - \tilde{\eta}_i^{\mathbb{P}}\right)S\frac{\partial \vec{F}^i}{\partial S} + \frac{1}{2}\left(\sigma^i\right)^2 S^2 \frac{\partial^2 \vec{F}^i}{\partial S^2}\right) \;+\; rB\right] dt \\
&\quad +\; \sigma^i S\left[-\frac{\partial V^i}{\partial S} \;+\; e \;+\; \vec{\phi}\cdot\frac{\partial \vec{F}^i}{\partial S}\right] dZ^{\mathbb{P}} \\
&\quad +\; \sum_{\substack{j=1 \\ j\neq i}}^{N}\left[-\Delta V^{ij} \;+\; e\Delta S^{ij} \;+\; \vec{\phi}\cdot\Delta\vec{F}^{ij}\right] dX_{ij}^{\mathbb{P}}\,, \qquad (4.25)
\end{aligned}$$

where the Δ notation represents the change in value due to the regime change $i \to j$: $\Delta S^{ij} = S(\eta^{ij} - 1)$, $\Delta V^{ij} = V^j(S\eta^{ij}, t) - V^i(S,t)$, and $\Delta\vec{F}^{ij} = \vec{F}^j(S\eta^{ij}, t) - \vec{F}^i(S,t)$.

The last two lines of (4.25) represent the random change in Π due to the Wiener process and Markov chain, respectively. If $\vec{F}$ contains $N-1$ instruments, then choosing hedge

weights that satisfy the $N \times N$ linear system

$$\underbrace{\begin{bmatrix} 1 & \frac{\partial F_1^i}{\partial S} & \frac{\partial F_2^i}{\partial S} & \cdots & \frac{\partial F_{N-1}^i}{\partial S} \\ \Delta S^{i1} & \Delta F_1^{i1} & \Delta F_2^{i1} & \cdots & \Delta F_{N-1}^{i1} \\ \Delta S^{i2} & \Delta F_1^{i2} & \Delta F_2^{i2} & \cdots & \Delta F_{N-1}^{i2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta S^{i(i-1)} & \Delta F_1^{i(i-1)} & \Delta F_2^{i(i-1)} & \cdots & \Delta F_{N-1}^{i(i-1)} \\ \Delta S^{i(i+1)} & \Delta F_1^{i(i+1)} & \Delta F_2^{i(i+1)} & \cdots & \Delta F_{N-1}^{i(i+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta S^{iN} & \Delta F_1^{iN} & \Delta F_2^{iN} & \cdots & \Delta F_{N-1}^{iN} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} e \\ \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_{N-3} \\ \phi_{N-2} \\ \phi_{N-1} \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} \frac{\partial V^i}{\partial S} \\ \Delta V^{i1} \\ \Delta V^{i2} \\ \vdots \\ \Delta V^{i(i-1)} \\ \Delta V^{i(i+1)} \\ \vdots \\ \Delta V^{iN} \end{bmatrix}}_{b} \tag{4.26}$$

will remove the instantaneous risk for state i, and therefore establish a perfect hedge.[5] If A has rank N, then a perfect hedge can be formed using $\vec{F}$. If $\text{rank}(A) < N$, then either: (i) no perfect hedge can be formed ($b \notin \text{columnspace}(A)$); or (ii) a perfect hedge can be formed ($b \in \text{columnspace}(A)$) and, furthermore, there are an infinite number of hedge portfolios that constitute a perfect hedge.

For the purposes of hedging, we define a redundant instrument as follows:

Definition 4.1 (Redundant Instrument). *An instrument F_{m^*} in $\vec{F}$ is said to be redundant if it can be statically replicated using the bond, underlying, and the other instruments in $\vec{F}$. That is, at time zero there exists a bond position B_0 and a set of coefficients $\vec{\alpha}$ such that, for all times t and asset values S,*

$$F_{m^*}(S,t) \;=\; B_0 e^{rt} \;+\; \alpha_0 S \;+\; \sum_{\substack{m=1 \\ m \neq m^*}}^{M} \alpha_m F_m(S,t)\,,$$

[5] The matrix in (4.26) has no row corresponding to the trivial transition $i \to i$.

where $\vec{F}$ has M instruments.

For example, put-call parity implies that a European put is redundant w.r.t. the bond, underlying, and a European call with the same strike.

When $\vec{F}$ has $N-1$ instruments, a redundant instrument will preclude the formation of a perfect hedge in an N-state market.

Theorem 4.2 (Absence of a Perfect Hedge). *Consider an N-state market and assume that $\vec{F}$ has $N-1$ instruments. If $\vec{F}$ contains a redundant instrument, then the $N \times N$ matrix in* (4.26) *is singular. As such, a perfect hedge cannot be formed using $\vec{F}$ and the underlying, unless $b \in$ columnspace(A).*

Proof of Theorem 4.2. Without loss of generality, assume that the last instrument in $\vec{F}$ (i.e. instrument $N-1$) is redundant. Consider the following set of column operations on the $N \times N$ matrix in (4.26), where COL_k denotes the k^{th} column of the matrix:

$$COL_N^{augment} \;=\; COL_N \;-\; \alpha_0\, COL_1 \;-\; \sum_{k=1}^{N-2} \alpha_k\, COL_{k+1}\,.$$

For the redundant instrument, we have from Definition 4.1 that

$$\frac{\partial F^i_{N-1}}{\partial S} \;=\; \alpha_0 \;+\; \sum_{k=1}^{N-2} \alpha_k \frac{\partial F^i_k}{\partial S}\;,$$

so entry $(1, N)$ of the column-augmented matrix is zero. The remaining entries in the last column of the augmented matrix are equal to $B_0 \mathrm{e}^{rt} - B_0 \mathrm{e}^{rt} = 0$, which again is a consequence of Definition 4.1. This column of zeros in the augmented matrix implies that the $N \times N$ matrix in (4.26) is singular. Hence, the system in (4.26) will not have a solution unless $b \in$ columnspace(A). □

Remark 4.2 (Knowledge of the Current Regime). *We will assume throughout that the current regime is known, e.g. from calibration. If the current regime is not known,*

then delta neutrality may be imposed for each possible regime, and the hedger can insulate themselves against any possible regime change. However, this will require $\mathcal{O}(N^2)$ hedging instruments.

Imposing delta neutrality

$$-\frac{\partial V^i}{\partial S} \;+\; e \;+\; \vec{\phi}\cdot\frac{\partial \vec{F}^i}{\partial S} \;=\; 0 \tag{4.27}$$

in (4.25) eliminates the diffusion risk, and leads to an explicit representation of the regime-switching risk. To this end, substituting the system of forward-time option-pricing PDEs into (4.25), with delta neutrality imposed, gives[6]

$$\begin{aligned} d\Pi \;&=\; r\Pi\,dt \\ &+\underbrace{\sum_{\substack{j=1\\ j\neq i}}^{N}\lambda_{ij}^{\mathbb{Q}}\Big[\Delta V^{ij}-\Big(e\Delta S^{ij}+\vec{\phi}\cdot\Delta\vec{F}^{ij}\Big)\Big]\,dt \;+\; \sum_{\substack{j=1\\ j\neq i}}^{N}\Big[-\Delta V^{ij}+\Big(e\Delta S^{ij}+\vec{\phi}\cdot\Delta\vec{F}^{ij}\Big)\Big]\,dX_{ij}^{\mathbb{P}}}_{\text{instantaneous regime-switching risk}}. \end{aligned} \tag{4.28}$$

The first part of the regime-switching risk is a deterministic drift while the second component is stochastic, as it depends on whether the Markov chain transitions out of state i over dt.

4.2.2 Previous Work Related to Hedging under Regime-Switching

Naik [58] explored the initial composition of a perfect hedge under a two-state regime-switching lognormal process; recall, this process is characterized by a low-volatility state

[6] In a forthcoming chapter, a detailed derivation of the *jump risk* for a jump-diffusion model will be presented. Since the regime-switching risk is a discrete analogue of the jump risk, we defer details until Chapter 6.

and a high-volatility state (with no jumps in asset price):

$$dS \;=\; \begin{cases} \alpha_L^{\mathbb{P}} S\,dt \;+\; \sigma^L S\,dZ^{\mathbb{P}} & \text{if the Markov chain is in its low-volatility state} \\ \alpha_H^{\mathbb{P}} S\,dt \;+\; \sigma^H S\,dZ^{\mathbb{P}} & \text{if the Markov chain is in its high-volatility state .} \end{cases} \tag{4.29}$$

However, Naik [58] did not carry out any simulation experiments to determine how well the hedging strategy performs when the portfolio is discretely rebalanced.

A local risk-minimizing trading strategy,[7] using just the underlying and bond, is considered in [55] for an N-state regime-switching model with no jumps in the underlying. With this strategy, the optimal weight in the underlying is the delta of the target option, which may seem surprising since this is nothing but a Black–Scholes-type delta hedge that simply eliminates the diffusion risk. However, because the underlying does not change value when the regime changes, it cannot be used to mitigate the regime-switching risk. Again, no tests of the hedging procedure are performed. The dynamic hedging strategy for regime-switching models discussed in [4] utilizes control theory to ensure the hedged position will break even on average and that the standard deviation of the profit and loss is minimized. The procedure assumes knowledge of the real-world measure, which is linked to the pricing measure through the Sharpe ratio.

4.2.3 Hedging in a Two-State Regime-Switching Lognormal Model

We are interested in hedging under the two-state process (4.29) considered by Naik [58]. Let V be the target option and $eS + \phi F$ the hedge portfolio, consisting of e units in the underlying and ϕ units of the hedging option F. To establish a perfect hedge at time t and

[7] Such a strategy aims to minimize the mean-squared hedging error over the next instant [23].

asset price S, the linear system (4.26) is trivially solved to yield the hedge weights

$$\phi = \frac{V^H(S,t) - V^L(S,t)}{F^H(S,t) - F^L(S,t)} \qquad e = \left.\frac{\partial V^L}{\partial S}\right|_{(S,t)} - \phi \times \left.\frac{\partial F^L}{\partial S}\right|_{(S,t)}$$

if the Markov chain is in its low-volatility state, and

$$\phi = \frac{V^L(S,t) - V^H(S,t)}{F^L(S,t) - F^H(S,t)} \qquad e = \left.\frac{\partial V^H}{\partial S}\right|_{(S,t)} - \phi \times \left.\frac{\partial F^H}{\partial S}\right|_{(S,t)}$$

if it is in the high-volatility state.

To study the behaviour of a discretely rebalanced perfect hedge, we consider the real-world market represented by the parameters of Table 4.2. A half-year call with a strike of \$100, that is initially at the money, is to be hedged over its lifetime with the underlying and a one-year call[8] that has a strike of \$110. The delta hedge

$$\phi = 0 \qquad e = \left.\frac{\partial V^i}{\partial S}\right|_{(S,t)},$$

which is the local risk-minimizing strategy in the incomplete market [55], is also carried out.

As a first example, we assume there is no risk adjustment when pricing, such that the transition intensities are the same under $\mathbb{P}$ and $\mathbb{Q}$. The initial regime for each of the 500,000 simulations is the low-volatility state. We vary the rebalancing interval Δt, and in Table 4.3 give the mean, standard deviation, and 99% value-at-risk (VaR) of the (discounted) relative

[8]It is more natural to use short-term options to hedge, as they tend to be more liquid than longer-term contracts. Furthermore, the effect of regimes plays a diminishing role as option maturity increases. We hedge with a longer-term contract to avoid complications due to rolling over hedging options. In subsequent examples, short-term instruments will be used to hedge longer-term contracts.

Parameter	Low-Volatility State	High-Volatility State
Volatility	$\sigma^L = 0.2$	$\sigma^H = 0.4$
Expected Return	$\alpha_L^{\mathbb{P}} = 0.14$	$\alpha_H^{\mathbb{P}} = 0.07$
Transition Intensity	$\lambda_{LH}^{\mathbb{P}} = 0.5$	$\lambda_{HL}^{\mathbb{P}} = 2.0$
Interest Rate	$r = 0.05$	

TABLE 4.2: *Financial parameters for a two-state regime-switching lognormal market that will be used to study various hedging strategies.*

Number of Rebalancings	Delta Hedge			Perfect Hedge		
	Mean	StdDev	99% VaR	Mean	StdDev	99% VaR
5	-0.01	0.36	-1.23	0.00	0.19	-0.59
25	0.00	0.23	-0.82	0.00	0.10	-0.31
50	0.00	0.20	-0.73	0.00	0.07	-0.23
100	0.00	0.19	-0.68	0.00	0.05	-0.16
200	0.00	0.18	-0.65	0.00	0.04	-0.12

TABLE 4.3: *Hedging results, in terms of the relative profit and loss, when hedging a half-year European call option over its lifetime in the two-state regime-switching market of Table 4.2 (with no risk adjustment). The initial regime is the low-volatility state. The target call has a strike of \$100, while the one-year European call used in the perfect hedge has a strike of \$110. The 99% VaR is defined as the 1% quantile of the relative profit and loss.*

profit and loss

$$\mathrm{e}^{-rT}\frac{\Pi(T)}{V^i(S_0,0)}\,. \tag{4.30}$$

Note that the α% VaR is defined as the $(100-\alpha)$% quantile of the relative profit and loss.

The discretely rebalanced perfect hedge behaves as expected—the mean is zero, and the standard deviation decreases as the rebalancing becomes more frequent. The delta hedge does not perform too badly, even though it ignores the regime-switching risk: since there are no jumps in the underlying, the only instrument in the overall hedged position $-V+eS+B$ that changes value during a transition is the target option, and in this case the change tends

not to be too large (e.g. $||V^H(S,0) - V^L(S,0)||_\infty = \$3.27$). However, while the standard deviation of the perfect hedge is approximately $\mathcal{O}(\sqrt{\Delta t})$, the improvement in the standard deviation of the delta hedge levels off as Δt decreases. This is because the regime-switching risk remains unhedged regardless of how often the delta hedge is rebalanced.

The results for the discretely rebalanced perfect hedge are similar regardless of the initial state. The distribution of the relative profit and loss for both initial states is plotted in the left panel of Figure 4.2 for the case of 100 rebalances, and are seen to be very alike. On the other hand, the behaviour of the delta hedge is highly dependent on the initial state. For any two-state lognormal model, the instantaneous change in the simple delta-hedged position is, by (4.28),

$$d\Pi \;=\; \begin{cases} r\Pi\,dt \;+\; \lambda^{\mathbb{Q}}_{LH}\Big(V^H - V^L\Big)dt \;-\; \Big(V^H - V^L\Big)dX^{\mathbb{P}}_{LH} & \text{in low-volatility state} \\ r\Pi\,dt \;+\; \lambda^{\mathbb{Q}}_{HL}\Big(V^L - V^H\Big)dt \;-\; \Big(V^L - V^H\Big)dX^{\mathbb{P}}_{HL} & \text{in high-volatility state}\,. \end{cases} \tag{4.31}$$

Assume the initial regime is the low-volatility state. If no regime change occurs over the life of the option, then $d\Pi$ in (4.31) is a pure-drift process with positive drift, since $V^H > V^L$ (in general). However, if the regime changes just once, there is a single large drop in the value of Π. These two possibilities are the major influence on the relative profit and loss for the specific delta hedge example considered here—the distributions are plotted in the right panel of Figure 4.2. The pronounced left tail of the density for the low-volatility initial state is due primarily to the simulations where there are regime changes, while the positivity of the central mass is a result of the positive drifts that accumulate when no regime changes occur. On the other hand, when the initial regime is the high-volatility state, there is a positive 'hump' since regime changes are more frequent when the initial state is the high-volatility regime.

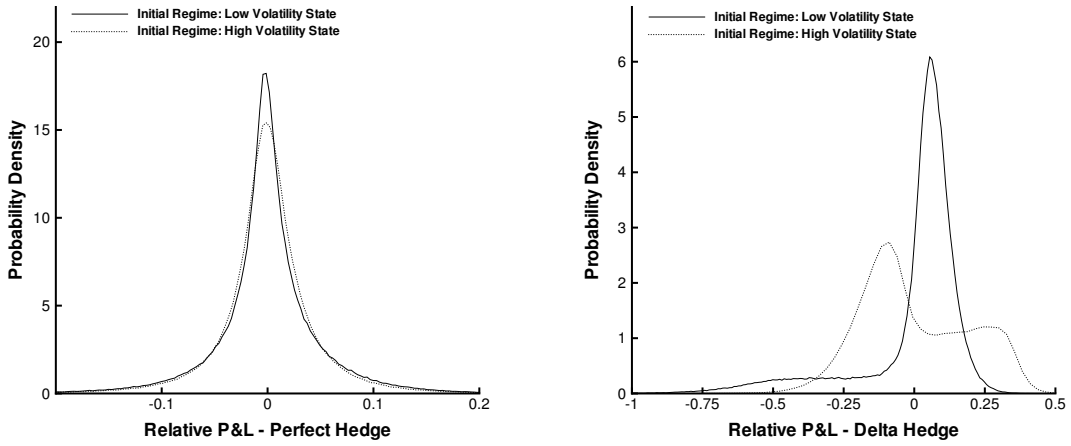


FIGURE 4.2: *Distributions of relative profit and loss when hedging a half-year European call option over its lifetime in the two-state regime-switching market of Table 4.2 (with no risk adjustment). The left panel contains the distribution for the perfect hedge, while the right panel is for the delta hedge. The hedge is rebalanced a total of 100 times. The target call has a strike of \$100, while the one-year European call used in the perfect hedge has a strike of \$110.*

By taking the $\mathbb{P}$-expected value of $d\Pi$ in (4.31), an ODE for the expected value of the simple delta-neutral position can be obtained, namely

$$\begin{aligned}\frac{d\mathbb{E}^{\mathbb{P}}\big[\Pi_t\big]}{dt} &= r\mathbb{E}^{\mathbb{P}}\big[\Pi_t\big] \\ &+ \left(\lambda^{\mathbb{Q}}_{LH}\mathbb{E}^{\mathbb{P}}\Big[V^H(S_t,t) - V^L(S_t,t)\Big] - \lambda^{\mathbb{P}}_{LH}\mathbb{E}^{\mathbb{P}}\Big[V^H(S_t,t) - V^L(S_t,t)\Big]\right) P(X_t = L) \\ &+ \left(\lambda^{\mathbb{Q}}_{HL}\mathbb{E}^{\mathbb{P}}\Big[V^L(S_t,t) - V^H(S_t,t)\Big] - \lambda^{\mathbb{P}}_{HL}\mathbb{E}^{\mathbb{P}}\Big[V^L(S_t,t) - V^H(S_t,t)\Big]\right) P(X_t = H)\end{aligned} \tag{4.32}$$

with initial condition $\mathbb{E}^{\mathbb{P}}\big[\Pi_0\big] = 0$, where $P(X_t = i)$ is the probability the Markov chain is in state i at time t. If there is no risk adjustment, then the transition intensities are the same under $\mathbb{P}$ and $\mathbb{Q}$, and the ODE (4.32) simplifies greatly, with solution $\mathbb{E}^{\mathbb{P}}\big[\Pi_t\big] = 0$: on average, the small ever-present drift compensates for the large changes that occur when the regime changes. The hedging results in Table 4.3 are for a market with no risk adjustment,

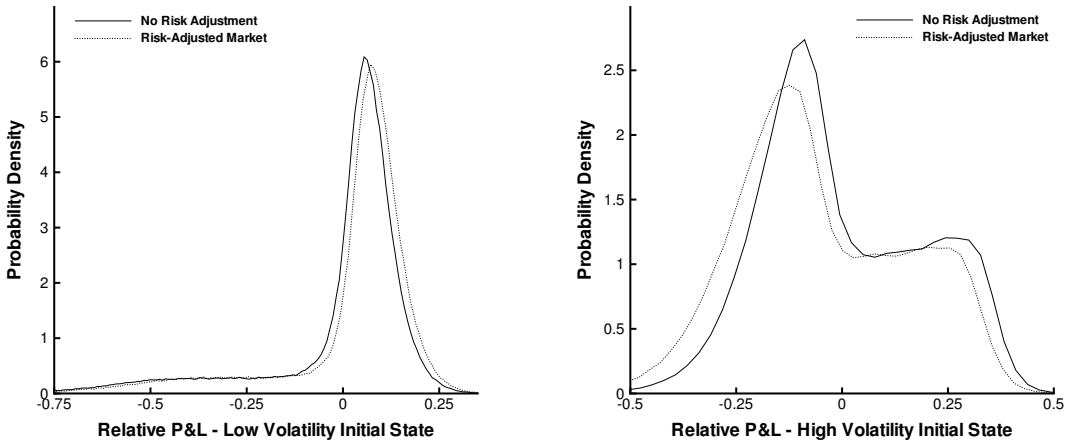


FIGURE 4.3: *Distributions of relative profit and loss when delta hedging a half-year European call option over its lifetime in a two-state regime-switching lognormal model, for both a risk-adjusted market and a market with no risk adjustment. The left panel contains the distribution for the low-volatility initial state, while the right panel is for the high-volatility initial state. The hedge is rebalanced a total of 100 times. The target call has a strike of \$100.*

so a zero mean for the profit and loss of the delta hedge is the expected outcome.

In a risk-adjusted market, the ODE (4.32) will maintain its non-trivial form, such that the mean of the delta hedge is expected to be non-zero. Consider a market where the real-world intensities are as in Table 4.2, but the risk-adjusted intensities[9] are higher, with $\lambda^{\mathbb{Q}}_{LH} = \frac{3}{4}$ and $\lambda^{\mathbb{Q}}_{HL} = 3$. The distributions of the relative profit and loss for the delta hedge, in both the original and risk-adjusted pricing environments, are presented in Figure 4.3. The mean for the delta hedge with the low-volatility initial state is now positive, while the high-volatility initial state has a negative mean. For the discretely rebalanced perfect hedge, the risk adjustment has little effect—the mean of the profit and loss remains zero with a small standard deviation.

[9]Naik [58] provides an argument for the risk adjustment of the transition intensities in a two-state lognormal model.

4.2.4 Hedging in a Three-State Model with Jumps in the Underlying

Consider a situation where there are many regimes, with jumps in the underlying during changes in regime. Since the underlying will now change value when a state transition occurs, the delta hedge is expected to perform worse than in a market with no jumps. However, if there is only one possible jump size per transition, the perfect hedge (which uses the underlying and $N-1$ additional hedging instruments) can still be constructed, and should continue to perform well when rebalancing is done discretely.

The specific three-state market we will consider comes from [4], and is the same as used in Section 4.1.4 for the pricing example. Although not available from calibration, the real-world drift rates are required to simulate the asset paths—the values $\vec{\alpha}^{\mathbb{P}} = [0.05, 0.08, 0.12]$ are used. Furthermore, we assume the risk adjustment is such that $\Lambda^{\mathbb{P}} = \frac{3}{4}\Lambda^{\mathbb{Q}}$, with $\Lambda^{\mathbb{Q}}$ the matrix of risk-adjusted intensities in (4.23).

We consider the same target option as before, and again use the longer-dated call as a hedging instrument. Since two hedging options are required to establish the perfect hedge, a 1.5-year European put with strike \$90 is also used. The hedge is rebalanced 100 times and 500,000 simulations are carried out. The results for the perfect and delta hedge, for each of the initial states, is given in Table 4.4, with the distributions of profit and loss plotted in Figure 4.4.

The perfect hedge again works well while the delta hedge remains inadequate, with tail events that are much worse than observed for the delta hedge in the two-state lognormal model. The distributions of the relative profit and loss for the delta hedge exhibit very rich behaviour, with secondary tails due to multiple regime changes over the life of the option. Due to the risk adjustment in the market, the means of the profit and loss for the delta hedge are non-zero.

Two of the distributions in Figure 4.4 call for comment. For the delta hedge with regime two the initial state, there is a pronounced double-hump structure: the positive hump

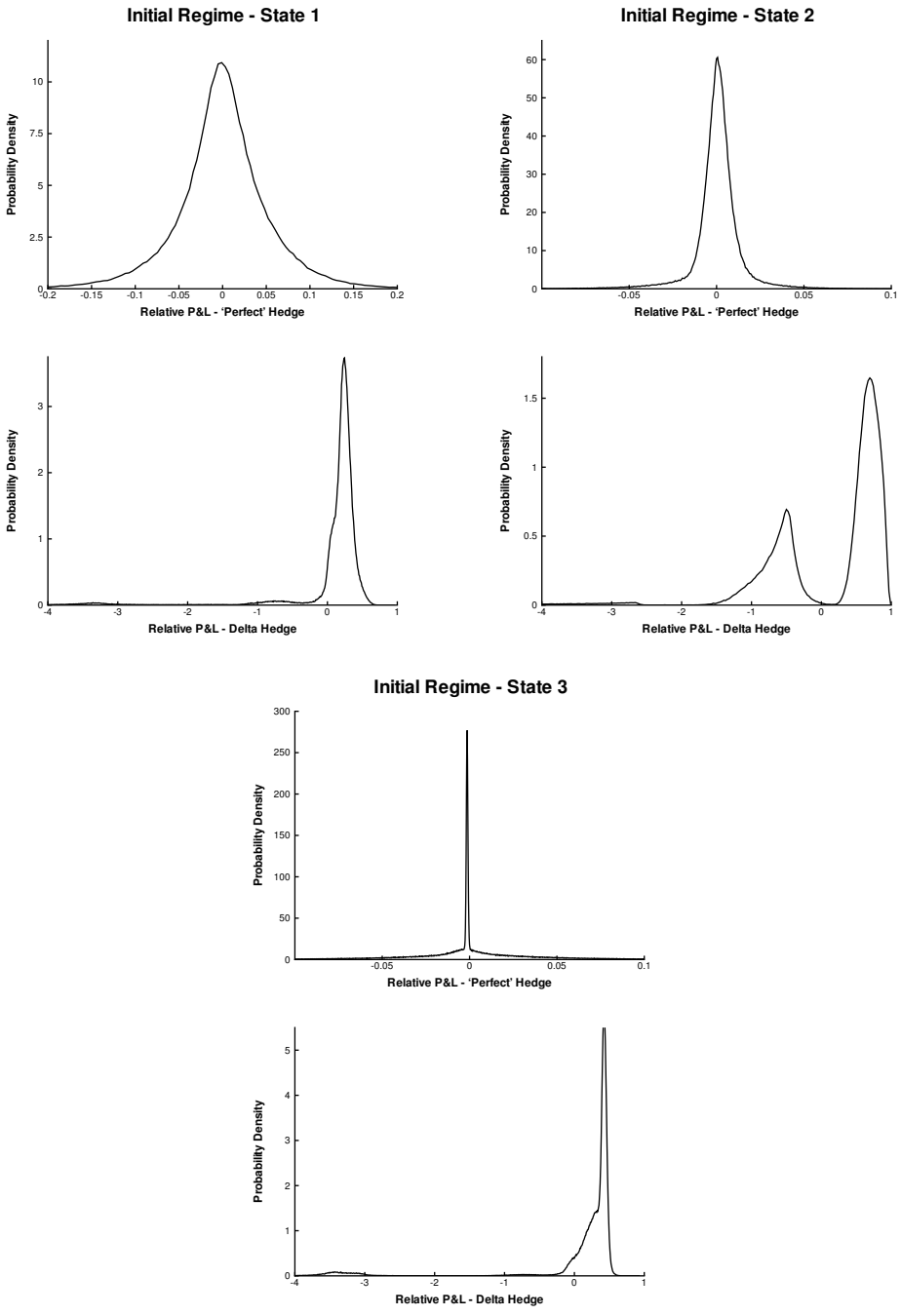


FIGURE 4.4: *Distributions of relative profit and loss when hedging a half-year European call option over its lifetime in the three-state regime-switching market from [4]. Each initial state is considered, and the hedge is rebalanced a total of 100 times. The target European call has a strike of $100, the one-year European call used in the perfect hedge has a strike of $110, and the 1.5-year European put used in the perfect hedge has a strike of $90.*

Initial State	Delta Hedge			Perfect Hedge		
	Mean	StdDev	99% VaR	Mean	StdDev	99% VaR
1	0.07	0.70	-3.49	0.00	0.06	-0.16
2	0.14	0.87	-3.53	0.00	0.02	-0.05
3	0.10	0.90	-3.62	0.00	0.05	-0.14

TABLE 4.4: *Hedging results, in terms of the relative profit and loss, when hedging a half-year European call option over its lifetime in the three-state regime-switching market from [4]. Each initial state is considered, and the hedge is rebalanced a total of 100 times. The target call has a strike of $100, the one-year European call used in the perfect hedge has a strike of $110, and the 1.5-year European put used in the perfect hedge has a strike of $90.*

corresponds to those simulations where the regime does not change, while the negative hump results from the simulations where the regime does change. Similar behaviour was observed for the delta hedge in the two-state lognormal model—see Section 4.2.3 for the analysis in that case. For the perfect hedge with regime three the initial state, the spike in the density at zero corresponds to those simulations where there is no change in regime. Since the volatility in regime three is quite small ($\sigma_3 = 0.0241$), there is little diffusion risk in this case.

4.2.5 Hedging by Local Risk Minimization

For an N-state market, the perfect hedge performs well when rebalancing is done discretely; however, it may be prohibitively expensive to maintain due to transaction costs. On the other hand, the delta hedge—which is relatively cheap to execute—is inadequate due to the fact it ignores the regime-switching risk. We now consider the intermediate situation of hedging with a total of M instruments (the underlying plus $M-1$ additional hedging instruments), where $M < N$. In this case, there are not enough hedging instruments to totally eliminate the risk, but there are degrees of freedom available to reduce it.

Imposing Delta Neutrality

Assume that regime i is the current state. The instantaneous change in the delta-neutral hedged position, given by (4.28), may be expressed as

$$d\Pi \;=\; a\,dt \;+\; \sum_{\substack{j=1\\ j\neq i}}^{N} b_{ij}\,dX_{ij}^{\mathbb{P}}$$

with

$$b_{ij} \;=\; \Big[-\Delta V^{ij} + \Big(e\Delta S^{ij} + \vec{\phi}\cdot\Delta\vec{F}^{ij}\Big)\Big].$$

The variance of $d\Pi$ is given by

$$\text{Variance}\big[d\Pi\big] \;\approx\; \sum_{\substack{j=1\\ j\neq i}}^{N} \big(\lambda_{ij}^{\mathbb{P}}\,dt\big)\,b_{ij}^2 \;=\; dt\sum_{\substack{j=1\\ j\neq i}}^{N} \lambda_{ij}^{\mathbb{P}}\Big[-\Delta V^{ij} + \Big(e\Delta S^{ij} + \vec{\phi}\cdot\Delta\vec{F}^{ij}\Big)\Big]^2, \tag{4.33}$$

where terms of $\mathcal{O}(dt^2)$ have been dropped.

If there are N (non-redundant) hedging instruments (i.e. $N-1$ options in $\vec{F}$, plus the underlying), the expression for Variance$\big[d\Pi\big]$ in (4.33) can be made zero by imposing

$$e\Delta S^{ij} \;+\; \vec{\phi}\cdot\Delta\vec{F}^{ij} \;=\; \Delta V^{ij} \qquad \begin{array}{c} j=1,2,\ldots,N \\ j\neq i \end{array}$$

in addition to the delta-neutral constraint (4.27). Of course, this is nothing but the perfect hedge found by solving the linear system (4.26). One way to choose the hedge portfolio weights when there are only $1 < M < N$ total hedging instruments (i.e. the underlying plus $M-1$ hedging options in $\vec{F}$) is by minimizing the expression for Variance$\big[d\Pi\big]$ in (4.33), while simultaneously imposing delta neutrality. However, the $\lambda_{ij}^{\mathbb{P}}$'s will generally not be

known—recall, it is the risk-adjusted intensities $\lambda_{ij}^{\mathbb{Q}}$ that are available from calibration.

The above discussion motivates

$$R^i \;=\; \sum_{\substack{j=1\\ j\neq i}}^{N} \zeta_{ij}\Big[-\Delta V^{ij} + \Big(e\Delta S^{ij} + \vec{\phi}\cdot\Delta\vec{F}^{ij}\Big)\Big]^2 \tag{4.34}$$

as a measure of the regime-switching risk for state i, where $\zeta_{ij} \geq 0$ is the weighting factor for each possible transition out of state i. With $\zeta_{ij} = \lambda_{ij}^{\mathbb{P}}$, minimizing (4.34) is equivalent to a local minimization of the variance. Again, we note that the real-world intensities $\lambda_{ij}^{\mathbb{P}}$ are generally not available, and would have to be estimated. Another choice is to set $\zeta_{ij} = 1$, which encapsulates a lack of information concerning which transitions out of state i are more likely than others.

The optimality conditions for minimizing (4.34) are

$$\begin{aligned}
\frac{\partial R^i}{\partial e} &= \sum_{\substack{j=1\\ j\neq i}}^{N} \zeta_{ij}\Big[-\Delta V^{ij} + \Big(e\Delta S^{ij} + \vec{\phi}\cdot\Delta\vec{F}^{ij}\Big)\Big]\Delta S^{ij} \;=\; 0 \\
\frac{\partial R^i}{\partial \phi_k} &= \sum_{\substack{j=1\\ j\neq i}}^{N} \zeta_{ij}\Big[-\Delta V^{ij} + \Big(e\Delta S^{ij} + \vec{\phi}\cdot\Delta\vec{F}^{ij}\Big)\Big]\Delta F_k^{ij} \;=\; 0 \qquad \text{for } k = 1,\ldots,M-1\,,
\end{aligned}$$

where the underlying and the $M-1$ instruments in $\vec{F}$ are used to hedge. Defining the

$(N-1) \times M$ matrix

$$\Phi = \begin{bmatrix} \Delta S^{i1} & \Delta F_1^{i1} & \Delta F_2^{i1} & \cdots & \Delta F_{M-1}^{i1} \\ \Delta S^{i2} & \Delta F_1^{i2} & \Delta F_2^{i2} & \cdots & \Delta F_{M-1}^{i2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta S^{i(i-1)} & \Delta F_1^{i(i-1)} & \Delta F_2^{i(i-1)} & \cdots & \Delta F_{M-1}^{i(i-1)} \\ \Delta S^{i(i+1)} & \Delta F_1^{i(i+1)} & \Delta F_2^{i(i+1)} & \cdots & \Delta F_{M-1}^{i(i+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta S^{iN} & \Delta F_1^{iN} & \Delta F_2^{iN} & \cdots & \Delta F_{M-1}^{iN} \end{bmatrix},$$

the $(N-1) \times 1$ vector

$$\vec{b} = \begin{bmatrix} \Delta V^{i1} \\ \vdots \\ \Delta V^{i(i-1)} \\ \Delta V^{i(i+1)} \\ \vdots \\ \Delta V^{iN} \end{bmatrix},$$

and the $(N-1) \times (N-1)$ matrix

$$\Omega = \begin{bmatrix} \zeta_{i1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \zeta_{i2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \zeta_{i(i-1)} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \zeta_{i(i+1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \zeta_{iN} \end{bmatrix},$$

the hedge weights $\vec{w} = [e, \phi_1, \phi_2, \ldots, \phi_{M-1}]^T$ that minimize (4.34) are found by solving

$$\Phi^T \, \Omega \, \Phi \, \vec{w} \;\; = \;\; \Phi^T \, \Omega \, \vec{b} \, . \tag{4.35}$$

Note that neither Φ, $\vec{b}$ or Ω have a row corresponding to the trivial transition $i \to i$, since this outcome does not need to enter into the risk metric (4.34). Using the method of Lagrange multipliers, the system (4.35) is easily augmented to impose delta neutrality. Specifically, the linear system

$$\left[\begin{array}{ccccc|c} & & \Phi^T \, \Omega \, \Phi & & & \begin{array}{c} 1 \\ \frac{\partial F_1^i}{\partial S} \\ \frac{\partial F_2^i}{\partial S} \\ \vdots \\ \frac{\partial F_{M-1}^i}{\partial S} \end{array} \\ \hline 1 & \frac{\partial F_1^i}{\partial S} & \frac{\partial F_2^i}{\partial S} & \cdots & \frac{\partial F_{M-1}^i}{\partial S} & 0 \end{array}\right] \left[\begin{array}{c} \vec{w} \\ \hline \Psi \end{array}\right] = \left[\begin{array}{c} \Phi^T \, \Omega \, \vec{b} \\ \hline \dfrac{\partial V^i}{\partial S} \end{array}\right] \tag{4.36}$$

is solved, where Ψ is a Lagrange multiplier.

As a specific example, we again consider hedging the half-year European call (as treated in the previous section) in the three-state market from [4]. However, only one hedging option is to be used along with the underlying: either the one-year call or the 1.5-year put. Therefore, the linear system (4.36) is solved to find the optimal hedge composition, using both $\zeta_{ij} = \lambda_{ij}^{\mathbb{P}}$ ($\mathbb{P}$ weighting) and $\zeta_{ij} = 1$ (uniform weighting). The hedging results with regime one as the initial state are given in Table 4.5. The standard deviation and 99% VaR of the relative profit and loss are much better than the results for the delta hedge, but not as good as the perfect hedge (see the first row of Table 4.4). In this case the put appears to be the superior hedging instrument. In Table 4.6, the hedging results are presented for each

of the initial states, where the long-dated put is used as the hedging option. The scenario with regime two the initial state is the hardest to hedge against. This is not surprising, as a transition out of regime two is always accompanied by a large upwards jump in the underlying.

Hedging Option	$\mathbb{P}$ Weighting			Uniform Weighting		
	Mean	StdDev	99% VaR	Mean	StdDev	99% VaR
1-year call	0.01	0.23	-0.54	0.01	0.17	-0.25
1.5-year put	0.01	0.16	-0.37	0.02	0.11	-0.26

TABLE 4.5: *Hedging results, in terms of the relative profit and loss, for the three-state regime-switching market from [4] using the local risk minimization with delta neutrality imposed. The target option is a half-year European call with strike $100. The hedge portfolio contains the underlying and one option. The initial state is regime one, and the hedge is rebalanced a total of 100 times over the lifetime of the option.*

Initial State	$\mathbb{P}$ Weighting			Uniform Weighting		
	Mean	StdDev	99% VaR	Mean	StdDev	99% VaR
1	0.01	0.16	-0.37	0.02	0.11	-0.26
2	0.07	0.43	-0.92	0.19	0.20	-0.09
3	0.02	0.14	-0.33	0.02	0.11	-0.29

TABLE 4.6: *Hedging results, in terms of the relative profit and loss, for the three-state regime-switching market from [4] using local risk minimization with delta neutrality imposed. The target option is a half-year European call with strike $100. The hedge portfolio contains the underlying and the 1.5-year put. Each initial state is considered, and the hedge is rebalanced a total of 100 times over the lifetime of the option.*

The hedging results from Table 4.5 and Table 4.6 would lead us to believe that uniform weighting is superior to $\mathbb{P}$ weighting in this case. However, the difference in the results is due mainly to the Truncated Singular Value Decomposition (TSVD) [38] used to solve the linear system (4.36) (the TSVD is introduced in Section 3.1). Numerical investigation reveals that the results we have presented for the perfect hedges are invariant to how the linear

system is solved (i.e. the TSVD with a cutoff parameter of 10^{-6} vs. an LU decomposition). However, this is not so for the local risk minization examples of Table 4.5 and Table 4.6. Experimentation suggests that, by employing the standard cutoff parameter of 10^{-6} in the TSVD, the hedge weights determined using uniform weighting tend to not exactly satisfy the delta-neutral constraint (which is imposed in the linear system (4.36)); instead, the weights offer more protection against the regime-switching risk. In the case of $\mathbb{P}$ weighting, the linear system (4.36) tends to be solved exactly by the TSVD. Since the hedging results with $\mathbb{P}$ weighting are not as good as those for uniform weighting, expending a degree of freedom to impose delta neutrality may not be the best approach. We explore this notion below.

Omitting the Delta-Neutral Constraint

Motivated by the results of the previous section concerning the TSVD, it appears that, instead of imposing delta neutrality, we should minimize an objective that captures both the diffusion and regime-switching risk. Consider the representation of $d\Pi$ in (4.25): the variance in this case, to $\mathcal{O}(dt^2)$, is

$$\text{Variance}\big[d\Pi\big] \;\approx\; dt\Bigg[\Big(\sigma^i S\Big)^2\Big(-\frac{\partial V^i}{\partial S} \;+\; e \;+\; \vec{\phi}\cdot\frac{\partial \vec{F}^i}{\partial S}\Big)^2 \;+\; \sum_{\substack{j=1\\ j\neq i}}^{N}\lambda_{ij}^{\mathbb{P}}\Big[-\Delta V^{ij}+\Big(e\Delta S^{ij}+\vec{\phi}\cdot\Delta\vec{F}^{ij}\Big)\Big]^2\Bigg]. \quad (4.37)$$

If we define a vector containing the deltas of S and $\vec{F}$ as

$$\vec{d} \;=\; \begin{bmatrix} 1 \\ \frac{\partial F_1^i}{\partial S} \\ \frac{\partial F_2^i}{\partial S} \\ \vdots \\ \frac{\partial F_{M-1}^i}{\partial S} \end{bmatrix},$$

then the expression for the variance of $d\Pi$ in (4.37) can be minimized by solving for the hedge weights $\vec{w} = [e, \phi_1, \phi_2, \ldots, \phi_{M-1}]^T$ in

$$\left[\Phi^T \Omega \Phi \;+\; \left(\sigma^i S\right)^2 \vec{d}\,\vec{d}^{\,T}\right] \vec{w} \;=\; \Phi^T \Omega \vec{b} \;+\; \left(\sigma^i S\right)^2 \frac{\partial V^i}{\partial S}\,\vec{d}, \tag{4.38}$$

where the general weighting factors ζ_{ij} are used in place of the transition intensities $\lambda_{ij}^{\mathbb{P}}$.

The hedging simulation in the three-state market is re-run for regime two the initial state (using the 1.5-year put as the hedging option), and the linear system (4.38) is solved when rebalancing. The results are

$$\text{Mean} = 0.17,\;\; \text{StdDev} = 0.20,\;\; 99\%\,\text{VaR} = -0.07$$

for $\mathbb{P}$ weighting and

$$\text{Mean} = 0.18,\;\; \text{StdDev} = 0.20,\;\; 99\%\,\text{VaR} = -0.07$$

when uniform weighting is used. In the case of $\mathbb{P}$ weighting, the results are a clear improvement over the implementation of local risk minization where delta neutrality is imposed (see the second row in Table 4.6).

4.3 Summary

In this chapter, we presented a stable discretization of the option-pricing PDEs for a regime-switching market with shifts in volatility, and developed a convergent fixed point iteration scheme for solving the associated discrete equations. This numerical treatment was required in order to investigate hedging. For an N-state market, we derived a linear system whose solution yields the hedge weights that eliminate the instantaneous risk. In both a two-state model without jumps and a three-state model with jumps, the perfect hedge was found to perform well when rebalanced at discrete times. However, the delta hedge is not an adequate strategy due to the fact it ignores the regime-switching risk. In a many-state market, using the perfect hedge may be prohibitively expensive due to transaction costs. We investigated, in the context of a three-state market, two forms of hedging that use local risk minization—the results are, in general, better than the delta hedge but not as good as the perfect hedge. Precise knowledge of the $\mathbb{P}$ measure is not required for these techniques, a fact we will exploit in subsequent chapters.

Chapter 5

Pricing and Hedging Swing Options

In the previous chapter, we developed a numerical scheme for solving the system of option-pricing PDEs used when the underlying is tradeable and volatility is the only regime-dependent parameter. This chapter extends the numerical scheme to price under a stochastic process where, in addition to a regime-dependent volatility, the drift is both regime and time dependent. In particular, we consider a regime-switching process for the risk-adjusted spot price of natural gas, where the parameters are obtained by calibrating to the derivatives market. The corresponding PDEs are used to price swing options, a type of derivative used by energy producers and consumers as a means of obtaining protection from both volatility in the spot market and uncertain demand. We also investigate the hedging of swing options, which is a more involved problem than the hedging examples from the previous chapter. Swing options are path dependent and, in addition, because the underlying is typically not used to hedge, forward contracts are employed when forming a hedge portfolio.

5.1 Background

Swing options are a type of derivative traded in energy markets that provide flexibility in the timing and amount of delivery [63]. They can include a variety of features, but we shall focus on a standard contract: the holder of the swing option has a series of rights—but not the obligation—to buy or sell, at certain discrete times, fixed quantities of the commodity at a set price. Swing options provide a level of protection comparable to a series of calls and puts, but at a lower cost.

Executing an *up-swing* allows the holder of the option to buy $V_u > 0$ units of the commodity at a cost of K_u per unit. Much like a call, the payoff for an up-swing is

$$V_u \max(S - K_u, 0)\,, \tag{5.1}$$

where S is the spot price.[1] Similar to a put, the payoff from exercising a *down-swing* is

$$V_d \max(K_d - S, 0)\,. \tag{5.2}$$

The option is endowed with a finite number of up-swing (N_u) and down-swing (N_d) rights that may be exercised over the life of the contract. Furthermore, the opportunities to execute these rights occur at pre-determined, discrete times (i.e. the exercise rights are Bermudan).

Penalties may be imposed at the end of the contract if the net amount involved in the swings is too high or too low. To avoid penalties at expiry, the holder of the swing option

[1] This formulation assumes that the holder of the swing option is a *profit maximizing agent*, in that their consumption requirements do not play a role in the exercise decision. The seller of the option takes this view when pricing, as they must consider the 'worst-case scenario' where the holder of the swing option acts to maximize the value of the contract.

should ensure their consumption pattern satisfies

$$V_{\min} \;\leq\; n_u V_u - n_d V_d \;\leq\; V_{\max}\,, \tag{5.3}$$

where $n_u \leq N_u$ and $n_d \leq N_d$ are the number of swing rights actually exercised over the life of the contract, and $[V_{\min}, V_{\max}]$ are bounds set at the beginning. Different types of penalties can be levied at expiry: a fixed amount could be charged for any violation of the consumption limits, or the penalty could be proportional to the number of units outside the bounds.

Lari-Lavassani et al. [52] use a "forest" of binomial trees to compute the value of a standard swing option—their methodology accommodates both one and two-factor processes for the spot price. A similar tree-based approach is considered in [44], where the volume involved in a swing is not fixed but may be chosen from a finite interval. If there are no penalties applied at expiry, then in this case the optimal strategy at each swing time is to either buy/sell the maximum allowed amount or do nothing (i.e. *bang-bang* exercise). In fact Barrera-Esteve et al. [6] show that, when $K_u = K_d$ and the penalty function satisfies certain conditions (of which smoothness is the most restrictive), the optimal consumption strategy is always of bang-bang type. However, we only consider the case of buying and selling a fixed amount of the commodity.

In general, a numerical PDE approach is to be preferred over tree methods. A PDE framework is used in [74] to price a standard swing option, where the spot price process is a one-factor mean-reverting diffusion with seasonality incorporated via time-dependent parameters. Dahlgren [28] considers pricing a more exotic swing option under a similar mean-reverting process, namely the one-factor model of Schwartz [70]. For this option, a swing can be exercised at any time, with the proviso that a refraction time of at least τ_R elapse before executing another right. The price of the swing option is found using a

Hamilton–Jacobi–Bellman (HJB) formulation.

The uncertain nature of demand within the natural gas markets make swing options an ideal risk-mitigating instrument in this setting [71]. To price these contracts, a realistic stochastic model of the underlying spot price is required. One-factor mean-reverting diffusions, which are typically used to model the spot price of natural gas, do not capture the long-term behaviour of the forward curve [44]. In Chen and Forsyth [21], a two-state regime-switching GBM model (with seasonality) is calibrated to natural gas futures and options on futures, and is found to give a good fit to the forward curve for both short and long-term contracts. Furthermore, solving the system of pricing PDEs for this regime-switching model is less computationally intensive than the numerical treatment required for the multi-dimensional PDE that would result from using a two or three-factor model. As such, we will use the regime-switching GBM model of [21] to price and hedge swing options on natural gas.

We note that regime-switching models have also been used for the valuation of derivatives on electrical power [see, e.g., 47], as they can capture the spiking behaviour that is common in the electricity markets. In Davison et al. [31], the spot price of electricity is modelled as a two-state regime-switching process in discrete time, where the price is drawn from a high-price or low-price distribution, depending on the current state (i.e. spike vs. non-spike). The probability of being in the 'spike' state at time t is a function of the demand and capacity at t, and does not depend on the state at the previous timestep—this is in contrast to the regime-switching processes we employ. Davison and Anderson [30] use a variant of the spot price model in [31] to approximate the early exercise boundary for a swing option; with this approximation, the valuation of swing options via Monte Carlo simulation is feasible.

Parameter	Description	Value	
A	Annual constant time adjustment	0.483	
SA	Semi-annual constant time adjustment	0.196	
λ_{12}	Transition intensity for state 1 to 2	0.025	
λ_{21}	Transition intensity for state 2 to 1	4.840	
r	Risk-free rate	0.05	
		Regime 1	Regime 2
σ^i	Diffusive volatility	0.160	0.799
α^i	Constant drift without seasonality	-0.101	1.170
β^i_A	Annual seasonality parameter	0.663	0.400
β^i_{SA}	Semi-annual seasonality parameter	0.332	0.860

TABLE 5.1: *Parameters and data for the two-state regime-switching process* (5.4) *used to model the risk-adjusted ($\mathbb{Q}$ measure) spot price of natural gas [21]. The initial state is regime one.*

5.2 Modelling the Spot Price for Natural Gas

Under the risk-adjusted measure, the spot price for natural gas (in \$/mmBtu) is assumed to follow the two-state regime-switching process from Chen and Forsyth [21]:

$$dS \;=\; \mu_i^{\mathbb{Q}}(t) S\, dt \;+\; \sigma^i S\, dZ \qquad\qquad i = 1, 2\,. \tag{5.4}$$

The drift coefficient

$$\mu_i^{\mathbb{Q}}(t) \;=\; \alpha^i \;+\; \beta^i_A \sin\big(2\pi(t + Dt_0 + A)\big) \;+\; \beta^i_{SA} \sin\big(4\pi(t + Dt_0 + SA)\big) \tag{5.5}$$

is time dependent in order to incorporate seasonality effects. The description of the parameters, plus the specific values from [21] obtained through calibration, are presented in Table 5.1. We use the convention that Dt_0 is the distance (in years) from January 1 to the initial time, and that $t \geq 0$ is measured with respect to this initial time. The drift rate (5.5) is plotted in Figure 5.1 for both regimes, with $Dt_0 = 0.25$.

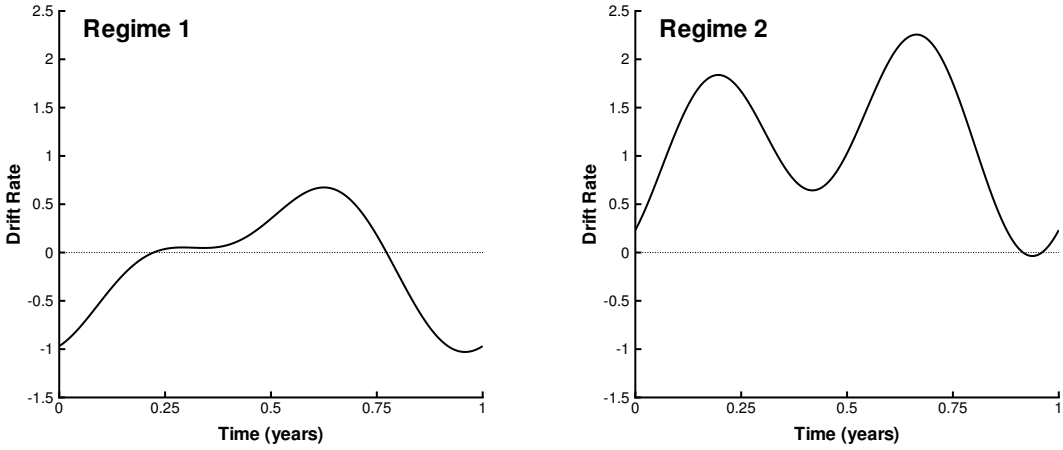


FIGURE 5.1: *Time and regime-dependent drift rate* (5.5) *used in modelling the risk-adjusted spot price of natural gas* (5.4). *The financial data from Table 5.1 is used, with* $Dt_0 = 0.25$.

Appendix A derives the system of pricing PDEs for a general one-factor regime-switching process. For the process (5.4), the system of PDEs is

$$\begin{aligned}
\frac{\partial U^1}{\partial \tau} &= \frac{1}{2}(\sigma^1)^2 S^2 \frac{\partial^2 U^1}{\partial S^2} + \mu_1^{\mathbb{Q}}(t) S \frac{\partial U^1}{\partial S} - \left(r + \lambda_{12}^{\mathbb{Q}}\right) U^1 + \lambda_{12}^{\mathbb{Q}} U^2 \\
\frac{\partial U^2}{\partial \tau} &= \frac{1}{2}(\sigma^2)^2 S^2 \frac{\partial^2 U^2}{\partial S^2} + \mu_2^{\mathbb{Q}}(t) S \frac{\partial U^2}{\partial S} - \left(r + \lambda_{21}^{\mathbb{Q}}\right) U^2 + \lambda_{21}^{\mathbb{Q}} U^1 ,
\end{aligned} \tag{5.6}$$

where $U^1 = U^1(S, \tau)$ and $U^2 = U^2(S, \tau)$, with $\tau = T - t$.

5.3 Pricing Swing Options

5.3.1 Mathematical Specification of the Contract

The swing option has a total of N_u up-swing and N_d down-swing rights that may be exercised over the lifetime T of the contract. These swing rights, whose payoffs are given by (5.1)

and (5.2), may be executed at the discrete times

$$0 \;<\; t_E^1 \;<\; t_E^2 \;<\; t_E^3 \;<\; \ldots \;<\; t_E^{\mathcal{N}-1} \;<\; t_E^{\mathcal{N}} \;\leq\; T\,,$$

where $\mathcal{N}$ is the total number of swing opportunities. Note that at most one swing right can be exercised at each of these times.

Penalties are imposed at time T if the net consumption of the option holder does not lie within $[V_{\min}, V_{\max}]$. For a charge of A_u per net unit traded over $V_{\max}$ and A_d per net unit traded below $V_{\min}$, the penalty is

$$\max\Big(A_u\big[\{n_uV_u - n_dV_d\} - V_{\max}\big], 0\Big) \;+\; \max\Big(A_d\big[V_{\min} - \{n_uV_u - n_dV_d\}\big], 0\Big)\,, \tag{5.7}$$

where n_u and n_d are the number of up-swing and down-swing rights, respectively, actually exercised over the lifetime of the swing option.

Let $U^i_{\{n_u,n_d\}}(S,\tau)$ denote the swing option value for a spot price S at (backward) time τ in regime i, where n_u up-swing and n_d down-swing rights have been exercised. Away from the exercise dates, the system of PDEs (5.6) governs the swing option values—there is a distinct system for each possible (n_u, n_d) pair. The initial conditions for the PDEs in (5.6) are set according to the penalty provision (5.7), with

$$\begin{aligned} U^i_{\{n_u,n_d\}}(S,\tau=0) \;=\; -\Bigg[&\max\Big(A_u\big[\{n_uV_u - n_dV_d\} - V_{\max}\big], 0\Big) \\ &+\; \max\Big(A_d\big[V_{\min} - \{n_uV_u - n_dV_d\}\big], 0\Big)\Bigg] \end{aligned}$$

for $n_u = 0, 1, \ldots, N_u$; $n_d = 0, 1, \ldots, N_d$; and $i = 1, 2$. As $S \to 0$, the system (5.6) reduces to

$$\begin{aligned} \frac{\partial U^1_{\{n_u,n_d\}}}{\partial \tau} &= -\big(r + \lambda^{\mathbb{Q}}_{12}\big)U^1_{\{n_u,n_d\}} + \lambda^{\mathbb{Q}}_{12}U^2_{\{n_u,n_d\}} \\ \frac{\partial U^2_{\{n_u,n_d\}}}{\partial \tau} &= -\big(r + \lambda^{\mathbb{Q}}_{21}\big)U^2_{\{n_u,n_d\}} + \lambda^{\mathbb{Q}}_{21}U^1_{\{n_u,n_d\}}\,. \end{aligned} \tag{5.8}$$

As $S \to \infty$, we assume that

$$\frac{\partial^2 U^1_{\{n_u,n_d\}}}{\partial S^2} \simeq 0 \quad \text{and} \quad \frac{\partial^2 U^2_{\{n_u,n_d\}}}{\partial S^2} \simeq 0\,. \tag{5.9}$$

Across each exercise date t_E, the no-arbitrage principle establishes the jump condition

$$U^i_{\{n_u,n_d\}}(S, t_E^-) = \max\begin{cases} U^i_{\{n_u+1,n_d\}}(S, t_E^+) + V_u \max(S - K_u, 0) & \text{(swing up)} \\ U^i_{\{n_u,n_d+1\}}(S, t_E^+) + V_d \max(K_d - S, 0) & \text{(swing down)} \\ U^i_{\{n_u,n_d\}}(S, t_E^+) & \text{(no swing)}\ . \end{cases} \tag{5.10}$$

If $n_u = N_u$ and/or $n_d = N_d$, then an up-swing and/or down-swing, respectively, cannot be executed.

5.3.2 Numerical Solution

Recall from Chapter 4 that, in order to price a simple contract under an N-state regime-switching process, N meshes with domain $[0, S_{\max}] \times [0, T]$ are required to numerically solve the system of coupled PDEs—these meshes may be thought of as forming a cube with domain $[0, S_{\max}] \times [0, T] \times [1, 2, \ldots, N]$. When pricing a swing option, for each possible pair (n_u, n_d), there exists a computational cube corresponding to a swing option with that

combination of exercised rights. This requires a total of

$$\big(N_u + 1\big)\big(N_d + 1\big)$$

cubes, where N_u and N_d are the total number of allowed up-swings and down-swings, respectively.

On each (n_u, n_d) cube the system of PDEs (5.6) are discretized, and the boundary conditions (5.8) and (5.9) handled, in the same manner as the PDE system (4.2) (see Section 4.1.2). Away from the the exercise dates, the option values for each cube are found by numerically solving the system of PDEs. This continues until τ_E^- (t_E^+)—an instant after a swing right may have been executed—at which time the option holder has already decided whether or not to use a permissible swing opportunity. The jump condition (5.10) is applied at each node, relative to backward time, to yield

$$U^i_{\{n_u,n_d\}}(S,\tau_E^+) \;=\; \max\begin{cases} U^i_{\{n_u+1,n_d\}}(S,\tau_E^-) \;+\; V_u \max(S-K_u,0) & \text{(swing up)} \\ U^i_{\{n_u,n_d+1\}}(S,\tau_E^-) \;+\; V_d \max(K_d-S,0) & \text{(swing down)} \\ U^i_{\{n_u,n_d\}}(S,\tau_E^-) & \text{(no swing)} \;, \end{cases} \tag{5.11}$$

where $U^i_{\{n_u,n_d\}}(S,\tau_E^-)$ denotes the value computed from solving the PDEs before the jump condition is applied.

This process continues until $\tau = T$ ($t = 0$). The swing option value at inception is given by $U^{R_0}_{\{0,0\}}(S_0,T)$, where R_0 is the initial regime and S_0 is the initial spot price.

5.3.3 A Pricing Example

Using the financial data of Table 5.1, a swing option specified by the parameters in Table 5.2 is priced. Here, $\Delta t_E = 0.05$ means there are twenty equally-spaced swing opportunities

$T = 1$ year	$\Delta t_E = 0.05$ years	$N_u = 2$	$N_d = 2$	$K_u = \$5$	$K_d = \$5$
$V_u = 20$ mmBtu	$V_d = 20$ mmBtu	$A_u = 0$	$A_d = 0$		

TABLE 5.2: *Specification of a swing option contract.*

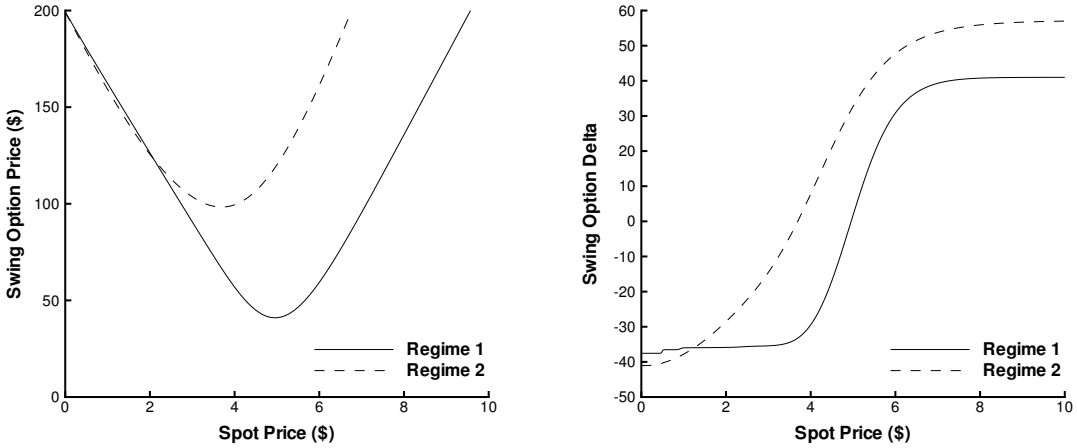


FIGURE 5.2: *Swing option value and delta at $t = 0$, as a function of spot price, for the financial data of Table 5.1 and the swing option specification in Table 5.2.*

(starting at $t = 0.05$ and ending at $t = 1.0$) over the one-year life of the contract. We use $Dt_0 = 0.25$, which means the initial time $t = 0$ roughly corresponds to April 1. The option price and delta at $t = 0$ for both regimes (with $n_u = 0$ and $n_d = 0$) are plotted in Figure 5.2.

The jump condition (5.11) yields the optimal exercise strategy. Consider the optimal down-swing regions presented in Figure 5.3: for a given n_d and exercise time, if the spot price is within the region then the option holder should execute a down-swing. The optimal exercise region clearly depends on the number of down-swings n_d that have already been executed; the region for $n_d = 0$ is always at least as big as the region corresponding to $n_d = 1$—since the option holder has more swing rights remaining when $n_d = 0$, they are quicker to "pull the trigger" and swing down. The extent of the down-swing regions is

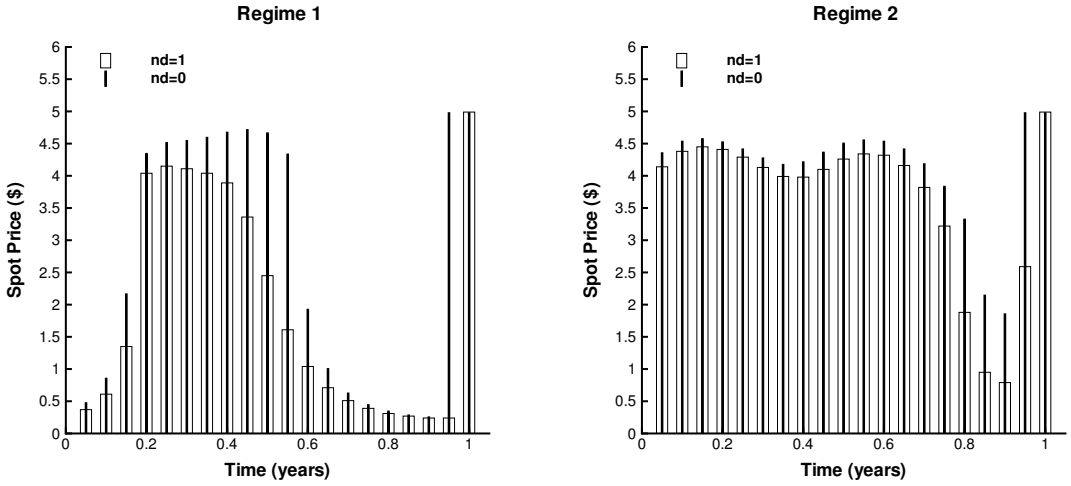


FIGURE 5.3: *Down-swing exercise regions for the swing option specified by the parameters of Table 5.2—left panel: regime one, right panel: regime two. For a given n_d and exercise time, if the spot price is within the region then a down-swing should be executed. The pricing parameters used are in Table 5.1.*

correlated to the risk-adjusted drift $\mu^{\mathbb{Q}}(t)$. By comparing Figures 5.3 and 5.1, it is seen the troughs and peaks of the exercise boundary roughly align with the troughs and peaks of $\mu^{\mathbb{Q}}(t)$.

The optimal swing-up regions are presented in Figure 5.4: for a given n_u and exercise time, the option holder should execute an up-swing if the spot price is within the region (if one exists). Note that for most exercise times it is never optimal to execute an up-swing, regardless of the spot's value.

Consider the delta at time zero for regime one, plotted in the right panel of Figure 5.2 as the solid curve. The irregular behaviour for small values of the spot price S is not a numerical artefact, but is due to the application of the jump condition (5.11) during the backward induction solution procedure. Consider the first swing opportunity at $t = 0.05$: during the backward induction, this is actually the last time the jump condition (5.11) will be applied, since it corresponds to $\tau = 0.95$. If we assume that the option value is continuous w.r.t. spot price before applying the jump condition (5.11), then it remains

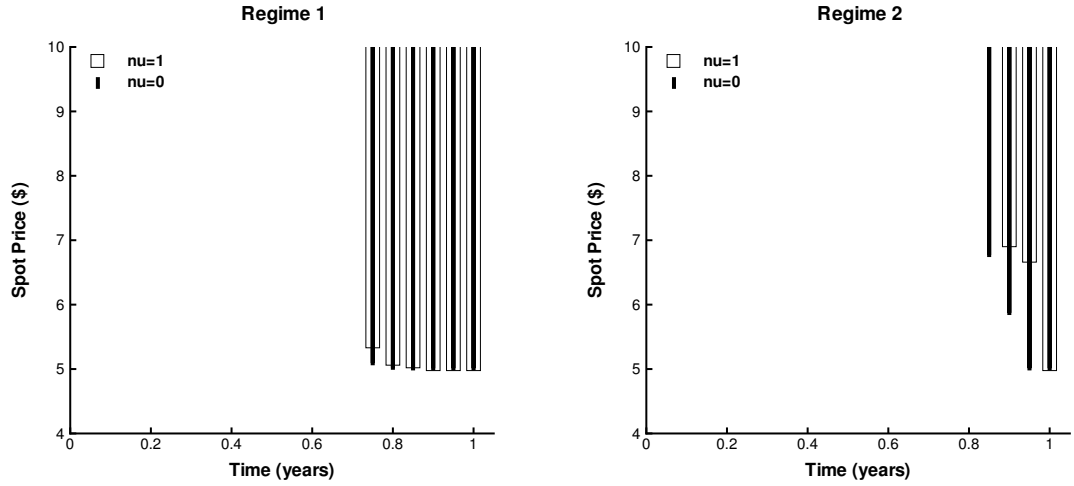


FIGURE 5.4: *Up-swing exercise regions for the swing option specified by the parameters of Table 5.2—left panel: regime one, right panel: regime two. For a given n_u and exercise time, if the spot price is within the region then an up-swing should be executed. The upper boundary of the regions is $S = \infty$. The pricing parameters used are in Table 5.1.*

continuous afterwards; i.e.

$$U^{\text{regime}=1}_{\{n_u=0,n_d=0\}}(S, t = 0.05) \tag{5.12}$$

is continuous w.r.t. spot price S. In general, however, the function in (5.12) will not be smooth at the boundary $S = S_{\text{X-DS}}$ of the region where a down-swing is optimal; in our example, $S_{\text{X-DS}} = 0.47$ at $t = 0.05$ ($\tau = 0.95$). Therefore, the first derivative of (5.12) (i.e. the delta) will be discontinuous at $S = 0.47$.

Figure 5.2 plots the option value and delta at time zero ($\tau = 1$), while the first swing opportunity—where a discontinuity in delta occurs—is at $t = 0.05$ ($\tau = 0.95$). However, the relatively low volatility in regime one ($\sigma^1 = 0.160$) means the discontinuity in delta at $\tau = 0.95$ ($t = 0.05$) will not have been substantially smoothed by the PDEs when $\tau = 1$ ($t = 0$) is reached. Indeed, near $S = 0.85$ there is another bump in the delta evident in Figure 5.2, which is a remnant of the discontinuity at $\tau = 0.9$ ($t = 0.1$) due to the application

of the jump condition at that time. By way of comparison, a similar discontinuity appears in the delta at $t = 0.05$ ($\tau = 0.95$) for regime two, but is quickly smoothed by the high volatility in that regime ($= 0.799$), such that it is not evident at $t = 0$ ($\tau = 1$).

5.4 Hedging a Swing Option

A swing option is path dependent, as the exercise decision at each swing opportunity is contingent on the spot price at that time and the number of swing rights already executed. We assume the seller of the option does not hedge with the underlying, so forward[2] contracts will be treated as the base security for hedging. However, since the swing option is priced in a regime-switching market, a perfect hedge cannot be constructed using just a forward.

Let $f(t_0, T)$ denote the forward price at time t_0 for a contract with delivery at T, and $U_F(S, t, f(t_0, T))$ be the value of a forward contact at (S, t) with forward price $f(t_0, T)$ (where $t_0 \leq t \leq T$). From the fundamental definition of the forward price,

$$U_F(S_0, t_0, f(t_0, T)) \;=\; 0\,.$$

At time t_m, the hedger enters into the prevailing forward contract (at no cost). At a time t_{m+1} later, she must rebalance the hedge portfolio. Consider two possible strategies:

1. Close out the position in the original forward contract, which is now worth $U_F(S_{m+1}, t_{m+1}, f(t_m, T))$, and enter into the prevailing forward contract (with forward price $f(t_{m+1}, T)$). This is essentially the *mark-to-market* process that occurs with futures trading;

[2] A long position in a forward contract obligates the holder to buy the underlying asset or commodity at some future *delivery time* T for the *forward price* f; the payoff of the contract at delivery is therefore $S_T - f$. The forward price f is chosen at inception such that it costs nothing to enter into the contract, and may be found using $f = \mathbb{E}^{\mathbb{Q}}[S_T|S_0]$. Conceptually, a forward contract can be thought of as the combination of a long position in a European call and a short position in a European put whose common strike f (the forward price) is chosen such that the initial value of the net position is zero.

2. Rebalance the position in the original forward contract.

In fact, since

$$S_T - f(t_{m+1}, T) \;=\; \Big[S_T - f(t_m, T)\Big] + \Big[f(t_m, T) - f(t_{m+1}, T)\Big]$$

we have that, at time t_{m+1} and spot price S_{m+1},

$$\overbrace{U_F\big(S_{m+1}, t_{m+1}, f(t_{m+1}, T)\big)}^{\text{New forward contract}}$$
$$= \underbrace{U_F\big(S_{m+1}, t_{m+1}, f(t_m, T)\big)}_{\text{Original forward contract}} + \underbrace{\mathrm{e}^{-r(T-t_{m+1})}\Big[f(t_m, T) - f(t_{m+1}, T)\Big]}_{\text{Cash}},$$

which means the value of the new and original forward contracts differ only by an amount of cash. Hence, the two above strategies are the same—we will employ the second for computational convenience. Therefore, the forward price at time zero is computed for delivery when the swing option expires at time T, such that $f_0 = f(0, T)$. A contract with payoff $S_T - f_0$ is used as the base hedging instrument.

We consider hedging the swing option specified by the parameters in Table 5.2. The initial state is regime one with $S_0 = 5.00$, such that the forward price for delivery in one year is $f_0 = \$4.55$. The same hedging strategies as Chapter 4 are employed: to eliminate just the diffusion risk, the forward is used to establish a delta-neutral hedge, while a 1.5-year European call option on the spot,[3] with a strike of \$5.0, is used along with the forward to hedge both the diffusion and regime-switching risk (i.e. a perfect hedge). For each hedging strategy, 250,000 simulations are carried out. It is assumed that the holder of the swing option acts optimally in terms of exercising their swing rights—they use the optimal exercise

[3]In general, put/call options in the energy markets are written on forward/future prices. However, for a forward contact with delivery time T, the forward price converges to the spot price as $t \to T$. Since options on natural gas forwards traded on the New York Mercantile Exchange expire the day before the underlying forward contract, using a European option written on the spot is a valid approximation.

Parameter	Description	Value	
A	Annual constant time adjustment	0.483	
SA	Semi-annual constant time adjustment	0.196	
λ_{12}	Transition intensity for state 1 to 2	0.5	
λ_{21}	Transition intensity for state 2 to 1	2.0	
r	Risk-free rate	0.05	
		Regime 1	Regime 2
σ^i	Diffusive volatility	0.160	0.799
α^i	Constant drift without seasonality	-0.040	0.268
β_A^i	Annual seasonality parameter	0.265	0.160
β_{SA}^i	Semi-annual seasonality parameter	0.133	0.344

TABLE 5.3: *Parameters and data for the two-state regime-switching process used to model the real-world ($\mathbb{P}$ measure) spot price of natural gas. The initial state is regime one.*

strategy captured by Figures 5.3 and 5.4.

The real-world spot price of natural gas, governed by the $\mathbb{P}$ measure, spends more time in the high-volatility state (regime two) than suggested by the $\mathbb{Q}$ measure [20]. We therefore make transitions out of regime one more frequent, and transitions out of regime two less frequent, under the $\mathbb{P}$ measure we use for simulation. To simulate the real-world spot price, the same process as (5.4) is used with the $\mathbb{P}$ measure data from Table 5.3. The time-dependent drift under the $\mathbb{P}$ measure, $\mu_i^{\mathbb{P}}(t)$, has the same fundamental shape as under the $\mathbb{Q}$ measure (see Figure 5.1), only it is vertically shifted and compressed. This is an admittedly ad-hoc approach, but our goal is to simulate a real-world process that is somewhat different from the risk-adjusted process.

The hedging results for various rebalancing frequencies are given in Table 5.4. The regime-switching risk ignored by the delta hedge manifests as a higher standard deviation and more negative 99% VaR than for the perfect hedge—treating the regime-switching risk by adding the call option to the hedge portfolio leads to a marked improvement in the

results. In addition, more frequent rebalancing leads to a lower standard deviation and less negative 99% VaR for the perfect hedge, as opposed to the delta hedge, where the improvement levels off.

For the case of 100 rebalancings, the distributions of profit and loss are presented in Figure 5.5. The mean of the P&L for the delta hedge is negative, such that the positive bias of the central mass is offset by a left tail (which is not visible); this left tail is absent with the perfect hedge

Number of Rebalancings	Forward Only ('delta' hedge)			Forward + Call ('perfect' hedge)		
	Mean	StdDev	99% VaR	Mean	StdDev	99% VaR
20	-0.40	0.75	-3.11	-0.01	0.22	-0.75
100	-0.37	0.69	-2.72	0.00	0.10	-0.33
200	-0.37	0.68	-2.66	0.00	0.07	-0.23

TABLE 5.4: *Statistical measures of the relative profit and loss when hedging the swing option specified by Table 5.2 under the real-world process described by the data in Table 5.3. Both the delta hedge and perfect hedge are considered. Regime one is the initial state, and 250,000 simulations are carried out. The 99% VaR is defined as the 1% quantile of the relative profit and loss.*

Hedging by Local Risk Minimization

Recall from Chapter 4 that, in a two-state regime-switching model where the underlying is tradeable but does not jump, the delta hedge (using the underlying asset) is the local risk-minimizing hedging strategy, even though it only treats the diffusion risk [55]. This is because the underlying's price does not change when the regime changes, and hence cannot be used to mitigate the regime-switching risk. However, the value of the forward contract in our natural gas market will change when there is a change in regime. Therefore, if the forward alone is used to hedge, we can consider a local risk-minimizing strategy that treats both the diffusion and regime-switching risk simultaneously (see Section 4.2.5).

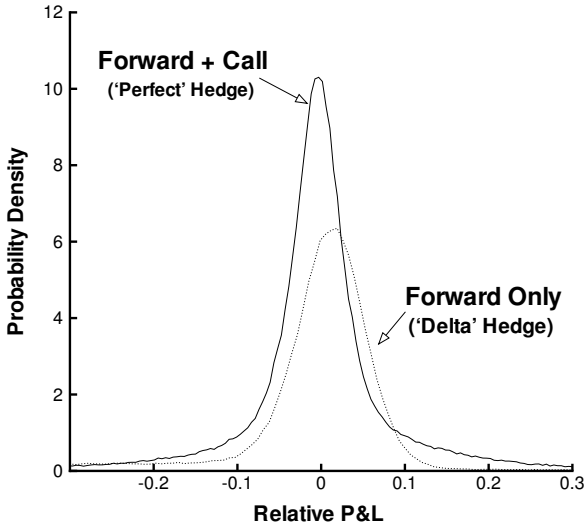


FIGURE 5.5: *Distributions of relative profit and loss when hedging the swing option specified in Table 5.2 under the real-world process described by the data in Table 5.3. Both the delta hedge and a perfect hedge are considered. The hedge is rebalanced at intervals of 0.01 years. Regime one is the initial state, and 250,000 simulations are carried out.*

Assuming the spot price is in regime i, the optimization to be solved when rebalancing is

$$\min_{\phi} \left(\left(\sigma^i S\right)^2 \left(- \frac{\partial V^i}{\partial S} + \phi \frac{\partial U_F^i}{\partial S} \right)^2 + \sum_{\substack{j=1 \\ j \neq i}}^{2} \zeta_{ij} \left[- \Delta V^{ij} + \phi \Delta U_F^{ij} \right) \right]^2 \right), \tag{5.13}$$

where ϕ is the weight in the forward contract U_F, and V is the value of the swing option. We again must choose the weighting factors ζ_{ij} for the transitions out of the current regime i. Using the $\mathbb{P}$ measure transition intensities $\lambda_{ij}^{\mathbb{P}}$ is the most natural choice, since (5.13) will correspond to a local minimization of Variance$[d\Pi]$, where Π is the value of the overall hedged position. Generally, however, these intensities are unknown. As before, setting $\zeta_{ij} = 1$ encapsulates a lack of knowledge concerning the rate of regime change.

Hedging results for both $\mathbb{P}$ and uniform weighting are presented in Table 5.5. With both weightings, the results are better than for the delta hedge (compare with Table 5.4),

although there is not much improvement as the rebalancing becomes more frequent.

Number of Rebalancings	$\mathbb{P}$ Weighting			Uniform Weighting		
	Mean	StdDev	99% VaR	Mean	StdDev	99% VaR
20	-0.09	0.55	-1.72	-0.03	0.58	-1.58
100	-0.08	0.51	-1.44	-0.02	0.52	-1.27
200	-0.08	0.50	-1.40	-0.02	0.52	-1.22

TABLE 5.5: *Statistical measures of the relative profit and loss for the local risk-minimizing hedge of the swing option specified by Table 5.2 under the real-world process described by the data in Table 5.3. Regime one is the initial state, and 250,000 simulations are carried out.*

5.5 Summary

In this chapter, we demonstrated how the numerical scheme for solving the option-pricing PDEs in a regime-switching model can be used to valuate swing options in a realistic market. The pricing procedure yields the optimal exercise strategy that the holder of such an option should execute. On the other hand, the seller of the swing option must hedge their short position. Using a forward contract to establish a delta-neutral hedge will not protect against the regime-switching risk. A local risk-minimizing hedge performs somewhat better, but the hedger is still subject to significant tail risk. However, in a two-state regime-switching market, the forward and an additional hedging instrument are sufficient to eliminate all sources of (instantaneous) risk. When this perfect hedge is used as the basis of a discretely rebalanced strategy, the results are much better than a delta or local risk-minimizing hedge that uses the forward only.

Chapter 6

Dynamic Hedging under Jump Diffusion: Theory

In the following, we derive an expression for the instantaneous jump risk in a jump-diffusion model. We then consider an idealized continuous trading environment and demonstrate that, by imposing delta neutrality and suitably bounding the jump risk and transaction costs at each instant, the variance of the terminal hedging error can be made arbitrarily small. This continuous-time treatment motivates an objective function that can be minimized at each rebalance time within a discrete hedging procedure.

6.1 Derivation of Jump Risk

Consider a jump-diffusion market whose evolution is governed by the stochastic process[1]

$$dS \;=\; \big(\alpha^{\mathbb{P}} - \lambda^{\mathbb{P}}\kappa^{\mathbb{P}}\big)S\,dt \;+\; \sigma S\,dZ^{\mathbb{P}} \;+\; \big(J-1\big)S\,d\pi^{\mathbb{P}}\,, \tag{6.1}$$

[1] Although dividends are not treated in this thesis, their addition would not complicate the hedging strategies.

where jumps arrive with intensity $\lambda^{\mathbb{P}}$. The jump size J is distributed according to $g^{\mathbb{P}}(J)$ and has a mean of $\kappa^{\mathbb{P}}+1$. The option-pricing PIDE for derivatives in this market is

$$\frac{\partial V}{\partial t} \;+\; \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \;+\; \big(r-\lambda^{\mathbb{Q}}\kappa^{\mathbb{Q}}\big)S\frac{\partial V}{\partial S} \;-\; \big(r+\lambda^{\mathbb{Q}}\big)V \;+\; \lambda^{\mathbb{Q}}\int_0^\infty V(JS,t)g^{\mathbb{Q}}(J)\,dJ \;\;=\;\; 0\,. \tag{6.2}$$

To hedge a target option V, the standard hedge portfolio is used: it contains an amount B in cash, a long position in e units of the underlying, and a long position in N additional hedging instruments $\vec{I}$ with weights $\vec{\phi}$. The overall hedged position has value

$$\Pi \;=\; -V \;+\; \underbrace{eS \;+\; \vec{\phi}\cdot\vec{I} \;+\; B}_{\text{Hedge Portfolio } \mathcal{H}}\,.$$

To represent changes in the components of Π due to a jump of size J, we use $\Delta V = V(JS,t)-V(S,t)$, $\Delta S = S(J-1)$, and $\Delta\vec{I} = \vec{I}(JS,t)-\vec{I}(S,t)$.

If a change in the short position $-V$ is always precisely neutralized by the hedge portfolio $\mathcal{H}$, the hedge is considered perfect and Π will have zero variation over an instant dt. We must therefore consider the infinitesimal change in the overall hedged position value Π. Since we are concerned with the real-world evolution of this portfolio, the underlying jump-diffusion process of interest is governed by the $\mathbb{P}$ measure, and is given in (6.1). We have:

$$\begin{aligned}
dS &= \xi^{\mathbb{P}} S\,dt \;+\; \sigma S\,dZ^{\mathbb{P}} \;+\; \Delta S\,d\pi^{\mathbb{P}} \\
dV &= \left[\frac{\partial V}{\partial t}+\frac{\sigma^2S^2}{2}\frac{\partial^2 V}{\partial S^2}+\xi^{\mathbb{P}}S\frac{\partial V}{\partial S}\right]dt \;+\; \sigma S\frac{\partial V}{\partial S}\,dZ^{\mathbb{P}} \;+\; \Delta V\,d\pi^{\mathbb{P}} \\
d\vec{I} &= \left[\frac{\partial \vec{I}}{\partial t}+\frac{\sigma^2S^2}{2}\frac{\partial^2 \vec{I}}{\partial S^2}+\xi^{\mathbb{P}}S\frac{\partial \vec{I}}{\partial S}\right]dt \;+\; \sigma S\frac{\partial \vec{I}}{\partial S}\,dZ^{\mathbb{P}} \;+\; \Delta\vec{I}\,d\pi^{\mathbb{P}} \\
dB &= rB\,dt\,,
\end{aligned}$$

where $\xi^{\mathbb{P}} = \alpha^{\mathbb{P}} - \lambda^{\mathbb{P}}\kappa^{\mathbb{P}}$ and Itô's lemma has been used to obtain dV and $d\vec{I}$. The above implies that the instantaneous change in the value of the overall hedged position is

$$\begin{aligned} d\Pi &= -dV \;+\; e\,dS \;+\; \vec{\phi}\cdot d\vec{I} \;+\; dB \\ &= -\left[\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\right] dt \;+\; \vec{\phi}\cdot\left[\frac{\partial \vec{I}}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 \vec{I}}{\partial S^2}\right] dt \\ &\quad +\; rB\,dt \;+\; \left[-\Delta V + \left(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right)\right] d\pi^{\mathbb{P}} \\ &\quad +\; \xi^{\mathbb{P}} S\left[-\frac{\partial V}{\partial S} + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right] dt \;+\; \sigma S\left[-\frac{\partial V}{\partial S} + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right] dZ^{\mathbb{P}}, \end{aligned} \tag{6.3}$$

where e and $\vec{\phi}$ are regarded as constant over dt as they must be set at the beginning of this instant.

If the portfolio is delta neutral, then

$$-\frac{\partial V}{\partial S} \;+\; e \;+\; \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S} \;=\; 0\,. \tag{6.4}$$

Imposing delta neutrality within equation (6.3) eliminates the final two terms in the expression for $d\Pi$, including the one involving the Wiener process $dZ^{\mathbb{P}}$. The expression for $d\Pi$ consequently simplifies to

$$\begin{aligned} d\Pi \;=\; &-\left[\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\right] dt \;+\; \vec{\phi}\cdot\left[\frac{\partial \vec{I}}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 \vec{I}}{\partial S^2}\right] dt \\ &+\; rB\,dt \;+\; \left[-\Delta V + \left(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right)\right] d\pi^{\mathbb{P}}, \end{aligned} \tag{6.5}$$

indicating that $d\Pi$ is now a pure jump process with drift. Using an elementary rearrange-

ment, the pricing PIDEs (6.2) for V and $\vec{I}$ may be written as

$$\begin{aligned} \frac{\partial V}{\partial t} \;+\; \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} \;&=\; rV \;+\; \big\{\lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}(\Delta S) - rS\big\}\frac{\partial V}{\partial S} \;-\; \lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}(\Delta V) \\ \frac{\partial \vec{I}}{\partial t} \;+\; \frac{\sigma^2 S^2}{2}\frac{\partial^2 \vec{I}}{\partial S^2} \;&=\; r\vec{I} \;+\; \big\{\lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}(\Delta S) - rS\big\}\frac{\partial \vec{I}}{\partial S} \;-\; \lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}(\Delta \vec{I})\,, \end{aligned} \tag{6.6}$$

where $\mathbb{E}^{\mathbb{Q}}(\Delta S) = \mathbb{E}^{\mathbb{Q}}\big(S[J-1]\big) = S\,\mathbb{E}^{\mathbb{Q}}(J-1) = S\kappa^{\mathbb{Q}}$ and

$$\mathbb{E}^{\mathbb{Q}}(\Delta V) \;=\; \int_0^\infty \big(V(JS,t) - V(S,t)\big) g^{\mathbb{Q}}(J)\, dJ\,.$$

Substituting (6.6) into (6.5) yields

$$\begin{aligned} d\Pi \;&=\; -\left[rV + \big\{\lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}(\Delta S) - rS\big\}\frac{\partial V}{\partial S} - \lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}(\Delta V)\right] dt \\ &\quad +\; \vec{\phi}\cdot\left[r\vec{I} + \big\{\lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}(\Delta S) - rS\big\}\frac{\partial \vec{I}}{\partial S} - \lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}(\Delta \vec{I})\right] dt \\ &\quad +\; rB\,dt \;+\; \Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big] d\pi^{\mathbb{P}} \\ &=\; r\left[-V + \left(\frac{\partial V}{\partial S} - \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right) S + \vec{\phi}\cdot\vec{I} + B\right] dt \\ &\quad +\; \lambda^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}(\Delta V) - \left(\frac{\partial V}{\partial S} - \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right)\mathbb{E}^{\mathbb{Q}}(\Delta S) - \vec{\phi}\cdot\mathbb{E}^{\mathbb{Q}}(\Delta\vec{I})\right] dt \\ &\quad +\; \Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]\, d\pi^{\mathbb{P}}\,. \end{aligned} \tag{6.7}$$

Using the delta neutral constraint (6.4) in (6.7) gives

$$\begin{aligned} d\Pi &= r\Big[-V + eS + \vec{\phi}\cdot\vec{I} + B\Big]dt \;+\; \lambda^{\mathbb{Q}}\Big[\mathbb{E}^{\mathbb{Q}}(\Delta V) - e\mathbb{E}^{\mathbb{Q}}(\Delta S) - \vec{\phi}\cdot\mathbb{E}^{\mathbb{Q}}(\Delta\vec{I})\Big]dt \\ &\quad + \Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]d\pi^{\mathbb{P}} \\ &= r\Pi\,dt \\ &\quad + \lambda^{\mathbb{Q}}\,dt\,\mathbb{E}^{\mathbb{Q}}\Big[\Delta V - \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big] \;+\; d\pi^{\mathbb{P}}\Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]. \end{aligned} \tag{6.8}$$

Equation (6.8) indicates that the value of the overall hedged position grows at the risk-free rate (as usual, if the portfolio is delta neutral), but has additional terms due to the jump component:

$$\underbrace{\lambda^{\mathbb{Q}}\,dt\,\mathbb{E}^{\mathbb{Q}}\Big[\Delta V - \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big] \;+\; d\pi^{\mathbb{P}}\Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]}_{\text{instantaneous jump risk}}. \tag{6.9}$$

The first component of the jump risk is deterministic, while the second part is stochastic as it depends on whether a jump occurs over the instant dt. Note that if the jump processes under $\mathbb{P}$ and $\mathbb{Q}$ are the same, the real-world expected value of the instantaneous jump risk is zero.

Remark 6.1 (Regime-Switching Risk vs. Jump Risk). *The regime-switching risk in* (4.28) *and the jump risk in* (6.9) *are of the same form. Since the possible states in a regime-switching model come from a finite set, the expectation over all states involves a sum. On the other hand, when jumps in a jump-diffusion model are drawn from a continuum, expectations will involve an integral.*

6.2 Global Bound on the Hedging Error

Without using an infinite number of hedging instruments, it is impossible to totally eliminate the jump risk—and hence form a perfect hedge—in a jump-diffusion market with a continuum of possible jump amplitudes [1]. However in this section we demonstrate that, by forcing the jump risk and transaction costs to be sufficiently small at each instant of a continuously rebalanced, delta-neutral hedge, the variance of the terminal hedging error can be made arbitrarily small.

For future reference, let $\mathbb{E}_s[X_t]$ for $t \geq s$ denote the expected value of X_t conditioned on information known at time s. We define a proper weighting function in the following way:

Definition 6.1 (Proper Weighting Function). *A proper weighting function W with respect to g is such that*

$$\int_0^\infty f^2(J)g(J)\,dJ \;\leq\; \int_0^\infty f^2(J)W(J)\,dJ \;<\; \infty \tag{6.10}$$

for any function f satisfying $\displaystyle\int_0^\infty f^2(J)g(J)\,dJ \;<\; \infty$.

Also, to avoid complication w.r.t. imposing delta neutrality, we shall assume for the purposes of this chapter that the transaction cost for the underlying is zero.

For a hedge portfolio consisting of N instruments, the total transaction cost of rebalancing at time t is

$$\Upsilon_t \;=\; \sum_{k=1}^{N} \Upsilon_t^k \;\geq\; 0\,, \tag{6.11}$$

where Υ_t^k is the transaction cost for the k^{th} instrument. We will construct the hedging strategy so that the transaction cost is proportional to dt; that is,

$$\Upsilon_t \;=\; \vartheta_t\,dt\,, \tag{6.12}$$

with $\vartheta_t \geq 0$. This assumption is required in order to ensure finite transaction costs—we will indicate in the next section how this can be approximated in practice.

Over an instant, the value of the overall hedged position will decrease by an amount $\Upsilon_t = \vartheta_t dt$ due to the transaction costs, such that the time evolution of the hedged portfolio will be augmented by a negative drift, namely $-\vartheta_t dt$. The instantaneous change in the delta-neutral hedged position (6.8) may therefore be expressed as

$$d\Pi_t \;=\; \Big(r\Pi_t \;+\; \lambda^{\mathbb{Q}}\,\mathbb{E}_t^{\mathbb{Q}}\Big[-\Delta H_J(S_t,t)\Big] \;-\; \vartheta_t\Big)\,dt \;+\; \Delta H_J(S_t,t)\,d\pi^{\mathbb{P}}\,, \tag{6.13}$$

where

$$\Delta H_J(S_t,t) \;=\; -\Big(V(JS_t,t)-V(S_t,t)\Big) \;+\; eS_t\big(J-1\big) \;+\; \vec{\phi}\cdot\Big(\vec{I}(JS_t,t)-\vec{I}(S_t,t)\Big) \tag{6.14}$$

is the random jump component. Note that, as opposed to the derivation of jump risk in the previous section, here we make explicit the dependence on time. This allows us to more easily consider the discounted overall hedged position $\tilde{\Pi}_t = \mathrm{e}^{-rt}\,\Pi_t$, such that

$$d\tilde{\Pi}_t \;=\; \mathrm{e}^{-rt}\Big(-r\Pi_t\,dt \;+\; d\Pi_t\Big)\,. \tag{6.15}$$

Substituting (6.13) into (6.15) yields

$$d\tilde{\Pi}_t \;=\; \mathrm{e}^{-rt}\Bigg(\Big(\lambda^{\mathbb{Q}}\,\Theta_t^{\mathbb{Q}} \;-\; \vartheta_t\Big)\,dt \;+\; \Delta H_J(S_t,t)\,d\pi^{\mathbb{P}}\Bigg) \tag{6.16}$$

with

$$\Theta_t^{\mathbb{Q}} \;=\; \mathbb{E}_t^{\mathbb{Q}}\Big[-\Delta H_J(S_t,t)\Big]\,. \tag{6.17}$$

Our goal is to show that, by making the jump risk and transaction costs sufficiently small at each instant, the variance of the terminal hedging error Π_T can be made arbitrarily close to zero. To do this we will examine the expectation of $(\tilde{\Pi}_T)^2$, where $\tilde{\Pi}_t^2$ follows (by Itô's formula)

$$d\tilde{\Pi}_t^2 \;=\; 2\mathrm{e}^{-rt}\left(\lambda^{\mathbb{Q}}\,\Theta_t^{\mathbb{Q}} \;-\; \vartheta_t\right)\tilde{\Pi}_t\,dt \;+\; \left(\tilde{\Pi}_t^2(JS_t) - \tilde{\Pi}_t^2(S_t)\right)d\pi^{\mathbb{P}}\,. \tag{6.18}$$

Here, $\tilde{\Pi}_t^2(JS_t) - \tilde{\Pi}_t^2(S_t)$ represents the change in the value of $\tilde{\Pi}^2$ assuming an asset price jump of size J occurs at time t. We will need the following result related to this and other quantities in order to establish a bound on the hedging error.

Lemma 6.1 (Bounds on Expectations Involving $\tilde{\Pi}_t$). *Assume that, for time t, the following five conditions are met:*

(A1) *$W(J)$ is a proper weighting function with respect to both $g^{\mathbb{P}}(J)$ and $g^{\mathbb{Q}}(J)$;*

(A2) *Jump risk is made small:* $\quad \forall\, S_t > 0\,, \quad \displaystyle\int_0^\infty \Big[\Delta H_J(S_t,t)\Big]^2 W(J)\,dJ \;<\; \epsilon\,;$

(A3) *Transaction cost is made small:* $\quad \forall\, S_t > 0\,, \quad \Big(\vartheta_t(S_t)\Big)^2 <\; \varrho^2\,\epsilon, \quad$ *where* $\varrho > 0\,;$

(A4) *The second moment of $\tilde{\Pi}_t$ exists:* $\quad \mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big] \;<\; \infty\,;$

(A5) *Delta neutrality is imposed.*

Then

$$\left|\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2(JS_t) - \tilde{\Pi}_t^2(S_t)\Big]\right| \;<\; \epsilon \;+\; 2\sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]}\,, \tag{6.19}$$

$$\left|\mathbb{E}_0^{\mathbb{P}}\Big[\Theta_t^{\mathbb{Q}}\,\tilde{\Pi}_t\Big]\right| \;<\; \sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]}\,, \tag{6.20}$$

and

$$\left|\mathbb{E}_0^{\mathbb{P}}\Big[\vartheta_t\,\tilde{\Pi}_t\Big]\right| \;<\; \varrho\sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]}\,. \tag{6.21}$$

Remark 6.2. *Conditions (A2) and (A3) imply that, for time t, we have a well-defined procedure for rebalancing the hedge portfolio (based on S_t and the existing weights) such that our measure of jump risk and the transaction cost can be made arbitrarily small.*

Remark 6.3. *The expression $\mathbb{E}_0^{\mathbb{P}}\big[\tilde{\Pi}_t^2\big]$ represents the expected squared value of the discounted overall hedged position at time t. The expression $\mathbb{E}_0^{\mathbb{P}}\big[\tilde{\Pi}_t^2(JS_t) - \tilde{\Pi}_t^2(S_t)\big]$ represents the expected change in the value of $\tilde{\Pi}_t^2$ assuming a jump occurs at time t. Note that both of these expectations are taken at time zero, such that the only information known is S_0.*

Proof of Lemma 6.1. The value of the discounted overall hedged position at time t (after rebalancing) is

$$\tilde{\Pi}_t(S_t) \;=\; \mathrm{e}^{-rt}\Big(-V(S_t,t) \;+\; eS_t \;+\; \vec{\phi}\cdot\vec{I}(S_t,t) \;+\; B(t)\Big)\,, \tag{6.22}$$

while after a jump at time t it has value

$$\tilde{\Pi}_t(JS_t) \;=\; \mathrm{e}^{-rt}\Big(-V(JS_t,t) \;+\; eJS_t \;+\; \vec{\phi}\cdot\vec{I}(JS_t,t) \;+\; B(t)\Big)\,. \tag{6.23}$$

Solving for the bank account in (6.22) and substituting into (6.23) yields

$$\tilde{\Pi}_t(JS_t) \;=\; \tilde{\Pi}_t(S_t) \;+\; \mathrm{e}^{-rt}\Delta H_J(S_t,t)\,,$$

where the definition of $\Delta H_J(S_t,t)$ in (6.14) is used. Therefore, we have

$$\begin{aligned}\left|\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t^2(JS_t)-\tilde{\Pi}_t^2(S_t)\right]\right| &= \left|\mathbb{E}_0^{\mathbb{P}}\left[\left(\tilde{\Pi}_t(S_t)+\mathrm{e}^{-rt}\Delta H_J(S_t,t)\right)^2 \;-\; \tilde{\Pi}_t^2(S_t)\right]\right| \\ &= \left|\mathbb{E}_0^{\mathbb{P}}\left[\left(\mathrm{e}^{-rt}\Delta H_J(S_t,t)\right)^2 \;+\; 2\mathrm{e}^{-rt}\tilde{\Pi}_t(S_t)\,\Delta H_J(S_t,t)\right]\right| \\ &\leq \mathbb{E}_0^{\mathbb{P}}\left[\left(\Delta H_J(S_t,t)\right)^2\right] \;+\; 2\left|\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t(S_t)\,\Delta H_J(S_t,t)\right]\right|. \end{aligned} \tag{6.24}$$

Consider the first expectation on the right-hand side of (6.24), which only depends on the asset price at time t. By conditioning on S_t:

$$\begin{aligned}\mathbb{E}_0^{\mathbb{P}}\left[\left(\Delta H_J(S_t,t)\right)^2\right] &= \int_0^\infty \left[\int_0^\infty \left(\Delta H_J(S_t,t)\right)^2 g^{\mathbb{P}}(J)\,dJ\right] p(S_t|S_0)\,dS_t \\ &\leq \int_0^\infty \left[\int_0^\infty \left(\Delta H_J(S_t,t)\right)^2 W(J)\,dJ\right] p(S_t|S_0)\,dS_t \qquad \big(\text{by } (A1)\big) \\ &< \int_0^\infty \epsilon\, p(S_t|S_0)\,dS_t \qquad \big(\text{by } (A2)\big) \\ &= \epsilon\,, \end{aligned} \tag{6.25}$$

where $p(S_t|S_0)$ is the transition density under $\mathbb{P}$. For the second expectation on the right-hand side of (6.24), since $\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t^2(S_t)\right] < \infty$ by $(A4)$ and $\mathbb{E}_0^{\mathbb{P}}\left[\left(\Delta H_J(S_t,t)\right)^2\right] < \epsilon$ by (6.25), the Cauchy–Schwarz inequality gives

$$\begin{aligned}\left|\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t(S_t)\,\Delta H_J(S_t,t)\right]\right| &\leq \sqrt{\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t^2\right]\mathbb{E}_0^{\mathbb{P}}\left[\left(\Delta H_J(S_t,t)\right)^2\right]} \\ &< \sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t^2\right]}\,. \end{aligned} \tag{6.26}$$

Using the upper bounds (6.25) and (6.26) in (6.24) yields

$$\left|\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2(JS_t) - \tilde{\Pi}_t^2(S_t)\Big]\right| \;<\; \epsilon \;+\; 2\sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]}\,,$$

which is the first implication (6.19) of this Lemma.

For the second part of the Lemma, consider

$$\begin{aligned}
\mathbb{E}_0^{\mathbb{P}}\Big[\big(\Theta_t^{\mathbb{Q}}\big)^2\Big] &= \mathbb{E}_0^{\mathbb{P}}\Big[\Big(\mathbb{E}_t^{\mathbb{Q}}\big[-\Delta H_J(S_t,t)\big]\Big)^2\Big] && \big(\text{by def. of } \Theta_t^{\mathbb{Q}} \text{ in (6.17)}\big)\\
&\le \mathbb{E}_0^{\mathbb{P}}\Big[\mathbb{E}_t^{\mathbb{Q}}\Big[\big(\Delta H_J(S_t,t)\big)^2\Big]\Big] && \big(\text{since } \mathbb{E}[X]^2 \le \mathbb{E}[X^2]\big)\\
&= \int_0^\infty \left[\int_0^\infty \big(\Delta H_J(S_t,t)\big)^2 g^{\mathbb{Q}}(J)\,dJ\right] p(S_t|S_0)\,dS_t\\
&\le \int_0^\infty \left[\int_0^\infty \big(\Delta H_J(S_t,t)\big)^2 W(J)\,dJ\right] p(S_t|S_0)\,dS_t && \big(\text{by } (A1)\big)\\
&< \int_0^\infty \epsilon\, p(S_t|S_0)\,dS_t && \big(\text{by } (A2)\big)\\
&= \epsilon\,. && (6.27)
\end{aligned}$$

As such, $\mathbb{E}_0^{\mathbb{P}}\Big[\big(\Theta_t^{\mathbb{Q}}\big)^2\Big]$ exists and, since $\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big] < \infty$ by $(A4)$,

$$\begin{aligned}
\left|\mathbb{E}_0^{\mathbb{P}}\Big[\Theta_t^{\mathbb{Q}}\,\tilde{\Pi}_t\Big]\right| &\le \sqrt{\mathbb{E}_0^{\mathbb{P}}\Big[\big(\Theta_t^{\mathbb{Q}}\big)^2\Big]\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]}\\
&< \sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]}\,,
\end{aligned}$$

where the Cauchy–Schwarz inequality and the bound in (6.27) are employed.

Finally, condition $(A3)$ guarantees that ϑ_t^2 is always bounded by $\varrho^2\,\epsilon$, which means $\mathbb{E}_0^{\mathbb{P}}\Big[\big(\vartheta_t\big)^2\Big] \;<\; \varrho^2\,\epsilon$. Similar to the above, an application of the Cauchy–Schwarz inequality gives the result (6.21). □

Before proving our main theorem, we need the following result.

Theorem 6.1. *Assume that both $f(t)$ and $g(t)$ are continuous, and satisfy*

$$\begin{aligned} f'(t) &< A\sqrt{f(t)} + B \\ g'(t) &= A\sqrt{g(t)} + B \end{aligned}$$

for all $t \geq 0$, with A and B positive constants. If $f(0) = g(0) = 0$, then $g(t) > f(t)$ for all $t > 0$.

Proof of Theorem 6.1. First, note that derivatives at $t = 0$ correspond to the right-hand (i.e. a one-sided) derivative. Due to the fact $g'(0) = B > f'(0)$, we have that $g'(0) - f'(0) > 0$. Define $h(t) = g(t) - f(t)$, which means $h(0) = 0$. Since $h'(t) = g'(t) - f'(t)$ implies $h'(0) = g'(0) - f'(0) > 0$, the definition of the one-sided derivative yields $g(t) > f(t)$ for some domain $(0, t_H)$, where $t_H > 0$.

Now, assume that, for some $\hat{t}$, $f(\hat{t}) > g(\hat{t})$. This means there exists a $t^* \in [t_H, \hat{t})$ such that $f(t^*) = g(t^*)$ and $f'(t^*) > g'(t^*)$. However, from the assumptions of the theorem, we have

$$f'(t^*) \;<\; A\sqrt{f(t^*)} + B \;=\; A\sqrt{g(t^*)} + B \;=\; g'(t^*)\,,$$

which leads to a contradiction. The case $f(\hat{t}) = g(\hat{t})$ is handled similarly. This implies $t_H = \infty$, and hence the theorem is proven. □

We are now in a position to prove the main result of this section.

Theorem 6.2 (Variance of the Hedging Error Can be Made Arbitrarily Small). *If, for all times t in the investment period $[0, T]$ the conditions $(A1) - (A5)$ of Lemma 6.1 hold and $\mathbb{E}_0^{\mathbb{P}}\left[\Pi_t^2\right]$ is assumed continuous, then*

$$\mathbb{E}_0^{\mathbb{P}}\left[\Pi_T^2\right] \;<\; \epsilon\, \frac{e^{2rT}\,(\lambda^{\mathbb{P}})^2}{4\,(\lambda^{\mathbb{P}} + \lambda^{\mathbb{Q}} + \varrho)^2} \left[W_{-1}\left(-\exp\left\{-\frac{2(\lambda^{\mathbb{P}} + \lambda^{\mathbb{Q}} + \varrho)^2}{\lambda^{\mathbb{P}}} T \;-\; 1\right\}\right) + 1\right]^2 , \tag{6.28}$$

where $W_{-1}(x)$ is the -1 branch of the Lambert W function [27].[2] *That is, $\mathbb{E}_0^{\mathbb{P}}\Big[\Pi_T^2\Big] \leq \mathcal{C}\epsilon$, where $\mathcal{C}$ is a constant that only depends on $\lambda^{\mathbb{P}}$, $\lambda^{\mathbb{Q}}$, ϱ, r, and T.*

Proof of Theorem 6.2. The stochastic differential equation (6.18) may be used to derive an ODE relating the moments of $\tilde{\Pi}_t$, namely

$$\frac{d\,\mathbb{E}_0^{\mathbb{P}}[\tilde{\Pi}_t^2]}{dt} \;=\; 2\mathrm{e}^{-rt}\lambda^{\mathbb{Q}}\,\mathbb{E}_0^{\mathbb{P}}\Big[\Theta_t^{\mathbb{Q}}\,\tilde{\Pi}_t\Big] \;-\; 2\mathrm{e}^{-rt}\,\mathbb{E}_0^{\mathbb{P}}\Big[\vartheta_t\,\tilde{\Pi}_t\Big] \;+\; \lambda^{\mathbb{P}}\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2(JS_t) - \tilde{\Pi}_t^2(S_t)\Big] \tag{6.29}$$

with initial condition $\mathbb{E}_0^{\mathbb{P}}\big[\tilde{\Pi}_0^2\big] = 0$. Equation (6.29) yields

$$\frac{d\,\mathbb{E}_0^{\mathbb{P}}[\tilde{\Pi}_t^2]}{dt} \;\leq\; 2\lambda^{\mathbb{Q}}\left|\mathbb{E}_0^{\mathbb{P}}\Big[\Theta_t^{\mathbb{Q}}\,\tilde{\Pi}_t\Big]\right| \;+\; 2\left|\mathbb{E}_0^{\mathbb{P}}\Big[\vartheta_t\,\tilde{\Pi}_t\Big]\right| \;+\; \lambda^{\mathbb{P}}\left|\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2(JS_t) - \tilde{\Pi}_t^2(S_t)\Big]\right|,$$

which allows us to use the bounds established in Lemma 6.1 to set up the differential inequality

$$\frac{d\,\mathbb{E}_0^{\mathbb{P}}[\tilde{\Pi}_t^2]}{dt} \;<\; 2\lambda^{\mathbb{Q}}\sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]} \;+\; 2\varrho\sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]} \;+\; \lambda^{\mathbb{P}}\left(\epsilon + 2\sqrt{\epsilon\,\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]}\right);$$

the above may be written more succinctly as

$$\frac{d\,\mathbb{E}_0^{\mathbb{P}}[\tilde{\Pi}_t^2]}{dt} \;<\; 2\sqrt{\epsilon}\big(\lambda^{\mathbb{P}} + \lambda^{\mathbb{Q}} + \varrho\big)\sqrt{\mathbb{E}_0^{\mathbb{P}}\Big[\tilde{\Pi}_t^2\Big]} \;+\; \epsilon\,\lambda^{\mathbb{P}}\,. \tag{6.30}$$

We define a *bounding function* $\beta(t)$, which satisfies the ODE

$$\frac{d\beta}{dt} \;=\; 2\sqrt{\epsilon}\big(\lambda^{\mathbb{P}} + \lambda^{\mathbb{Q}} + \varrho\big)\sqrt{\beta(t)} \;+\; \epsilon\,\lambda^{\mathbb{P}} \tag{6.31}$$

[2]The Lambert W function is defined as the inverse of $y = x\mathrm{e}^x$. For any $y \in (-\mathrm{e}^{-1}, 0)$, there are two real values x^* such that $y = x^*\mathrm{e}^{x^*}$, with one root $x_1^* \in (-1, 0)$ and the other root $x_2^* \in (-\infty, -1)$. The branch W_{-1} corresponds to those x^* that are less than or equal to -1. Therefore, $W_{-1}(-\mathrm{e}^{-1}) = -1$ and $W_{-1}(0) = -\infty$.

with initial condition $\beta(0) = 0$. This initial value problem has the exact solution

$$\beta(t) \;=\; \epsilon\frac{(\lambda^{\mathbb{P}})^2}{4\left(\lambda^{\mathbb{P}}+\lambda^{\mathbb{Q}}+\varrho\right)^2}\left[W_{-1}\left(-\exp\left\{-\frac{2(\lambda^{\mathbb{P}}+\lambda^{\mathbb{Q}}+\varrho)^2}{\lambda^{\mathbb{P}}}t\;-\;1\right\}\right)+1\right]^2, \tag{6.32}$$

where $W_{-1}(x)$ is the -1 branch of the Lambert W function. Consider the relationship between $\beta(t)$ and $\mathbb{E}_0^{\mathbb{P}}\big[\tilde{\Pi}_t^2\big]$. Both are non-negative quantities with an initial value of zero. Theorem 6.1 implies $\mathbb{E}_0^{\mathbb{P}}[\tilde{\Pi}_t^2] \;<\; \beta(t)$, which means

$$\mathbb{E}_0^{\mathbb{P}}\big[\tilde{\Pi}_t^2\big] \;<\; \epsilon\frac{(\lambda^{\mathbb{P}})^2}{4\left(\lambda^{\mathbb{P}}+\lambda^{\mathbb{Q}}+\varrho\right)^2}\left[W_{-1}\left(-\exp\left\{-\frac{2(\lambda^{\mathbb{P}}+\lambda^{\mathbb{Q}}+\varrho)^2}{\lambda^{\mathbb{P}}}t\;-\;1\right\}\right)+1\right]^2.$$

Using the fact that $\Pi_t = \mathrm{e}^{rt}\tilde{\Pi}_t$ and considering the above at $t = T$ yields the bound on $\mathbb{E}_0^{\mathbb{P}}[\Pi_T^2]$ in (6.28). □

6.3 A Discrete Hedging Strategy

Theorem 6.2 formalizes the idea that, in order to make the terminal hedging error small, the jump risk and transaction cost should be made small at each instant of a continuously rebalanced hedge (assuming that the instantaneous diffusion risk is eliminated by imposing delta neutrality). In practice, of course, a hedge cannot be continuously rebalanced—the rebalancing interval Δt will be of non-infinitesimal length. Nonetheless, our discrete hedging strategy will treat

$$\int_0^\infty \Big[\Delta H_J(S_t,t)\Big]^2 W(J)\,dJ \tag{6.33}$$

as the measure of jump risk that should be made as small as possible when the hedge is rebalanced. Recall that Υ_t^k is the transaction cost for rebalancing the k^{th} instrument at

time t. Therefore, by minimizing

$$\sum_{k=1}^{N} \left(\Upsilon_t^k \right)^2 , \tag{6.34}$$

the total transaction cost (6.11) due to rebalancing at time t can be made small. We use the 'sum of squares' representation in (6.34) for computational reasons, as the expression can be minimized by the method of Lagrange multipliers.

Minimizing jump risk and reducing transaction costs are competing goals, so this problem falls under the rubric of multi-objective optimization. One typical way of handling such a problem is to weight the objectives by a set of coefficients that sum to unity. If ξ is the weight on the jump risk exposure (6.33) and $1 - \xi$ is the weight on the transaction cost component (6.34), then the objective we minimize at each rebalance time is

$$\xi \left\{ \int_0^\infty \left[\Delta H_J(S_t, t) \right]^2 W(J)\, dJ \right\} \quad + \quad (1 - \xi) \sum_{k=1}^{N} \left(\Upsilon_t^k \right)^2 , \tag{6.35}$$

where $\xi \in [0, 1]$. The instantaneous diffusion risk is eliminated by simultaneously imposing delta neutrality via a linear constraint.

The following Lemma relates a bound on the objective (6.35) to bounds required to establish the continuous-time result in Theorem 6.2 for $\mathbb{E}^{\mathbb{P}}[\Pi_T^2]$, i.e. when $\Delta t \to 0$.

Lemma 6.2 (Bounds Dictated by the Hedging Objective). *Assume that, for the current time t and asset price S_t, the condition*

$$\xi \left\{ \int_0^\infty \left[\Delta H_J(S_t, t) \right]^2 W(J)\, dJ \right\} \; + \; (1 - \xi) \sum_{k=1}^{N} \left(\Upsilon_t^k \right)^2 \quad < \quad \epsilon^* \, \Delta t^2 \tag{6.36}$$

holds for some $\xi \in (0, 1)$. In that case,

$$\int_0^\infty \Big[\Delta H_J(S_t,t)\Big]^2 W(J)\,dJ \;<\; \frac{\epsilon^*\,\Delta t^2}{\xi} \tag{6.37}$$

and

$$\Big(\Upsilon_t\Big)^2 \;=\; \left[\sum_{k=1}^{N}\Upsilon_t^k\right]^2 \;<\; \frac{\epsilon^*\,\Delta t^2}{1-\xi}N\,, \tag{6.38}$$

where Υ_t *the total transaction cost due to rebalancing at time* t*.*

Proof of Lemma 6.2. The first result (6.37) is a simple consequence of the fact that both the jump risk and transaction cost component of the objective are non-negative. Similarly

$$\sum_{k=1}^{N}\Big(\Upsilon_t^k\Big)^2 \;<\; \frac{\epsilon^*\,\Delta t^2}{1-\xi}\,.$$

Now consider

$$\begin{aligned}
\left[\sum_{k=1}^{N}\Upsilon_t^k\right]^2 &= \sum_{l=1}^{N}\sum_{m=1}^{N}\Upsilon_t^l\Upsilon_t^m \\
&\le \sum_{l=1}^{N}\sum_{m=1}^{N}\frac{\big(\Upsilon_t^l\big)^2+\big(\Upsilon_t^m\big)^2}{2} \\
&= \sum_{l=1}^{N}\sum_{m=1}^{N}\frac{\big(\Upsilon_t^l\big)^2}{2} \;+\; \sum_{l=1}^{N}\sum_{m=1}^{N}\frac{\big(\Upsilon_t^m\big)^2}{2} \\
&= \frac{N}{2}\sum_{l=1}^{N}\Big(\Upsilon_t^l\Big)^2 \;+\; \frac{N}{2}\sum_{m=1}^{N}\big(\Upsilon_t^m\big)^2 \\
&\le \frac{N}{2}\frac{\epsilon^*\,\Delta t^2}{1-\xi} \;+\; \frac{N}{2}\frac{\epsilon^*\,\Delta t^2}{1-\xi} \\
&= \frac{\epsilon^*\,\Delta t^2}{1-\xi}N\,,
\end{aligned}$$

which is the desired result in (6.38). □

Consider an idealized trading environment: by making the objective (6.35) arbitrarily small (in a manner to be described below) for a discretely rebalanced hedge, the variance of the terminal hedging error will be small as $\Delta t \to 0$. To this end assume that, given ϵ, ϱ and Δt, the trading environment allows ϵ^*, C, N and ξ to be chosen such that:

1. $\epsilon^* = C\epsilon$ in (6.36) at any rebalance time t_r and for any S_{t_r}, where N is the maximum number of hedging instruments required to do this;

2. $\xi = C\Delta t^2$.

In that case, the results (6.37) and (6.38) of Lemma 6.2 imply that

$$\int_0^\infty \Big[\Delta H_J(S_t, t)\Big]^2 W(J)\, dJ \;<\; \epsilon \tag{6.39}$$

and

$$\begin{aligned}\Big(\Upsilon_t\Big)^2 \;&<\; N\, C\, \epsilon\, \Delta t^2 \;+\; \mathcal{O}(\Delta t^4) \\ &<\; \varrho^2\, \epsilon\, \Delta t^2 \;+\; \mathcal{O}(\Delta t^4)\end{aligned}$$

where $\rho = \sqrt{NC}$, for any rebalance time t_r with any S_{t_r}.

This process can be carried out for any $\Delta t < T$. Let $\Upsilon_t = \vartheta_t\, \Delta t$, so that

$$\vartheta_t^2 \;<\; \varrho^2\, \epsilon \;+\; \mathcal{O}(\Delta t^2)\,. \tag{6.40}$$

In the limit as $\Delta t \to 0$, conditions $(A2)$ and $(A3)$ of Theorem 6.2 are satisfied from (6.39) and (6.40), respectively, and we obtain the bound (6.28) on $\mathbb{E}^{\mathbb{P}}\big[\Pi_T^2\big]$ (assuming, of course, that $(A1)$, $(A4)$ and $(A5)$ in Theorem 6.2 also hold). Hence, in the limit, our discrete

hedging strategy will yield a terminal hedging error whose variance can be made arbitrarily small.

Remark 6.4 (Hedging in Practice). *It is always possible to satisfy the transaction cost objective (A3) by trading only in the underlying, although in this case the jump risk is expected to be substantial. Indeed, we may not be able to make both the jump risk and transaction cost small in the manner required by* (6.36)*. As a practical matter, at each rebalance time we simply attempt to make the objective* (6.35) *as small as possible for a given* ξ*.*

Remark 6.5 (Choice of ξ). *If we fix* ξ *and minimize the objective* (6.35)*, the above analysis implies that the best choice of* ξ *should be* $\mathcal{O}(\Delta t^2)$*.*

Remark 6.6 (The Form of the Transaction Cost Penalty Function). *If the transaction cost penalty function in the objective function* (6.35) *is of the form*

$$\left[\sum_{k=1}^{N} \Upsilon_t^k\right]^2,$$

then ρ *in* (6.40) *will be independent of* N*. However, this penalty function is not amenable to minimization via Lagrange multipliers. In any event, since in practice rebalancing is done discretely and only a small number of hedging instruments are used, the theory of this chapter provides a justification for our real-world strategy.*

6.4 Summary

In this chapter, we derived a representation of the instantaneous jump risk in a jump-diffusion model, and defined an integral which is a measure of the exposure to this risk. We showed that, if this measure of jump risk and the transaction cost can be made sufficiently small at each instant of a continuously rebalanced, delta-neutral hedge, then the terminal

hedging error can be made arbitrarily small. This analysis led to an objective function, which is the weighted sum of a jump risk and transaction cost penalty function, that can be minimized at each rebalancing time of discretely rebalanced hedge. We also demonstrated that, in an idealized trading environment, the hedging error from this discrete strategy can be made arbitrarily small as the rebalancing interval goes to zero.

Chapter 7

Dynamic Hedging under Jump Diffusion: No Transaction Costs

In order to focus on the effect of jumps when hedging, this chapter considers a jump-diffusion market in which there are no transaction costs. Therefore, the objective minimized at each rebalance time consists simply of a measure of the jump risk. The analysis of Chapter 6 implies that, if the jump risk at each rebalance time is made small, then the global hedging error should also be small. We will examine the particulars of the optimization problem in this case, including the selection of a weighting function and the mathematical details of its solution. Hedging simulations of the discrete strategy are then carried out for both a European and American target option. In reality, a hedger will not have exact knowledge of the $\mathbb{P}$ and $\mathbb{Q}$ measures that govern the market, so we investigate the effect of using estimates of these measures.

7.1 Minimizing Jump Risk

Consider the writer of an option who wants to use the hedge portfolio

$$\mathcal{H} \;=\; eS \;+\; \vec{\phi}\cdot\vec{I} \;+\; B$$

to insulate their position from diffusion and jump risk over the next instant. If the jump amplitudes are drawn from a finite set of size M, the instantaneous jump risk can be eliminated by introducing M hedging instruments into $\vec{I}$. The linear system

$$\begin{aligned} -\Big[V(J_iS,t)-V(S,t)\Big] \;+\; eS\Big[J_i-1\Big] \;+\; \vec{\phi}\cdot\Big[\vec{I}(J_iS,t)-\vec{I}(S,t)\Big] \;&=\; 0 \qquad i=1,\ldots,M \\ -\frac{\partial V}{\partial S} \;+\; e \;+\; \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S} \;&=\; 0 \end{aligned} \tag{7.1}$$

ensures that, for the current time t and asset price S, the diffusion risk is removed and the overall hedged position is invariant to jumps of size $J^* \in \{J_1, J_2, \ldots, J_M\}$.

Since it is most often assumed that jumps are drawn from a continuum, the goal when rebalancing the hedge is to minimize the jump risk using a practical number of instruments, while simultaneously eliminating the diffusion risk. Therefore, at each rebalance time, we compute the hedge weights $\{e^*, \vec{\phi}^*\}$ that solve the constrained optimization

$$\min_{\{e,\vec{\phi}\}} \int_0^\infty \Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]^2 W(J)\,dJ$$

subject to

$$e \;+\; \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S} \;=\; \frac{\partial V}{\partial S}\,. \tag{7.2}$$

Other constraints may also be imposed. For example, since the hedge will in practice be

rebalanced discretely, the gamma-neutral constraint

$$\vec{\phi} \cdot \frac{\partial^2 \vec{I}}{\partial S^2} = \frac{\partial^2 V}{\partial S^2} \tag{7.3}$$

may be added to hedge the higher orders of diffusion risk.

The weighting function $W(J)$ in the objective of (7.2) is set by the hedger. One possible choice corresponds to the distribution $g^{\mathbb{P}}(J)$ of jump amplitudes observed in the market, since in this case the optimization problem (7.2) is a local variance minimization of $d\Pi$. In a model with lognormally distributed jumps, for example, one would use the probability density function associated with a lognormal$(\mu^{\mathbb{P}}, \gamma^{\mathbb{P}})$ random variable. However, since this requires knowledge of the $\mathbb{P}$ measure, the parameters would often have to be approximated.

In order to ensure that the jump risk can become small through bounding the integral in (7.2), $W(J)$ must be a proper weighting function as described in Definition 6.1. However, since only a small number of hedging instruments are used in practice, $W(J)$ need not be a proper weighting function. Therefore a more practical weighting function would be a uniform density, set to non-zero for the range of jumps deemed likely.

The numerical examples of this chapter will at times employ the uniform-like weighting function plotted in Figure 7.1, which is constant between $\frac{1}{5} \leq J \leq \frac{9}{5}$ and extends linearly down to zero outside this range on either side. The triangular tails ensure that the weighting function is continuous, which results in better numerical behaviour within the FFT procedure we use to compute the integrals necessary for solving the optimization (this procedure is discussed in Appendix B). This weighting function encapsulates a lack of knowledge pertaining to the jumps under the $\mathbb{P}$ measure: with no information about which jump sizes are more likely than others, broad protection is sought for $J \in [0, 2]$. In general, when a uniform-like weighting function with support $[J_{\min}, J_{\max}]$ is used, $J_{\min}$ should be chosen close to zero in order to protect against all downward jumps. Selecting $J_{\max}$ is a

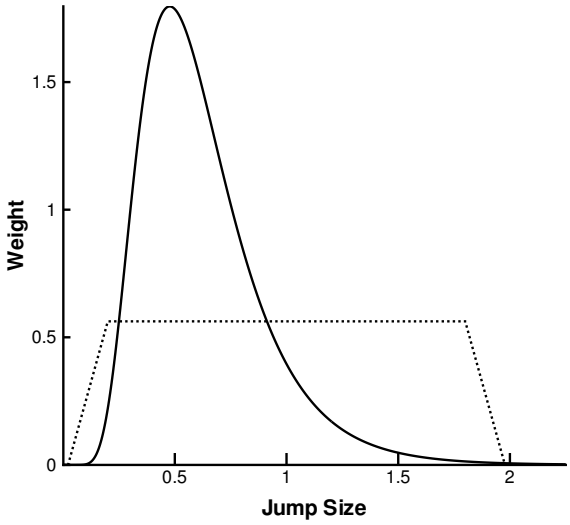


FIGURE 7.1: *A lognormal and uniform-like weighting function. Both functions have a unit mass.*

bit more difficult. The uniform-like weighting function of Figure 7.1 implies jumps of size $J > 2$ will not be taken into consideration when the hedge portfolio is rebalanced—if it is suspected that $g^{\mathbb{P}}(J)$ has a slowly decaying right tail, a higher value of $J_{\max}$ may be needed.

The objective in (7.2) may be expressed as

$$\min_{\{e,\vec{\phi}\}} \underbrace{\int_0^\infty \left[\frac{1}{2}\left(-\Delta V + \left\{e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right\}\right)^2\right] W(J)\,dJ}_{F(e,\vec{\phi})} . \tag{7.4}$$

The objective function is quadratic in the unknowns e and $\vec{\phi}$; as such, optimality requires

$$\begin{aligned}
\frac{\partial F}{\partial e} &= \int_0^\infty \left[\left(-\Delta V + \left\{e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right\}\right)(\Delta S)\right] W(J)\,dJ = 0 \\
\frac{\partial F}{\partial \phi_k} &= \int_0^\infty \left[\left(-\Delta V + \left\{e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right\}\right)(\Delta I_k)\right] W(J)\,dJ = 0 \qquad \text{for } k = 1,\ldots,N\,,
\end{aligned} \tag{7.5}$$

where N is the number of hedging instruments in $\vec{I}$. Any desired linear equality constraints, such as delta neutrality, may be included via Lagrange multipliers. Therefore, the optimal solution $\{e^*, \vec{\phi}^*\}$ to the optimization problem (7.2) is found by solving the linear system

$$\begin{bmatrix} (\Delta S, \Delta S) & (\Delta S, \Delta I_1) & \cdots & (\Delta S, \Delta I_N) & 1 \\ (\Delta I_1, \Delta S) & (\Delta I_1, \Delta I_1) & \cdots & (\Delta I_1, \Delta I_N) & \frac{\partial I_1}{\partial S} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (\Delta I_N, \Delta S) & (\Delta I_N, \Delta I_1) & \cdots & (\Delta I_N, \Delta I_N) & \frac{\partial I_N}{\partial S} \\ 1 & \frac{\partial I_1}{\partial S} & \cdots & \frac{\partial I_N}{\partial S} & 0 \end{bmatrix} \begin{bmatrix} e^* \\ \phi_1^* \\ \vdots \\ \phi_N^* \\ \Psi \end{bmatrix} = \begin{bmatrix} (\Delta V, \Delta S) \\ (\Delta V, \Delta I_1) \\ \vdots \\ (\Delta V, \Delta I_N) \\ \frac{\partial V}{\partial S} \end{bmatrix}, \tag{7.6}$$

where Ψ is a Lagrange multiplier associated with the delta-neutral constraint.

The linear system entries of the form $(\Delta X, \Delta Y)$ arise from the optimality conditions in (7.5), and can be expressed as

$$\begin{aligned} (\Delta X, \Delta Y) &= \int_0^\infty \big[\Delta X \Delta Y\big] W(J)\, dJ \\ &= \int_0^\infty \Big[X(JS,t) - X(S,t)\Big]\Big[Y(JS,t) - Y(S,t)\Big] W(J)\, dJ \\ &= X(S,t)Y(S,t)\int_0^\infty W(J)\, dJ \;+\; \int_0^\infty X(JS,t)Y(JS,t)W(J)\, dJ \\ &\quad -\; Y(S,t)\int_0^\infty X(JS,t)W(J)\, dJ \;-\; X(S,t)\int_0^\infty Y(JS,t)W(J)\, dJ\,. \end{aligned} \tag{7.7}$$

Within their numerical solution of the option-pricing PIDE (2.9), d'Halluin et al. [34] use the FFT to efficiently compute the integral term. Appendix B gives an overview of the numerical solution to the PIDE [34], and discusses how the integrals in (7.7) can be computed using the same FFT techniques. The linear system (7.6) may be poorly conditioned in certain situations so, as described in Section 3.1, we use a Truncated Singular Value Decomposition

(TSVD) [38] to solve it.

Consider the following scenario: An option writer, who has just sold a one-year at-the-money American straddle[1] with a strike of \$100, wants to form a hedge portfolio that eliminates the instantaneous diffusion risk and reduces her exposure to jump risk. The option writer has access to a range of 90-day European call options, with strikes in increments of \$10, which may be used in $\vec{I}$. The uniform-like weighting function plotted in Figure 7.1 is used for $W(J)$, and the optimization (7.2) is solved for different compositions of $\vec{I}$. The left panel of Figure 7.2 presents the jump risk profile ΔH_J, defined in (6.14), for the optimized hedge portfolios resulting from various hedging strategies. These curves represent the change in the value of the overall hedged position as a function of jump size.

There is a large potential loss to a portfolio that ignores jumps, which is evident in the *Delta Hedge Only* curve. This is due to the fact that a delta hedge short a convex option will always suffer a loss when a jump occurs [59]. The intercepts where the curves cross the zero line are fortuitous points, as there is no change in the overall hedged position for jumps corresponding to these amplitudes. As more hedging instruments are included, the curves more closely hug the zero line and the number of intercepts increases. In fact, for the case of ten hedging options, the curve remains so close to zero that its structure is not clearly seen; an enlargement is provided in the right panel of Figure 7.2.

Minimizing the jump risk while simultaneously eliminating the instantaneous diffusion risk is the basis of our hedging strategy. At each rebalance time, the hedge weights are determined by solving the optimization problem (7.2), which itself entails solving the linear system (7.6). The integrals necessary to form this linear system can be precomputed and, during the simulations, the necessary quantities are found by table lookup and interpolation. We refer the reader to Chapter 3 where these computational details are discussed. Before presenting some hedging examples, we examine the existing studies related to dynamic

[1] A straddle is the combination of a call and put of the same strike.

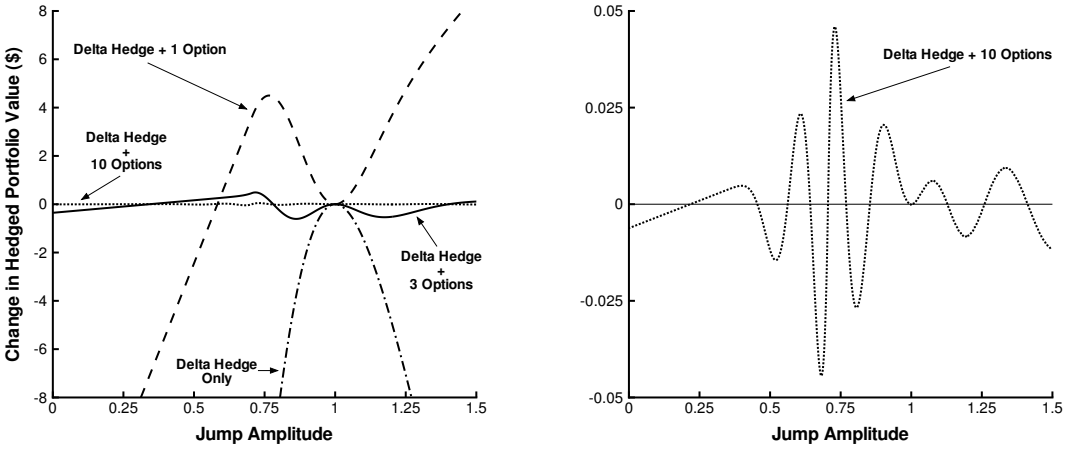


FIGURE 7.2: *Change in the value of the overall hedged position resulting from a jump. The left panel shows the performance of several hedging strategies, while the right panel is an enlargement of the curve when ten options are used to hedge. The target option is a one-year American straddle with strike \$100. The jumps are lognormally distributed, with financial parameters* $r = 0.05$, $\sigma = 0.2$, $\lambda^{\mathbb{Q}} = 0.1$, $\mu^{\mathbb{Q}} = -0.92$, *and* $\gamma^{\mathbb{Q}} = 0.425$. *Strikes of the European calls used to hedge:* $K = 100$, $K = [90, 100, 110]$, $K = [50, 60, 70, 80, 90, 100, 110, 120, 130, 140]$.

hedging under jump diffusion.

7.2 Dynamic Hedging: Relationship to Previous Work

Merton [56] considered the characteristics of a simple delta hedge within the context of the jump-diffusion model. Since this strategy eliminates the diffusion risk, the return process for the overall position is a pure jump process with drift. Consequently, most of the time the change in the portfolio is small, but when a jump occurs, the variation can be quite large. With Merton's assumption of diversifiable jump risk, the jump processes under $\mathbb{P}$ and $\mathbb{Q}$ are the same, so the real-world expected value of the instantaneous jump risk for the delta hedge is zero: the many small changes and the infrequent large changes will, on average, balance each other over a long time period. In this case, as noted in [23], the delta hedge accounts for the average effect of jumps. However, in general the jump processes

under $\mathbb{P}$ and $\mathbb{Q}$ will not be the same, such that the mean of the delta hedge is expected to be non-zero.

A delta hedge based on option values found by ignoring jumps (i.e. using the Black–Scholes formula) was studied by Naik and Lee [59] in a model with lognormally distributed amplitudes, $\log(J) \sim N(\mu, \gamma)$. The volatility used for pricing in the Black–Scholes formula is initially set to σ, which comes from the diffusive component of the jump-diffusion process. If no jump event occurs over the investment period, the continuously rebalanced delta hedge based on these values is perfect, since the market behaves as if it exists in the Black–Scholes universe. However, a jump will lead to failure of the hedge as cash inflow/outflow is needed to maintain replication. The deficiencies remain when a higher value of volatility that captures some of the added variability due to jumps is used for pricing.[2]

Delta hedging is clearly insufficient, so other instruments besides the underlying must be used to hedge. The idea of using a finite number of options as part of a dynamic hedging strategy to minimize jump risk has also been suggested by Bates [7] and Andersen and Andreasen [1]. Note that neither of these papers provided any tests of the strategy. In Bates [7], the hedge portfolio is selected so that (for infinitesimal hedging intervals) the diffusion risk is identically zero and the expected value of the local jump risk is minimized. In our notation, this amounts to

$$\{e^*, \vec{\phi}^*\} = \underset{\{e,\vec{\phi}\}}{\arg\min}\ \mathbb{E}\left[-\Delta V + \left(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right)\right]^2 , \tag{7.8}$$

subject to the delta-neutral constraint. Bates [7] suggests that the expectation in equation (7.8) can be taken w.r.t. either $\mathbb{P}$ or $\mathbb{Q}$. A similar procedure is suggested in Andersen and Andreasen [1], but they recommend using $\mathbb{E}^{\mathbb{P}}$ in equation (7.8), along with the addi-

[2]The total volatility of a jump-diffusion process with lognormally distributed jumps is $\sqrt{\sigma^2 + \lambda(\gamma^2 + \mu^2)}$. However, Naik and Lee [59] use $\sqrt{\sigma^2 + \lambda\gamma^2}$ in the Black–Scholes formula to capture the increased variability due to jumps. Although Naik and Lee do not explicitly refer to this quantity as the total volatility, it is a common error in the literature for μ to be omitted from the calculation [60].

tional constraint

$$\mathbb{E}^{\mathbb{P}}\left[-\Delta V + \left(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right)\right] = 0\,. \tag{7.9}$$

Both of these approaches are similar in spirit to our strategy: the jump risk is minimized in some sense, subject to the delta-neutral constraint that eliminates the instantaneous diffusion risk.

Cont et al. [25] consider the quadratic program

$$\underset{\{e_t,\vec{\phi}_t,\nu_0\}}{\arg\min}\ \mathbb{E}^{\mathbb{Q}}\left[\mathcal{H}_T - V_T\right]^2,$$

where $\mathcal{H}_T$ is the terminal value of the hedge portfolio, V_T is the payoff of the option being hedged, and $\{e_t, \vec{\phi}_t\}$ is the trading strategy over $[0,T]$ that minimizes the objective. Also, ν_0 is the initial capital. This program employs a global criterion, as its objective function is an expectation involving the terminal hedging error. However, since this expectation is taken w.r.t. the pricing measure $\mathbb{Q}$, the solution reduces to a local risk minimization [23]. In our notation (assuming finite activity jumps[3]), the optimal hedging weights at each instant are given by

$$\{e^*, \vec{\phi}^*\} \;=\; \underset{\{e,\vec{\phi}\}}{\arg\min}\left[\lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}\left[-\Delta V + \left(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right)\right]^2 + \sigma^2 S^2\left[-\frac{\partial V}{\partial S} + e + \vec{\phi}\cdot\frac{\partial\vec{I}}{\partial S}\right]^2\right], \tag{7.10}$$

and $\nu_0 = V(S_0, 0)$ (the option price) [25]. Our hedging procedure is similar to that of Cont et al. [25], except that in the optimization (7.2) the diffusion risk is explicitly eliminated and a weighting function is employed to minimize the $\mathbb{P}$ measure local risk. If the delta-neutral condition is imposed in (7.10) and the weighting function in (7.2) is set to $g^{\mathbb{Q}}(J)$, then the

[3]With infinite activity jump processes, an infinite number of jumps with infinitesimal size can occur over a finite time period. We will consider hedging under such processes in Chapter 9.

two procedures are identical.

If we minimize the expression in (7.10) then, for infrequent jumps ($\lambda^{\mathbb{Q}} \ll 1$), the diffusion component in (7.10) may swamp the jump component, such that the protection against jumps may be lacking. On the other hand, by explicitly enforcing delta neutrality in (7.2), the jump protection is the same regardless of how often the jump events occur; no estimate of the jump intensity is required. Furthermore, only the weighting function has to be set: we can use $g^{\mathbb{P}}(J)$, $g^{\mathbb{Q}}(J)$, or even a uniform-like density as in Figure 7.1. Conversely, by not explicitly enforcing the delta-neutral constraint in (7.10), there exists an extra degree of freedom, which may be important if using a small number of hedging instruments. For example, if only the underlying is used in the hedge, the optimization (7.2) will trivially yield the delta-neutral position, while the weight in the underlying computed by solving (7.10) will take into account both the diffusion and jump risk.

7.3 Hedging a European Option

7.3.1 The Market Setup

We will use a jump-diffusion model with lognormally distributed jump amplitudes, $\log(J) \sim N(\mu, \gamma)$; the values that characterize the real-world and pricing measures are reported in Table 7.1. For the $\mathbb{Q}$ measure, we use values quite similar to those reported by Andersen and Andreasen [1], which were found by calibrating to observed prices of S&P 500 index options. To obtain the $\mathbb{P}$ measure parameters, we transform the $\mathbb{Q}$ measure using the power

Probability Measure	λ	μ	γ	σ	α
Risk-adjusted ($\mathbb{Q}$)	0.1000	-0.9200	0.4250	0.2000	0.0500
Real-world ($\mathbb{P}$)	0.0228	-0.5588	0.4250	0.2000	0.1779

TABLE 7.1: *The pricing $\mathbb{Q}$ measure and real-world $\mathbb{P}$ measure that characterize the jump-diffusion model, where $\log(J) \sim N(\mu, \gamma)$. The dividend yield $q = 0$ and $\alpha^{\mathbb{Q}} = r = 0.05$.*

utility equilibrium model of Naik and Lee [59] and Bates [8]. The relations are

$$\begin{aligned}
\sigma^{\mathbb{P}} &= \sigma^{\mathbb{Q}} \\
\gamma^{\mathbb{P}} &= \gamma^{\mathbb{Q}} \\
\mu^{\mathbb{P}} &= \mu^{\mathbb{Q}} + (1-\beta)(\gamma^{\mathbb{Q}})^2 \\
\lambda^{\mathbb{P}} &= \lambda^{\mathbb{Q}} \exp\left\{ \big(1-\beta\big)\big(\mu^{\mathbb{Q}} + \frac{1}{2}(1-\beta)(\gamma^{\mathbb{Q}})^2\big) \right\} \\
\alpha^{\mathbb{P}} &= r + (1-\beta)\sigma^2 + (\kappa^{\mathbb{P}}\lambda^{\mathbb{P}} - \kappa^{\mathbb{Q}}\lambda^{\mathbb{Q}}),
\end{aligned} \tag{7.11}$$

where $1-\beta$ is the coefficient of relative risk aversion; the values in Table 7.1 are found using $1-\beta = 2$. This market will experience a jump event, on average, once every 43.9 years with a mean jump size of -37.4%.

We take as the target option a one-year European straddle with strike \$100. Since this option is a combination of a call and put, the payoff is convex and a short position will be highly sensitive to jumps of any direction. The hedging horizon is half a year, and three-month call options with strikes in intervals of \$5 are used as hedging instruments. These options are assumed to exist for $t \in [0.0, 0.25]$ and $t \in [0.25, 0.5]$. When the hedge portfolio is rebalanced, the choice of options is based on liquidity considerations as the active instruments should ideally have strikes in $\pm 10\%$ increments of the underlying's current value. For example, if five options are to be held in the hedge at inception, those with strikes closest to $[0.8S_0, 0.9S_0, 1.0S_0, 1.1S_0, 1.2S_0]$ should be used. At $t = 0.25$, a new collection of three-month options become available.

7.3.2 Description of the Simulation Experiments

The hedge portfolio will contain five hedging options and the underlying. The hedge is rebalanced once every 0.0125 years, a total of forty times over the half-year hedging horizon. A total of 500,000 simulations are carried out: with a mean arrival rate of $\lambda^{\mathbb{P}} = 0.0228$, there are expected to be about 5,700 jump events.

In addition to assessing the efficacy of our hedging strategy, the simulations will also explore the consequences of using an estimated pricing measure $\mathbb{Q}'$. It is standard industry practice to take a set of option prices quoted in the market and "back-out" the pricing parameters for a particular model that yield the best fit to these observed market prices. However, calibrating the pricing measure from market prices in this way is an ill-posed problem for the jump-diffusion model [40, 24]. As such, many different 4-tuples $(\lambda^{\mathbb{Q}'}, \mu^{\mathbb{Q}'}, \gamma^{\mathbb{Q}'}, \sigma')$ exist that, when used for pricing, produce close agreement to the values of vanilla options generated by the $\mathbb{Q}$ measure of Table 7.1. In the experiments to follow, the estimated pricing measure $\mathbb{Q}'$ is set to

$$(\lambda^{\mathbb{Q}'}, \mu^{\mathbb{Q}'}, \gamma^{\mathbb{Q}'}, \sigma') = (0.1077, -0.8639, 0.4906, 0.1991)\,.$$

In He et al. [40], this estimate was obtained by calibrating to a synthetic market whose prices were generated from the $\mathbb{Q}$ measure of Table 7.1. The associated real-world measure $\mathbb{P}'$, found using the relations in (7.11), is $(\lambda^{\mathbb{P}'}, \mu^{\mathbb{P}'}, \gamma^{\mathbb{P}'}, \sigma') = (0.0310, -0.3825, 0.4906, 0.1991)$.

Six simulation sets are carried out. In all cases, the underlying asset evolves according to a synthetic $\mathbb{P}$ measure and the market prices of options are determined by a synthetic $\mathbb{Q}$ measure. These two measures are given in Table 7.1 and, in general, the hedger does not have exact knowledge of these parameters. However, in the first set of simulations, the hedger does have perfect knowledge of $\mathbb{P}$ and $\mathbb{Q}$—these simulations may be regarded as the baseline case as they only involve the true model parameters of Table 7.1. The $\mathbb{Q}$ measure

is used by the hedger to compute all option quantities (e.g. prices, deltas) required to find the hedge weights, while the real-world probability measure $\mathbb{P}$ is used to form the weighting function—in this case, a lognormal density with parameters $\mu^{\mathbb{P}}$ and $\gamma^{\mathbb{P}}$.

The second set of simulations is almost identical to the first, only now the estimated pricing measure $\mathbb{Q}'$ is used by the hedger; that is, the option prices found from the $\mathbb{Q}'$ measure are used to establish the linear system (7.6) that determines the hedge weights. It is important to note that it is the true option values under $\mathbb{Q}$ that are used when rebalancing the hedge portfolio, as these are the prices observed in the market. If $\mathbb{Q}$ and $\mathbb{Q}'$ give similar prices, the error introduced by using $\mathbb{Q}'$ is expected to be small.

The third set of simulations gauges the effect of utilizing the estimated pricing measure in a potentially unintelligent way. As with simulation set #2, $\mathbb{Q}'$ is used by the hedger to price the options. However, we assume the coefficient of relative risk aversion is known, and is used in combination with $\mathbb{Q}'$ to estimate the objective measure $\mathbb{P}'$. This estimated measure is in turn used to establish $W(J)$.

The intention of the final simulation sets (#4, #5 and #6) is to explore further the choice of weighting function used in constructing the hedge portfolio. For simulation set #4, the true pricing measure $\mathbb{Q}$ is utilized for this purpose, such that the weighting function is a lognormal density with parameters $\mu^{\mathbb{Q}}$ and $\gamma^{\mathbb{Q}}$. For simulation set #5, a highly unreasonable parameter set is used as an estimate for the distribution of jump amplitudes, which in turn is employed for $W(J)$. The last simulation set uses the uniform-like weighting function plotted in Figure 7.1. The simulation experiments are summarized in Table 7.2; note that the same 500,000 paths are used for each.

7.3.3 Hedging Results

Statistical measures of the relative P&L (3.1) are provided in Table 7.3. In addition to the simulation experiments summarized in Table 7.2, results for the delta hedge are also given.

	Measures for Determining the Hedge Portfolio	
Simulation Set #	Option Quantities	Weighting Function
1	$\mathbb{Q}$	$\mathbb{P}$
2	$\mathbb{Q}'$	$\mathbb{P}$
3	$\mathbb{Q}'$	$\mathbb{P}'$
4	$\mathbb{Q}$	$\mathbb{Q}$
5	$\mathbb{Q}$	Lognormal $(\mu = 0.5, \gamma = 0.1)$
6	$\mathbb{Q}$	Uniform-like density of Figure 7.1

TABLE 7.2: *Summary of the simulation experiments. For all simulations, the real-world $\mathbb{P}$ measure is used to generate asset price paths. Moreover, all observable prices in the synthetic market are derived using the true pricing $\mathbb{Q}$ measure.*

The outliers of the distributions are very important, as it is these potentially catastrophic events one tries to prevent by hedging. Various outliers (in the form of the 0.2% and 99.8% quantiles) are given in the table.

The first row of Table 7.3 gives the summary statistics for the delta hedge. The mean of the distribution is positive, which agrees with theory: since the $\mathbb{Q}$ measure is more "pessimistic" than the $\mathbb{P}$ measure (i.e. the $\mathbb{Q}$ measure parameters imply more frequent jumps with, on average, larger drops in the underlying), the expected value of the instantaneous jump risk (6.9) is positive. However, this positive mean is offset by the losses that result when a jump event takes place. The corresponding outliers may be quite large: for about 20% of the jump events, the option seller loses approximately three times the value of the initial option premium.

Simulation sets #1 through #4 demonstrate that the dynamic hedging strategy works well for this straddle. The means are all approximately zero, with standard deviations that are fairly small. More importantly, there are no significantly negative outliers, as was the case with delta hedging. The use of the estimated pricing measure has little effect, as the results for simulation sets #1 and #2 are almost identical. This is expected since both $\mathbb{Q}$ and $\mathbb{Q}'$ produce similar option prices.

Simulation Set	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
Delta Hedge	0.12	0.24	-2.99	0.22
1	0.00	0.05	-0.17	0.28
2	0.00	0.05	-0.17	0.27
3	0.00	0.04	-0.14	0.22
4	0.00	0.07	-0.28	0.43
5	0.08	0.16	-2.11	0.15
6	0.00	0.04	-0.13	0.20

TABLE 7.3: *Statistical measures of the relative profit and loss when hedging a one-year European straddle ($K = \$100$) under jump diffusion. The hedge is rebalanced forty times, equally interspersed over the half-year hedging horizon. For the $\alpha\%$ quantile, approximately $\alpha\%$ of the 500,000 simulations resulted in a relative P&L less than the reported amount.*

As for the choice of the weighting function, there appears to be a certain amount of leeway. Using a $W(J)$ motivated by $\mathbb{P}$ and $\mathbb{P}'$ gives similar results, though there is a slight degradation for the weighting function suggested by $\mathbb{Q}$. The salient conclusion from simulation set #5, which has poor summary statistics, is that the use of a totally inappropriate weighting function will not give good results. On the other hand, the results from using a uniform-like weighting function (simulation set #6) are encouraging: this $W(J)$ encapsulates a lack of knowledge concerning the distribution of the jump amplitudes, but nonetheless yields good results.

Imposing delta neutrality eliminates the instantaneous diffusion risk. However, because the hedge is rebalanced discretely, the protection will degrade as time passes. We now consider two ways to better hedge the diffusion risk within our strategy: (i) intermediate delta rebalancing where, between the "complete" rebalances achieved by solving (7.2), the weight of the underlying is alone adjusted to reimpose delta neutrality; and (ii) imposing the gamma-neutral constraint (7.3) in the optimization. Simulation sets #1 ($\mathbb{P}$ weighting) and #6 (uniform-like weighting) are re-run using these two strategies. For the case of

Simulation Set	Hedging Strategy	Mean	Standard Deviation	Quantiles	
				0.2%	99.8%
$\mathbb{P}$ Weighting	Original	0.00	0.05	-0.17	0.28
	Intermediate Delta Rebalancing	0.00	0.03	-0.12	0.15
	Delta + Gamma Neutral	0.00	0.01	-0.03	0.02
Uniform-like Weighting	Original	0.00	0.04	-0.13	0.20
	Intermediate Delta Rebalancing	0.00	0.02	-0.09	0.11
	Delta + Gamma Neutral	0.00	0.01	-0.03	0.03

TABLE 7.4: *Statistical measures of the relative profit and loss when hedging a one-year European straddle ($K = \$100$) under jump diffusion, using (i) the original strategy; (ii) intermediate delta rebalancing; and (iii) the imposition of gamma neutrality. The hedge is rebalanced forty times over the half-year hedging horizon. For intermediate delta rebalancing, the underlying is adjusted three times between each "complete" rebalance. In the case of gamma neutrality, both delta and gamma neutrality are imposed when the jump risk is minimized.*

intermediate delta rebalancing, the underlying is adjusted three times (equally spaced) between each of the forty complete rebalances.

The results are presented in Table 7.4. For intermediate delta rebalancing, the standard deviation is smaller as compared to the original strategy. By adjusting only the underlying between complete rebalances, the protection against the jump risk will be adversely affected. This does not seem to have a major effect in this example, perhaps due to the infrequent arrival of jumps. The results when gamma neutrality is imposed exhibit a marked improvement over the other two strategies. Even though this extra constraint will expend an additional degree a freedom, a hedge with five options will still be able to provide adequate jump protection. However, if only the underlying and one other instrument are used to hedge under jump diffusion, the use of a delta-gamma hedge is not advisable.

Remark 7.1 (Using the TSVD). *For the most part, the results of this section are in-*

variant to how the linear system (7.6) *is solved (i.e. the TSVD vs. an LU decomposition). However, if the standard TSVD cutoff of* 10^{-6} *is used in the case where gamma neutrality is imposed, the results are only slightly better than the original strategy; the results in Table 7.4 corresponding to the imposition of gamma neutrality were found using an LU decomposition. Furthermore, using a cutoff of* 10^{-8} *in the TSVD for this delta-gamma–neutral example yields the same results as the LU decomposition. This demonstrates that the cutoff parameter used in the TSVD should be chosen with care and via experimentation, as suggested in Press et al. [65].*

7.4 Hedging an American Option

This section applies the hedging strategy to an option with American-style early exercise features. The target contract to be hedged is an American put with a half-year maturity. The hedging horizon is three months, and American puts and calls are used as the hedging instruments (these all have the same maturity as the hedging horizon). Similar hedge portfolio selection rules apply as in the European case, only now out-of-the-money options are (ideally) used, as well as the underlying asset. For instance, if five options are to be held in the hedge portfolio at inception, those puts with strikes closest to $[0.8S_0, 0.9S_0, 1.0S_0]$ and calls with strikes nearest $[1.1S_0, 1.2S_0]$ are used.

An American put in the hedge portfolio may be optimally exercised before the end of the hedging horizon. Therefore, the early exercise regions of all active American puts are monitored, and if the asset price path enters such a region, it is assumed both long and short positions in the corresponding put are exercised optimally. In this case, the hedging option is suitably replaced with a put of lower strike, and the hedge is rebalanced. If the target half-year put is exercised early, the hedge portfolio is immediately liquidated to cover the payoff.

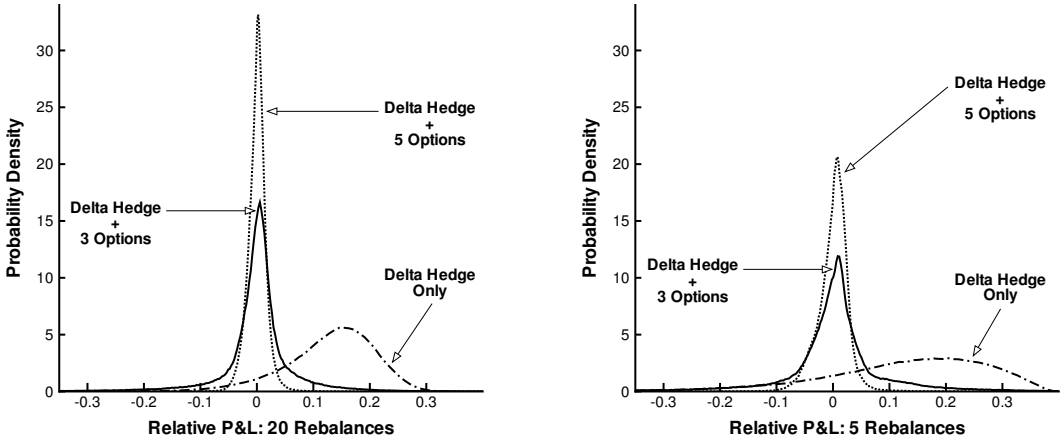


FIGURE 7.3: *Distributions of the relative profit and loss when hedging a half-year American put under jump diffusion. Left panel: twenty rebalances, right panel: five rebalances. The put has a strike of $100, and the hedging horizon is three months.*

The same market parameters as in the European case are used, except that the mean arrival rate of jumps in the market is set about four times higher, to $\lambda^{\mathbb{P}} = 0.1$. The uniform-like weighting function of Figure 7.1 is used, and the number of hedging instruments and rebalancing frequency is varied. Table 7.5 provides statistics for the distributions of relative P&L. Figure 7.3 plots the distributions for the cases of twenty and five rebalances.

The results are consistent with the European case. When a simple delta hedge is employed, the seller of the put makes money both most of the time and on average. However, this strategy may again result in catastrophic losses when a jump occurs. The use of other instruments within our dynamic hedge mitigates this jump risk, as is evident with the 0.2% quantile. The use of more options in the hedge portfolio leads to better performance, which is not surprising. Note that the standard deviation and negative outlier for the case of delta hedging are essentially the same, regardless of the rebalancing frequency. This should reinforce the idea that the ignored jump risk inherent to the simple delta hedge will manifest itself regardless of the frequency of rebalancing.

Number of Rebalances	Hedging Options	Mean	Standard Deviation	Quantiles	
				0.2%	99.8%
20	Delta Hedge	0.04	0.69	-7.18	0.30
	3	0.00	0.08	-0.49	0.31
	5	0.00	0.02	-0.07	0.08
10	Delta Hedge	0.04	0.69	-7.12	0.34
	3	0.00	0.10	-0.64	0.37
	5	0.00	0.02	-0.09	0.11
5	Delta Hedge	0.04	0.69	-6.95	0.38
	3	0.00	0.12	-0.75	0.46
	5	0.00	0.03	-0.13	0.14
3	Delta Hedge	0.03	0.69	-6.80	0.39
	3	0.00	0.12	-0.69	0.57
	5	0.00	0.04	-0.17	0.16

TABLE 7.5: *Statistical measures of the relative profit and loss when hedging a half-year American put under jump diffusion. The put has a strike of $100, and the hedging horizon is three months.*

7.5 Summary

A delta hedge, which only treats the diffusion risk, will often experience a large loss if a jump occurs. In this chapter, we investigated a dynamic hedging strategy that, at each rebalance time, removes the instantaneous diffusion risk and minimizes a measure of the jump risk. A hedge portfolio consisting of five options and the underlying, whose weights are chosen in this way, is effective in reducing the exposure to jump risk. In addition, we demonstrated that the strategy is able to handle both European and American-style claims. The experiments included an investigation of estimation errors in the parameters, assuming that a hedger would clearly not use the exact $\mathbb{Q}$ and $\mathbb{P}$ parameters. Using a pricing measure obtained from calibration, and a uniform-like weighting function that encapsulates a lack of knowledge about the jumps, leads to good results.

Chapter 8

Dynamic Hedging under Jump Diffusion in the Presence of Transaction Costs

This chapter considers a jump-diffusion market where transaction costs are present in the form of a relative bid-ask spread. We develop an optimization problem that selects the hedge portfolio to simultaneously (i) eliminate the instantaneous diffusion risk by imposing delta neutrality; and (ii) minimize an objective function that is a linear combination of a jump risk and transaction cost penalty function. The properties of this optimization are investigated, including how to appropriately weight the jump risk and transaction cost components of the objective function. Hedging simulations are carried out for both European and American-style claims, using a constant relative bid-ask spread and a more realistic bid-ask model drawn from market data.

8.1 Transaction Costs: Relationship to Previous Work

In the previous chapter, the discretely-rebalanced dynamic strategy represented by the optimization (7.2) was shown to provide good results when hedging a longer term European straddle and American put using short-term calls and puts. However, no consideration was given to the effect of transaction costs, as they were assumed absent. For the frequent rebalancing necessitated by dynamic hedging, these costs need to be taken into account, otherwise they can make the procedure prohibitively expensive.

For one-factor diffusion models where the underlying is the only traded instrument, there is an extensive body of literature that incorporates transaction costs into the pricing and hedging framework. When the underlying asset evolves according to GBM and transaction costs are proportional to the value of the transaction in the asset, a (discrete) hedging argument similar to that used for deriving the Black–Scholes equation yields a nonlinear option-pricing PDE [42]. When pricing individual calls and puts, the nonlinearity is absent due to the single-signed gamma, such that only a volatility adjustment is required in the Black–Scholes PDE. Leland [53] considers this problem, and develops a delta hedging strategy based on the delta $\frac{\partial \hat{V}}{\partial S}$ of the option values $\hat{V}$ found using the augmented volatility: the hedge is rebalanced at intervals of Δt so that the amount of stock in the hedge after rebalancing is $\frac{\partial \hat{V}}{\partial S}$. The hedging error in Leland's model as $\Delta t \to 0$ is a non-trivial function, and almost surely negative [45, 37].

For Merton's original jump-diffusion model (i.e. lognormally distributed jumps with no risk adjustment), Mocioalca [57] derives a volatility adjustment for pricing calls and puts that is analogous to Leland's. However, the delta hedge motivated by this analysis has the same drawback as in a jump model with no transaction costs, namely that it does not provide adequate protection against jump risk.

The hedging approach motivated by a nonlinear pricing equation, as outlined above,

is local in time. Global-in-time methods that use utility indifference pricing can also be employed. Davis et al. [29] consider a GBM model (with proportional transaction costs) where the underlying can be traded continuously. Within this utility framework, a HJB system can be solved to establish the optimal hedging strategy. The solution defines a buy, sell and no-trade region: when the underlying enters either the buy or sell region, the hedger performs the necessary transaction that brings the position in the underlying back onto the boundary of the no-trade region.

An analogous stochastic control program could be set up when hedging under jump diffusion with transaction costs. In this case, more than the underlying is needed to hedge, so the solution would define a complex set of trading regions. Keppo and Peura [46] consider a problem similar to this (i.e. where more than one hedging instrument is used), only within a GBM model. The authors use approximations to yield an augmented formulation, not involving HJB equations, that can be treated numerically. The system of HJB equations would involve the values of the target option and hedging instruments, as well as the controls that yield the optimal hedging strategy. For example, a hedge portfolio which consists of the underlying and L additional instruments would result in a $(L+1)$-dimensional HJB–PIDE system. Clearly, any numerical solution of these equations would be computationally intractable for any more than a few hedging instruments.

8.2 Incorporating Transaction Costs

The hedging strategy represented by the optimization (7.2) does not take into account transaction costs, so it may yield hedge portfolio weights that require expensive trading to implement. In this section we show how the objective function in (7.2) can be modified so that these costs are taken into consideration when rebalancing. For our purposes, transaction costs refer to the difference between the bid/ask price and the theoretical value of the

security (underlying asset or option). Brokerage commissions and other fees are ignored.

We assume the following scenario: using the linear pricing equation (2.9), a hedger fits option pricing parameters to the midpoint option values observed in the market. Then, a hedging strategy is constructed using a simple market model of bid-ask spreads. This approach preserves the property that the prices are linear in the numbers bought/sold, and it makes minimal assumptions about a model of bid-ask spreads. In contrast, a nonlinear pricing equation [e.g. 42, 57] in effect attempts to predict the bid-ask spread for options. Furthermore, for nonlinear pricing equations, the value of the overall hedged position $-V + eS + \vec{\phi} \cdot \vec{I} + B$ is not the same as the sum of the values of the individual components. Linear pricing rules are the market standard for the simple contracts we use for hedging [23].

We incorporate transaction costs using a relative bid-ask spread. We will assume that the options to be used for hedging can be characterized by a single strike price K—this will include vanilla puts and calls. This assumption will allow the relative bid-ask spread for a range of options to be modelled as a function of moneyness K/S. The dollar spread is assumed to be symmetric around the theoretical option value found using the pricing equation (2.9). Furthermore, the (relative) bid-ask spread will be quoted as a fraction of the option price. For example, with a relative bid-ask spread of 0.10 on an option with theoretical value \$5.00, the bid price is \$4.75 and the ask price is \$5.25. In other words, the hedger will have to pay \$5.25 to purchase the option with theoretical value \$5.00, and would receive \$4.75 if the option were to be sold. If the weight of an instrument is $\rho(t_{n-1})$ before rebalancing and $\rho(t_n)$ after rebalancing, the total cost of the transaction is

$$\left|\rho(t_n) - \rho(t_{n-1})\right| \times \frac{\text{BA}}{2} \times \text{ Instrument Value}\,,$$

where BA denotes the relative bid-ask spread.

The objective function

$$\int_0^\infty \left[-\Delta V + \left(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right)\right]^2 W(J)\,dJ \tag{8.1}$$

from the optimization problem (7.2) is a measure of the exposure to jump risk. Since it is quadratic in the hedge weights $\{e, \vec{\phi}\}$, it facilitates a straightforward application of Lagrange multipliers (see Section 7.1). Therefore, a similar quadratic representation for handling transaction costs is desirable. One suitable candidate is the sum of the squares of transaction costs

$$\left[\Big(e(t_n) - e(t_{n-1})\Big) \times \frac{\mathrm{BA}_S}{2} \times S\right]^2 + \sum_{j=1}^{N}\left[\Big(\phi_j(t_n) - \phi_j(t_{n-1})\Big) \times \frac{\mathrm{BA}_j}{2} \times \mathrm{OptVal}_j\right]^2, \tag{8.2}$$

where BA_S is the relative bid-ask spread for the underlying, and ϕ_j is the weight in the j^{th} hedging option, which has value OptVal_j and relative bid-ask spread BA_j.

Consider an optimization problem whose objective function is a linear combination of the jump risk measure (8.1) and the transaction cost penalty function (8.2), and which eliminates the instantaneous diffusion risk via a linear constraint. The resulting program is

$$\begin{aligned}\min_{\{e(t_n),\vec{\phi}(t_n)\}} \quad & \xi\left\{\int_0^\infty \left[-\Delta V + \left(e(t_n)\Delta S + \vec{\phi}(t_n)\cdot\Delta\vec{I}\right)\right]^2 W(J)\,dJ\right\} \\ & + \ (1-\xi)\left\{\left[\Big(e(t_n) - e(t_{n-1})\Big) \times \frac{\mathrm{BA}_S}{2} \times S\right]^2 \right. \\ & \qquad \left. + \ \sum_{j=1}^{N}\left[\Big(\phi_j(t_n) - \phi_j(t_{n-1})\Big) \times \frac{\mathrm{BA}_j}{2} \times \mathrm{OptVal}_j\right]^2\right\}\end{aligned}$$

subject to

$$e(t_n) + \vec{\phi}(t_n)\cdot\frac{\partial\vec{I}}{\partial S} = \frac{\partial V}{\partial S}, \tag{8.3}$$

where $0 \leq \xi \leq 1$ is an influence parameter that determines the weighting of the two components in the objective.

In Section 6.3 we showed that, for an idealized trading environment, by making the objective function in (8.3) sufficiently small for a delta-neutral and discretely rebalanced hedge, the variance of the terminal hedging error can be made arbitrarily small as the rebalancing interval $\Delta t \to 0$. Since it may not be possible to make the objective function small, and because continuous rebalancing is impossible in practice, our aim is to make the objective function as small as possible at each rebalancing time. The analysis in Chapter 6 also showed that, under ideal conditions, ξ should be $\mathcal{O}(\Delta t^2)$. The optimization problem and the associated hedging simulations will be considered for a range of ξ values.

With the influence parameter ξ and the weighting function $W(J)$ specified by the hedger, the optimization problem (8.3) may be solved using Lagrange multipliers. Indeed, the form for the transaction cost penalty function (8.2) was chosen because it is amenable to this treatment. At each hedge rebalance time, the linear system is established using table lookup and interpolation of the precomputed data, and will yield $e(t_n)$ and $\vec{\phi}(t_n)$, the positions that the hedge portfolio should have after rebalancing at t_n. Before considering the hedging simulations, we will investigate the behaviour of the optimization problem (8.3) that lies at the heart of the strategy.

8.3 A Representative Optimization Problem

At each rebalance time, our goal is simple: choose hedge portfolio weights that impose delta neutrality and reduce the jump risk, while keeping transaction costs as small as possible. The tradeoff between the minimization of jump risk and the reduction of transaction costs is controlled by the influence parameter ξ: for $\xi = 1$ we are only concerned with jump risk, while $\xi = 0$ corresponds to total concentration on transaction costs. A solution x^* to a

Instrument	Initial Maturity	Strike	Value at $t = 0$, $S = 100$	Weight at $t = 0$, $S = 100$	Value at $t = 0.05$, $S = 106.5$
Straddle	1 year	\$100.00	\$21.41	-1.00	\$24.05
Underlying	n.a.	n.a.	\$100.00	-0.64	\$106.50
Put	0.25 years	\$80.00	\$0.91	1.29	\$0.67
Put	0.25 years	\$90.00	\$1.53	-0.94	\$0.91
Call	0.25 years	\$100.00	\$5.34	1.92	\$9.38
Call	0.25 years	\$110.00	\$1.50	-0.93	\$3.23
Call	0.25 years	\$120.00	\$0.28	0.60	\$0.69

TABLE 8.1: *Instruments in the overall hedged position used to investigate the optimization problem* (8.3) *that considers both jump risk and transaction costs. All options have European-style exercise rights.*

multi-objective optimization problem is said to be Pareto optimal if any perturbation of x^* required to improve one of the component objectives can only be made at the expense of another objective. The collection of all such solutions is the Pareto optimal set, and the associated set of component objective values $\vec{F}(x^*)$ is the Pareto front. The most well-known example of a Pareto front in finance is the efficient frontier of the Markowitz model.

We consider a specific rebalancing example as a means to study the behaviour of the objective function in (8.3) for different values of the influence parameter ξ. Assume a financial institution has sold a one-year, at-the-money European straddle, where $S_0 = 100$. To hedge its exposure, positions are taken in the underlying and five put and call options with three months until expiry—the instruments in the overall hedged position are given in Table 8.1. We use the $\mathbb{Q}$ measure from Table 7.1 for pricing. The initial hedge portfolio weights are found by solving (8.3) with $\xi = 1$. At $t = 0.05$ with the underlying at $S = 106.5$, the hedge portfolio is to be rebalanced: the optimization problem (8.3) is solved for the range of influence parameter values $\xi \in [0, 1]$. The relative bid-ask spread is fixed at 0.10 for all hedging options and 0.002 for the underlying, and the uniform-like weighting function in Figure 7.1 is used for $W(J)$.

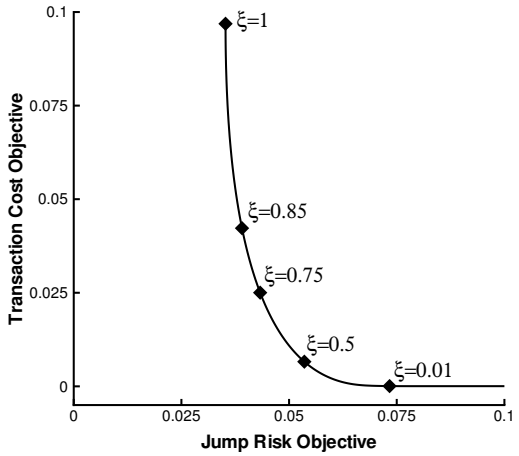


FIGURE 8.1: *The Pareto optimal front for the optimization* (8.3) *when the hedge is rebalanced at* $t = 0.05$, $S = 106.5$. *The target option is a one-year European straddle, and the instruments used in the hedge portfolio are given in Table 8.1.*

The hedge portfolio weights at $t = 0.05$ found from solving the optimization problem with varying ξ are used to compute the jump risk objective function (8.1) and the transaction cost objective function (8.2), and these are plotted together in Figure 8.1. The exposure to jump risk is smallest for $\xi = 1$, but gets larger as the influence parameter decreases. The opposite behaviour is observed for the transaction cost objective. This curve displays how the "best" solution is subjective: the interested party should opt for a solution from the Pareto optimal set, but the specific choice will be based on other considerations.

The value from the jump risk objective function (8.1) is a rather blunt statistic, as it condenses the risk from a continuum of possible jumps into a single number. Nonetheless, the protection afforded against jump risk by a hedge portfolio may be visualized by considering a plot of ΔH_J. Recall that ΔH_J in (6.14) represents the change in the overall hedged position due to a jump of size J. Figure 8.2 presents the jump risk profiles for the hedge portfolios found at $t = 0.05$ using three values of the influence parameter ξ. The best possible curve in terms of minimizing jump risk is for $\xi = 1$, but the associated hedge

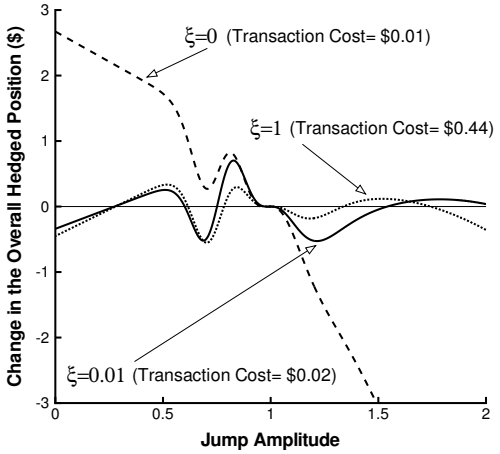


FIGURE 8.2: *Change in the value of the overall hedged position due to a jump, assuming the jump occurs an instant after rebalancing at $t = 0.05$, $S = 106.5$. The curves correspond to the hedge portfolio weights found using three different values of the influence parameter ξ. The rebalancing cost of forming the hedge portfolio associated with each profile is given by "Transaction Cost". The target option is a one-year European straddle ($K = \$100$), and the instruments used in the hedge portfolio are given in Table 8.1.*

portfolio weights are selected in a manner that ignores transaction costs. For $\xi = 0$ only transaction costs are considered; these costs are quite low, but the protection against jumps is not very good. The third curve, corresponding to $\xi = 0.01$, offers the middle ground we seek, namely low transaction costs with good protection against jumps. The conclusion that may be drawn from Figure 8.2 is that hedge portfolio weights can be found which do a good job of adequately satisfying both objectives.

8.4 Hedging Simulations

The hedge portfolio is formed and rebalanced using the general procedure outlined in Section 2.4, with the weights chosen by solving the optimization (8.3). The relative profit and loss (3.1) at exercise/expiry T^* is again used as our metric for the hedging error. When liquidating the hedge portfolio to cover the short position $-V$, we must take into account

transaction costs. The value of the overall hedged position at exercise/expiry T^* is given by

$$\begin{aligned}\Pi(T^*) = &-V(S_{T^*}, T^*) + B(t')\mathrm{e}^{r(T^*-t')} + e(t')S_{T^*} + \vec{\phi}(t') \cdot \vec{I}(S_{T^*}, T^*) \\ &- \Big|e(t')\Big|\Big(\frac{\mathrm{BA}_S}{2}\Big)S_{T^*} - \sum_{j=1}^{N}\Big|\phi_j(t')\Big|\Big(\frac{\mathrm{BA}_j}{2}\Big)I_j(S_{T^*}, T^*)\,,\end{aligned}$$

where t' is the time of the last rebalancing. When hedging a European option, this liquidation will most likely coincide with the expiry of shorter term options, such that the transaction costs will usually only come from disposing of the underlying.

When using a model of bid and ask prices, one must be cognizant of arbitrage opportunities that can arise through violations of put-call parity, especially when using an idealized trading environment that does not encompass any frictions that tend to damp these out. However, our bid and ask prices are determined by adding a proportional transaction cost to the values obtained by solving the option-pricing PIDE (2.9) (for European options) or the variational inequality (2.10) (for American options). Since these option values obey a put-call parity relationship, and the transaction costs always act as a sink, there will never be any arbitrage opportunities.

To illustrate this point, we consider a European call and put of the same strike K and expiry T, and assume no dividends are paid by the stock. We take a long position in the call and hold cash of amount Ke^{-rT}, while shorting the put and underlying asset—this is one of the two standard portfolios used to take advantage of violations to put-call parity. The money needed to form this portfolio is $C^{ask} - P^{bid} - S^{bid} + Ke^{-rT}$, such that it will offer an instantaneous risk-free profit if

$$C^{ask} - P^{bid} - S^{bid} + Ke^{-rT} < 0 \tag{8.4}$$

since the net payoff at expiry is zero (ignoring the liquidation costs). Now, the bid and ask prices can be expressed as

$$C^{ask} = V_C + \frac{BA_C}{2}V_C, \quad P^{bid} = V_P - \frac{BA_P}{2}V_P, \quad S^{bid} = S - \frac{BA_S}{2}S\,,$$

where V_C and V_P are the option values obtained from solving the PIDE (2.9), and BA_j are the associated relative bid-ask spreads. By substituting the above into (8.4), an arbitrage opportunity will result if

$$\begin{aligned} V_C + \frac{BA_C}{2}V_C - \left(V_P - \frac{BA_P}{2}V_P\right) - \left(S - \frac{BA_S}{2}S\right) + Ke^{-rT} &< 0 \\ \longrightarrow \qquad \frac{BA_C}{2}V_C + \frac{BA_P}{2}V_P + \frac{BA_S}{2}S &< 0\,, \end{aligned}$$

where put-call parity

$$V_C \;-\; V_P \;-\; S \;+\; Ke^{-rT} \;=\; 0$$

has been employed in the simplification. Therefore, provided the relative bid-ask spreads are non-negative, there will never be any arbitrage opportunities. If the opposite position is considered (i.e. $-C^{bid} + P^{ask} + S^{ask} - Ke^{-rT}$), the same result is obtained.

We now consider a variety of hedging experiments for target contracts with European and American-style exercise rights, using both a constant relative bid-ask spread and a model of bid-ask spreads drawn from market data.

8.4.1 A Simple Hedging Example: Five Hedging Options

To provide a simple illustration of the hedging strategy, we extend the example of Section 8.3. A one-year European straddle is to be hedged over its lifetime. Initially, the underlying along with puts of strike $K = [80, 90]$ and calls with strike $K = [100, 110, 120]$

Hedging Strategy	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
Delta hedge	0.25	0.38	-3.89	0.55
Five hedging options	0.00	0.02	-0.06	0.05

TABLE 8.2: *Hedging results, in terms of the relative profit and loss, for the European example when there are no transaction costs. The hedge weights are chosen by solving the optimization* (7.2)*. The target option is a one-year European straddle (*$K = \$100$*), and the instruments used to establish the hedge are given in Table 8.1. The hedge is rebalanced every 0.025 years. For the* $\alpha\%$ *quantile, approximately* $\alpha\%$ *of the 250,000 simulations resulted in a relative P&L less than the reported amount.*

are used, where all of the hedging options have three months until maturity. These options are traded until they expire, at which time new options are purchased, and these new options have the same strikes and time to maturity as the initial set. The $\mathbb{P}$ and $\mathbb{Q}$ measures employed are in Table 7.1. Each simulation set consists of 250,000 individual simulations, meaning a total of about 5,700 jumps are expected over each set.

No Transaction Costs. We first investigate a hedging example in which financial instruments may be traded without incurring transaction costs. Consider the case where only the underlying is used in a delta hedge, which is rebalanced every 0.025 years. The results, in the form of statistical measures for the relative P&L, are contained in the first row of Table 8.2. The 0.2% quantile is very negative in this case, corresponding to the large losses that often result when a jump occurs. When five options are included in the hedge along with the underlying, the weights are chosen using the hedging strategy represented by the optimization in (7.2) that is employed when transaction costs are absent. Compared to the delta hedge, this procedure dramatically reduces the exposure to jump risk, as demonstrated by the results in the second row of Table 8.2.

Transaction Costs Present, but Ignored. We next consider a set of simulations where transaction costs are present, with a constant relative bid-ask spread of 0.10 for the hedging

Hedging Strategy	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
Delta hedge	0.22	0.38	-3.90	0.53
Five hedging options	-0.38	0.27	-1.41	-0.06

TABLE 8.3: *Hedging results, in terms of the relative profit and loss, for the European example when transaction costs are present but ignored. The hedge weights are chosen by solving the optimization* (8.3) *with* $\xi = 1$. *The target option is a one-year European straddle (*$K = \$100$*), and the instruments used to establish the hedge are given in Table 8.1. The hedge is rebalanced every 0.025 years.*

options and 0.002 for the underlying. The results for the delta hedge, presented in the first row of Table 8.3, are very similar to those when no transaction costs are incurred (see the first row of Table 8.2), although there is a small negative change in the mean due to the cost of trading. This negative shift is also evident in the distribution of relative P&L presented in the left panel of Figure 8.3.

When five options are included in the hedge along with the underlying, the weights are chosen in a manner that ignores transaction costs (i.e. solving (7.2) or, equivalently, (8.3) with $\xi = 1$). This strategy of using five hedging options yields results that are very poor—it could be argued the delta hedge is better in this instance, as indicated by the statistics in Table 8.3. Even though the five-option hedge will help protect the overall position when a jump occurs, the cost of the required transactions tends to be high. For this example, our original hedging strategy with five options performs quite well until transaction costs are introduced, at which point it becomes essentially useless.

Transaction Costs Present, and Taken into Account. We now carry out the simulations again, only this time using the optimization problem (8.3) that takes into account both jump risk and transaction costs. The simulations are performed using the influence

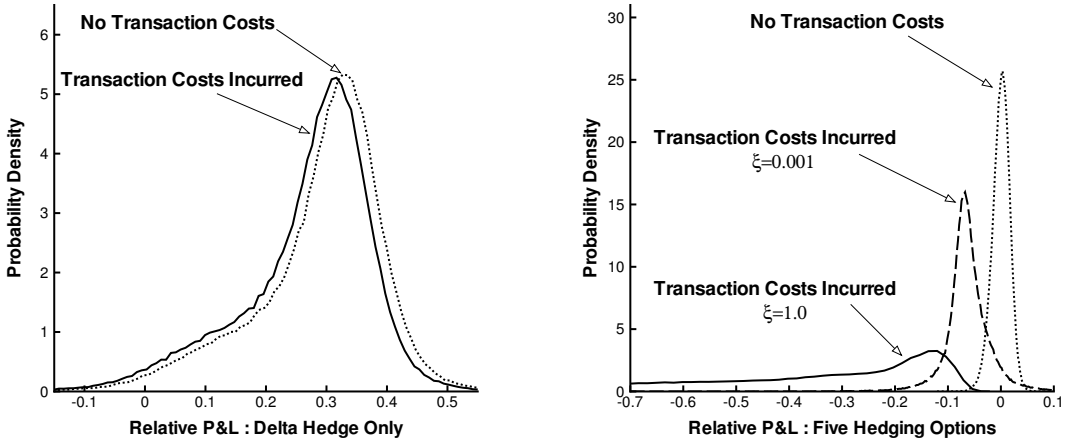


FIGURE 8.3: *Distributions of the relative profit and loss for the European hedging example. When $\xi = 1.0$, the transaction costs are ignored in the optimization (8.3), while for $\xi = 0.001$ both transaction costs and jump risk are taken into account. The target option is a one-year European straddle ($K = \$100$), and the instruments used to establish the hedge are given in Table 8.1. The hedge is rebalanced every 0.025 years.*

parameters

$$\xi \in \left[0, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 0.0025, 0.005, 0.0075, 0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 0.9, 0.95, 1\right]. \quad (8.5)$$

For each of the 250,000 simulations, the same value of the influence parameter will be used throughout; for example, if forty rebalances are carried out over the entire hedging horizon, the same value of ξ is used for each of the forty individual optimization problems. The criteria used for deciding which ξ is best for hedging will be the mean and standard deviation of the relative P&L.

For each value of the influence parameter ξ in (8.5), the mean and standard deviation of the relative P&L is plotted in Figure 8.4. The results for the extreme values of $\xi = 0$ and $\xi = 1$ are poor. If the standard deviation is used as the sole criterion to select the best

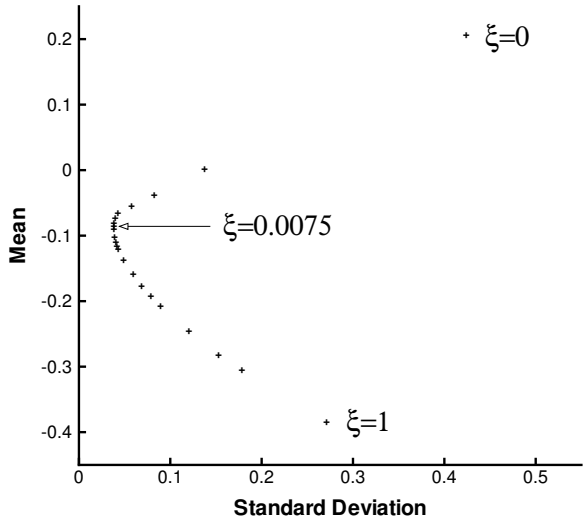


FIGURE 8.4: *Mean and standard deviation of the relative profit and loss, with varying ξ, for the European hedging example. When $\xi = 1$, the transaction costs are ignored in the optimization (8.3), while for $\xi = 0$ the jump risk is ignored. When $0 < \xi < 1$, both transaction costs and jump risk are taken into account. The target option is a one-year European straddle ($K = \$100$), and the instruments used to establish the hedge are given in Table 8.1. The hedge is rebalanced every 0.025 years.*

value of the influence parameter, the choice is $\xi = 0.0075$. This corresponds to a standard deviation of 0.0384 and a mean of -0.0843. However, for $\xi = 0.001$ the standard deviation is only slightly higher at 0.0429, while the mean of -0.0632 is much better. The distribution of the relative P&L for this case is the middle density in the right panel of Figure 8.3. A mean of -0.0632 translates to \$1.35 in monetary terms: if the hedger charges this as a premium over and above the theoretical option price of \$21.41, the simulations will have a zero mean. This is slightly higher than the 5% premium (i.e. half of the relative bid-ask spread of 10%) assumed for quarter-year vanilla options in the hedging portfolio.

In many cases, only a small amount of buying and selling occurs when the hedge is rebalanced. The left panel of Figure 8.5 presents a sample price path over the first quarter year (right axis), and the corresponding evolution of the position in the short-term hedging

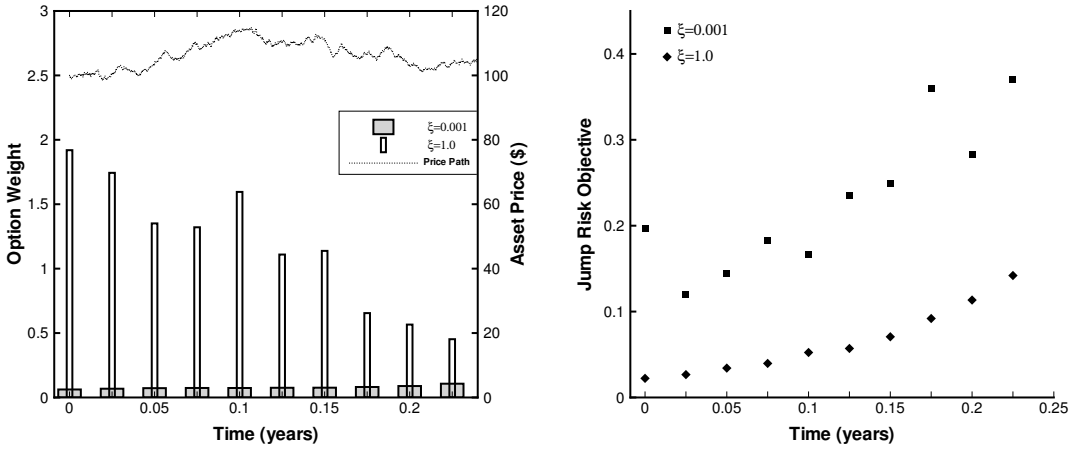


FIGURE 8.5: *Characteristics of the hedge portfolio for a specific asset price path and differing values of the influence parameter ξ. The left panel depicts the sample price path over the first quarter year (right axis), and the corresponding evolution of the position in the short-term hedging call option with a strike of \$100 (left axis). The right panel is a plot of the jump risk measure (8.1) after each rebalance.*

call option with strike \$100 (left axis). Since the value of the stock is above $S = 100$ for most of the time, this hedging option is predominantly in the money, with its value at rebalancing times ranging between \$3.92 and \$15.57. When transaction costs are ignored ($\xi = 1$), the initial position and the subsequent changes can be significant, resulting in large transaction costs. On the other hand, when transaction costs are taken into account (e.g. $\xi = 0.001$), the option position remains small. Although we are controlling for transaction costs, the exposure to jump risk, as represented by the objective function (8.1), is not much greater than the minimum value corresponding to $\xi = 1$. This is demonstrated in the right panel of Figure 8.5.

We may conclude that, for the hedging example explored in this section, it is possible to select hedge weights that provide sufficient protection against jump risk while not incurring large transaction costs.

Remark 8.1 (Using the TSVD). *If a relatively high cutoff (e.g. 10^{-3}) is used within*

the TSVD when $\xi = 1$, the strategy may produce reasonable results, even though transaction costs are not taken into account when choosing the hedge weights. In this case, the TSVD solution procedure returns a vector of weights with a small norm—an ideal way to keep transaction costs down—while still providing adequate protection against jump risk. In general, any strategy which uses a regularization method for determining the portfolio weights will tend to keep transaction costs under control (e.g. the strategy in Cont et al. [25]).

8.4.2 Varying the Rebalancing Frequency and Number of Options

Up to this point the hedge portfolio has consisted of five options and the underlying, and has been rebalanced every 0.025 years. We now vary the hedge portfolio composition—using three, five and seven options—and the frequency of rebalancing. When seven options are employed, those with strikes $K = [70, 80, 90, 100, 110, 120, 130]$ are used, while the three-option portfolio utilizes the middle strikes. The hedge is rebalanced a total of twenty, forty and eighty times over the one-year investment horizon. The hedging simulations are carried out under the same guidelines as before, and the results are presented in Table 8.4.

The "best" value of the influence parameter indicated in Table 8.4 comes from the simulation set that yields the smallest standard deviation. Consider a fixed number of hedging options: we find (i) the best (i.e. lowest) standard deviation gets smaller as the rebalancing becomes more frequent, and the mean does not become considerably more negative; and (ii) in general, as the rebalancing becomes more frequent, the best result is achieved by putting more weight on the transaction cost component of the objective (i.e. ξ gets smaller).

In Chapter 6, we examined an idealized continuous trading environment and demonstrated that, by making the jump risk objective function (8.1) and transaction cost objective function (8.2) sufficiently small at each instant, the variance of the terminal hedging error may be made small. In relation to the discrete framework, the bound on the transaction

No. of Hedging Options	Rebalance Interval	Best ξ	Standard Deviation	Mean
3	0.0125	0.001	0.05	-0.10
	0.025	0.001	0.05	-0.08
	0.05	0.0025	0.06	-0.07
5	0.0125	0.0025	0.03	-0.09
	0.025	0.0075	0.04	-0.08
	0.05	0.05	0.05	-0.09
7	0.0125	0.005	0.04	-0.10
	0.025	0.02	0.04	-0.10
	0.05	0.1	0.05	-0.10

TABLE 8.4: *Hedging results, in terms of the relative profit and loss, for the European example with different rebalancing frequencies and a varying number of options in the hedge portfolio. The influence parameter from the discrete set (8.5) that yields the lowest standard deviation is termed the best ξ. The target option is a one-year European straddle ($K = \$100$).*

cost objective should be $\mathcal{O}(\Delta t^2)$ as $\Delta t \to 0$ in order to ensure finite transaction costs. Within our objective function in (8.3), the appropriate tradeoff between the jump risk and transaction cost can be achieved with an influence parameter ξ that is $\mathcal{O}(\Delta t^2)$. In other words, as the rebalancing becomes more frequent, more and more weight should be put on the transaction cost component of the objective function. The results of Table 8.4 are generally consistent with the theory. For example in the case of seven hedging options, which is our closest approximation to an idealized trading environment, the best value of the influence parameter is approximately $\mathcal{O}(\Delta t^2)$.

8.4.3 Using Calls and Puts with the Same Strike

For a given strike price, both calls and puts are typically available in the market, so limiting the hedge portfolio to holding either one or the other is not realistic. We therefore consider an augmented version of the example in Section 8.4.1 by doubling the number of available hedging options, such that there is now access to all European calls and puts of strike

$K = [80, 90, 100, 110, 120]$. Note that all other settings, such as the constant relative bid-ask spread of 0.10 for options, remain the same.

When only jump risk is considered ($\xi = 1$), the linear system resulting from the application of Lagrange multipliers is singular. This is due to put-call parity, as the redundancy inherent in this relationship manifests as a rank-deficient matrix. For ξ close to unity, we expect the matrix to be ill conditioned. As noted before, using a TSVD is a common way to deal with an ill-conditioned linear system. Nonetheless, the (usually low) range of influence parameters that give the best hedging results tend to quell the ill conditioning. This is due to the fact the transaction cost component of the objective is not susceptible to degeneracy problems resulting from put-call parity.

For hedging with the ten options above, the influence parameter value $\xi = 0.0025$ yields a mean of -0.0581 and a standard deviation of 0.0228, which are better results than can be achieved when only the five original options are used. In this case, put-call parity implies that anything achievable with the put, call and underlying can be accomplished with any two of these three instruments. Consequently, the hedger should include both puts and calls of similar strikes in the hedge portfolio, as this will tend to reduce transaction costs.

8.4.4 Utilizing a More Realistic Model of Bid-Ask Spreads

The constant relative bid-ask spread assumption is clearly deficient: out-of-the-money options tend to have higher relative bid-ask spreads than in-the-money options. Consider, for example, the 22Oct2005 option prices for Amazon.com, Inc. (AMZN) taken during trading on August 10, 2005; this data is presented in Table 8.5. We use this Amazon.com option data to create a representative model of the relative bid-ask spread as a function of moneyness. The relative bid-ask spread for the puts and calls is found via

$$\text{Relative Bid-Ask Spread} = \frac{\text{Dollar Spread}}{\text{Midpoint Price}} = 2 \times \frac{\text{Ask - Bid}}{\text{Bid + Ask}}\,.$$

Strike	Moneyness $\frac{K}{S}$	Calls			Puts		
		Bid	Ask	Relative Spread	Bid	Ask	Relative Spread
\$20.00	0.4437	\$25.10	\$25.30	0.0079	\$ 0.05	\$ 0.05	0.0000
\$22.50	0.4991	\$22.70	\$22.80	0.0044	\$ 0.05	\$ 0.05	0.0000
\$25.00	0.5546	\$20.20	\$20.30	0.0049	\$ 0.05	\$ 0.05	0.0000
\$27.50	0.6100	\$17.70	\$17.80	0.0056	\$ 0.05	\$ 0.05	0.0000
\$30.00	0.6655	\$15.30	\$15.40	0.0065	\$ 0.05	\$ 0.10	0.6667
\$32.50	0.7209	\$12.80	\$13.00	0.0155	\$ 0.10	\$ 0.15	0.4000
\$35.00	0.7764	\$10.40	\$10.50	0.0096	\$ 0.15	\$ 0.20	0.2857
\$37.50	0.8319	\$ 8.10	\$ 8.20	0.0123	\$ 0.35	\$ 0.40	0.1333
\$40.00	0.8873	\$ 6.00	\$ 6.20	0.0328	\$ 0.70	\$ 0.75	0.0690
\$42.50	0.9428	\$ 4.10	\$ 4.30	0.0476	\$ 1.30	\$ 1.35	0.0377
\$45.00	0.9982	\$ 2.60	\$ 2.65	0.0190	\$ 2.20	\$ 2.30	0.0444
\$47.50	1.0537	\$ 1.50	\$ 1.55	0.0328	\$ 3.60	\$ 3.70	0.0274
\$50.00	1.1091	\$ 0.75	\$ 0.85	0.1250	\$ 5.40	\$ 5.50	0.0183
\$55.00	1.2201	\$ 0.20	\$ 0.25	0.2222	\$ 9.90	\$10.00	0.0101
\$60.00	1.3310	\$ 0.05	\$ 0.10	0.6667	\$14.80	\$15.00	0.0134

TABLE 8.5: *Option price data used to generate the relative bid-ask spread curves of Figure 8.6. The data is for 22Oct2005 puts and calls on Amazon.com, Inc. (AMZN), taken during trading on August 10, 2005. The spot value of the underlying is* \$45.08.

The discrete data is first smoothed using a simple moving average method, and the complete relative bid-ask spread curves for the calls and puts in Figure 8.6 are formed by linear interpolation and nearest neighbour extrapolation of this smoothed data. Note that the first four data points for the puts are discarded.

We have flat-topped the curves of Figure 8.6 in order to avoid unrealistically large values for the relative bid-ask spread. Options with large spreads will not be selected for the hedge portfolio (or will have very small weights), so as long as the transaction costs are included in the objective function, the precise form of the bid-ask spread for far out-of-the-money options should not be very important.

Up to this point, only options of strike $K = [70, 80, 90, 100, 110, 120, 130]$ have been

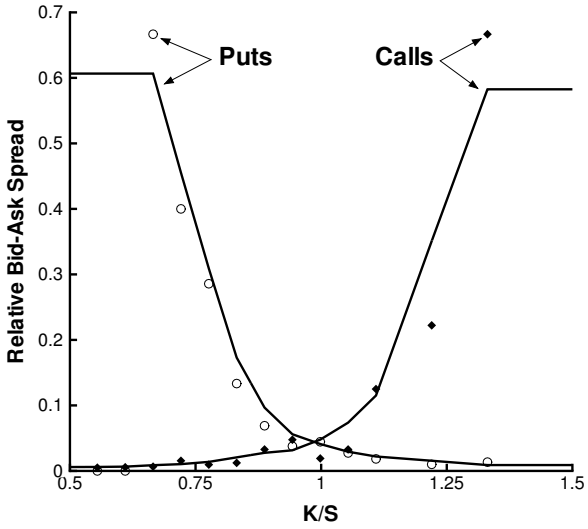


FIGURE 8.6: *Relative bid-ask spread curves drawn from market data. The option price data used to generate the curves is for 22Oct2005 puts and calls on Amazon.com, Inc. (AMZN), taken during trading on August 10, 2005. This data is in Table 8.5.*

used to hedge the one-year European straddle, regardless of the value of the underlying. More realistically, a wide range of options may be available for hedging. Of course, some of these will not contribute to a significant reduction of jump risk, and so will not be used in substantial amounts. Furthermore, the transaction costs associated with certain option positions may be prohibitive, and consequently these would be avoided. The objective in our optimization problem is designed to deal with these two facets of hedging.

Consider the following scenario: in addition to using the relative bid-ask curves of Figure 8.6, access to a wider range of options is allowed, namely all puts and calls with strikes from \$10 to \$200 in increments of \$10 are available. As such, the linear system used to determine the hedge portfolio weights has dimension 42×42. Table 8.6 contains representative results for different numbers of options in the hedging portfolio. Note that six hedging options corresponds to puts and calls of strike $[90, 100, 110]$, ten represents strikes of $[80, 90, 100, 110, 120]$, while fourteen represents $[70, 80, 90, 100, 110, 120, 130]$. We find that

No. of Hedging Options	Influence Parameter ξ	Mean	Standard Deviation	Quantiles	
				0.2%	99.8%
6	0.0001	-0.06	0.05	-0.22	0.21
10	0.001	-0.06	0.02	-0.13	0.02
14	0.0075	-0.07	0.02	-0.12	-0.01
40	0.02	-0.08	0.02	-0.13	-0.02

TABLE 8.6: *Hedging results, in terms of the relative profit and loss, for the European example with the relative bid-ask curves of Figure 8.6. The hedge weights are chosen by solving the optimization* (8.3) *for the given ξ. The target option is a one-year European straddle. The hedge is rebalanced every 0.025 years.*

a portfolio with a large number of hedging instruments does not outperform a hedge with a smaller number, as long as the smaller hedge contains those short-term options that are best at replicating the target straddle position, i.e. calls and puts with strikes near \$100. The hedging results of Table 8.6 indicate our procedure can indeed successfully handle a more realistic model of bid-ask spreads.

8.4.5 An American Example

To demonstrate the applicability of our proposed technique to path-dependent options, we consider hedging an American put over its lifetime. The same $\mathbb{P}$ and $\mathbb{Q}$ parameters as in Table 7.1 are employed, except that a higher jump arrival rate of $\lambda^{\mathbb{P}} = 0.1$ is used in the simulations. The American put to be hedged has a strike of \$100, a half-year maturity, and is initially at the money. American calls, with an initial maturity of three months and strikes from $K = \$10$ to $K = \$200$ in increments of \$10, are available as hedging instruments. Quarter-year American puts, with strikes from $K = \$10$ to $K = \$100$ in increments of \$10, are also used. At $t = 0.25$ all hedging options are replaced. As with the American hedging example from the previous chapter, we must be mindful of the possibility of early exercise for the American puts.

A total of 250,000 simulations are carried out. The hedge is regularly rebalanced at

No. of Hedging Options	Influence Parameter ξ	Mean	Standard Deviation	Quantiles	
				0.2%	99.8%
0 (delta hedge)	n.a.	0.05	1.05	-8.94	0.71
5	0.0075	-0.06	0.02	-0.14	0.03
8	0.03	-0.06	0.02	-0.14	-0.01
11	0.2	-0.07	0.02	-0.14	-0.02
30	0.05	-0.08	0.02	-0.16	-0.02

TABLE 8.7: *Hedging results, in terms of the relative profit and loss, when hedging a half-year American put ($K = \$100$) over its lifetime. The hedge is rebalanced every 0.025 years, with weights that are chosen by solving the optimization* (8.3) *for the given ξ. The relative bid-ask curves of Figure 8.6 are used to set the transaction costs.*

intervals of 0.025 years, and is also rebalanced if a hedging put is removed due to early exercise. To incorporate transaction costs, the bid-ask spread model of Figure 8.6 is used. Some representative results are presented in Table 8.7. Similar to the previous example, using all thirty available options does not outperform a hedge that contains fewer instruments. Note the hedge compositions are the same as in Table 8.6, only now puts with strikes above \$100 are excluded due to early exercise provisions. For the case of eight hedging options, the distribution of the relative P&L is presented in Figure 8.7.

The hedging results demonstrate that, for this American put, we can simultaneously protect our position from jumps without incurring prohibitive transaction costs.

Remark 8.2 (Optimality of the Hedging Strategy). *Our hedging strategy is local in time, as it is only concerned with the instantaneous state of the overall hedged position. In general, it will not be globally optimal. As noted previously, though, solving the full stochastic control problem would be computationally infeasible. We expect that our hedging results can certainly be improved upon. However, even our (non-optimal) hedging results clearly demonstrate that the use of a dynamic hedge containing traded options is a viable technique for minimizing jump risk.*

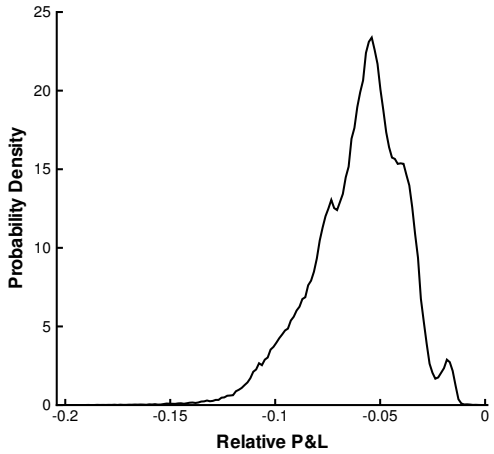


FIGURE 8.7: *Distribution of the relative profit and loss for the American hedging example when eight options are used to hedge the half-year American put over its lifetime. The influence parameter $\xi = 0.03$, and the hedge is rebalanced every 0.025 years. The relative bid-ask curves of Figure 8.6 are used to set the transaction costs.*

8.5 Summary

We solve an optimization problem at each hedge rebalance time to minimize a linear combination of a jump risk and transaction cost penalty function, while simultaneously imposing delta neutrality. Under our assumptions, simulations of our dynamic hedging strategy show the following:

- If the hedge portfolio is determined solely on the basis of minimizing jump risk (and ignoring transaction costs), the results are worse than simple delta hedging (which is itself quite poor). This is in accordance with conventional wisdom, which states that hedging with options is too expensive.

- On the other hand, if both jump risk and transaction costs are included in the objective function, our dynamic strategy is effective. In many cases only a small amount of buying and selling takes place while, at the same time, the overall position is protected

against jumps. The standard deviation of the relative P&L is much reduced compared to simple delta hedging.

Using a bid-ask spread model which captures the gross features of observed market prices (i.e. out-of-the-money options have larger relative spreads than near-the-money options) forces our strategy to reduce trading costs, while still minimizing jump risk.

Chapter 9

Hedging under a Lévy Process

In this chapter, we show how our dynamic procedure for hedging under jump diffusion may be extended to a more general Lévy process: imposing delta neutrality will eliminate the instantaneous diffusion risk and hedge the small jumps, while the large jumps are treated as before by minimizing an integral.

9.1 Introduction

Over the course of the last three chapters we have been concerned with hedging under the jump-diffusion process, where the return on the underlying asset is the superposition of a deterministic drift, Brownian motion, and compound Poisson process. Although this is considered a representative stochastic process for finance applications, it is merely one member of a broader class of stochastic processes called *Lévy processes.* The defining characteristic of a Lévy process is that they have independent and identically distributed increments. While Brownian motion and the Poisson process are the most well-known examples, other stochastic processes that have recently been employed in a financial setting, such as in the CGMY [16, 73] and log stable [18] models, are also Lévy processes.

Some Lévy processes exhibit an infinite number of jumps over every time period (almost surely). It is therefore not clear if our hedging method developed for use with jump diffusion—where the number of jumps in any time interval is finite—can be adapted for hedging under such a process. However, Lévy processes may be accurately approximated by a jump diffusion [3], which suggests our hedging strategy may be applicable. This chapter will demonstrate that this is indeed the case.

We will first introduce the main concepts associated with Lévy processes; two of the fundamental results, namely the Lévy-Itô Decomposition and Itô's formula, are covered in Appendix C. Next, assuming the return of the underlying asset follows a Lévy process, we will derive a representation of the instantaneous risk for a hedged position. Using this representation, the mean-variance hedging procedure developed by Cont et al. [26] will be derived. Finally, we demonstrate how our dynamic strategy for hedging under jump diffusion may be adapted to hedge under any Lévy process.

9.2 Background on Lévy Processes

In finance applications, the return on an asset is often modelled as a Lévy process X_t, such that

$$S_t = S_0 \mathrm{e}^{X_t}. \tag{9.1}$$

In this section, we present the fundamentals of Lévy processes that will be needed to develop a dynamic hedging strategy based on an underlying that evolves according to (9.1). This material is drawn from Cont and Tankov [23], Protter [66], Sato [69], and Papapantoleon [62].

Lévy processes are stochastic processes whose increments are independent and identically distributed. This specification is made more precise with the following definition.

Definition 9.1. (Lévy Process) *A real-valued stochastic process* $\{X_t : t \geq 0\}$ *with* $X_0 = 0$ *is referred to as a Lévy process if:*

1. *It has independent increments: for all* $n \geq 1$ *with* $0 \leq t_0 < t_1 < t_2 < \ldots < t_n$*, the random variables* X_{t_0}, $X_{t_1} - X_{t_0}$, $X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ *are independent;*

2. *It has identically distributed (stationary) increments: the distribution of* $\{X_{s+t} - X_s : t \geq 0\}$ *does not depend on* s*;*

3. *It is stochastically continuous (discontinuities are not predictable): for all* $\epsilon > 0$*,* $P\big[|X_{t+s} - X_s| > \epsilon\big] \to 0$ *as* $t \to 0$*;*

4. *It is* **càdlàg***: right-continuous with left limits;*

Although a Lévy process can assume values in $\mathbb{R}^d$, we shall only be concerned with one-dimensional processes. Every Lévy process is the superposition of a linear drift, Brownian motion and jump component (formally stated as the Lévy–Itô Decomposition; see Appendix C), and is identified by its characteristic triplet (ψ, σ, ν). The deterministic drift is given by ψt, while σZ_t represents the Brownian motion. The Lévy measure ν governs the distribution of jumps.

Definition 9.2. (Lévy Measure) *For a Lévy process* X_t*, the Lévy measure* $\nu(\cdot) \in \mathbb{R}$ *is defined as*

$$\nu(A) \;=\; \mathbb{E}\Big[\#\big\{t \in [0,1] \;\; \textit{such that} \;\; \Delta X_t \in A \;\; \textit{and} \;\; \Delta X_t \neq 0\big\}\Big],$$

where $A \subset \mathbb{R}$ *and* ΔX_t *represents the jump size.*

The Lévy measure gives the average number of jumps over unit time whose amplitudes are in A, where A is a subset of the real line. The condition $\nu(\mathbb{R}) < \infty$ characterizes a Lévy

process as having *finite activity*, and in this case

$$g(y) = \frac{\nu(y)}{\nu(\mathbb{R})}$$

is the probability density of the jump amplitudes. If $\nu(\mathbb{R}) = \infty$, the Lévy process exhibits *infinite activity*: there are expected to be an infinite number of infinitesimally small jumps (and a finite number of non-infinitesimal jumps) over any time interval. The Lévy measure must satisfy[1]

$$\int_{|y|\leq 1} y^2 \nu(y)\, dy \;<\; \infty \qquad \text{and} \qquad \int_{|y|\geq 1} \nu(y)\, dy \;<\; \infty\,. \tag{9.2}$$

The conditions in (9.2) are important for many theoretical results, such as the Lévy-Itô Decomposition.

Each realized path of a Lévy process may be considered as an event ω drawn from the set of all possible events Ω. The Lévy measure ν encapsulates the average behaviour of jumps over many realized paths. For example, after many simulations of a finite activity Lévy process, the average number of observed jumps per unit time will approach $\nu(\mathbb{R})$. On the other hand, the jump measure $J_X(\omega, A \times [0,t])$ is a *random measure*, as it depends on the specific realization ω from Ω. For a given $\omega \in \Omega$ (which corresponds to a specific path X), $J_X(\omega, A \times [0,t])$ provides the number of jumps that occur between $[0,t]$ whose size is in A. Furthermore, quantities involving jumps may be represented by integrating against the jump measure. For instance, consider the following example from Cont and Tankov

[1]The cutoff value $\Upsilon = 1$ is arbitrary, and may in fact be set to any $\Upsilon \in (0, \infty)$. The first condition ensures that a singularity at $y = 0$ will nonetheless integrate y^2. The second condition guarantees that, over any finite time period, the Lévy process does not exhibit an infinite number of jumps of non-infinitesimal size.

[23]: the expression

$$[X,X]_T^{dc} \;=\; \sum_{t\in[0,T]} [\Delta X_t]^2 \tag{9.3}$$

represents the sum of the squared jump sizes (known as the discontinuous quadratic variation) over $[0,T]$ for a specific path. With respect to the jump measure, (9.3) can be expressed as

$$\int_0^T \int_{\mathbb{R}} y^2 \, J_X(dydt)\,, \tag{9.4}$$

where reference to ω is dropped in the interest of notational simplicity. The integral in (9.4) gives the discontinuous quadratic variation for one specific path, while

$$\int_0^T \int_{\mathbb{R}} y^2 \, \nu(y)dydt$$

gives the average over many paths. As stated in [23], the jump measure J_X: "contains all information about the discontinuities (jumps) of the process X: it tells when the jumps occur and how big they are."

The compensated jump measure along a specific path ω,

$$\tilde{J}_X(dydt) \;=\; J_X(dydt) - \nu(y)dydt\,,$$

is important for infinite activity processes, due to the fact that the sum of the infinitely many small jumps may not converge. By integrating against the compensated jump measure, convergence is obtained. Note that if the jump process is of *finite variation*, i.e.

$$\int_{|y|\le 1} |y| \, \nu(y)dy < \infty\,,$$

the use of the compensated measure is not required.

9.3 The Instantaneous Risk of a Hedged Position

The asset price is assumed to follow

$$S_t \;=\; S_0 \mathrm{e}^{X_t}\,, \tag{9.5}$$

where the return X_t is a Lévy process. If we assume that $\mathbb{E}[X_t] < \infty$, then, using the Lévy–Itô Decomposition from Theorem C.1 in Appendix C, dX_t may be expressed as

$$dX_t \;=\; \psi dt \;+\; \sigma dZ_t \;+\; \int_{\mathbb{R}} y\, \tilde{J}_X(dydt)\,,$$

where (ψ, σ, ν) is the characteristic triplet associated with this decomposition.

We will assume that

$$\int_{|y|\geq 1} \mathrm{e}^{2y}\, \nu(y) dy \;<\; \infty\,, \tag{9.6}$$

which is equivalent to assuming $\mathbb{E}\big[S_t^2\big] < \infty$ [23]. Applying Itô's formula from Theorem C.2 in Appendix C to (9.5) gives[2]

$$dS_t \;=\; \underbrace{\sigma S_{t^-}\, dZ_t \;+\; \int_{\mathbb{R}} S_{t^-}(\mathrm{e}^y - 1)\, \tilde{J}_X(dydt)}_{\text{martingale: } \mathbb{E}[\cdot]=0} \;+\; \underbrace{\hat{\psi} S_{t^-}\, dt}_{\substack{\text{continuous, finite} \\ \text{variation process}}}\,, \tag{9.7}$$

where t^- is the instant immediately before time t, and y is the size of a jump in the return.

[2]Condition (9.6) allows us to express the jump component of S_t in terms of the compensated jump measure only. This is known as the martingale-drift decomposition [23].

The deterministic constant

$$\hat{\psi} \;=\; \psi + \frac{\sigma^2}{2} + \int_{\mathbb{R}} \left(\mathrm{e}^y - 1 - y\right) \nu(y) dy \,, \tag{9.8}$$

is associated with the drift, where ψ comes from the characteristic triplet of the Lévy process X_t.

To derive the instantaneous risk under a Lévy process, we proceed as in Section 6.1 when a jump diffusion was treated, only now the more general jump measure is used. Again consider the overall hedged position

$$\Pi \;=\; -V(S,t) \;+\; eS \;+\; \vec{\phi} \cdot \vec{I}(S,t) \;+\; B(t) \,, \tag{9.9}$$

where $S = S_0 \mathrm{e}^{X_t}$ is the asset price at time t; therefore, both V and $\vec{I}$ are functions of the return X_t. We assume that all payoff functions $h(S_T)$ are Lipschitz, which means there exists a $\mathcal{C} > 0$ such that $|h(x) - h(y)| \leq \mathcal{C}|x - y|$ for all x, y.

The instantaneous change of Π w.r.t. the $\mathbb{P}$ measure is

$$d\Pi \;=\; -dV \;+\; edS \;+\; \vec{\phi} \cdot d\vec{I} \;+\; dB \,. \tag{9.10}$$

We assume that V is twice differentiable in S and once differentiable in t, and apply Itô's formula from Theorem C.2 of Appendix C to obtain

$$\begin{aligned} dV \;=\; \left[\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2}\right] dt \;+\; \left[\hat{\psi}^{\mathbb{P}} S \frac{\partial V}{\partial S} + \int_{\mathbb{R}} \left(\Delta V - \Delta S \frac{\partial V}{\partial S}\right) \nu^{\mathbb{P}}(y) dy\right] dt \\ +\; \sigma S \frac{\partial V}{\partial S} \, dZ^{\mathbb{P}} \;+\; \int_{\mathbb{R}} \Delta V \, \tilde{J}_X^{\mathbb{P}}(dydt) \,. \end{aligned} \tag{9.11}$$

where $\Delta S = S(\mathrm{e}^y - 1)$ and $\Delta V = V(S\mathrm{e}^y, t) - V(S,t)$, with S the current asset price and y the jump in the return.

The pricing PIDE for a European option under a Lévy process [23]

$$\frac{\partial V}{\partial t} \;+\; \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} \;=\; rV \;-\; rS\frac{\partial V}{\partial S} \;-\; \int_{\mathbb{R}} \left[\Delta V - \Delta S \frac{\partial V}{\partial S}\right] \nu^{\mathbb{Q}}(y)dy$$

can be substituted into (9.11) to yield

$$\begin{aligned} dV \;=\; & \left[rV - rS\frac{\partial V}{\partial S} - \int_{\mathbb{R}} \left(\Delta V - \Delta S \frac{\partial V}{\partial S}\right) \nu^{\mathbb{Q}}(y)dy\right] dt \\ & + \;\left[\hat{\psi}^{\mathbb{P}} S\frac{\partial V}{\partial S} + \int_{\mathbb{R}} \left(\Delta V - \Delta S \frac{\partial V}{\partial S}\right) \nu^{\mathbb{P}}(y)dy\right] dt \\ & + \;\sigma S\frac{\partial V}{\partial S}\, dZ^{\mathbb{P}} \;+\; \int_{\mathbb{R}} \Delta V \, \tilde{J}_X^{\mathbb{P}}(dydt)\,. \end{aligned} \tag{9.12}$$

When condition (9.6) holds and the payoff function $g(S_T)$ of the option is Lipschitz, Cont and Tankov [23] show that

$$\left(\sigma S\frac{\partial V}{\partial S}\right)^2 \;<\; \infty \tag{9.13}$$

and

$$\int_{\mathbb{R}} \left|\Delta V\right|^2 \nu^{\mathbb{P}}(y)dy \;=\; \int_{\mathbb{R}} \left|V(S\mathrm{e}^y, t) - V(S,t)\right|^2 \nu^{\mathbb{P}}(y)dy \;<\; \infty\,, \tag{9.14}$$

which implies both processes in the last line of (9.12) are martingales. Furthermore, conditions (9.13) and (9.14) also imply these processes satisfy the isometry relations [23]

$$\mathbb{E}^{\mathbb{P}}\left[\left(\sigma S\frac{\partial V}{\partial S} dZ^{\mathbb{P}}\right)^2\right] \;=\; \left(\sigma S\frac{\partial V}{\partial S}\right)^2 dt \tag{9.15}$$

and

$$\mathbb{E}^{\mathbb{P}}\left[\left|\int_{\mathbb{R}} \Delta V \, \tilde{J}_X^{\mathbb{P}}(dydt)\right|^2\right] \;=\; dt \int_{\mathbb{R}} \left|\Delta V\right|^2 \nu^{\mathbb{P}}(y)dy\,. \tag{9.16}$$

A representation similar to (9.12) holds for all of the hedging options in $\vec{I}$. Substituting the expressions for dV, $d\vec{I}$, and the martingale-drift decomposition of dS into (9.10) gives

$$\begin{aligned} d\Pi &= rdt\left[-V + \vec{\phi}\cdot\vec{I} + B\right] + rSdt\left[\frac{\partial V}{\partial S} - \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right] \\ &+ dt\int_{\mathbb{R}}\left[-\Delta V + \left(\frac{\partial V}{\partial S} - \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right)\Delta S + \vec{\phi}\cdot\Delta\vec{I}\right]\left(\nu^{\mathbb{P}}(y) - \nu^{\mathbb{Q}}(y)\right)dy \\ &+ \sigma S dZ^{\mathbb{P}}\left[-\frac{\partial V}{\partial S} + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right] + \hat{\psi}^{\mathbb{P}}Sdt\left[-\frac{\partial V}{\partial S} + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right] \\ &+ \int_{\mathbb{R}}\Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]\tilde{J}_X^{\mathbb{P}}(dydt)\,. \end{aligned} \tag{9.17}$$

The above is the instantaneous change of the overall hedged position, and thus represents the instantaneous risk of the position.

9.4 Mean-Variance Hedging with $\mathbb{E}^{\mathbb{P}} = \mathbb{E}^{\mathbb{Q}}$

The goal of mean-variance hedging is to solve

$$\min_{\{e_t,\vec{\phi}_t,v_0\}} \mathbb{E}^{\mathbb{P}}\big[\mathcal{H}_T - V_T\big]^2\,, \tag{9.18}$$

where $\mathcal{H}_T$ is the terminal value of the hedge portfolio, V_T is the payoff of the option being hedged, and $\{e_t, \vec{\phi}_t\}$ is the trading strategy over $[0,T]$ that minimizes the objective function. Also, v_0 is the initial capital. In general, mean-variance hedging is a difficult mathematical problem, as its objective is a global criterion achieved through continuous rebalancing. However, if we minimize the objective function in (9.18) w.r.t. $\mathbb{E}^{\mathbb{Q}}$ instead of $\mathbb{E}^{\mathbb{P}}$, this procedure simplifies to a local risk minimization, where the expected square of the instantaneous (i.e. local) hedging error is minimized. This case was considered by Cont et al. [26], who show that, when the expectation is taken w.r.t. $\mathbb{E}^{\mathbb{Q}}$, solving (9.18) is equivalent to solving

the local optimization

$$\{e^*, \vec{\phi}^*\} \;=\; \underset{\{e,\vec{\phi}\}}{\arg\min} \left\{ \sigma^2 S^2 \left[-\frac{\partial V}{\partial S} + e + \vec{\phi} \cdot \frac{\partial \vec{I}}{\partial S} \right]^2 \right. \\ \left. + \int_{\mathbb{R}} \left[-\Delta V + \left(e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \right) \right]^2 \nu^{\mathbb{Q}}(y) dy \right\} \quad (9.19)$$

at each instant. Furthermore, the initial capital

$$v_0 \;=\; \mathrm{e}^{-rT} \mathbb{E}^{\mathbb{Q}} \left[V_T | S_0 \right],$$

which is the option price $V(S_0, 0)$.

We now show that the objective in (9.19) can be obtained by minimizing the variance of $d\Pi$ in (9.17) w.r.t. $\mathbb{E}^{\mathbb{Q}}$. If $\mathbb{P} = \mathbb{Q}$ in (9.17), then the integral in the second line of (9.17) immediately vanishes. Also, since $\mathrm{e}^{-rt} S_t$ is a martingale under $\mathbb{Q}$, the constant $\hat{\psi}$ from (9.8) must be equal to the risk-free rate, i.e. $\hat{\psi}^{\mathbb{P}} = \hat{\psi}^{\mathbb{Q}} = r$. Consequently, the instantaneous change in the overall hedged position (9.17) may be expressed in the simplified form

$$d\tilde{\Pi} \;=\; \sigma S \left[-\frac{\partial V}{\partial S} + e + \vec{\phi} \cdot \frac{\partial \vec{I}}{\partial S} \right] dZ^{\mathbb{Q}} \;+\; \int_{\mathbb{R}} \left[-\Delta V + \left(e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \right) \right] \tilde{J}_X^{\mathbb{Q}}(dydt), \quad (9.20)$$

where the discounted portfolio value $\tilde{\Pi} = \mathrm{e}^{-rt}\Pi$ is used. Note our hedging procedure always assumes the initial capital is $V(S_0, 0)$, as this ensures the starting value of the overall hedged position is zero.

By assuming that all option payoffs are Lipschitz and (9.6) holds, it is straightforward to show (see [23])

$$\sigma^2 S^2 \left[-\frac{\partial V}{\partial S} + e + \vec{\phi} \cdot \frac{\partial \vec{I}}{\partial S} \right]^2 < \infty \quad \text{and} \quad \int_{\mathbb{R}} \left[-\Delta V + \left(e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \right) \right]^2 \nu^{\mathbb{Q}}(y) dy < \infty,$$

which implies that both processes in (9.20) are martingales and, furthermore, obey isometry relations analogous to (9.15) and (9.16). Note that

$$\text{VARIANCE}^{\mathbb{Q}}\big[d\tilde{\Pi}\big] \;=\; \mathbb{E}^{\mathbb{Q}}\big[d\tilde{\Pi}^2\big] - \underbrace{\Big(\mathbb{E}^{\mathbb{Q}}\big[d\tilde{\Pi}\big]\Big)^2}_{=0} \;=\; \mathbb{E}^{\mathbb{Q}}\big[d\tilde{\Pi}^2\big]\,.$$

Now, consider

$$\begin{aligned}
&\text{VARIANCE}^{\mathbb{Q}}\big[d\tilde{\Pi}\big] \;=\; \mathbb{E}^{\mathbb{Q}}\big[d\tilde{\Pi}^2\big] \\
&= 2\mathbb{E}^{\mathbb{Q}}\left[\left(\sigma S\left[-\frac{\partial V}{\partial S} + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right]dZ^{\mathbb{Q}}\right)\left(\int_{\mathbb{R}}\Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]\tilde{J}_X^{\mathbb{Q}}(dydt)\right)\right] \\
&\qquad + \;\mathbb{E}^{\mathbb{Q}}\left[\left(\sigma S\left[-\frac{\partial V}{\partial S} + e + \vec{\phi}\cdot\vec{I}_S\right]dZ^{\mathbb{Q}}\right)^2\right] \\
&\qquad + \;\mathbb{E}^{\mathbb{Q}}\left[\left(\int_{\mathbb{R}}\Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]\tilde{J}_X^{\mathbb{Q}}(dydt)\right)^2\right]\,. \qquad (9.21)
\end{aligned}$$

The cross term in (9.21) is zero since the components are independent and both martingales. Therefore, using the appropriate isometry relations, we get

$$\begin{aligned}
\text{VARIANCE}^{\mathbb{Q}}\big[d\tilde{\Pi}\big] \;=\; dt\Bigg(&\sigma^2S^2\left[-\frac{\partial V}{\partial S} + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\right]^2 \;+ \\
&\int_{\mathbb{R}}\Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]^2\nu^{\mathbb{Q}}(y)dy\Bigg)\,.
\end{aligned}$$

One can minimize $\text{VARIANCE}^{\mathbb{Q}}\big[d\tilde{\Pi}\big]$ over the next instant by a suitable choice of hedge weights $\{e, \vec{\phi}\}$, as in [26]. The optimal weights can be found by solving a linear system.

9.5 Hedging a Finite Activity Lévy Process

Assume the return process X_t for the asset has a diffusive component ($\sigma > 0$) and finite activity jumps ($\nu(\mathbb{R}) < \infty$). This process is a jump diffusion (as we have used the term), since the Lévy measure provides the intensity $\lambda = \nu(\mathbb{R})$ and distribution of amplitudes $g(y) = \frac{\nu(y)}{\nu(\mathbb{R})}$ that govern the compound Poisson process. We will show that, in this case, the instantaneous change in the overall hedged position (9.17) can be made equivalent to $d\Pi$ in (6.8), which was derived for a jump-diffusion process. As such, the procedure developed in Chapter 6 to hedge under jump diffusion may be used, without alteration, for any finite activity Lévy process.

We first impose delta neutrality to eliminate the instantaneous diffusion risk. Therefore, substituting the delta-neutral condition

$$-\frac{\partial V}{\partial S} \; + \; e \; + \; \vec{\phi} \cdot \frac{\partial \vec{I}}{\partial S} \; = \; 0 \tag{9.22}$$

into (9.17) eliminates both terms in the third line. Furthermore, the delta-neutral condition implies

$$rSdt \left(\frac{\partial V}{\partial S} - \vec{\phi} \cdot \frac{\partial \vec{I}}{\partial S} \right) \; = \; reSdt \, ,$$

which simplifies the first line of (9.17) to $r\Pi dt$. A similar substitution within the integral of the second line yields $dt \int_{\mathbb{R}} \left[-\Delta V + \left(e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \right) \right] \left(\nu^{\mathbb{P}}(y) - \nu^{\mathbb{Q}}(y) \right) dy$. Consequently, the instantaneous change in the value of the delta-neutral overall hedged position may now be expressed as

$$\begin{aligned} d\Pi \; = \; r\Pi dt \; + \; dt \int_{\mathbb{R}} \Big[-\Delta V + \left(e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \right) \Big] \Big(\nu^{\mathbb{P}}(y) - \nu^{\mathbb{Q}}(y) \Big) \, dy \\ + \; \int_{\mathbb{R}} \Big[-\Delta V + \left(e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \right) \Big] \tilde{J}_X^{\mathbb{P}}(dydt) \, . \end{aligned} \tag{9.23}$$

Since the jump component has finite activity (and hence finite variation), the compensated jump measure is not required. Substituting $\tilde{J}_X^{\mathbb{P}}(dydt) = J_X^{\mathbb{P}}(dydt) - \nu^{\mathbb{P}}(y)dydt$ into (9.23) gives

$$\begin{aligned} d\Pi \;=\; r\Pi dt \;+\; dt \int_{\mathbb{R}} \Big[\Delta V - \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]\,\nu^{\mathbb{Q}}(y)dy \\ +\; \int_{\mathbb{R}} \Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]\,J_X^{\mathbb{P}}(dydt)\,. \end{aligned}$$

Furthermore, $\nu^{\mathbb{Q}}(\mathbb{R}) = \lambda^{\mathbb{Q}}$ and $\nu^{\mathbb{Q}}(y) = \lambda^{\mathbb{Q}} g^{\mathbb{Q}}(y)$, where $\lambda^{\mathbb{Q}}$ is the risk-adjusted jump intensity and $g^{\mathbb{Q}}$ is the risk-adjusted distribution of jump amplitudes. Therefore

$$\begin{aligned} d\Pi \;=\; r\Pi dt \;+\; \lambda^{\mathbb{Q}}\,dt\,\mathbb{E}^{\mathbb{Q}}\Big[\Delta V - \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big] \\ +\; \int_{\mathbb{R}} \Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]J_X^{\mathbb{P}}(dydt)\,. \end{aligned} \tag{9.24}$$

We may interpret $J_X^{\mathbb{P}}$ as the random measure that governs the compound Poisson process controlling the arrival time of the jumps in the return and their associated size. Using the shorthand notation introduced in Section 2.3, $d\Pi$ in (9.24) may be expressed as

$$\begin{aligned} d\Pi \;=\; & r\Pi dt \\ & +\; \lambda^{\mathbb{Q}}\,dt\,\mathbb{E}^{\mathbb{Q}}\Big[\Delta V - \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big] \;+\; d\pi^{\mathbb{P}}\Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big], \end{aligned} \tag{9.25}$$

which is clearly the same expression as $d\Pi$ in (6.8), namely the instantaneous change in the overall hedged position under a jump-diffusion process.

A finite activity Lévy process may be hedged using the same procedure as developed in Chapter 6: delta neutrality is imposed to eliminate the diffusion risk, while the integral

$$\int_{\mathbb{R}} \Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]^2\, w(y)dy$$

is minimized at each rebalance time, where $w(y)$ is a weighting function.

9.6 Hedging an Infinite Activity Lévy Process

9.6.1 A Measure of the Jump Risk

Let X_t be an infinite activity Lévy process with triplet (ψ, σ, ν), where $S_t = S_0 \mathrm{e}^{X_t}$. We denote the change in portfolio value due to a jump $S \to S\mathrm{e}^y$ at time t as

$$\begin{aligned} \hbar_{\{S,t\}}(y) &= -\Delta V + e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \\ &= -\Big(V(S\mathrm{e}^y, t) - V(S,t)\Big) \; + \; eS\big(\mathrm{e}^y - 1\big) \; + \; \vec{\phi} \cdot \Big(\vec{I}(S\mathrm{e}^y, t) - \vec{I}(S,t)\Big)\,, \end{aligned}$$

where S is the current asset price and y is the jump in the return. Let $w(y)$ be a proper weighting function w.r.t. both $\nu^{\mathbb{P}}(y)$ and $\nu^{\mathbb{Q}}(y)$ (as defined in Definition 6.1 with $y = \log(J)$), and assume that

$$\sigma_w^2(\delta) \; = \; \int_{-\delta}^{\delta} y^2 \, w(y) dy \; < \; \infty \tag{9.26}$$

for $0 < \delta < \infty$.

We consider the expression

$$\int_{\mathbb{R}} \Big[-\Delta V + \big(e\Delta S + \vec{\phi} \cdot \Delta \vec{I}\big) \Big]^2 w(y) dy \; = \; \int_{\mathbb{R}} \big[\hbar_{\{S,t\}}(y)\big]^2 \; w(y) dy \,. \tag{9.27}$$

Analogous to the jump-diffusion case, the integral in (9.27) is a measure of the jump risk: forcing it to be zero will eliminate this risk. In most cases, however, this will be impossible, so the goal is to minimize it. To this end, the expression in (9.27) is broken up into the sum

of two integrals:

$$\int_{\mathbb{R}} \Big[-\Delta V + \big(e\Delta S + \vec{\phi}\cdot\Delta\vec{I}\big)\Big]^2 w(y)dy \;=\; \int_{|y|>\delta} \big[\hbar_{\{S,t\}}(y)\big]^2 w(y)dy + \int_{-\delta}^{\delta} \big[\hbar_{\{S,t\}}(y)\big]^2 w(y)dy\,. \quad (9.28)$$

The first integral on the right-hand side of (9.28), which can be expressed in terms of the truncated weighting function as

$$\int_{\mathbb{R}} \big[\hbar_{\{S,t\}}(y)\big]^2 w(y)\mathbf{1}_{|y|>\delta}\,dy\,, \quad (9.29)$$

is a measure of the risk associated with the non-infinitesimal jumps (of which there are a finite number over any instant). These jumps arrive at a rate of

$$\lambda^{\mathbb{P}}_{\text{trun}} \;=\; \nu^{\mathbb{P}}\big((-\infty,-\delta)\cup(\delta,\infty)\big) \quad (9.30)$$

with amplitudes distributed according to

$$f^{\mathbb{P}}_{\text{trun}}(y) \;=\; \frac{1}{\lambda^{\mathbb{P}}_{\text{trun}}}\nu^{\mathbb{P}}(y)\mathbf{1}_{|y|>\delta}\,. \quad (9.31)$$

Therefore, as with any jump diffusion, the "large" jumps are hedged by minimizing the integral in (9.29).

The second integral on the right-hand side of (9.28),

$$\int_{-\delta}^{\delta} \big[\hbar_{\{S,t\}}(y)\big]^2 w(y)dy\,, \quad (9.32)$$

is a measure of the risk associated with the small jumps (of which there are infinitely many

over any instant). The Maclaurin polynomial for $\hbar_{\{S,t\}}(y)$ on $[-\delta,\delta]$ is

$$\begin{aligned}\hbar_{\{S,t\}}(y) &= \hbar_{\{S,t\}}(0) + \hbar'_{\{S,t\}}(0)\cdot y + \frac{1}{2}\hbar''_{\{S,t\}}(\xi)\cdot y^2 \\ &= S\left(-\frac{\partial V}{\partial S}\bigg|_S + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\bigg|_S\right) y + \frac{1}{2}\hbar''_{\{S,t\}}(\xi)\cdot y^2\,, \end{aligned} \tag{9.33}$$

where ξ is a function of y, with $|\xi| < |y|$. By imposing the delta-neutral condition

$$-\frac{\partial V}{\partial S}\bigg|_S + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\bigg|_S = 0$$

in (9.33) and substituting the resulting Maclaurin polynomial

$$\hbar_{\{S,t\}}(y) = \frac{1}{2}\hbar''_{\{S,t\}}(\xi)\cdot y^2 \tag{9.34}$$

into (9.32), we get

$$\int_{-\delta}^{\delta}\left[\hbar_{\{S,t\}}(y)\right]^2 w(y)dy = \frac{1}{4}\int_{-\delta}^{\delta}\left(\hbar''_{\{S,t\}}(\xi)\right)^2 y^4\, w(y)dy\,. \tag{9.35}$$

The second derivative of $\hbar_{\{S,t\}}(y)$ is

$$\begin{aligned}\hbar''_{\{S,t\}}(y) = (S\mathrm{e}^y)\left(-\frac{\partial V}{\partial S}\bigg|_{S\mathrm{e}^y} + e + \vec{\phi}\cdot\frac{\partial \vec{I}}{\partial S}\bigg|_{S\mathrm{e}^y}\right) \\ + (S\mathrm{e}^y)^2\left(-\frac{\partial^2 V}{\partial S^2}\bigg|_{S\mathrm{e}^y} + \vec{\phi}\cdot\frac{\partial^2 \vec{I}}{\partial S^2}\bigg|_{S\mathrm{e}^y}\right). \end{aligned} \tag{9.36}$$

We have already assumed that $\frac{\partial V}{\partial S}$, $\frac{\partial \vec{I}}{\partial S}$, $\frac{\partial^2 V}{\partial S^2}$ and $\frac{\partial^2 \vec{I}}{\partial S^2}$ are continuous (when Itô's formula was used to derive (9.17)). Therefore, since any continuous function on a closed interval is bounded, $|\hbar''_{\{S,t\}}(y)|$ will be bounded on $[-\delta,\delta]$; we denote this upper bound as $H_{\{S,t\}}(\delta)$. Using the fact that $y^2 \le \delta^2$ on $[-\delta,\delta]$, a bound may be established for the integral in (9.35),

namely

$$\begin{aligned}\int_{-\delta}^{\delta}\left[\hbar_{\{S,t\}}(y)\right]^2 w(y)dy &\leq \frac{1}{4}\delta^2 H^2_{\{S,t\}}(\delta)\int_{-\delta}^{\delta} y^2\, w(y)dy \\ &= \frac{1}{4}\delta^2 H^2_{\{S,t\}}(\delta)\cdot\sigma_w^2(\delta)\,. \end{aligned} \tag{9.37}$$

Thus, assuming delta neutrality is imposed, the second integral on the right-hand side of (9.28) can be made small by an appropriate choice of $\delta > 0$.

The Lévy–Itô Decomposition tells us that the processes controlling the small jumps and large jumps are independent—imposing delta neutrality will hedge the small jumps (and eliminate the instantaneous diffusion risk), while the integral in (9.29) is minimized directly to hedge the large jumps.

9.6.2 Global Bound on the Hedging Error

Assume there are no transaction costs present. By imposing delta neutrality and making the jump risk sufficiently small at each instant, the terminal hedging error can be made arbitrarily small. The steps for establishing this global bound are analogous to those in Section 6.2, where a bound was established for the variance of the terminal hedging error under jump diffusion. Therefore, after stating the theorem, we will only outline the proof.

Theorem 9.1 (Variance of the Hedging Error Can be Made Arbitrarily Small). *Assume there exists a $\delta^* > 0$ such that, for all times t and all possible asset values S, the hedge portfolio weights $\{e, \vec{\phi}\}$ can be chosen to satisfy:*

(B1) *$w(y)$ is a proper weighting function with respect to both $\nu^{\mathbb{P}}(y)$ and $\nu^{\mathbb{Q}}(y)$;*

(B2) *The function $\hbar''_{\{S,t\}}(y)$ in (9.36) is bounded: $\hbar''_{\{S,t\}}(y) < H_{\{S,t\}}(\delta^*) < H^* < \infty$ for $y \in [-\delta^*, \delta^*]$;*

(B3) $\sigma_w^2(\delta^*) = \displaystyle\int_{-\delta^*}^{\delta^*} y^2\, w(y)dy < \epsilon_1$;

(B4) *The "large" jump risk is made small:* $\int_{\mathbb{R}} \left[\hbar_{\{S,t\}}(y)\right]^2 w(y)\mathbf{1}_{|y|>\delta^*}\, dy \;<\; \epsilon_2$*;*

(B5) *The moments* $\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t^4\right]$ *and* $\mathbb{E}_0^{\mathbb{P}}\left[S_t^6\right]$ *exist;*

(B6) *All option payoffs are Lipschitz;*

(B7) *Delta neutrality is imposed;*

Then, $\mathbb{E}\left[\tilde{\Pi}_T^2\right]$ *can be made small, i.e.* $\mathbb{E}\left[\tilde{\Pi}_T^2\right] = \mathcal{O}(\epsilon_1 + \epsilon_2)$.

Outline of Proof for Theorem 9.1. Itô's formula (in its integral form) applied to $\left(\tilde{\Pi}_t\right)^2$ yields

$$\begin{aligned}\tilde{\Pi}_t^2 \;=\; 2\int_0^t \mathrm{e}^{-ru}\Theta_u^{\mathbb{Q}}\tilde{\Pi}_u\, du \;+\; \int_0^t\int_{\mathbb{R}} \left(\tilde{\Pi}_u^2(S_u e^y) - \tilde{\Pi}_u^2(S_u)\right)\nu^{\mathbb{P}}(y)dy\, du \\ +\; \int_0^t\int_{\mathbb{R}} \left(\tilde{\Pi}_u^2(S_u e^y) - \tilde{\Pi}_u^2(S_u)\right)\tilde{J}_X^{\mathbb{P}}(dydu)\,, \quad (9.38)\end{aligned}$$

with

$$\Theta_u^{\mathbb{Q}} \;=\; \int_{\mathbb{R}} -\hbar_{\{S,t\}}(y)\,\nu^{\mathbb{Q}}(y)dy \quad (9.39)$$

and

$$\tilde{\Pi}_u^2(S_u e^y) - \tilde{\Pi}_u^2(S_u) \;=\; \mathrm{e}^{-2ru}\left[\hbar_{\{S,t\}}(y)\right]^2 + 2\mathrm{e}^{-ru}\tilde{\Pi}_u(S_u)\hbar_{\{S,t\}}(y)\,; \quad (9.40)$$

imposing delta neutrality has eliminated the diffusion term. Now, using assumptions $(B5)$ and $(B6)$, we can show

$$\mathbb{E}_0^{\mathbb{P}}\left[\int_0^t\int_{\mathbb{R}} \left(\tilde{\Pi}_u^2(S_u e^y) - \tilde{\Pi}_u^2(S_u)\right)^2 \nu^{\mathbb{P}}(y)dy\, du\right] \;<\; \infty\,,$$

which means that the last term in (9.38) is a martingale. In this case, the differential

equation

$$\frac{d\,\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t^2\right]}{dt} \;=\; 2\mathrm{e}^{-rt}\mathbb{E}_0^{\mathbb{P}}\left[\Theta_t^{\mathbb{Q}}\tilde{\Pi}_t\right] \;+\; \mathbb{E}_0^{\mathbb{P}}\left[\int_{\mathbb{R}}\left(\tilde{\Pi}_t^2(S_te^y)-\tilde{\Pi}_t^2(S_t)\right)\nu^{\mathbb{P}}(y)dy\right] \tag{9.41}$$

can be obtained.

As in Section 6.2, we can establish upper bounds on the expectations in (9.41). First, consider

$$\begin{aligned}
\mathbb{E}_0^{\mathbb{P}}\left[\left(\Theta_t^{\mathbb{Q}}\right)^2\right] &= \mathbb{E}_0^{\mathbb{P}}\left[\left(\int_{\mathbb{R}} -\hbar_{\{S,t\}}(y)\,\nu^{\mathbb{Q}}(y)dy\right)^2\right] \\
&= \mathbb{E}_0^{\mathbb{P}}\left[\left(\int_{-\delta^*}^{\delta^*}\hbar_{\{S,t\}}(y)\,\nu^{\mathbb{Q}}(y)dy \;+\; \int_{\mathbb{R}}\hbar_{\{S,t\}}(y)\nu^{\mathbb{Q}}(y)\mathbf{1}_{|y|>\delta^*}\,dy\right)^2\right] \\
&= \mathbb{E}_0^{\mathbb{P}}\left[\left(\int_{-\delta^*}^{\delta^*}\hbar_{\{S,t\}}(y)\,\nu^{\mathbb{Q}}(y)dy\right)^2\right] \;+\; \mathbb{E}_0^{\mathbb{P}}\left[\left(\int_{\mathbb{R}}\hbar_{\{S,t\}}(y)\nu^{\mathbb{Q}}(y)\mathbf{1}_{|y|>\delta^*}\,dy\right)^2\right] \\
&\quad +\; 2\mathbb{E}_0^{\mathbb{P}}\left[\left(\int_{-\delta^*}^{\delta^*}\hbar_{\{S,t\}}(y)\,\nu^{\mathbb{Q}}(y)dy\right)\left(\int_{\mathbb{R}}\hbar_{\{S,t\}}(y)\nu^{\mathbb{Q}}(y)\mathbf{1}_{|y|>\delta^*}\,dy\right)\right].
\end{aligned} \tag{9.42}$$

For the first expectation in (9.42), we use the Maclaurin polynomial (9.34) for $\hbar_{\{S,t\}}(y)$:

$$\begin{aligned}
\mathbb{E}_0^{\mathbb{P}}\left[\left(\int_{-\delta^*}^{\delta^*}\hbar_{\{S,t\}}(y)\,\nu^{\mathbb{Q}}(y)dy\right)^2\right] &= \mathbb{E}_0^{\mathbb{P}}\left[\left(\frac{1}{2}\int_{-\delta^*}^{\delta^*}\hbar''_{\{S,t\}}(\xi)y^2\,\nu^{\mathbb{Q}}(y)dy\right)^2\right] \\
\text{Bound on } |\hbar''_{\{S,t\}}(y)| \longrightarrow \quad &\leq \mathbb{E}_0^{\mathbb{P}}\left[\frac{1}{4}H_{\{S,t\}}^2(\delta^*)\left(\int_{-\delta^*}^{\delta^*}y^2\,\nu^{\mathbb{Q}}(y)dy\right)^2\right] \\
w(y) \text{ a proper weighting function, } (B1) \longrightarrow \quad &\leq \mathbb{E}_0^{\mathbb{P}}\left[\frac{1}{4}H_{\{S,t\}}^2(\delta^*)\left(\int_{-\delta^*}^{\delta^*}y^2\,w(y)dy\right)^2\right] \\
\text{Definition of } \sigma_w^2 \text{ in (9.26)} \longrightarrow \quad &= \mathbb{E}_0^{\mathbb{P}}\left[\frac{1}{4}\left(H_{\{S,t\}}(\delta^*)\sigma_w^2(\delta^*)\right)^2\right] \\
\begin{array}{c}\text{Conditioning on } S \\ \text{and assumptions } \textit{(B2)} \text{ and } \textit{(B3)}\end{array} \longrightarrow \quad &\leq \frac{1}{4}\left(H^*\epsilon_1\right)^2.
\end{aligned} \tag{9.43}$$

For the second expectation in (9.42), we recognize $\frac{1}{\lambda^{\mathbb{Q}}_{\text{trun}}}\nu^{\mathbb{Q}}(y)\mathbf{1}_{|y|>\delta^*}$ as a probably density, with $\lambda^{\mathbb{Q}}_{\text{trun}} = \nu^{\mathbb{Q}}\big((-\infty,-\delta^*)\cup(\delta^*,\infty)\big)$. Therefore

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}_0\left[\left(\int_{\mathbb{R}}\hbar_{\{S,t\}}(y)\nu^{\mathbb{Q}}(y)\mathbf{1}_{|y|>\delta^*}\,dy\right)^2\right] &= \mathbb{E}^{\mathbb{P}}_0\left[\left(\lambda^{\mathbb{Q}}_{\text{trun}}\int_{\mathbb{R}}\hbar_{\{S,t\}}(y)\left\{\frac{1}{\lambda^{\mathbb{Q}}_{\text{trun}}}\nu^{\mathbb{Q}}(y)\mathbf{1}_{|y|>\delta^*}\right\}dy\right)^2\right] \\
\mathbb{E}[X]^2\le\mathbb{E}[X^2]\longrightarrow \quad &\le \mathbb{E}^{\mathbb{P}}_0\left[(\lambda^{\mathbb{Q}}_{\text{trun}})^2\int_{\mathbb{R}}\big(\hbar_{\{S,t\}}(y)\big)^2\left\{\frac{1}{\lambda^{\mathbb{Q}}_{\text{trun}}}\nu^{\mathbb{Q}}(y)\mathbf{1}_{|y|>\delta^*}\right\}dy\right] \\
&= \mathbb{E}^{\mathbb{P}}_0\left[\lambda^{\mathbb{Q}}_{\text{trun}}\int_{\mathbb{R}}\big(\hbar_{\{S,t\}}(y)\big)^2\nu^{\mathbb{Q}}(y)\mathbf{1}_{|y|>\delta^*}\,dy\right] \\
\begin{array}{c}w(y)\text{ a proper}\\ \text{weighting function, }\textit{(B1)}\end{array}\longrightarrow \quad &\le \mathbb{E}^{\mathbb{P}}_0\left[\lambda^{\mathbb{Q}}_{\text{trun}}\int_{\mathbb{R}}\big(\hbar_{\{S,t\}}(y)\big)^2 w(y)\mathbf{1}_{|y|>\delta^*}\,dy\right] \\
\begin{array}{c}\text{Conditioning on } S\\ \text{and assumption }\textit{(B4)}\end{array}\longrightarrow \quad &\le \lambda^{\mathbb{Q}}_{\text{trun}}\epsilon_2\,. \qquad (9.44)
\end{aligned}$$

The results (9.43) and (9.44) can be used in (9.42), along with the Cauchy–Schwarz inequality, to yield

$$\mathbb{E}^{\mathbb{P}}_0\left[\left(\Theta^{\mathbb{Q}}_t\right)^2\right] \le \frac{1}{4}(H^*\epsilon_1)^2 + \lambda^{\mathbb{Q}}_{\text{trun}}\epsilon_2 + H^*\epsilon_1\sqrt{\lambda^{\mathbb{Q}}_{\text{trun}}\epsilon_2}\,.$$

Hence, the first expectation in the differential equation (9.41) is bounded by

$$\begin{aligned}
2\mathrm{e}^{-rt}\mathbb{E}^{\mathbb{P}}_0\left[\Theta^{\mathbb{Q}}_t\tilde{\Pi}_t\right] &\le 2\sqrt{\mathbb{E}^{\mathbb{P}}_0\left[\left(\Theta^{\mathbb{Q}}_t\right)^2\right]\mathbb{E}^{\mathbb{P}}_0\big[\tilde{\Pi}^2_t\big]} \\
&\le 2\sqrt{\frac{1}{4}(H^*\epsilon_1)^2 + \lambda^{\mathbb{Q}}_{\text{trun}}\epsilon_2 + H^*\epsilon_1\sqrt{\lambda^{\mathbb{Q}}_{\text{trun}}\epsilon_2}}\cdot\sqrt{\mathbb{E}^{\mathbb{P}}_0\big[\tilde{\Pi}^2_t\big]}\,. \qquad (9.45)
\end{aligned}$$

A similar analysis can be carried out on the second term in the differential equa-

tion (9.41), where

$$\begin{aligned}\mathbb{E}_0^{\mathbb{P}}\left[\int_{\mathbb{R}}\left(\tilde{\Pi}_t^2(S_te^y)-\tilde{\Pi}_t^2(S_t)\right)\nu^{\mathbb{P}}(y)dy\right] &= \mathbb{E}_0^{\mathbb{P}}\left[\int_{\mathbb{R}}\left(\mathrm{e}^{-rt}\hbar_{\{S,t\}}(y)\right)^2\nu^{\mathbb{P}}(y)dy\right] \\ &\quad + \mathbb{E}_0^{\mathbb{P}}\left[2\mathrm{e}^{-rt}\tilde{\Pi}_t(S_t)\int_{\mathbb{R}}\hbar_{\{S,t\}}(y)\nu^{\mathbb{P}}(y)dy\right],\end{aligned} \tag{9.46}$$

is obtained using the representation in (9.40). For the first integral on the right-hand side of (9.46):

$$\begin{aligned}\mathbb{E}_0^{\mathbb{P}}\left[\int_{\mathbb{R}}\left(\mathrm{e}^{-rt}\hbar_{\{S,t\}}(y)\right)^2\nu^{\mathbb{P}}(y)dy\right] &\leq \mathbb{E}_0^{\mathbb{P}}\left[\int_{\mathbb{R}}\left[\hbar_{\{S,t\}}(y)\right]^2 w(y)dy\right] \\ \text{Break integral into two} \longrightarrow &= \mathbb{E}_0^{\mathbb{P}}\left[\int_{-\delta^*}^{\delta^*}\left[\hbar_{\{S,t\}}(y)\right]^2 w(y)dy\right] \\ &\quad + \mathbb{E}_0^{\mathbb{P}}\left[\int_{\mathbb{R}}\left[\hbar_{\{S,t\}}(y)\right]^2 w(y)\mathbf{1}_{|y|>\delta^*}\,dy\right] \\ \begin{array}{c}\text{By bound in (9.37)}\\ \text{and assumption } \textit{(B4)}\end{array} \longrightarrow &\leq \mathbb{E}_0^{\mathbb{P}}\left[\frac{1}{4}(\delta^*)^2 H_{\{S,t\}}^2(\delta^*)\sigma_w^2(\delta^*) \;+\; \epsilon_2\right] \\ \begin{array}{c}\text{Conditioning on } S\\ \text{and assumptions } \textit{(B2)} \text{ and } \textit{(B3)}\end{array} \longrightarrow &\leq \frac{1}{4}(\delta^*)^2(H^*)^2\epsilon_1 \;+\; \epsilon_2\,.\end{aligned} \tag{9.47}$$

For the second integral on the right-hand side of (9.46), the bound is established exactly as for the first expectation in the differential equation (9.41):

$$\begin{aligned}\mathbb{E}_0^{\mathbb{P}}\left[2\mathrm{e}^{-rt}\tilde{\Pi}_u(S_t)\int_{\mathbb{R}}\hbar_{\{S,t\}}(y)\nu^{\mathbb{P}}(y)dy\right] &\leq \\ 2\sqrt{\frac{1}{4}\left(H^*\epsilon_1\right)^2 \;+\; \lambda_{\text{trun}}^{\mathbb{P}}\epsilon_2 \;+\; H^*\epsilon_1\sqrt{\lambda_{\text{trun}}^{\mathbb{P}}\epsilon_2}}\cdot\sqrt{\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t^2\right]}\end{aligned} \tag{9.48}$$

Therefore, the second term in the differential equation (9.41) is bounded by

$$\mathbb{E}_0^{\mathbb{P}}\left[\int_{\mathbb{R}}\left(\tilde{\Pi}_t^2(S_te^y)-\tilde{\Pi}^2(S_t)\right)\nu^{\mathbb{P}}(y)dy\right] \;\leq\; \frac{1}{4}(\delta^*)^2(H^*)^2\epsilon_1 \;+\; \epsilon_2 \\ +\; 2\sqrt{\frac{1}{4}\left(H^*\epsilon_1\right)^2 \;+\; \lambda^{\mathbb{P}}_{\text{trun}}\epsilon_2 \;+\; H^*\epsilon_1\sqrt{\lambda^{\mathbb{P}}_{\text{trun}}\epsilon_2}}\cdot\sqrt{\mathbb{E}_0^{\mathbb{P}}\left[\tilde{\Pi}_t^2\right]} \tag{9.49}$$

The bounds (9.45) and (9.49) can be used in the differential equation (9.41) to yield a differential inequality. We can then proceed as in Section (6.2) to show the variance of Π_T can be made arbitrarily small. □

9.7 Summary

In this chapter, we have laid down the theoretical framework for hedging under any Lévy process. Finite activity processes are essentially jump diffusions, and hence may be hedged using the dynamic strategy developed in Chapter 6. For infinite activity processes, imposing delta neutrality will not only eliminate the diffusion risk, but also hedge the risk associated with small jumps. The larger jumps, which are governed by a truncated Lévy measure of the form $\nu^{\mathbb{P}}(x)\mathbf{1}_{|x|>\delta}$, are handled in a similar way as the jumps in the jump-diffusion model. This means that, essentially, the same procedure can be used to hedge both finite and infinite activity processes. This result is not surprising, since infinite activity jump processes can be approximated to arbitrary precision as the sum of a finite activity jump process and a diffusion [3].

Chapter 10

Semi-Static Hedging Under Jump Diffusion

With semi-static hedging, the hedge portfolio weights are chosen with the goal of replicating the value of the target option at a future time (which usually corresponds to the expiry of the shorter term options used for hedging). It is a buy-and-hold strategy that requires infrequent rebalancing. In this chapter, we motivate the optimization that is performed at each rebalance time of the semi-static procedure, and show how the objective function to be minimized is a simple modification of the objective function that lies at the heart of our dynamic strategy for hedging under jump diffusion. We carry out simulations of the semi-static strategy for a European option, and investigate the use of different weighting functions within the optimization.

10.1 The Semi-Static Objective Function

If an option $V(S,t)$ and the portfolio $\mathcal{H}(S,t)$ have the same value in all future states at time T, i.e.

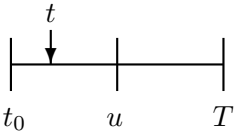
$$\forall\, S_T \quad V(S_T,T) \;=\; \mathcal{H}(S_T,T)\,,$$

then the no-arbitrage principle implies that the two positions will have the same value for all times before T and, as such, are indistinguishable: $\mathcal{H}(S,t)$ is said to replicate $V(S,t)$. Semi-static hedging is based on the principle that a continuum of calls and puts may be used to replicate a more complex or longer-dated option [13, 17], and hence form a perfect hedge. Since the use of a continuum of options is clearly impossible, the idea is to choose a finite number of option positions that match the contract being hedged in some best sense. With semi-static hedging, positions are updated infrequently.

Carr and Wu [19] derive a relationship in which a long-term European option may be replicated by holding a continuum of shorter-term options of the same type. For European call options, this spanning relationship is

$$\underbrace{C^*(S,t;K,T)}_{\substack{\text{Long-term call}\\ \text{with strike } K}} \;=\; \int_0^\infty w(\kappa)\,\underbrace{C(S,t;\kappa,u)}_{\substack{\text{Short-term}\\ \text{call with}\\ \text{strike } \kappa}}\,d\kappa\,, \tag{10.1}$$

where C^* and C are the respective time-t option values corresponding to asset price S, $T \geq t_0$ is the maturity of the long-term call option C^*, and $u \in [t_0, T]$ is the expiry of the short-term call options. The relationship between the various times is represented in the following schematic:



The required weight to hold in each of the short-term options is given by

$$w(\kappa) = \frac{\partial^2}{\partial S^2} C^*(S, u; K, T)\bigg|_{S=\kappa},$$

which is the gamma of the long-term call option C^* at time u; these weights are independent of t and hence may be readily calculated at some initial time t_0. Furthermore, with these weights the spanning relation (10.1) holds for all times in the interval $t \in [t_0, u]$, such that at time u the gamma-weighted portfolio of short-term options completely replicates the European call C^*. Therefore, to hedge the long-term call option C^* over the time period $[t_0, u]$, the gamma weighted portfolio of shorter-term calls is formed at time t_0 and held until they expire at u. To hedge beyond time u, a new set of short-term options should be used to form the replicating portfolio. Note the spanning relationship (10.1) holds for any Markovian pricing model (i.e. where dS only depends on the current value of S), as it is derived from the Breeden–Litzenberger result [13]

$$p^{\mathbb{Q}}(S_T|S_t) \;=\; \mathrm{e}^{-r(T-t)} \frac{\partial^2 C}{\partial K^2}(S_t, t; K, T)$$

that links the risk-adjusted transition density $p^{\mathbb{Q}}$ to the partial derivatives of the call value.

A continuum of options obviously cannot be employed, so the authors appeal to Gauss-Hermite quadrature to find the hedge portfolio weights ω_j and strikes κ_j that allow a finite number N of short-term call options to approximate the replicating portfolio. This quadrature rule is

$$C^*(S, t; K, T) \;=\; \int_0^\infty w(\kappa) C(S, t; \kappa, u)\, d\kappa \;\approx\; \sum_{j=1}^{N} \omega_j C(S, t; \kappa_j, u)\,,$$

where an explicit relationship exists between the weights and strikes $\{\omega_j, \kappa_j\}_{j=1}^N$ and the

standard Gauss-Hermite quadrature weights and nodes $\{v_j, x_j\}_{j=1}^N$ from the canonical form

$$\int_{-\infty}^{\infty} f(x)e^{-x^2}dx \approx \sum_{j=1}^{N} v_j f(x_j)\,.$$

Carr and Wu [19] provide evidence supporting the efficacy of this method for hedging under jump diffusion.

For a fixed number of hedging options N, the procedure of Carr and Wu [19] has $2N$ degrees of freedom, as both the strikes and corresponding weights of these hedging options are set by the quadrature rule. However, since only a finite number of call options are available in the market, the procedure may compute required strikes that cannot be purchased. As such, it is more practical to impose the strikes that can be used, and then calculate the associated hedge weights that best replicate—according to some criterion—the contract being hedged. We consider this variant of semi-static hedging.

We assume the target option $V(S,t)$ is European, and consider a market with no transaction costs. The hedge portfolio again consists of a cash component $B(t)$, e units in the underlying, and $\vec{\phi}$ units in N hedging options $\vec{I}$ whose characteristics (e.g. strikes) are completely specified beforehand. The value of the overall hedged position at time t for an asset value of S_t is

$$\Pi(S_t,t) \;=\; -\,V(S_t,t) \;+\; eS_t \;+\; \vec{\phi}\cdot\vec{I}(S_t,t) \;+\; B(t)\,.$$

The cost of establishing the overall hedged position time at t is

$$B(t) \;=\; V(S_t,t) \;-\; eS_t \;-\; \vec{\phi}\cdot\vec{I}(S_t,t)\,, \tag{10.2}$$

which is invested ($B(t) > 0$) or borrowed ($B(t) < 0$) at the risk-free rate r. At a time $t+\delta t$ in the future, where δt is a non-infinitesimal amount of time, the value of the overall hedged

position is

$$\Pi(S_{t+\delta t}, t+\delta t) \;=\; -V(S_{t+\delta t}, t+\delta t) \;+\; eS_{t+\delta t} \;+\; \vec{\phi}\cdot\vec{I}(S_{t+\delta t}, t+\delta t) \;+\; \mathrm{e}^{r\delta t}B(t)\,. \tag{10.3}$$

Now, we would like $\Pi(S_{t+\delta t}, t+\delta t)$ to be zero for all possible future states $S_{t+\delta t}$, as this signifies a perfect hedge exists over δt. This will most likely be impossible, so

$$\min_{\{e,\vec{\phi}\}} \mathbb{E}^{\mathbb{P}}\Big[\big(\Pi(S_{t+\delta t}, t+\delta t)\big)^2\Big] \tag{10.4}$$

is solved. When the expression for $B(t)$ in (10.2) is substituted into (10.3), the optimization problem (10.4) may be expressed as

$$\min_{\{e,\vec{\phi}\}} \mathbb{E}^{\mathbb{P}}\left[F(e,\vec{\phi})^2\right] \tag{10.5}$$

with

$$\begin{aligned} F(e,\vec{\phi}) \;=\; &-\Big(V(S_{t+\delta t}, t+\delta t) - \mathrm{e}^{r\delta t}V(S_t,t)\Big) \;+\; e\Big(S_{t+\delta t} - \mathrm{e}^{r\delta t}S_t\Big) \\ &+\; \vec{\phi}\cdot\Big(\vec{I}(S_{t+\delta t}, t+\delta t) - \mathrm{e}^{r\delta t}\vec{I}(S_t,t)\Big)\,. \end{aligned} \tag{10.6}$$

The expectation in (10.5) is taken with respect to the transition probability density $p(S_{t+\delta t}|S_t)$ under the $\mathbb{P}$ measure. He et al. [40] explore this strategy, and show it is effective for hedging under a jump-diffusion model with lognormally distributed jumps. In this case, the transition density can be expressed in closed form [51] as

$$p(S_T|S_0) \;=\; \frac{\mathrm{e}^{-\lambda^{\mathbb{P}}T}}{\sqrt{2\pi}S_T}\sum_{n=0}^{\infty}\frac{\left(\lambda^{\mathbb{P}}T\right)^n \exp\left\{-\frac{\left(\log\left(\frac{S_T}{S_0}\right)-\left(\alpha^{\mathbb{P}}-\frac{\sigma^2}{2}-\lambda^{\mathbb{P}}\kappa^{\mathbb{P}}\right)T-n\mu^{\mathbb{P}}\right)^2}{2\left(\sigma^2T+n(\gamma^{\mathbb{P}})^2\right)}\right\}}{n!\sqrt{\sigma^2T+n\left(\gamma^{\mathbb{P}}\right)^2}}\,; \tag{10.7}$$

this is referred to as Merton's transition density. When $\lambda^{\mathbb{P}} = 0$ (no jumps), the distribution in (10.7) collapses to the lognormal density that characterizes the Black–Scholes model. The expectation in (10.5) effectively enters into the optimization as a weighting function. Analogous to the dynamic hedging strategy, lack of knowledge pertaining to the $\mathbb{P}$ measure will often make it impossible to use the associated transition density as the weighting function.

10.2 The Link Between the Dynamic and Semi-Static Strategies

Recall that the objective for our dynamic hedging procedure may be written as

$$\min_{\{e,\vec{\phi}\}} \int_0^\infty \left[-\Delta V + \left(e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \right) \right]^2 W(J)\, dJ\,, \tag{10.8}$$

where $\Delta V = V(JS) - V(S)$, $\Delta \vec{I} = \vec{I}(JS) - \vec{I}(S)$ and $\Delta S = S(J-1)$ represent the changes due to a jump of size J. Since the goal of the dynamic procedure is to minimize the instantaneous jump risk, all option values in (10.8) come from the time of the rebalance. The weighting function $W(J)$ is tied to our knowledge (or lack of knowledge) of the distribution of jump sizes under the $\mathbb{P}$ measure.

Instead of considering J as the instantaneous jump size, for the purposes of semi-static hedging we define it as the relative change of the asset price between times t and $t + \delta t$; that is

$$J = \frac{S_{t+\delta t}}{S_t}\,.$$

This new interpretation of J allows the function (10.6) used in the semi-static procedure to be expressed in the same form as in (10.8), such that the optimization to be solved at a

rebalance time t is

$$\min_{\{e,\vec{\phi}\}} \int_0^\infty \left[-\Delta V + \left(e\Delta S + \vec{\phi} \cdot \Delta \vec{I} \right) \right]^2 W_{se}(J)\, dJ \tag{10.9}$$

with

$$\Delta V = V_{t+\delta t}(JS_t) - e^{r\delta t} V_t(S_t), \quad \Delta \vec{I} = \vec{I}_{t+\delta t}(JS_t) - e^{r\delta t} \vec{I}_t(S_t), \quad \Delta S = S_t(J - e^{r\delta t}) \tag{10.10}$$

and $W_{se}(J)$ the weighting function. Note the presence of the subscripts related to time in (10.10): since a semi-static hedge is concerned with instrument values both now and in the future, this time-stamp is necessary. The role of the weighting function $W_{se}(J)$ in (10.9) is now (ideally) played by the transition density from S_t to $S_{t+\delta t}$ (defined in terms of $J = \frac{S_{t+\delta t}}{S_t}$) under the $\mathbb{P}$ measure.

As illustrated above, it is essentially the same optimization problem that is the basis for both the dynamic and semi-static hedging procedures. As such, the technology developed to implement the dynamic hedging strategy may be used, with only minor changes, for the semi-static procedure. To this end, the method of Lagrange multipliers is employed to find the optimal solution to (10.9), although this usually will be an unconstrained optimization since linear equality constraints, such as the imposition of delta neutrality, are generally not present. Analogous to the dynamic strategy (see (7.6)), the linear system resulting from

the application of Lagrange multipliers will have entries of the form $(\Delta X, \Delta Y)$, given by

$$\begin{aligned}
(\Delta X, \Delta Y) &= \int_0^\infty \big[\Delta X \Delta Y\big] W_{se}(J)\, dJ \\
&= \int_0^\infty \Big[X_{t+\delta t}(JS_t) - e^{r\delta t} X_t(S_t)\Big]\Big[Y_{t+\delta t}(JS_t) - e^{r\delta t} Y_t(S_t)\Big] W_{se}(J)\, dJ \\
&= e^{2r\delta t} X_t(S_t) Y_t(S_t) \int_0^\infty W_{se}(J)\, dJ \;+\; \overbrace{\int_0^\infty X_{t+\delta t}(JS_t) Y_{t+\delta t}(JS_t) W_{se}(J)\, dJ}^{\text{precomputed for time } t+\delta t} \\
&\qquad - \; e^{r\delta t} Y_t(S_t) \underbrace{\int_0^\infty X_{t+\delta t}(JS_t) W_{se}(J)\, dJ}_{\text{precomputed for time } t+\delta t} \\
&\qquad - \; e^{r\delta t} X_t(S_t) \underbrace{\int_0^\infty Y_{t+\delta t}(JS_t) W_{se}(J)\, dJ}_{\text{precomputed for time } t+\delta t} \;.
\end{aligned}$$

The correlation-type integrals are precomputed as before (see Appendix B); this should entail considerably less work as compared to the dynamic strategy, since the integrals need only be calculated for the typically infrequent times when the semi-static hedge is to be rolled over.

When computing the semi-static hedge weights at the rebalance time t, it is simply a matter of forming the linear system in the same fashion as for the dynamic hedge. The only differences are: (i) the correlation-type integrals come from the time $t + \delta t$ in the future; (ii) the hedging instrument values from time t are adjusted by the exponential factor $\mathrm{e}^{r\delta t}$; and (iii) the delta-neutral constraint is not imposed

10.3 Hedging Simulations

To illustrate the semi-static procedure, the same hedging example as in Section 7.3 is considered: the target option is a one-year European straddle with strike \$100 that exists in a jump-diffusion market with lognormally distributed jumps. The pricing $\mathbb{Q}$ measure

Semi-static Hedging Test	Mean	Standard Deviation	Quantiles	
			0.2%	99.8%
Augmented Dynamic (A)	0.00	0.01	-0.05	0.06
Augmented Dynamic (B)	0.00	0.01	-0.05	0.06
He et al. [40]	0.00	0.01	-0.05	0.05

TABLE 10.1: *Statistical measures of the relative profit and loss when hedging a one-year European straddle ($K = \$100$) under jump diffusion using the semi-static procedure. The weighting function used in the optimization* (10.9) *is the transition density under* $\mathbb{P}$*, given in* (10.7)*. The hedge is formed at* $t = 0$ *and then rolled over at* $t = 0.25$*; the hedging horizon is half a year. The results corresponding to (A) and (B) are found using two different sets of 500,000 paths. For the* $\alpha\%$ *quantile, approximately* $\alpha\%$ *of the 500,000 simulations resulted in a relative P&L less than the reported amount.*

and real-world $\mathbb{P}$ measure that characterize the market are given in Table 7.1. The target option is initially at the money (i.e. $S_0 = 100$), and is to be hedged over a half-year time horizon. The underlying, plus three-month call options with strikes in intervals of \$5, are used as the hedging instruments. At each rebalance/rollover time t, the call options with strikes closest to $[0.8, 0.9, 1.0, 1.1, 1.2]S_t$ are considered as the most liquid and thus used in the hedge. The hedge is formed at $t_0 = 0$ by solving (10.9) with $t = 0$ and $t + \delta t = 0.25$. This hedge portfolio is held until $t_1 = 0.25$, at which point it is rolled over: the hedge weights for the new three-month options are computed by solving (10.9) with $t = 0.25$ and $t + \delta t = 0.5$.

This example is the same as the semi-static test case considered in He et al. [40]; their simulation set #1 corresponds to using the transition density (10.7) as the weighting function $W_{se}(J)$. Our first goal is to carry out this specific hedging experiment in order to confirm that our dynamic hedging framework can indeed accommodate the semi-static procedure. We execute our "augmented dynamic" version of semi-static hedging twice, using two different sets of 500,000 paths. The results, along with those for simulation set #1 from He et al. [40], are presented in Table 10.1 in the form of statistical measures of the relative profit and loss.

The results in Table 10.1 are very close to each other, indicating that the technology developed for the dynamic procedure can indeed be used to carry out semi-static hedging. Now, to investigate the effect of the weighting function on the profit & loss, we consider hedging the straddle over the same 500,000 asset paths, with the only difference being the $W_{se}(J)$ employed in the optimization (10.9). The six simulation experiments are summarized in Table 10.2 and the results are presented in Table 10.3. Note that plots of the various weighting functions can be found in the bottom graph of each panel in Figure 10.1—interpreting the top graph of these panels is the subject of the following section.

The hedging results for the first three cases are quite similar, which is not surprising since the weighting functions are much alike—the volatility is the primary influence on the shape of these transition densities, and this value is close for all three cases, i.e.

$$(1)\,\sigma^{\mathbb{P}} = 0.2, \quad (2)\,\sigma = 0.226, \quad (3)\,\sigma^{\mathbb{Q}} = 0.2\,.$$

The hedging results for the first three tests are quite good, exhibiting a zero mean and small standard deviation. The Black–Scholes transition density used as the weighting function in the fourth test case is totally inappropriate for our market, as the drift is set unreasonably high and the extremely low volatility leads to a highly peaked weighting function. The results are very poor for this case, as protection is only afforded for a relatively narrow band of future asset prices.

If the weighting function utilized in the fourth test case was our best guess for the transition density under the $\mathbb{P}$ measure, we may pay a significant price for being wrong, as losses from the left tail of the P&L distribution are substantial (e.g. the 0.2% quantile is -2.08). As with the dynamic strategy, using a uniform-like weighting function encapsulates a lack of knowledge pertaining to the $\mathbb{P}$ measure; an example of a uniform-like weighting function is plotted in Figure 7.1. In the case of semi-static hedging, the uniform-like weighting function

Test Number	Weighting Function	Summary
(1)	Merton Transition Density Under $\mathbb{P}$	The transition density from S_t to $S_{t+\delta t}$ for a jump-diffusion model with lognormally distributed jumps, under the $\mathbb{P}$ measure. The density is given by (10.7) with $J = \log\left(\frac{S_{t+\delta t}}{S_t}\right)$.
(2)	Black–Scholes Transition Density	The transition density from S_t to $S_{t+\delta t}$ for the Black–Scholes model (i.e. the density in (10.7) with $\lambda^{\mathbb{P}} = 0$). The drift $\alpha^{\mathbb{P}} = 0.18$ and volatility $\sigma = 0.226$ come from the $\mathbb{P}$ measure (see caption).
(3)	Merton Transition Density Under $\mathbb{Q}$	The transition density from S_t to $S_{t+\delta t}$ for a jump-diffusion model with lognormally distributed jumps, under the $\mathbb{Q}$ measure.
(4)	BAD Black–Scholes Transition Density	The transition density from S_t to $S_{t+\delta t}$ for the Black–Scholes model. The drift is set to $\alpha = 1.0$ and the volatility is $\sigma = 0.05$.
(5)	Uniform-like Density	90% of the mass is uniformly distributed between $J = 0.3$ and $J = 1.7$. The remaining 10% is in triangular tails that extend linearly down to zero.
(6)	BAD Uniform-like Density	90% of the mass is uniformly distributed between $J = 0.7$ and $J = 0.8$. The remaining 10% is in triangular tails that extend linearly down to zero.

TABLE 10.2: *Description of the different weighting functions used in the semi-static hedging experiments. The target option is a one-year European straddle with strike \$100. The value used as σ in test number (2) is the total volatility of the jump-diffusion process under $\mathbb{P}$: $\sqrt{(\sigma^{\mathbb{P}})^2 + \lambda^{\mathbb{P}}((\gamma^{\mathbb{P}})^2 + (\mu^{\mathbb{P}})^2)}$.*

Test Number	Weighting Function	Mean	Standard Deviation	Quantiles	
				0.2%	99.8%
(1)	Transition Density Under $\mathbb{P}$	0.00	0.01	-0.05	0.06
(2)	Black–Scholes Transition Density	0.00	0.01	-0.05	0.05
(3)	Transition Density Under $\mathbb{Q}$	0.00	0.01	-0.06	0.06
(4)	BAD Black–Scholes Transition Density	0.09	0.18	-2.08	0.20
(5)	Uniform-like Density	0.00	0.02	-0.18	0.08
(6)	BAD Uniform-like Density	-0.03	0.23	-1.20	0.35

TABLE 10.3: *Statistical measure of the relative profit and loss for hedging a one-year European straddle ($K = \$100$) under jump diffusion using the semi-static procedure. A variety of weighting functions are used, and are summarized in Table 10.2. The hedge is formed at $t = 0$ and then rolled over at $t = 0.25$; the hedging horizon is half a year.*

is set to non-zero for those future asset prices deemed plausible. Such a density is used for the fifth test case and the hedging results are good, especially given the fact this weighting function represents no knowledge of the real-world measure. The final test case employs a highly peaked uniform-like density and, similar to the use of the "bad" Black–Scholes transition density, the results are poor. The use of a highly-peaked weighting function is a not advisable unless one is sure it accurately captures the dynamics of the market: it is better to be over-cautious than over-confident.

10.4 Analysis of the Overall Hedged Position Value

For each of the 500,000 simulations, the same hedge portfolio is used for the first three months. It is therefore instructive to examine the value of the overall hedged position at $t = 0.25$ as a function of the relative change in the asset price $J = \frac{S_{0.25}}{100}$, as these plots provide a graphical representation of the protection against movements in the underlying. These plots also furnish us with a tangible way of viewing the effect of the various weighting functions. Six dual-plots are presented in Figure 10.1, corresponding to the simulation experiments summarized in Table 10.2. The top graph of each panel represents the value of

the overall hedged position at $t = 0.25$ as a function of $J = \frac{S_{0.25}}{100}$, while the bottom graph is the associated weighting function. The desired behaviour of the future value profile (i.e. the curve in the top graph) is for it to closely hug the zero-line.

The most striking characteristic of the top three panels is how similar the weighting functions—and hence the future value profiles—are. This is because the effect of the log-normal jumps on Merton's transition density, for both the $\mathbb{P}$ and $\mathbb{Q}$ measures we use, is small compared to the influence of the drift α and diffusive volatility σ. The left tail of the density corresponding to the $\mathbb{Q}$ measure is visible due to the more severe jump parameters (e.g. $\lambda^{\mathbb{Q}} = 0.1$ vs. $\lambda^{\mathbb{P}} = 0.023$), but is still insignificant compared to the central mass. Nonetheless, this more massive tail does provide somewhat better protection against downward jumps than is admitted by the $\mathbb{P}$ measure, as the jump-to-ruin (i.e. when $S_{0.25} = 0$) under $\mathbb{P}$ is a loss of \$3 while under $\mathbb{Q}$ it is approximately \$2.

For each of the first three cases, the protection for future asset values falling within the central mass of the weighting functions, approximately $75 \leq S_{0.25} \leq 135$, is very good. Furthermore, this protection persists for asset values assigned a weight close to zero. The fourth panel corresponds to a weighting function of similar shape, only it is more highly peaked with considerably less breadth. There is good protection for future asset values that lie within the central mass of the density, but this protection rapidly deteriorates.

If sufficient breadth of the weighting function is a desired characteristic, then a uniform-like weighting function can be used. Furthermore, detailed knowledge of the $\mathbb{P}$ measure is not required to construct such a weighting function. The future value profile in the fifth panel demonstrates the overall protection is quite good. There is more protection for the extreme values of $S_{0.25}$ than in cases (1), (2), and (3); however, this comes at the expense of a deterioration in the protection for the most likely asset values, $75 \leq S_{0.25} \leq 135$. The sixth panel demonstrates how the use of a narrow uniform-like density has the same pernicious effect as the highly-peaked lognormal density treated in the fourth panel.

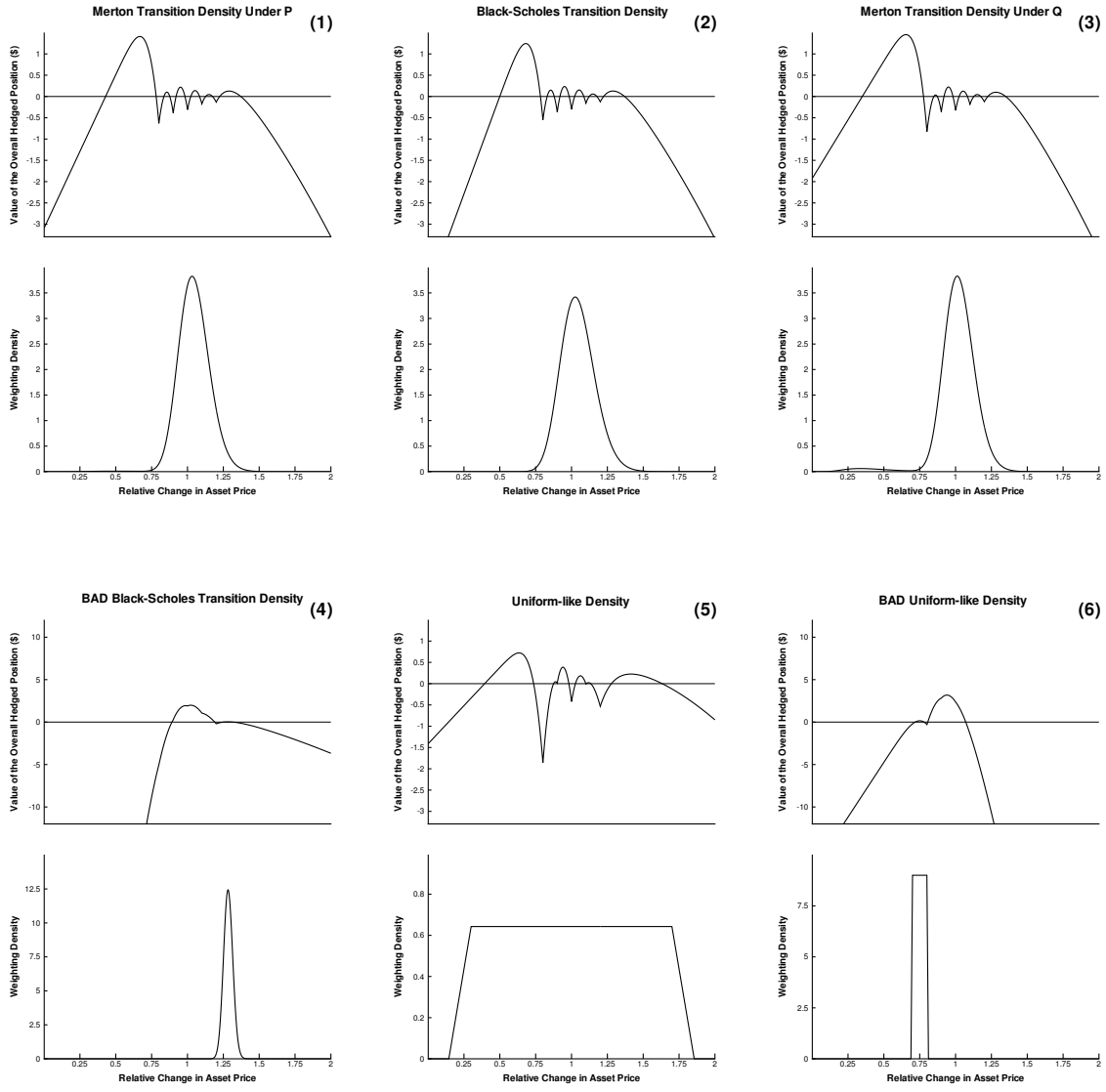


FIGURE 10.1: *TOP: Value of the overall hedged position at time* $t = 0.25$ *as a function of the relative change in the asset price* $J = \frac{S_{0.25}}{100}$. *BOTTOM: Associated weighting function. All six of the weighting functions have the properties of a probability density.*

10.5 Summary

In this chapter, we demonstrated how the optimization problem for a particular implementation of semi-static hedging is a modification of the optimization that is solved by our dynamic procedure for hedging under jump diffusion. This semi-static procedure was shown to be effective when hedging a one-year European straddle in a jump-diffusion market with lognormally distributed jumps. Furthermore, the hedging results are good for a variety of reasonable weighting functions, including a uniform-like density that assumes no knowledge of the $\mathbb{P}$ measure.

Chapter 11

Conclusions

The seller of a contingent claim must hedge the risk associated with their short position. In a market with jumps, delta hedging is not an adequate strategy for managing this risk: although a delta hedge will eliminate the diffusion risk related to the Brownian component of the underlying asset, it ignores the risk due to jumps. Moreover, this risk is often substantial, with the hedger subject to large losses if a jump occurs. In this thesis we have demonstrated that, to properly hedge a contingent claim in a market with jumps, more than the underlying asset must be used. In particular:

- For a regime-switching market, a finite number of hedging instruments are sufficient to eliminate all instantaneous risk from an option position—this perfect hedge leads to small hedging error when used as the basis of a discretely rebalanced strategy.

- In a jump-diffusion market, the continuum of possible jump sizes means the instantaneous risk can only be eliminated by using an infinite number of hedging instruments. Nonetheless, the jump risk can be treated in a systematic way, using a finite and practical number of hedging instruments: at each rebalance time, the hedge portfolio weights are chosen to simultaneously (i) eliminate the instantaneous diffusion risk by

imposing delta neutrality; and (ii) minimize an objective that is a linear combination of a jump risk and transaction cost penalty function. Hedging simulations indicate that this strategy provides sufficient protection against the diffusion and jump risk while not incurring large transaction costs.

- Our dynamic strategy for hedging under jump diffusion possesses many positive attributes, including:
  - It is easily applied to options with path-dependent features, including those with American-style early exercise provisions;
  - Exact knowledge of the $\mathbb{Q}$ and $\mathbb{P}$ measures is not required. In particular, we can use an estimate of the pricing measure obtained through calibration, and the weighting function employed to minimize the jump risk need not depend on the distributional properties of jump amplitudes under the real-world measure;
  - It can be used to hedge under any Lévy process.

Due to the transaction costs associated with a hedge portfolio containing several hedging options, it is a common belief that jumps cannot be dynamically hedged in practice. We hope that our results concerning a jump-diffusion market with transaction costs helps dispel the notion that hedging contingent claims with other options, as a means to treat the jump risk, is too expensive.

11.1 Future Work

Some directions for future research are:

- The unification of the regime-switching and jump-diffusion models considered in this thesis would be a stochastic volatility model where both the underlying asset and

volatility can jump, with the amplitudes for both jump types drawn from a continuum. It would be of interest to extend our hedging methodology to such a model.

- Since our hedging strategy treats the instantaneous risk, it is local in time. It would be desirable to develop globally optimal strategies by employing stochastic control theory.

Appendix A

Derivation of the Option Pricing PDEs for a Regime-Switching Model

A.1 Continuous Markov Chains

Consider a continuous-time stochastic process X_t that may assume values in the state space $\{1, 2, 3, \ldots, N\}$. The process X_t is said to be a continuous N-state Markov chain if it obeys the Markov property; i.e.

$$Prob\big[X_{t+s} = n|\mathcal{F}_t\big] \;=\; Prob\big[X_{t+s} = n|X_t\big] \qquad \text{for } \; s > t\,,$$

where $\mathcal{F}_t$ is the filtration (history) of the process. For every pair of states (i, j), there exists a transition probability

$$p_{ij}(t) \;=\; Prob\big[X_t = j|X_0 = i\big]\,,$$

such that $p_{ij}(t)$ is the probability the Markov chain is in state j at time t given it is initially in state i. If the Markov chain is homogeneous in time, then $Prob\big[X_{t+s} = j | X_t = i\big] = p_{ij}(s)$.

The continuous N-state Markov chain X_t is completely specified by its $N \times N$ rate matrix (generator) Q; we will treat the entries as independent of time [72]. All off-diagonal entries of Q, known as the transition intensities, are greater than or equal to zero, while the diagonal entries are given by

$$q_{ii} \;=\; -\sum_{\substack{j=1 \\ j\neq i}}^{N} q_{ij}\,.$$

For each state i, the transition probabilities $\vec{p}_i(t) = \big[p_{i1}(t), p_{i2}(t), \ldots, p_{iN}(t)\big]^T$ obey the system of differential equations

$$\begin{aligned} \frac{\mathrm{d}\vec{p}_i}{\mathrm{dt}} \;&=\; Q\,\vec{p}_i \\ \vec{p}_i(0) &= \hat{e}_i\,, \end{aligned} \tag{A.1}$$

where $\hat{e}_i$ is the i^{th} Euclidean basis vector. The solution to (A.1) is given by

$$\vec{p}_i(t) \;=\; \big[\exp(Qt)\big]^T \hat{e}_i \;=\; \left[I + \sum_{n=1}^{\infty} \frac{(Qt)^n}{n!}\right]_{i^{th}\,row}\,.$$

For small time Δt the infinite series may be truncated, yielding

$$p_{ij}(\Delta t) \;=\; \left\{ \begin{array}{ll} q_{ij}\Delta t + \mathcal{O}(\Delta t^2) & i \neq j \\ 1 + q_{ii}\Delta t + \mathcal{O}(\Delta t^2) & i = j \end{array} \right\}\,.$$

If the Markov chain is in state i, the probability of a transition to a new state occurring over time Δt is $v_i \Delta t + \mathcal{O}(\Delta t^2)$, where $v_i = -q_{ii}$. If a transition does occur, the probability of it taking place from state i to j is $\frac{q_{ij}}{v_i}$. These facts may be used to simulate a continuous Markov chain.

The increment of the Markov chain, $dX_t = X_{t+dt} - X_t$, will be a non-zero integer if there is a change in state over dt, and zero otherwise. For ease of exposition, we will use the term dX_{ij} to indicate the transition of the Markov chain over an instant dt, with

$$dX_{ij} \;=\; \begin{cases} 1 & \text{If the Markov chain transitions from state } i \text{ to } j \text{ over } dt \\ 0 & \text{Otherwise} \end{cases}$$

and

$$dX_{ij} \;=\; \begin{cases} 1 & \text{with probability} \quad q_{ij}dt + \delta_{ij} \\ 0 & \text{with probability} \quad 1 - q_{ij}dt - \delta_{ij} \;, \end{cases}$$

with the understanding that there can be only one transition over an instant; i.e.

$$Prob\big[\,(dX_{il} = 1) \cap (dX_{ik} = 1)\,\big] \;=\; 0 \qquad l \neq k\,.$$

A.2 Pricing in a Regime-Switching Market

A.2.1 Jumps May Only Occur with a Change in Regime

To derive the system of PDEs, we will consider the general N-state process

$$dS \;=\; a^i(S,t)\,dt \;+\; b^i(S,t)\,dZ \;+\; \sum_{j=1}^{N} S\big(\eta^{ij} - 1\big)\,dX_{ij}\,, \tag{A.2}$$

where the superscripts denote that the drift and diffusion coefficients are regime dependent. In addition, η^{ij} is the size of the jump in the underlying $S \to S\eta^{ij}$ that occurs when the Markov chain transitions from state i to state j, and is assumed to be a deterministic constant. It is also assumed that $\eta^{ii} = 1$, such that the asset price can only jump when there is a change in regime. Jumps in the underlying that occur without a change in regime

are easily incorporated by including a compound Poisson process in (A.2)—the system of pricing PDEs for such an example is developed in the following section.

A target option V is to be hedged with a portfolio of N instruments $\{F_n\}_{n=1}^N$ whose values depend on the underlying S, where the underlying may not be tradeable. A bank account B that grows at the risk-free rate r is also available. Consequently, the overall hedged position has value

$$\Pi \;=\; -V \;+\; \sum_{n=1}^{N} w_n F_n \;+\; B\,, \tag{A.3}$$

where w_n is the weight in each of the hedging instruments. Assuming the Markov chain is in state i, the instantaneous change in the value of the target option over an instant dt is found using Itô's lemma:

$$dV \;=\; \hat{\mu}^i\,dt \;+\; \hat{\sigma}^i\,dZ \;+\; \sum_{\substack{j=1\\ j\neq i}}^{N} \Delta V^{ij}\,dX_{ij} \tag{A.4}$$

where

$$\hat{\mu}^i \;=\; \left[\frac{\partial V^i}{\partial t} + a^i\frac{\partial V^i}{\partial S} + \frac{1}{2}(b^i)^2\frac{\partial^2 V^i}{\partial S^2}\right], \quad \hat{\sigma}^i \;=\; b^i\frac{\partial V^i}{\partial S}\,, \quad \Delta V^{ij} = V^j - V^i \tag{A.5}$$

and $V^k = V^k(S\eta^{ik}, t)$. Similarly for each of the hedging instruments:

$$dF_n \;=\; \bar{\mu}_n^i\,dt \;+\; \bar{\sigma}_n^i\,dZ \;+\; \sum_{\substack{j=1\\ j\neq i}}^{N} \Delta F_n^{ij}\,dX_{ij} \tag{A.6}$$

with

$$\bar{\mu}_n^i \;=\; \left[\frac{\partial F_n^i}{\partial t} + a^i\frac{\partial F_n^i}{\partial S} + \frac{1}{2}(b^i)^2\frac{\partial^2 F_n^i}{\partial S^2}\right], \quad \bar{\sigma}_n^i \;=\; b^i\frac{\partial F_n^i}{\partial S}\,, \quad \Delta F_n^{ij} = F_n^j - F_n^i \tag{A.7}$$

and $F_n^k = F_n^k(S\eta^{ik}, t)$.

Using the expressions for dV in (A.4) and dF_n in (A.6), the instantaneous change in the overall hedged position is

$$\begin{aligned} d\Pi &= -dV + \sum_{n=1}^{N} w_n\, dF_n + dB \\ &= \left[-\hat{\mu}^i + \sum_{n=1}^{N} w_n \bar{\mu}_n^i + rB\right] dt + \left[-\hat{\sigma}^i + \sum_{n=1}^{N} w_n \bar{\sigma}_n^i\right] dZ \\ &\qquad + \sum_{\substack{j=1 \\ j\neq i}}^{N} \left[-\Delta V^{ij} + \sum_{n=1}^{N} w_n \Delta F_n^{ij}\right] dX_{ij}\,. \end{aligned} \tag{A.8}$$

To eliminate the diffusion risk and the regime-switching risk from (A.8), a total of N linear equations need to be satisfied, namely

$$\left.\begin{array}{lll} \displaystyle\sum_{n=1}^{N} w_n \bar{\sigma}_n^i = \hat{\sigma}^i & & \\ \end{array}\right\} \begin{array}{c}\text{One Equation} \\ \text{(Diffusion Risk)}\end{array}$$

$$\left.\begin{array}{lll} \displaystyle\sum_{n=1}^{N} w_n \Delta F_n^{ij} = \Delta V^{ij} & j = 1, 2, \ldots, N \\ & j \neq i \end{array}\right\} \begin{array}{c} N-1 \text{ Equations} \\ \text{(Regime-Switching Risk)}\,. \end{array}$$

Since all the risk has been eliminated from the overall hedged position, its value must grow at the risk-free rate, implying $d\Pi = r\Pi dt$. Using the (deterministic) $d\Pi$ obtained from (A.8) and the expression for Π in (A.3), the linear equation

$$\left.\begin{array}{lll} \displaystyle\sum_{n=1}^{N} w_n \left(\bar{\mu}_n^i - rF_n^i\right) = \left(\hat{\mu}^i - rV^i\right) & & \end{array}\right\} \begin{array}{c}\text{One Equation} \\ \text{(No-Arbitrage Condition)}\end{array}$$

must be satisfied in order to prevent the existence of arbitrage opportunities. As such, we

have a linear system consisting of $N+1$ equations in N unknowns:

$$\begin{bmatrix} \bar{\sigma}_1^i & \bar{\sigma}_2^i & \bar{\sigma}_3^i & \cdots & \bar{\sigma}_N^i \\ \Delta F_1^{i1} & \Delta F_2^{i1} & \Delta F_3^{i1} & \cdots & \Delta F_N^{i1} \\ \Delta F_1^{i2} & \Delta F_2^{i2} & \Delta F_3^{i2} & \cdots & \Delta F_N^{i2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta F_1^{i(i-1)} & \Delta F_2^{i(i-1)} & \Delta F_3^{i(i-1)} & \cdots & \Delta F_N^{i(i-1)} \\ \Delta F_1^{i(i+1)} & \Delta F_2^{i(i+1)} & \Delta F_3^{i(i+1)} & \cdots & \Delta F_N^{i(i+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta F_1^{iN} & \Delta F_2^{iN} & \Delta F_3^{iN} & \cdots & \Delta F_N^{iN} \\ \bar{\mu}_1^i - rF_1^i & \bar{\mu}_2^i - rF_2^i & \bar{\mu}_3^i - rF_3^i & \cdots & \bar{\mu}_N^i - rF_N^i \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ \vdots \\ w_{N-3} \\ w_{N-2} \\ w_{N-1} \\ w_N \end{bmatrix} = \begin{bmatrix} \hat{\sigma}^i \\ \Delta V^{i1} \\ \Delta V^{i2} \\ \vdots \\ \Delta V^{i(i-1)} \\ \Delta V^{i(i+1)} \\ \vdots \\ \Delta V^{iN} \\ \hat{\mu}^i - rV^i \end{bmatrix} \tag{A.9}$$

The system in (A.9) is over specified: in order for it to have a solution, the equations must be linearly dependent. Consider the linear combination

$$\Lambda^i \times EQ_1 \; - \; \lambda_{i1}^{\mathbb{Q}} \times EQ_2 \; - \; \lambda_{i2}^{\mathbb{Q}} \times EQ_3 \; - \; \ldots \; - \; \lambda_{iN}^{\mathbb{Q}} \times EQ_N \; - \; EQ_{N+1}\,, \tag{A.10}$$

where EQ_j refers to the j^{th} row of the matrix in (A.9). The expression in (A.10) must be zero component-wise for some choice of Λ^i and $\{\lambda_{ij}^{\mathbb{Q}}\}_{j=1,j\neq i}^{N}$. This gives rise to a new linear

system:

$$\begin{bmatrix} \bar{\sigma}_1^i & \Delta F_1^{i1} & \Delta F_1^{i2} & \cdots & \Delta F_1^{i(i-1)} & \Delta F_1^{i(i+1)} & \cdots & \Delta F_1^{iN} \\ \bar{\sigma}_2^i & \Delta F_2^{i1} & \Delta F_2^{i2} & \cdots & \Delta F_2^{i(i-1)} & \Delta F_2^{i(i+1)} & \cdots & \Delta F_2^{iN} \\ \bar{\sigma}_3^i & \Delta F_3^{i1} & \Delta F_3^{i2} & \cdots & \Delta F_3^{i(i-1)} & \Delta F_3^{i(i+1)} & \cdots & \Delta F_3^{iN} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \bar{\sigma}_N^i & \Delta F_N^{i1} & \Delta F_N^{i2} & \cdots & \Delta F_N^{i(i-1)} & \Delta F_N^{i(i+1)} & \cdots & \Delta F_N^{iN} \\ \hat{\sigma}^i & \Delta V^{i1} & \Delta V^{i2} & \cdots & \Delta V^{i(i-1)} & \Delta V^{i(i+1)} & \cdots & \Delta V^{iN} \end{bmatrix} \begin{bmatrix} \Lambda^i \\ -\lambda_{i1}^{\mathbb{Q}} \\ -\lambda_{i2}^{\mathbb{Q}} \\ \vdots \\ -\lambda_{i(i-1)}^{\mathbb{Q}} \\ -\lambda_{i(i+1)}^{\mathbb{Q}} \\ \vdots \\ -\lambda_{i(N-1)}^{\mathbb{Q}} \\ -\lambda_{iN}^{\mathbb{Q}} \end{bmatrix} = \begin{bmatrix} \bar{\mu}_1^i - rF_1^i \\ \bar{\mu}_2^i - rF_2^i \\ \bar{\mu}_3^i - rF_3^i \\ \vdots \\ \bar{\mu}_N^i - rF_N^i \\ \hat{\mu}^i - rV^i \end{bmatrix} \tag{A.11}$$

Each linear equation in (A.11) is independent of the other instruments; without loss of generality, consider the last one:

$$\Lambda^i \hat{\sigma}^i \; - \; \sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} \Delta V^{ij} \; = \; \hat{\mu}^i \; - \; rV^i \; .$$

Using the known expressions for $\hat{\mu}^i$, $\hat{\sigma}^i$ and ΔV^{ij} from (A.5) yields

$$\frac{\partial V^i}{\partial t} \; + \; \frac{1}{2}(b^i)^2 \frac{\partial^2 V^i}{\partial S^2} \; + \; \big(a^i - \Lambda^i b^i\big) \frac{\partial V^i}{\partial S} \; - \; rV^i \; + \; \sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} \big(V^j - V^i\big) \; = \; 0 \, .$$

The system of PDEs is therefore

$$\text{for } i = 1, 2, \ldots, N:$$

$$\frac{\partial V^i}{\partial t} + \frac{1}{2}(b^i)^2 \frac{\partial^2 V^i}{\partial S^2} + \left(a^i - \Lambda^i b^i\right) \frac{\partial V^i}{\partial S} - \left(r - \lambda_{ii}^{\mathbb{Q}}\right) V^i + \sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} V^j = 0 \tag{A.12}$$

where

$$\lambda_{ii}^{\mathbb{Q}} = -\sum_{\substack{j=1 \\ j \neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} . \tag{A.13}$$

The Λ^i term is a market price of risk, with $a^i - \Lambda^i b^i$ the risk-adjusted drift. A simple no-arbitrage argument shows that the $\lambda_{ij}^{\mathbb{Q}}$'s must be non-negative [35]. The $\lambda_{ij}^{\mathbb{Q}}$ $(j \neq i)$ term can be interpreted as a risk-adjusted transition intensity of the Markov chain. If the regime-switching risk is diversifiable, then the $\lambda_{ij}^{\mathbb{Q}}$'s are equal to the real-world intensities (i.e. the $\lambda_{ij}^{\mathbb{P}}$'s) of the Markov chain—that is, there is no risk adjustment. In general, this will not be the case. On the other hand, the jumps η^{ij} in the pricing equations are those that would be observed in the real world. By definition, the $\mathbb{P}$ and $\mathbb{Q}$ measures must have the same null and almost-sure events. It is therefore impossible for a jump amplitude to exist under one measure and not the other.

If the underlying is tradeable, the market price of risk Λ^i may be eliminated from the system of PDEs in (A.12). Assume that F_1 is the underlying: in this case $F_1 = S$, $\bar{\mu}_1^i = a^i$,

$\bar{\sigma}_1^i = b^i$, and $\Delta F_1^{ij} = S(\eta^{ij} - 1)$. The first equation from the system in (A.11) becomes

$$\Lambda_i b^i \;-\; \sum_{\substack{j=1\\ j\neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} S\big(\eta^{ij} - 1\big) \;=\; a^i - rS \qquad \longrightarrow \qquad \Lambda_i b^i - S\tilde{\eta}_i^{\mathbb{Q}} \;=\; a^i - rS$$

$$\longrightarrow \qquad a^i - \Lambda_i b^i \;=\; (r - \tilde{\eta}_i^{\mathbb{Q}})S\,, \quad \text{(A.14)}$$

with

$$\tilde{\eta}_i^{\mathbb{Q}} \;=\; \sum_{\substack{j=1\\ j\neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} \left(\eta^{ij} - 1\right) \;=\; \sum_{j=1}^{N} \lambda_{ij}^{\mathbb{Q}} \eta^{ij}$$

obtained using (A.13) and the fact $\eta^{ii} = 1$. Hence, by substituting (A.14) into (A.12), the PDE for regime i may be written as

$$\frac{\partial V^i}{\partial t} \;+\; \frac{1}{2}(b^i)^2 \frac{\partial^2 V^i}{\partial S^2} \;+\; \big(r - \tilde{\eta}_i^{\mathbb{Q}}\big) S \frac{\partial V^i}{\partial S} \;-\; \big(r - \lambda_{ii}^{\mathbb{Q}}\big) V^i \;+\; \sum_{\substack{j=1\\ j\neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} V^j \;=\; 0\,. \quad \text{(A.15)}$$

A.2.2 Jumps May Occur without a Change in Regime

The above derivations assume that the asset price can only jump when there is a regime change. We now relax this assumption by allowing jumps in the underlying that can occur without a change of state. Within each of the N regimes, a jump of deterministic size ϑ^i may occur, with the arrival of this jump governed by the Poisson process π_i. The stochastic differential equation for the evolution of the underlying asset is now

$$dS \;=\; a^i(S,t)\,dt \;+\; b^i(S,t)\,dZ \;+\; S(\vartheta^i - 1)\,d\pi_i \;+\; \sum_{j=1}^{N} S\big(\eta^{ij} - 1\big)\,dX_{ij}\,.$$

The possibility of a jump absent a regime change adds another random factor, and as such a new hedging instrument F_{N+1} is required to eliminate the associated risk. In this

case, the instantaneous change in the overall hedged position, with regime i the current state, is

$$\begin{aligned} d\Pi \;=\; & \left[-\hat{\mu}^i + \sum_{n=1}^{N+1} w_n \bar{\mu}_n^i + rB\right] dt \;+\; \left[-\hat{\sigma}^i + \sum_{n=1}^{N+1} w_n \bar{\sigma}_n^i\right] dZ \\ & + \;\left[-\Delta V^i + \sum_{n=1}^{N+1} w_n \Delta F_n^i\right] d\pi_i \;+\; \sum_{\substack{j=1 \\ j\neq i}}^{N} \left[-\Delta V^{ij} + \sum_{n=1}^{N+1} w_n \Delta F_n^{ij}\right] dX_{ij}\,, \end{aligned}$$

where $\Delta V^i = V^i(S\vartheta^i, t) - V^i(S,t)$, $\Delta F_n^i = F_n^i(S\vartheta^i, t) - F_n^i(S,t)$, and all other notations as before.

Following the same risk-elimination and no-arbitrage arguments as in the previous section, we arrive at the system of pricing PDEs, which contains new terms:

for $i = 1, 2, \ldots, N$:

$$\begin{aligned} & \frac{\partial V^i}{\partial t} \;+\; \frac{1}{2}(b^i)^2 \frac{\partial^2 V^i}{\partial S^2} \;+\; \big(a^i - \Lambda^i b^i\big)\frac{\partial V^i}{\partial S} \;-\; \big(r - \lambda_{ii}^{\mathbb{Q}} + \lambda_i^{\mathbb{Q}}\big)V^i \\ & \qquad\qquad + \;\lambda_i^{\mathbb{Q}} V^i(S\vartheta^i, t) \;+\; \sum_{\substack{j=1 \\ j\neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} V^j(S\eta^{ij}, t) \;=\; 0\,. \end{aligned}$$

Here, $V^i = V^i(S,t)$ and $\lambda_i^{\mathbb{Q}}$ is the risk-adjusted intensity corresponding to the jump $S \to \vartheta^i S$ in regime i. If F_1 is the underlying and is tradeable, then the pricing PDE for regime i is

$$\begin{aligned} & \frac{\partial V^i}{\partial t} \;+\; \frac{1}{2}(b^i)^2 \frac{\partial^2 V^i}{\partial S^2} \;+\; \big(r - \tilde{\eta}_i^{\mathbb{Q}} - \tilde{\vartheta}_i^{\mathbb{Q}}\big)S\frac{\partial V^i}{\partial S} \;-\; \big(r - \lambda_{ii}^{\mathbb{Q}} + \lambda_i^{\mathbb{Q}}\big)V^i \\ & \qquad\qquad + \;\lambda_i^{\mathbb{Q}} V^i(S\vartheta^i, t) \;+\; \sum_{\substack{j=1 \\ j\neq i}}^{N} \lambda_{ij}^{\mathbb{Q}} V^j(S\eta^{ij}, t) \;=\; 0\,, \end{aligned}$$

with $\tilde{\vartheta}_i^{\mathbb{Q}} = \lambda_i^{\mathbb{Q}}(\vartheta^i - 1)$.

Appendix B

Numerical Implementation when Hedging Under Jump Diffusion

B.1 Numerical Solution of the Option-Pricing PIDE

An efficient numerical solution of the option-pricing PIDE

$$\frac{\partial V}{\partial \tau} \;=\; \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \;+\; \big(r - \lambda^{\mathbb{Q}}\kappa^{\mathbb{Q}}\big) S \frac{\partial V}{\partial S} \;-\; \big(r + \lambda^{\mathbb{Q}}\big) V \;+\; \lambda^{\mathbb{Q}} \int_0^\infty V(JS,\tau) g^{\mathbb{Q}}(J)\, dJ \tag{B.1}$$

was developed in [34]; here, we provide a brief overview. The partial derivatives are discretized in the same manner as the derivatives in the system of PDEs for the regime-switching model (see Section 4.1.2). As for the integral term

$$\int_0^\infty V(JS,\tau) g^{\mathbb{Q}}(J)\, dJ\,, \tag{B.2}$$

by applying the change of variables

$$y = \log(JS)\,,$$

it is transformed to

$$\int_{-\infty}^{\infty} V(e^y,\tau) f\left(\frac{e^y}{S}\right)\, dy\,, \tag{B.3}$$

where the shorthand notation $f(u) = u.g^{\mathbb{Q}}(u)$ has been introduced.

When the further change of variables

$$x = \log(S), \qquad \bar{V}(x,\tau) = V(e^x,\tau) = V(S,\tau), \qquad \bar{f}\big(\log(y)\big) = f(y)$$

is established, the integral in (B.3) becomes

$$I(x) \;=\; \int_{-\infty}^{\infty} \bar{V}(x+y)\bar{f}(y)\, dy\,, \tag{B.4}$$

which is a correlation integral. In discrete form, the correlation integral (B.4) is

$$I(x_i) \;=\; \sum_{j=-\frac{N}{2}+1}^{j=\frac{N}{2}} \bar{V}_{i+j}\bar{f}_j \Delta y \;+\; \mathcal{O}(\Delta y^2)\,, \tag{B.5}$$

where $x_i = i\Delta x$, $\bar{V}_j = \bar{V}(j\Delta x)$, $\bar{f}_j = \bar{f}(j\Delta x)$, and $\Delta x = \Delta y$. Furthermore, N is chosen such that contributions to the sum are negligible for nodes outside the index set of summation.

The discrete correlation integral (B.5) involves an equally spaced grid in $x = \log(S)$ space. As such, the option values $V(S,\tau)$ are interpolated onto this equally spaced $\log(S)$ grid to yield $\bar{V}$. The FFT is then employed to efficiently compute (B.5) for all nodes on the x-grid, and interpolation is used to relate these values to the integral term (B.2) of the

PIDE. d'Halluin et al. [34] also develop techniques for dealing with the problem of FFT wrap-around effects by suitably extending the x-grid used in (B.5).

To numerically solve the PIDE, a fixed point iteration is employed which uses option values from a previous iteration to compute the correlation integrals. Under several standard assumptions, the discretization of the PIDE is unconditionally stable and the fixed point iteration will converge [34]. Furthermore, with Crank–Nicolson timestepping, the numerical procedure should yield quadratic convergence. The penalty method for pricing American options can also be incorporated [33].

B.2 Precomputing Quantities Required for Hedging

Consider the hedge portfolio

$$\mathcal{H} \;=\; eS \;+\; \vec{\phi}\cdot\vec{I} \;+\; B\,.$$

The quantities required for determining the hedge weights $\{e, \vec{\phi}\}$, given in Table B.1, are precomputed and stored in a series of S-t grids. When $\vec{I}$ has N hedging instruments, $\mathcal{O}(N^2)$ grids of precomputed data will be needed. First, the option values V and $\vec{I}$ are found by numerically solving the PIDE (B.1), where the mesh used within the option pricer will generally be finer than the grid for the precomputed data. Once the option values are determined, the deltas $\frac{\partial V}{\partial S}$ and $\frac{\partial \vec{I}}{\partial S}$ are calculated via numerical differentiation.

The other required quantities are the integrals (involving V and $\vec{I}$) necessary to compute the entries of the linear system (7.6). These integrals, given in Table B.1, are calculated using the techniques explained above for treating the integral term (B.2) of the PIDE [34]. Therefore, during the hedging simulations, the linear system (7.6) used to solve the optimization problem is formed by table lookup and interpolation.

Type of Data	Representation	Number of Grids Required
Option Values	V and $\vec{I}$	$N+1$
Option Deltas	$\frac{\partial V}{\partial S}$ and $\frac{\partial \vec{I}}{\partial S}$	$N+1$
Integrals	$\int_0^\infty V(JS)I^j(JS)W(J)\,dJ \qquad j=1,\ldots,N$	N
Integrals	$\int_0^\infty I^j(JS)W(J)\,dJ \qquad j=1,\ldots,N$	N
Integrals	$\int_0^\infty V(JS)W(J)\,dJ$	1
Integrals	$\int_0^\infty I^j(JS)I^k(JS)W(J)\,dJ \quad (j,k)=1,\ldots,N$	$\frac{1}{2}N(N+1)$
Integrals	$\int_0^\infty \big[JS\big]I^j(JS)W(J)\,dJ \qquad j=1,\ldots,N$	N
Integrals	$\int_0^\infty \big[JS\big]V(JS)W(J)\,dJ$	1
Integrals	$\int_0^\infty \big[JS\big]W(J)\,dJ$	1
Integrals	$\int_0^\infty \big[JS\big]^2W(J)\,dJ$	1
Integrals	$\int_0^\infty W(J)\,dJ$	1

TABLE B.1: *The number and type of precomputed grids required for calculating the hedge portfolio weights when $\vec{I}$ contains N instruments. Note that full grids are not required for the final three quantities, as there is an absence of time and/or asset price dependence.*

Appendix C

Standard Results for Lévy Processes

C.1 Lévy–Itô Decomposition

The characteristic triplet (ψ, σ, ν) of a Lévy process is intimately linked to the Lévy–Itô Decomposition, which formally states that any Lévy process is the superposition of a drift, Brownian motion and jump component [23].

Theorem C.1. (Lévy–Itô Decomposition) *Consider a Lévy process X_t with Lévy measure ν satisfying $\int_{|y|\leq\Upsilon} y^2\,\nu(dy) < \infty$ and $\int_{|y|\geq\Upsilon} \nu(dy) < \infty$, where $\Upsilon \in (0,\infty)$. Furthermore, let J_X be the associated jump measure and $\tilde{J}_X$ the compensated jump measure. Then, the increment of the Lévy process may be written as*

$$dX_t \;=\; \psi\,dt \;+\; \sigma\,dZ_t \;+\; \int_{|y|\geq\Upsilon} y\,J_X(dydt) \;+\; \lim_{\epsilon\to 0}\int_{\epsilon\leq|y|<\Upsilon} y\,\tilde{J}_X(dydt)\,, \quad \text{(C.1)}$$

where $\psi \in \mathbb{R}$ and $\sigma \in \mathbb{R}$. Each of the four components are independent.

The first two terms in (C.1) represent the drift and Brownian motion, respectively. The

third term is a compound Poisson process with arrival rate

$$\lambda^* = \nu\big((-\infty, -\Upsilon] \cup [\Upsilon, \infty)\big)$$

and jump amplitude distribution

$$\frac{\nu(y)}{\lambda^*}\mathbf{1}_{|y|\geq\Upsilon}\,.$$

The final term is a *pure jump martingale*, where the compensated jump measure is required to obtain convergence.

The Lévy–Itô Decomposition in Theorem C.1 may be written in a more succinct form as [61]

$$dX_t \;=\; \psi\,dt \;+\; \sigma\,dZ_t \;+\; \int_{\mathbb{R}} y\,\bar{J}_X(dydt)\,, \tag{C.2}$$

where

$$\bar{J}_X(dydt) \;=\; \begin{cases} \tilde{J}_X(dydt) & \text{if } |y| < \Upsilon \\ J_X(dydt) & \text{if } |y| \geq \Upsilon\ . \end{cases} \tag{C.3}$$

The diffusion coefficient σ and Lévy measure ν are invariants of the decomposition, while the drift coefficient ψ is not, as it depends on the cutoff parameter Υ. This cutoff parameter can assume any value in $(0, \infty)$, but typically it is set to unity. If $\mathbb{E}[X_t] < \infty$, then we can use $\Upsilon = \infty$. If the jump process is of finite variation, i.e. $\int_{|y|\leq 1} |y|\,\nu(dy) < \infty$, then Υ may be chosen as zero. Setting $\Upsilon \ll 1$ is the basis of the procedure for approximating an infinite activity Lévy process as a jump diffusion [3]: the first three terms of (C.1) already comprise a jump diffusion, while the pure jump martingale is approximated as a Brownian motion.

There are two salient points that deserve to be stressed:

1. The Lévy–Itô Decomposition determines the characteristic triplet, as opposed to the

triplet specifying the decomposition. For Lévy processes whose characteristic triplet is not explicit (e.g. those obtained through subordinating a Brownian motion), Theorem C.1 guarantees such a triplet exists;

2. It is implicit in the fundamental structure of a Lévy process that small jumps are compensated. This will sometimes manifest as seemingly spurious terms.

C.2 Itô's Formula

Itô's formula provides the change of variables formula [61].

Theorem C.2. (Itô's formula) *Consider a process U_t which may be written as*

$$dU_t \;=\; \alpha(U_t,t)dt \;+\; \beta(U_t,t)dZ_t \;+\; \int_{\mathbb{R}} \gamma(U_t,t,y)\bar{J}_X(dydt)\,,$$

where $\bar{J}_X$ is the jump measure (C.3) *for some underlying Lévy process with cutoff parameter $\Upsilon \in [0,\infty]$. For a function $f(u,t)$ such that $f, \frac{\partial f}{\partial t}, \frac{\partial f}{\partial u}, \frac{\partial^2 f}{\partial u^2}$ are all continuous,*

$$\begin{aligned} df \;=\; & \frac{\partial f}{\partial t}(U_t,t)\,dt \;+\; \frac{\partial f}{\partial u}(U_t,t)\big[\alpha dt + \beta dZ_t\big] \;+\; \frac{1}{2}\beta^2\frac{\partial^2 f}{\partial u^2}(U_t,t)\,dt \\ & +\; dt\int_{|y|<\Upsilon}\Big[f(U_{t^-}+\gamma,t) - f(U_{t^-},t) - \gamma\frac{\partial f}{\partial u}(U_{t^-},t)\Big]\,\nu(y)dy \\ & +\; \int_{\mathbb{R}}\Big[f(U_{t^-}+\gamma,t) - f(U_{t^-},t)\Big]\,\bar{J}_X(dydt)\,, \end{aligned} \tag{C.4}$$

where y represents a jump in the underlying Lévy process.

Bibliography

[1] L. Andersen and J. Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4:231–262, 2000.

[2] L. Andersen, J. Andreasen, and D. Eliezer. Static replication of barrier options: Some general results. *Journal of Computational Finance*, 5(4):1–25, 2002.

[3] S. Asmussen and J. Rosiński. Approximations of small jumps of Lévy processes with a view to simulation. *Journal of Applied Probability*, 38:482–493, 2001.

[4] E. Ayache, P. Henrotte, S. Nasser, and X. Wang. Can anyone solve the smile problem? *Wilmott Magazine*, pages 78–96, January 2004.

[5] G. Barles. Convergence of numerical schemes for degenerate parabolic equations arising in finance. In L. C. G. Rogers and D. Talay, editors, *Numerical Methods in Finance*, pages 1–21. Cambridge University Press, Cambridge, 1997.

[6] C. Barrera-Esteve, F. Bergeret, C. Dossal, E. Gobet, A. Meziou, R. Munos, and D. Reboul-Salze. Numerical methods for the pricing of swing options: A stochastic control approach. *Methodology and Computing in Applied Probability*, 8:517–540, 2006.

[7] D. S. Bates. Pricing options under jump-diffusion processes. Working paper 37-88,

Rodney L. White Center for Financial Research, The Wharton School, University of Pennsylvania, 1988.

[8] D. S. Bates. The crash of '87: Was it expected? The evidence from options markets. *Journal of Finance*, 46:1009–1044, 1991.

[9] N.H. Bingham and R. Kiesel. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. Springer-Verlag, London, 2000.

[10] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:639–659, 1973.

[11] P.P. Boyle. Options: a Monte Carlo approach. *Journal of Financial Economics*, 4(3): 323–338, 1977.

[12] P.P. Boyle and T. Draviam. Pricing exotic options under regime switching. *Insurance: Mathematics and Economics*, 40(2):267–282, 2007.

[13] D. Breeden and R. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51:621–652, 1978.

[14] J. Buffington and R.J. Elliot. American options with regime switching. *International Journal of Theoretical and Applied Finance*, 5:497–514, 2002.

[15] P. Carr, K. Ellis, and V. Gupta. Static hedging of exotic options. *Journal of Finance*, 53:1165–91, 1998.

[16] P. Carr, H. Geman, D.B. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75(2):305–332, 2002.

[17] P. Carr and D.B. Madan. Optimal positioning in derivative securities. *Quantitative Finance*, 1:19–37, 2001.

[18] P. Carr and L. Wu. The finite moment log stable process and option pricing. *Journal of Finance*, 58(2):753–778, 2003.

[19] P. Carr and L. Wu. Static hedging of standard options. Working paper, Courant Institute, New York University, 2004.

[20] Z. Chen. Private communication, 2007.

[21] Z. Chen and P. A. Forsyth. Stochastic models of natural gas prices and applications to natural gas storage valuation. Working paper, University of Waterloo, 2006.

[22] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[23] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, Boca Raton, 2004.

[24] R. Cont and P. Tankov. Non-parametric calibration of jump-diffusion option pricing models. *Journal of Computational Finance*, 7(2):1–49, 2004.

[25] R. Cont, P. Tankov, and E. Voltchkova. Hedging with options in models with jumps. *Stochastic Analysis and Applications*, forthcoming, 2005.

[26] R. Cont, P. Tankov, and E. Voltchkova. Hedging with options in models with jumps. *Stochastic Analysis and Applications (Abel Symposium 2005 in honor of Kiyosi Ito's 90th birthday)*, 2005. to appear.

[27] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5:329–359, 1996.

[28] M. Dahlgren. A continuous time model to price commodity-based swing options. *Review of Derivatives Research*, 8:27–47, 2005.

[29] M.H.A. Davis, V.G. Panas, and T. Zariphopoulou. European option pricing with transaction costs. *SIAM Journal of Optimization and Control*, 31(2):470–493, 1993.

[30] M. Davison and C.L. Anderson. Approximate recursive valuation of electricity swing options. Working paper, University of Western Ontario, 2002.

[31] M. Davison, C.L. Anderson, B. Marcus, and K. Anderson. Development of a hybrid model for electrical power spot prices. *IEEE Transactions on Power Systems*, 17(2): 257–264, 2002.

[32] E. Derman, D. Ergener, and I. Kani. Static options replication. *Journal of Derivatives*, 2(4):78–95, 1995.

[33] Y. d'Halluin, P. A. Forsyth, and G. Labahn. A penalty method for American options with jump diffusion processes. *Numerische Mathematik*, 97:321–352, 2004.

[34] Y. d'Halluin, P.A. Forsyth, and K.R. Vetzal. Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis*, 25:65–92, 2005.

[35] P.A. Forsyth and K. Vetzal. CS870 course notes, University of Waterloo. 2006.

[36] P.A. Forsyth and K.R. Vetzal. Quadratic convergence for valuing American options using a penalty method. *SIAM Journal on Scientific Computing*, 23:2096–2123, 2002.

[37] P. Grandits and W. Schachinger. Leland's approach to option pricing: The evolution of a discontinuity. *Mathematical Finance*, 11(3):347–355, 2001.

[38] P.C. Hansen. The truncated SVD as a method for regularization. *BIT*, 27:534–553, 1987.

[39] M. Hardy. A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(2):41–53, 2001.

[40] C. He, J.S. Kennedy, T.F. Coleman, P.A. Forsyth, Y. Li, and K. Vetzal. Calibration and hedging under jump diffusion. *Review of Derivatives Research*, 9(1):1–35, 2006.

[41] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

[42] T. Hoggard, E. Whalley, and P. Wilmott. Hedging option portfolios in the presence of transaction costs. *Advances in Futures and Options Research*, 7:21–35, 1994.

[43] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.

[44] P. Jaillet, E.I. Ronn, and S. Tompaidis. Valuation of commodity-based swing options. *Management Science*, 50(7):909–921, 2004.

[45] Y.M. Kabanov and M.M. Safarian. On Leland's strategy of option pricing with transactions costs. *Finance and Stochastics*, 1:239–250, 1997.

[46] J. Keppo and S. Peura. Optimal portfolio hedging with nonlinear derivatives and transaction costs. *Computational Economics*, 13(2):117–145, 1999.

[47] V.A. Kholodnyi. Valuation and hedging of European contingent claims on power with spikes: a non-Markovian approach. *Journal of Engineering Mathematics*, 49(3):233–252, 2004.

[48] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 1999.

[49] P.E. Kloeden, E. Platen, and H. Schurz. *Numerical Solution of SDE through computer experiments*. Springer, New York, 1994.

[50] S. Kou. A jump-diffusion model for option pricing. *Management Science*, 48:1086–1101, 2002.

[51] G. Labahn. Closed form PDF for Merton's jump diffusion model. Technical report, School of Computer Science, University of Waterloo, 2003.

[52] A. Lari-Lavassani, M. Simchi, and A. Ware. A discrete valuation of swing options. *Canadian Applied Mathematics Quarterly*, 1:35–74, 2001.

[53] H. E. Leland. Option pricing and replication with transaction costs. *Journal of Finance*, 40(5):1283–1301, 1985.

[54] R.S. Mamon and M.R. Rodrigo. Explicit solutions to European options in a regime-switching economy. *Operations Research Letters*, 33:581–586, 2005.

[55] G.B. Di Masi, Y.M Kabanov, and W.J. Runggaldier. Mean-variance hedging of options on stocks with Markov volatilities. *Theory of Probability and its Applications*, 39:172–182, 1994.

[56] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[57] O. Mocioalca. Jump diffusion option with transaction costs. Working paper, Kent State University, 2003.

[58] V. Naik. Option valuation and hedging strategies with jumps in the volatility of asset returns. *Journal of Finance*, 48(5):1969–1984, 1993.

[59] V. Naik and M. Lee. General equilibrium pricing of options on the market portfolio with discontinuous returns. *Review of Financial Studies*, 3:493–521, 1990.

[60] J.F. Navas. On jump diffusion processes for asset returns. Working paper, Instituto de Empresa, 2000.

[61] B. Øksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions*. Springer-Verlag, Berlin, 2005.

[62] A. Papapantoleon. An introduction to Lévy processes with applications in finance. Lecture notes, University of Freiburg, 2005.

[63] D. Pilopovic. *Energy Risk: Valuing and Managing Energy Derivatives*. McGraw-Hill, 1998.

[64] D.M. Pooley, K.R. Vetzal, and P.A. Forsyth. Convergence remedies for non-smooth payoffs in option pricing. *Journal of Computational Finance*, 6(4):25–40, 2003.

[65] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical Recipes in C*. Cambridge University Press, second edition, 1993.

[66] P. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, 1990.

[67] R. Rannacher. Finite element solution of diffusion problems with irregular data. *Numerische Mathematik*, 43:309–327, 1984.

[68] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, 2nd edition, 2003.

[69] K. Sato. *Lévy processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

[70] E.D. Schwartz. The stochastic behaviour of commodity prices: implications for valuation and hedging. *Journal of Finance*, 52:923–973, 1997.

[71] H. Simons. Everybody's got to swing. *Futures*, 29(9):42–44, September 2000.

[72] W.J. Stewart. *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press, Princeton, 1994.

[73] I.R. Wang, P.A. Forsyth, and J.W.L. Wan. Robust numerical valuation of European

and American options under the CGMY process. Working paper, University of Waterloo (to appear in the Journal of Computational Finance), 2007.

[74] T. Wegner. Swing options and seasonality of power prices. Master's thesis, University of Oxford, 2002.

[75] P. Wilmott. *Quantitative Finance*. John Wiley & Sons, Chichester, 1998.

[76] H. Windcliff, P. A. Forsyth, and K. Vetzal. Analysis of the stability of the linear boundary condition for the Black–Scholes equation. *Journal of Computational Finance*, 8(1):65–92, 2004.

[77] C. Yuan and X. Mao. Convergence of the Euler-Maruyama method for stochastic differential equations with Markovian switching. *Math. Comput. Simul.*, 64(2):223–235, 2004.

[78] D. Zwillinger. *Standard Mathematical Tables and Formulae*. CRC Press, Boca Raton, 1996.