HEC MONTRÉAL
École affiliée à l'Université de Montréal

Valuing financial derivatives under Lévy processes

par Rim Chérif

Thèse présentée en vue de l'obtention du grade de Ph. D. en administration
(option Méthodes quantitatives)

Décembre 2016

HEC MONTRÉAL
École affiliée à l'Université de Montréal

Cette thèse intitulée :

Valuing financial derivatives under Lévy processes

présentée par :

Rim Chérif

a été évaluée par un jury composé des personnes suivantes :

Georges Dionne
HEC Montréal
Président-rapporteur

Hatem Ben Ameur
HEC Montréal
Codirecteur de recherche

Bruno Rémillard
HEC Montréal
Codirecteur de recherche

Pascal François
HEC Montréal
Membre du jury

Matt Davison
Western University Canada
Examinateur externe

Marie Allard
HEC Montréal
Représentante du directeur de HEC Montréal

RÉSUMÉ

Cette thèse se concentre sur l'évaluation de contrats financiers dans le cadre d'un processus de Lévy. Elle se compose de trois essais.

Le premier essai propose une méthode numérique pour l'évaluation d'options bermudiennes lorsque le sous-jacent est modélisé selon un processus mixte de diffusions avec sauts de fréquences et amplitudes décrites selon un processus de Poisson composé. Notre modèle repose sur la programmation dynamique couplée avec des éléments finis. Nous proposons une preuve de convergence uniforme et présentons une étude numérique et empirique qui confirment la convergence et l'efficience de notre méthodologie.

Le deuxième essai se penche sur le problème d'évaluation d'options bermudiennes lorsque le sous-jacent est décrit selon un processus de sauts purs, plus particulièrement, le processus de variance gamma. Nous présentons une méthode numérique basée sur la programmation dynamique. Nous exposons une étude empirique sur des options américaines écrites sur un future sur S&P 500. Nos résultats confirment l'efficience de l'approche proposée.

Le troisième essai propose un modèle structurel pour l'évaluation de dettes corporatives risquées lorsque la valeur des actifs de la firme suit un processus de Lévy. Notre méthodologie est basée sur la programmation dynamique couplée avec des éléments finis. Notre modèle accommode le principe d'égalité de bilan, un portefeuille de dettes corporatives arbitraires, plusieurs classes de séniorité, les économies de taxes, ainsi que les coûts de faillite. Nos résultats confirment que la prise en compte du risque associé à d'éventuels sauts combinée à un défaut endogène a un impact significatif sur les écarts de crédits.

Mots clés : Évaluation d'options ; Programmation dynamique ; Éléments finis ; Processus de Lévy ; Calibration ; Maximum de vraisemblance ; Modèle structurel ; Actifs corporatifs ; Risque de crédit.

Méthodes de recherche : Modélisation mathématique ; recherche quantitative ; Méthodes statistiques et numériques.

ABSTRACT

This thesis focuses on valuing financial derivatives under Lévy processes. It comprises of three essays.

The first essay proposes a numerical methodology for valuing Bermudan options under Gaussian and double exponential jumps. Under an extended version of these processes, the pricing problem is addressed with dynamic programming coupled with finite elements. We also provide a proof of uniform convergence, and present numerical and empirical experiments that confirm this convergence and show the efficiency of our methodology.

The second essay focuses on valuing options under pure-jump Lévy processes, specifically the variance-gamma model. Under this setting, explicit solutions for complex derivative prices are unavailable, for instance for the valuation of Bermudan options. We develop a numerical approach based on dynamic programming under an extended version of the variance-gamma model. We also conduct a numerical investigation on American-style options on the S&P 500 futures contracts. Our numerical experiments show the efficiency of our methodology.

The third essay proposes a Lévy-type model for credit risk modeling and presents a general structural model for valuing corporate securities under various Lévy processes. We work with a flexible framework which accommodates the balance-sheet equality, arbitrary corporate debts, multiple seniority classes, tax benefits, and bankruptcy costs. While our approach applies to several Lévy processes, we compute the values of equity, debt, firm, and credit-spreads under Gaussian, double exponential, and variance-gamma-jump models. Our results show that jump risk and endogenous default have a significant impact on credit spreads.

Keywords: Options valuation; Dynamic programming; Finite elements; Lévy processes; Calibration; Maximum likelihood; Structural model; Corporate securities; Credit risk.

Research methods: Mathematical modeling; Quantitative research; Statistical and numerical methods.

CONTENTS

LIST OF PAPERS

1. Ben-Ameur, H., Chérif, R., and Rémillard, B. (2016). American-style options in jump-diffusion models: estimation and evaluation. *Quantitative Finance*, 16(8):1313.

2. Ben-Ameur, H., Chérif, R., and Rémillard, B. (2016). Dynamic programming for valuing american options under the Variance-Gamma process. *Les Cahiers du GERAD* G-2016-74, HEC Montréal, Montreal, QC, Canada, September 2016. Submitted.

3. Ben-Ameur, H., Chérif, R., and Rémillard, B. (2016). A dynamic program under Lévy processes for valuing corporate securities. To appear in *Les Cahiers du GERAD in January 2017.*

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

À mon père qui vivra à jamais dans mon coeur.

À mon grand amour, mon fils Adamo.

ACKNOWLEDGMENTS

La thèse de doctorat représente le travail d'une vie, accompli sur de longues années, un travail qui m'a forgée et a donné une meilleure version de moi-même. Plusieurs personnes sont intervenues pendant ce long périple, des personnes que j'aimerais mettre en avant dans ces remerciements.

Je tiens à exprimer mes plus sincères remerciements à Pr Hatem Ben Ameur et Pr Bruno Rémillard, mes professeurs et directeurs de thèse, pour leur soutien incessant, leur disponibilité, leur passion et pour tout le savoir qu'ils ont su me transmettre. Votre support sur tous les plans, votre rigueur et votre amour pour la recherche furent déterminants dans l'accomplissement de mon projet de thèse.

Je remercie aussi tous les membres du département des sciences de la décision, mes professeurs et mes collègues de travail pour leurs encouragements et leurs conseils. Mention spéciale à Diane pour son support et son amitié. Je remercie aussi tous les membres du programme de certificat de m'avoir chaleureusement accueilli dans leur équipe.

Un grand merci à mes parents pour tout leur amour et leurs sacrifices. Merci papa de m'avoir toujours rappelé qu'il n'y a pas de victoire sans sacrifices. Merci maman, pour ta tendresse et ton dévouement. Hommage à ma défunte grand-mère pour tout son amour. Un merci à mon frère, à ma soeur et à toute ma famille.

Les mots m'échappent pour remercier à juste valeur, mon conjoint pour ses sacrifices, son amour et son support psychologique dans les moments de doutes et de faiblesses et à travers les divers aléas de la vie. Merci d'avoir cru en moi et de m'avoir toujours valorisée. Merci d'avoir très bien pris soin de notre fils pendant mes moments d'absence, car tu as bien compris mon déchirement entre ma thèse et ma famille.

Le dernier remerciement, et non le moindre, est pour la soeur que le destin a oublié de me donner: Malek. Ma "tourterelle" comme dit si bien Bruno, avec laquelle j'ai partagé les meilleurs et les pires moments de ce long parcours et sans laquelle ce projet de thèse n'aurait pas eu le même goût, ni la même signification.

INTRODUCTION

In 2016, the International Monetary Fund (IMF) downgraded its global growth forecasts, warning against the Chinese slowdown and its impact on the global economy. Many experts predict that the global economy may soon collapse due to the chaotic state of the stock market. Thus, are we on the brink of a major new stock market crisis?

In light of these concerns, it is evident that modeling under continuous geometric-Brownian motion for stock returns is not realistically representative of the market. Indeed, the Gaussian model assumes that the probability of occurrence of jumps is zero. This modeling is therefore incapable of taking into account the possibility of a stock crash. Neglecting such a situation when modeling price processes will undoubtedly lead to evaluation errors and to poor risk management.

The idea of this thesis is to take into account these discontinuities in the returns dynamics and to model them according to a Lévy process. This class of processes has been met with great success, and has been validated by empirical investigations that consolidate its relevance and confirm the weaknesses of Brownian-type financial models. This success results in several theoretical developments in various financial and actuarial applications, including the valuation of derivatives and financial contracts such as options contracts. Various types of options contracts are extensively traded on the financial market; their complexity is closely linked to the offered clauses in the contract. In addition to European-style options, which are often easily evaluated, we are also interested in American-style options whose evaluation is complex even under the assumption of geometric-Brownian motion.

In this thesis, we are interested in the evaluation of financial products where the underlying assets are described according to an exponential Lévy process. To do this, we use dynamic programming coupled with finite elements approximations. The principle of this recursive approach states that the solution of a global problem can be obtained by dismantling the problem into sub-problems that are easier to solve.

A common way to incorporate discontinuities in the asset returns is to add a Poisson process to Brownian motion; such a modeling design is called jump-diffusion process. Examples from this class of finite-activity Lévy processes are Merton's (1976) and Kou's (2002) models, under which we present an extended version and value Bermudan options.

Secondly, we pursue the same aim, namely the pricing of Bermudan options, but we consider a time-changed Brownian motion, in particular the variance-gamma model belonging to the infinite-activity pure jump Lévy process. Unlike the previous process, the variance-gamma model generates purely discontinuous paths made by an infinite number of small jumps. The economic interpretation of the gamma time change is the passage from a calendar time to an economic activity time. Hence, this stochastic time change can have two effects: it can speed up calendar time and subject the market to turbulence, or slow down calendar time and maintain an unperturbed market.

Finally, in the situation of an economic crisis, the bondholders and shareholders are subject to significant losses in value, leading to the downgrading or the bankruptcy of the company. Given that credit risk assessment is a decisive task in risk management, we address the problem of valuing corporate securities when the firm's asset dynamic is described by a Lévy process. As we try to reproduce the observed features on the market, we value corporate debts and study the contribution of jumps to explain credit spreads within a structural model.

CHAPTER 1

AMERICAN-STYLE OPTIONS IN JUMP-DIFFUSION MODELS: ESTIMATION AND EVALUATION

Hatem Ben Ameur,[1] Rim, Chérif,[2] Bruno Rémillard[3]

Abstract

We propose a dynamic programming coupled with finite elements for valuing American-style options under Gaussian and double exponential jumps à la Merton (1976) and Kou (2002), and we provide a proof of uniform convergence. Our numerical experiments confirm this convergence result and show the efficiency of the proposed methodology. We also address the estimation problem and report an empirical investigation based on Home Depot. Jump-diffusion models outperform their pure-diffusion counterparts.

Keywords: American options; Jump-diffusion process; Dynamic programming; finite elements, Calibration; Maximum likelihood.

1.1 Introduction

Pure diffusion models, like that of Black and Scholes (1973), fail to capture stylized facts of the financial market, such as over/under-reactions, fat tails, and discontinuities in the underlying-asset returns. One way of solving this problem is to introduce a hybrid process made up of a Gaussian continuous component and a composite-Poisson-type discontinuous component.

While Gaussian jumps go back to Merton (1976), double-exponential jumps go back to Kou (2002). Both authors use Lévy processes, extend the geometric

1. Hatem Ben-Ameur is a Professor at HEC Montréal in the Department of Decision Sciences, and member of the GERAD.

2. Rim Chérif is a Ph.D. student at HEC Montréal.

3. Bruno Rémillard is a Professor at HEC Montréal in the Department of Decision Sciences, and member of the CRM and GERAD.

Brownian motion used in Black and Scholes (1973), and under certain financial assumptions, provide closed-form solutions for European vanilla options. We consider their extended versions in this paper.

Valuing American-style options under Lévy models is challenging due to their associated early exercise strategies. Several methodologies are proposed in the literature: the quasi-analytical approach (Bates, 1991, Gukhal, 2001, Kou and Wang, 2004), the binomial tree (Amin, 1993, Hilliard and Schwartz, 2005), partial integro-differential equations (Zhang, 1997, Andersen and Andreasen, 2000, Matache et al., 2003, 2005, Cont and Voltchkova, 2005, Almendral and Oosterlee, 2005, Toivanen, 2008, Mayo, 2008, Chiarella and Ziogas, 2009, Chockalingam and Muthuraman, 2010, Chan and Hubbert, 2014, Salmi and Toivanen, 2014), the Markov chain approximation (Simonato, 2011), and Monte-Carlo simulation (DiCesare and Mcleish, 2008, Levendorskii, 2004).

We propose a numerical procedure for valuing American-style options in jump-diffusion models. Our approach is based on dynamic programming coupled with finite elements. The value function under consideration is approximated by a piecewise polynomial at each decision date. We experiment with piecewise-constant, linear, and quadratic approximations. Higher-order polynomials are more accurate, but are also more time consuming. A compromise between accuracy and computing time must be found. Dynamic programming, coupled with piecewise quadratic interpolations, shows the highest degree of efficiency and competes well against alternative methodologies reported in the literature.

Under jump diffusions, the distribution of the underlying-asset return is a mixture, which results in a likelihood function with several modes. Press (1967), Bates (1991), Ait-Sahalia (2004), and Hanson and Zhu (2004) use a parametric approach (historical); while Bates (1991), Andersen and Andreasen (2000), He et al. (2006), and Cont and Tankov (2004b) use calibration. We use maximum likelihood to estimate Merton (1976) and Kou (2002) models and maximum likelihood coupled with calibration to estimate their extended versions.

The rest of this paper is organized as follows. Section 1.2 presents a general jump-diffusion model and outlines Merton (1976) and Kou (2002) settings as special cases. While Section 1.3 describes the dynamic programming, Section 1.4 presents a numerical investigation, Section 1.5 reports an empirical investigation. Section 6 concludes.

1.2 The jump-diffusion model

Let $\mathscr{M}$ be a frictionless market with a risk-free asset and a risky asset, a stock whose price S experiences jumps at random times, and let $(\Omega, A, \mathbb{P})$ be a complete probability space. Define a standard Brownian motion $(W_t)_{t\geq 0}$, a Poisson process $(N_t)_{t\geq 0}$ with a constant intensity λ, and a series of independent random variables $(\xi_n)_{n\geq 1}$ with distribution ν and density $f(\cdot)$, such that $\kappa = \mathbb{E}_{\mathbb{P}}\left[e^{\xi_n} - 1\right]$ is finite. The random variables $U_n = e^{\xi_n} - 1 = (S_{\tau_n} - S_{\tau_n -})/S_{\tau_n -} \in \,]-1, \infty[$ represent the jumps' relative amplitudes at random jump times τ_n, for $n \geq 1$.

The processes W, N, and ξ are assumed to be independent under $\mathbb{P}$. They are used to model the continuous part of the stock-price trajectory, the jump times, and the jump amplitudes, respectively. Further, denote by $\mathscr{F}_t$ the sigma-algebra generated by $\{W_u, N_u, \xi_1, \ldots, \xi_{N_t}$, for $0 \leq u \leq t\}$.

The stock-price process S is defined by

$$S_t = S_0 e^{X_t}, \quad \text{for } t \geq 0,$$

where the stock log-return X is a Lévy process (with stationary independent increments) defined by

$$X_t = \left(\mu - \bar{d} - \frac{\sigma^2}{2} - \lambda\kappa\right) t + \sigma W_t + \sum_{n=1}^{N_t} \xi_n, \quad \text{for } t \geq 0, \tag{1.1}$$

where μ is the instantaneous stock return, $\bar{d}$ the proportional dividend rate, σ the stock-return volatility conditional on no jumps, and N_t the number of jumps till time $t \in [0, T]$. Set the convention $\sum_{n=1}^{0} = 0$. The process X is discontinuous at jump times.

The jump-diffusion model in eq. (1.1) is arbitrage free but incomplete. There exist several equivalent martingale probabilities measures that provide multiple rational values for a given option contract, all of which are consistent with the no-arbitrage principle. We provide herein a simple change of probability measure.

Let $h(\cdot)$ be a positive function such that $a = \mathbb{E}_{\mathbb{P}}[h(\xi_n)] < \infty$ and define

$$\Lambda_t^{b,h} = e^{bW_t - tb^2/2 - \lambda t(a-1)} \prod_{n=1}^{N_t} h(\xi_n), \quad \text{for } t \geq 0. \tag{1.2}$$

The process $\Lambda^{b,h}$ is a positive $\mathbb{P}$-martingale and has an expectation of one. The risk-neutral probability measure $\mathbb{Q}^{b,h}$, associated to the constant b and the function $h(\cdot)$, is defined by the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}^{b,h}}{d\mathbb{P}}|_{\mathscr{F}_t} = \Lambda_t^{b,h}.$$

From Cont and Tankov (2004a, Proposition 9.8), $\tilde{W}_t = W_t - bt$ is a Brownian motion, $\tilde{N}_t = N_t$ is a Poisson process of intensity $\tilde{\lambda} = \lambda a$, and $\tilde{\xi}_n = \xi_n$ has the law $\tilde{\nu}$ with density $\tilde{f}(\cdot) = f(\cdot)h(\cdot)/a$ under $\mathbb{Q}^{b,h}$. Let $\tilde{\kappa} = \mathbb{E}_{\mathbb{Q}^{b,h}}\left[e^{\tilde{\xi}_n} - 1\right] \in \mathbb{R}$. One has

$$e^{-rt}S_t + \bar{d}\int_0^t e^{-ru}S_u du$$

is a $\mathbb{Q}^{b,h}$-martingale if, and only if,

$$b = \frac{r - \mu - (\tilde{\lambda}\tilde{\kappa} - \lambda\kappa)}{\sigma},$$

where r is the risk-free rate. Given eq. (1.2), the dynamics of S under the martingale probability measure $\mathbb{Q} \equiv \mathbb{Q}^{b,h}$ is

$$S_t = S_0 e^{\tilde{X}_t},$$

where $\tilde{X}$ is still a Lévy process represented by

$$\tilde{X}_t = \left(r - \bar{d} - \frac{\sigma^2}{2} - \tilde{\lambda}\,\tilde{\kappa}\right)t + \sigma\tilde{W}_t + \sum_{n=1}^{\tilde{N}_t}\tilde{\xi}_n, \quad \text{for } t \geq 0. \tag{1.3}$$

It's worth noting that the process $\{e^{-rt}S_t\}$ is a $\mathbb{Q}$-martingale with respect to the filtration $\mathscr{F}_t$.

The function $h(\cdot)$ is often chosen so that the law of jumps remains in the same family, both under the objective and martingale probabilities measures. To this end, we set

$$h(x) = \alpha e^{\beta x}, \quad \text{for } x \in \mathbb{R}, \tag{1.4}$$

where $\alpha \in \mathbb{R}_+^*$ and $\beta \in \mathbb{R}$. See (Rémillard, 2013). Appendix 1.A provides the density function $\tilde{f}(\cdot)$.

For extended Gaussian jumps, the random variables ξ_n, for $n \geq 1$, follow $\mathscr{N}(\gamma, \delta^2)$ under $\mathbb{P}$, and $\tilde{\xi}_n$ follow $\mathscr{N}(\tilde{\gamma}, \tilde{\delta}^2)$ under $\mathbb{Q}$, where $a = \alpha e^{\beta\gamma+\beta^2\delta^2/2}$, $\tilde{\lambda} = a\lambda$, $\tilde{\gamma} = \gamma + \beta\delta^2$, $\tilde{\delta} = \delta$, and $\tilde{\kappa} = e^{\tilde{\gamma}+\tilde{\delta}^2/2} - 1$. Merton's (1976) model is obtained with $\alpha = 1$ and $\beta = 0$, which results in $b = (r-\mu)/\sigma$, $a = 1$, $\tilde{\lambda} = \lambda$, $\tilde{\gamma} = \gamma$, $\tilde{\delta} = \delta$, and $\tilde{\kappa} = e^{\gamma+\delta^2/2} - 1$. To justify this particular choice of parameters, the author claims that jump risk is not systemic and can be offset through portfolio diversification.

For extended double-exponential jumps, the parameter $\beta \in]-\eta_2, \eta_1[$. The random variables ξ_n, for $n \geq 1$, follow an asymmetric double-exponential distribution with a density function under $\mathbb{P}$:

$$f(\xi) = p_1\,\eta_1 e^{-\eta_1\xi}\,\mathbb{I}_{\{\xi\geq 0\}} + p_2\,\eta_2 e^{\eta_2\xi}\,\mathbb{I}_{\{\xi<0\}}, \quad \text{for } \eta_1 > 1 \text{ and } \eta_2 > 0,$$

where $p_1 \geq 0$ and $p_2 \geq 0$ represent the probability of an upward and a downward jump and verify $p_1 + p_2 = 1$. The parameters η_1 and η_2 represent the mean sizes of an upward and a downward jump, respectively. The distribution of the random variables $\tilde{\xi}_n$, for $n \geq 1$, remains in the same family under $\mathbb{Q}$, where $a = \alpha p_1\eta_1/(\eta_1 - \beta) + \alpha p_2\eta_2/(\eta_2 + \beta)$, $\tilde{\lambda} = a\lambda$, $\tilde{\eta}_1 = \eta_1 - \beta$, $\tilde{\eta}_2 = \eta_2 + \beta$, $\tilde{p}_1 = p_1\eta_1\tilde{\eta}_2/(p_1\eta_1\tilde{\eta}_2 +$

$p_2\tilde{\eta}_1\eta_2)$, $\tilde{p}_2 = 1 - \tilde{p}_1$, and $\tilde{\kappa} = \tilde{p}_1\tilde{\eta}_1/(\tilde{\eta}_1 - 1) + \tilde{p}_2\tilde{\eta}_2/(\tilde{\eta}_2 + 1) - 1$. Kou's (2002) model is obtained with $\alpha = 1$ and $\beta = 1$, which results in $b = (r - \mu)/\sigma$, $a = p_1\eta_1/(\eta_1 - 1) + p_2\eta_2/(\eta_2 + 1)$, $\tilde{\lambda} = \lambda a$, $\tilde{\eta}_1 = \eta_1 - 1$, $\tilde{\eta}_2 = \eta_2 + 1$, $\tilde{p}_1 = p_1\eta_1(\eta_2 + 1)/(p_1\eta_1(\eta_2 + 1) + \eta_2 p_2(\eta_1 - 1))$, $\tilde{p}_2 = 1 - \tilde{p}_1$, and $\tilde{\kappa} = a - 1$. In support of its choice, the authors relies on the expected utility theory and shows that a rational price can be obtained via this particular change of parameters.

While Merton (1976) derives his formula directly from an arbitrage argument, Kou (2002) considers an equilibrium model to provide option values that are (also) consistent with the no-arbitrage principle.

1.3 Valuing American-style options

1.3.1 The option contract

An American option on a stock with a maturity T is characterized by its known exercise value $v_t^e(s)$, where $s = S_t$ is the stock price at time $t \in [0, T]$. For example, an American vanilla option is defined by

$$v_t^e(s) = \begin{cases} (s - K)^+ & \text{for a call option} \\ (K - s)^+ & \text{for a put option} \end{cases},$$

where K is the option's exercise price and $(x)^+ \equiv \max(0, x)$. We consider its Bermudan version, which admits a finite number of exercise opportunities, given by $t_0 = 0, \ldots, t_m, \ldots, t_M = T$. For simplicity, assume that $t_{m+1} - t_m = \Delta t$ is a fixed constant.

By the no-arbitrage evaluation principle, the option's holding value at t_m is

$$\begin{aligned} v_m^h(s) &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r\Delta t} v_{m+1}(S_{t_{m+1}}) \mid S_{t_m} = s\right], \\ &\quad \text{for } m = 0, \ldots, M, \end{aligned} \tag{1.5}$$

where $\mathbb{E}_{\mathbb{Q}}[\cdot \mid S_{t_m} = s]$ is the conditional expectation under $\mathbb{Q}$. Set the convention that $v_M^h(\cdot) = 0$ to say that the option must be exercised at maturity $t_M = T$. The option's overall value at time t_m is

$$v_m(s) = \max\left(v_m^e(s), v_m^h(s)\right), \quad \text{for } s > 0. \tag{1.6}$$

The associated European version is characterized by $v_m^e(\cdot) = 0$, for $m = 0, \ldots, M-1$.

1.3.2 The dynamic program

Let $\mathcal{G} = \{a_1, \ldots, a_p\}$ be the grid points, where $0 = a_0 < a_1 < \ldots < a_p < +\infty$ and $\Delta a_i = a_i - a_{i-1}$, $i = 1, \ldots, p$. The grid $\mathcal{G}$ must be selected so that $\text{mesh}(\mathcal{G}) = \max_{1 \le i \le p} \Delta a_i$, $\mathbb{Q}(S_t < a_1)$ and $\mathbb{Q}(S_t > a_p)$ all converge to 0 as $p \to \infty$, for $t \in \{t_1, \ldots, t_M\}$. We select the grid points to be the quantiles of S_T.

Suppose that an approximation $\tilde{v}_{m+1}(\cdot)$ of the value function $v_{m+1}(\cdot)$ is available at a given future date t_{m+1} on the grid $\mathcal{G}$. This is not a strong assumption since the value function $v_M(\cdot) = v_M^e(\cdot)$ is known in closed form.

We use a piecewise polynomial and extend the approximation $\tilde{v}_{m+1}(\cdot)$ from $\mathcal{G}$ to the overall state space $\mathbb{R}_+^*$, that is,

$$\hat{v}_{m+1}(s) = \sum_{i=0}^{p-1} \left(\beta_i^0 + \beta_i^1\, s + \cdots + \beta_i^d\, s^d\right) \mathbb{I}(a_i \le s < a_{i+1}), \quad \text{for } s > 0, \tag{1.7}$$

where d is the degree of the piecewise polynomial, whose local coefficients depend on the time step $m+1$. Eq. (1.5) and eq. (1.7) result in

$$\begin{aligned} \tilde{v}_m^h(a_k) &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r\Delta t}\hat{v}_{m+1}\left(S_{t_{m+1}}\right) | S_{t_m} = a_k\right] \\ &= e^{-r\Delta t} \sum_{i=0}^{p} \left(\beta_i^0\, T_{ki}^0 + \cdots + \beta_i^d\, T_{ki}^d\right), \end{aligned} \tag{1.8}$$

where

$$T_{ki}^{\nu} = \mathbb{E}_{\mathbb{Q}}\left[S_{t_{m+1}}^{\nu}\mathbb{I}\left(a_i \leq S_{t_{m+1}} < a_{i+1}\right)|S_{t_m} = a_k\right],$$

$$\text{for } \nu = 0,\ldots,d.$$

For example, T_{ki}^{0} is the conditional probability that the stock price at t_{m+1} falls in the interval $[a_i, a_{i+1})$ given that the stock price at time t_m is a_k. Key ingredients for the dynamic program (DP) to run are the transition tables T^{ν}, for $\nu = 0,\ldots,d$. We derive T^{ν}, for $\nu \in \{0,1,2\}$, in closed form under Merton's (1976) setting (see Appendix 1.B) and Kou's (2002) setting (see Appendix 1.C). We find that the piecewise-quadratic interpolation ($d = 2$) is the most efficient. To reach the pure American option, set Δt as small as possible. Ben-Ameur et al. (2002) use DP for valuing Asian options in the Black and Scholes' (1973) model.

The formula in eq. (1.8) separates the option holding value in two parts. The first is related to the dynamics of the state process (the transition parameters) and the second to the option contract (the interpolation coefficients). The transition parameters are a fixed cost for the DP procedure as long as the time step and the model's parameters remain constant. For a given experiment and a set of the model's parameters, about 90% of CPU time is used to compute the transition parameters. Thus, the effective CPU time to run a DP experiment is about 10% of what is reported in our tables. Unlike finite differences-based methods, DP does not need a time discretization since the transition parameters do respect the true dynamics of the state process. All components of eq. (1.8) are computed in closed form.

From eq. (1.6), the approximate value function at t_m is

$$\tilde{v}_m(a_k) = \max\left(v_m^e(a_k), \tilde{v}_m^h(a_k)\right), \quad \text{for } a_k \in \mathcal{G}, \tag{1.9}$$

and the approximate exercise policy at t_m is characterized by

$$\tilde{v}_m^h(a_k) < v_m^e(a_k), \quad \text{for } a_k \in \mathcal{G}.$$

The local coefficients of the piecewise-quadratic interpolation $\hat{v}_{m+1}(\cdot)$ in eq. (1.7) verify $\hat{v}_{m+1}(\cdot) = \tilde{v}_{m+1}(\cdot)$ on $\mathscr{G}$.

The dynamic program works as follows :

1. For $m = M-1$, set $\tilde{v}_{m+1}(\cdot) = v_{m+1}(\cdot) = v^e_{m+1}(.)$ on $\mathscr{G}$;
2. Use a piecewise-quadratic interpolation as in eq. (1.7), and extend $\tilde{v}_{m+1}(\cdot)$, defined on $\mathscr{G}$, to $\hat{v}_{m+1}(\cdot)$, defined on the overall state space $\mathbb{R}^*_+$;
3. By eq. (1.8), compute $\tilde{v}^h_m(\cdot)$, defined on $\mathscr{G}$;
4. By eq. (1.9), compute $\tilde{v}_m(\cdot)$, defined on $\mathscr{G}$;
5. Exercise the option at (t_m, a_k) if $\tilde{v}^h_m(a_k) < v^e_m(a_k)$;
6. If $m = 0$, stop; else set $m = m-1$, and go to step 2.

The next theorem, whose proof is given in Appendix 1.E, shows that the proposed methodology provides uniformly convergent approximations of the real prices over any given compact set at each time step. Before stating the result, we need to define the Lipschitz norm.

On a given interval $[a,c]$, the Lipschitz norm of a continuous function f is defined by

$$|f|_{Lip} = \sup_{x \neq y, x,y \in [a,c]} \frac{|f(x) - f(y)|}{|x-y|}.$$

It is easy to check that $\max(x-K,0)$ and $\max(K-x,0)$ are both Lipschitz with norm 1. Finally, for any closed interval I, set $\|f\|_I = \sup_{x\in I}|f(x)|$.

The following assumption on the interpolation method is essential in the proof:

Definition 1.3.1. *An interpolation method $\mathscr{I}$ satisfies Assumption $\mathscr{D}$ if there exists a constant D so that for any compact set I, any Lipschitz function g, and any grid $\mathscr{G}$ of I, one has $\|\mathscr{I}_{\mathscr{G}} g - g\|_I \leq D|g|_{Lip}\,\mathrm{mesh}(\mathscr{G})$. Here, $\mathscr{I}_{\mathscr{G}} g$ stands for the interpolation of g over grid $\mathscr{G}$.*

Remark 1.3.1. *It is shown in Appendix 1.D that Assumption $\mathscr{D}$ is met for the linear and quadratic interpolations we used in this paper.*

Theorem 1.3.2. *Suppose that the interpolation method satisfies Assumption $\mathscr{D}$. Suppose also that v_k^e are Lipschitz for all $k \in \{0, \ldots, n\}$. Set $\mathscr{R}_0 = [0, R_0]$ and let $\varepsilon > 0$ be given. Then one can find intervals $\mathscr{R}_1 = [0, R_1], \ldots, \mathscr{R}_{n-1} = [0, R_{n-1}]$, and grids $\mathscr{G}_0, \ldots, \mathscr{G}_{n-1}$ generating respectively $\mathscr{R}_0, \ldots, \mathscr{R}_{n-1}$, so that*

$$\|v_k - \hat{v}_k\|_{\mathscr{R}_k} < \varepsilon, \text{ for all } k \in \{0, \ldots, n\}.$$

1.4 Numerical investigation

The code is written in C, compiled with GCC, and executed under a standard laptop computer running Windows 7. For each set of results, we report the highest total CPU time in seconds. Our results are based on DP coupled with piecewise quadratic interpolations. Our CPU times cannot be directly compared to CPU times reported in the literature. Both are obtained under different hardware characteristics, operation systems, and programming languages. Despite this, we report them for general guidance purposes.

1.4.1 Gaussian jumps

To start with, Table 1.1 compares DP to the binomial tree of Amin (1993) for European put options, both with 200 time steps. It also reports Merton's values (Merton, 1976). Set S_0 = \$ 40, $\bar{d} = 0$, $\sigma^2 = 0.05$, $T = 0.25$ (years), $r = 0.08$, $\tilde{\lambda} = 5$ (jumps per year), $\tilde{\gamma} = -0.025$, and $\tilde{\delta}^2 = 0.05$. DP shows convergence and efficiency. It ensures accuracy to the fourth digit within a few seconds. DP agrees almost perfectly with Merton's values. Amin (1993) does not report his CPU times and provides only three digits of accuracy.

	DP with a grid size p					
K	50	100	200	400	Merton (1976)	Amin (1993)
30	0.6560	0.6711	0.6695	0.6697	0.6697	0.669
35	1.6875	1.6756	1.6728	1.6727	1.6727	1.674
40	3.5544	3.5857	3.5916	3.5920	3.5920	3.594
45	6.6251	6.6505	6.6550	6.6547	6.6547	6.656
50	10.5273	10.5413	10.5447	10.5445	10.5445	10.545
CPU	(0.00)	(0.00)	(0.03)	(0.08)		

Table 1.1: European put options

		DP with a grid size p						
T	K	50	100	200	400	Bates (1991)	Amin (1993)	Gukhal (2004)
.25	30	0.6608	0.6759	0.6743	0.6744	0.685	0.674	0.672
.25	35	1.7030	1.6902	1.6874	1.6873	1.708	1.688	1.680
.25	40	3.5811	3.6221	3.6281	3.6283	3.663	3.630	3.610
.25	45	6.7028	6.7277	6.7320	6.7318	6.787	6.734	6.695
.25	50	10.6797	10.6926	10.6957	10.6955	10.776	10.696	10.634
1	30	2.6826	2.7124	2.7175	2.7176	2.790	2.720	2.661
1	35	4.5434	4.5930	4.6000	4.6001	4.692	4.603	4.501
1	40	7.0471	7.0308	7.0249	7.0244	7.131	7.030	6.867
1	45	9.9715	9.9411	9.9477	9.9482	10.063	9.954	9.714
1	50	13.3252	13.3062	13.3123	13.3119	13.430	13.318	12.981
CPU		(0.00)	(0.01)	(0.03)	(0.11)			

Table 1.2: Bermudan put options

Next, Table 1.2 compares DP to Bates (1991), Amin (1993), and Gukhal (2004) for Bermudan put options. Their methodologies are based on quadratic approximations, binomial trees, and compound options, respectively. Both DP and Amin (1993) run with 200 time steps. Set $S_0 = \$\ 40$, $\bar{d} = 0$, $\sigma^2 = 0.05$, $r = 0.08$, $\tilde{\lambda} = 5$ (jumps per year), $\tilde{\gamma} = -0.025$, and $\tilde{\delta}^2 = 0.05$.

DP shows convergence and efficiency. DP values are always between the approximations of Bates (1991), Amin (1993), and Gukhal (2004). They are perfectly in line with the literature. These authors report only three digits of accuracy and do not report their CPU times.

Finally, Table 1.3 and Table 1.4 compare DP to Chiarella and Ziogas (2009) and Simonato (2011). Their methodologies are based on finite differences and on the Markov chain approximation, respectively. Chiarella and Ziogas (2009) use 10,000 time steps and 5,000 space steps. Both DP and Simonato (2011) run with a daily time step. While Simonato (2011) uses 10,001 space steps, DP uses at most $p = 400$ space steps.

The parameters of Table 1.3 are $S_0 = \$\ 50$, $\bar{d} = 0$, $K = \$\ 50$, $\sigma^2 = 0.04$, $r = 0.05$, $\tilde{\lambda} = 5$ (jumps per year), $\tilde{\gamma} = -0.105$, and $\tilde{\delta}^2 = 0.01$. Those of Table 1.4 are $\bar{d} = 0.05$, $K = \$\ 100$, $\sigma^2 = 0.0136$, $T = 0.5$ (years), $r = 0.03$, $\tilde{\lambda} = 1$ (jump per year), $\tilde{\gamma} = 0.0192$, and $\tilde{\delta}^2 = 0.04$.

Again, DP is in line with the literature. More importantly, DP reaches the same level of accuracy with a limited number of space steps and, consequently, exhibits very competitive CPU times.

	DP with a grid size p					
T	50	100	200	400	Merton (1976)	Simonato (2011)
10/365	1.0214	1.0222	1.0224	1.0224	1.0224	1.0224
30/365	2.0421	2.0475	2.0473	2.0474	2.0474	2.0474
60/365	3.0780	3.0899	3.0895	3.0895	3.0895	3.0896
90/365	3.8725	3.8849	3.8847	3.8847	3.8847	3.8847
270/365	7.0897	7.1263	7.1299	7.1299	7.1299	7.1302
CPU	(0.00)	(0.00)	(0.02)	(0.06)		

Table 1.3: European call options

1.4.2 Double-exponential jumps

Table 1.5 compares DP to Kou's (2002) formula for European call options. Set $S_0 = \$\ 100$, $\bar{d} = 0$, $\sigma^2 = 0.0256$, $T = 0.5$ (years), $r = 0.05$, $\tilde{\lambda} = 1$ (jump per year), $\tilde{\eta}_1 = 10$, $\tilde{\eta}_2 = 5$, and $\tilde{p}_1 = 0.4$. Again, DP shows convergence and efficiency. It ensures accuracy to the fourth digit within few seconds. Kou and Wang (2004) adapt the binomial tree of Amin (1993) and evaluate American-style options. DP and the binomial tree run with 1600 time steps. They also use the approximation

S_0	DP with a grid size p 50	100	200	400	Chiarella & Ziogas (2009)	Simonato (2011)
80	0.9717	0.9664	0.9650	0.9649	0.9648	0.9647
90	2.3178	2.3041	2.3065	2.3065	2.3063	2.3062
100	5.4076	5.3560	5.3603	5.3603	5.3603	5.3602
110	11.5481	11.5113	11.5073	11.5074	11.5079	11.5073
120	20.1237	20.1304	20.1323	20.1324	20.1333	20.1322
CPU	(0.00)	(0.00)	(0.02)	(0.09)		(25.169)

Table 1.4: Bermudan call options

K	DP with a grid size p 50	100	200	400	Kou (2002)
90	14.7530	14.8060	14.8116	14.8119	14.8119
95	11.1737	11.1041	11.1131	11.1133	11.1133
98	9.0386	9.1574	9.1469	9.1473	9.1473
100	8.0465	7.9694	7.9598	7.9594	7.9594
105	5.3778	5.4606	5.4521	5.4518	5.4518
110	3.5276	3.5915	3.5998	3.5996	3.5996
CPU	(0.02)	(0.08)	(0.33)	(1.26)	

Table 1.5: European call options - Kou (2002)

of Barone-Adesi and Whaley (1987) (BAW). Table 1.6 compares DP to Kou and Wang (2004) (KW) for Bermudan put options. Set S_0 = \$ 100, $\bar{d} = 0$, $\sigma^2 = 0.04$, $T = 0.25$ (years), $r = 0.05$, and $\tilde{p}_1 = 0.6$. While DP remains very accurate, the results provided by Kou and Wang (2004) are not, especially for out-of-the-money options. Some European-option values exceed their Bermudan counterparts. This abnormality might be due to the approximation of Barone-Adesi and Whaley (1987), which seems to give an error larger than the option's early exercise premium. Table 1.7 (FL) compares DP to Feng and Linetsky (2008), who use partial integro-differential equations and an extrapolation scheme based on sparse space steps for options valuation. Both procedures run with 252 time steps. Set $\bar{d} = 0.02$, K = \$ 100, $\sigma^2 = 0.01$, $T = 1$ (year), $r = 0.05$, $\tilde{\lambda} = 3$ (jumps per year), $\tilde{\eta}_1 = 40$, $\tilde{\eta}_2 = 12$, and $\tilde{p}_1 = 0.6$. They report a CPU time of 28 seconds for their second-order IMEX

				DP with a grid size p				KW (2004)		
K	$\tilde{\lambda}$	$\tilde{\eta}_1$	$\tilde{\eta}_2$	50	100	200	400	BAW	Tree	Europ.
110	3	25	25	10.5517	10.5698	10.5733	10.5738	10.43	10.48	10.1785
110	3	25	50	10.5451	10.5244	10.5193	10.5185	10.38	10.42	10.1146
110	3	50	25	10.4312	10.4517	10.4461	10.4465	10.31	10.36	9.9808
110	3	50	50	10.3771	10.3903	10.3943	10.3937	10.26	10.31	9.9151
110	7	25	25	10.9001	10.9366	10.9278	10.9287	10.79	10.81	10.6222
110	7	25	50	10.8255	10.7848	10.7912	10.7904	10.64	10.68	10.4758
110	7	50	25	10.6464	10.6203	10.6144	10.6136	10.47	10.51	10.1892
100	7	50	50	10.4634	10.4777	10.4807	10.4813	10.34	10.39	10.0337
90	3	25	25	0.7301	0.7571	0.7639	0.7746	0.76	0.75	0.7633
90	3	25	50	0.6599	0.6779	0.6823	0.6828	0.66	0.65	0.6739
90	3	50	25	0.7397	0.7044	0.7105	0.7098	0.69	0.68	0.6960
90	3	50	50	0.6468	0.6126	0.6185	0.6179	0.60	0.59	0.6067
90	7	25	25	1.0249	1.0668	1.0594	1.0603	1.04	1.03	1.0487
90	7	25	50	0.8228	0.8488	0.8552	0.8546	0.83	0.82	0.8474
90	7	50	25	0.8636	0.8951	0.8993	0.9000	0.88	0.87	0.8826
90	7	50	50	0.6589	0.6971	0.6911	0.6918	0.67	0.66	0.6589
CPU				(0.03)	(0.09)	(0.37)	(1.50)	(0.03)	(4029)	

Table 1.6: Bermudan put options - Kou and Wang (2004)

midpoint method, and 1.5 seconds for their extrapolation scheme. DP must be compared to the second-order IMEX midpoint method.

1.5 Empirical investigation

1.5.1 Estimation

We use maximum likelihood to estimate Merton (1976) and Kou (2002) models and maximum likelihood coupled with calibration to estimate their extended versions.

Let $R_t = \log\left(X_{th}/X_{(t-1)h}\right)$ be the daily stock log-return, for $t = 1,\ldots,M$, where $T = Mh$ is the last observation date. By eq. (1.1) and the properties of Lévy processes, the returns are independent and stationary, and they have the same distribution as

$$R_1 = \bar{\mu}h + \sigma W_h + \sum_{n=1}^{N_h} \xi_n,$$

European	DP with a grid size p				
S_0	50	100	200	400	Kou (2002)
85	13.6377	13.6349	13.6455	13.6462	13.6462
90	10.4557	10.4636	10.4527	10.4518	10.4518
95	7.9318	7.9051	7.9201	7.9223	7.9223
100	5.9855	5.9573	5.9788	5.9800	5.9801
105	4.5057	4.4927	4.5115	4.5132	4.5133
110	3.3841	3.3979	3.4119	3.4137	3.4137
115	2.5635	2.5993	2.5921	2.5909	2.5909
CPU	(0.01)	(0.06)	(0.25)	(1.12)	
Bermudan	DP with a grid size p				
S_0	50	100	200	400	FL (2008)
85	15.0843	15.0638	15.0663	15.0693	15.0695
90	11.3726	11.3656	11.3626	11.3661	11.3662
95	8.5604	8.5469	8.5476	8.5476	8.5479
100	6.4335	6.4160	6.4168	6.4169	6.4171
105	4.8314	4.8215	4.8223	4.8223	4.8225
110	3.6153	3.6338	3.6345	3.6348	3.6347
115	2.7374	2.7505	2.7504	2.7504	2.7505
CPU	(0.02)	(0.08)	(0.28)	(1.17)	

Table 1.7: Put options - Feng and Linetsky (2008)

where $\bar{\mu} = \mu - \bar{d} - \sigma^2/2 - \lambda\kappa$. The parameters to be estimated in Merton's (1976) model are $\bar{\mu}$, σ, λ, γ, and δ. The parameters to be estimated under Kou's (2002) model are $\bar{\mu}$, σ, λ, η_1, η_2, and p_1.

The log-likelihood of the random sample $R_1, \ldots, R_M$ under Merton's (1976) model is

$$\mathscr{L} = \sum_{t=1}^{M} \log\left[\sum_{n=0}^{\infty} \pi_n \varphi\left(R_t; \bar{\mu}h + n\gamma, \sigma^2 h + n\delta^2\right)\right],$$

where $\pi_n = P(N_h = n) = e^{-\lambda h}(\lambda h)^n/n!$, $\varphi(x; a, b)$ is the probability density function of a normal distribution with mean a and variance b, evaluated at x.

The log-likelihood of the random sample $R_1, \ldots, R_M$ under Kou's (2002) model is

$$\mathscr{L} = \sum_{t=1}^{M} \log\left(f\left(R_t | \bar{\mu}, \sigma, \lambda, \eta_1, \eta_2, p_1\right)\right),$$

where

$$
\begin{aligned}
f\left(R_t|\bar{\mu},\sigma,\lambda,\eta_1,\eta_2,p_1\right) &= \frac{e^{(\sigma\eta_1)^2h/2}}{\sigma\sqrt{2\pi h}}\sum_{n=1}^{\infty}\pi_n\sum_{k=1}^{n}P_{n,k}\left(\sigma\sqrt{h}\eta_1\right)^k \\
&\times\ e^{-\eta_1(R_t-\bar{\mu}h)}Hh_{k-1}\left(-\frac{R_t-\bar{\mu}h}{\sigma\sqrt{h}}+\sigma\sqrt{h}\eta_1\right) \\
&+\ \frac{e^{(\sigma\eta_2)^2h/2}}{\sigma\sqrt{2\pi h}}\sum_{n=1}^{\infty}\pi_n\sum_{k=1}^{n}Q_{n,k}\left(\sigma\sqrt{h}\eta_2\right)^k \\
&\times\ e^{\eta_2(R_t-\bar{\mu}h)}Hh_{k-1}\left(\frac{R_t-\bar{\mu}h}{\sigma\sqrt{h}}+\sigma\sqrt{h}\eta_2\right) \\
&+\ \frac{\pi_0}{\sigma\sqrt{h}}\varphi\left(-\frac{R_t-\bar{\mu}h}{\sigma\sqrt{h}}\right),
\end{aligned}
$$

and the function $\varphi(\cdot)$ is the normal density function. The functions $P_{n,k}(\cdot)$ and $Q_{n,k}(\cdot)$ are reported in Appendix 1.C, as defined in Kou (2002). The function $Hh(\cdot)$, which can be viewed as a generalization of the cumulative normal distribution function, is defined in Abramowitz and Stegun (1972).

1.5.2 Case study: Home Depot

We consider a set of American-put options on Home Depot, issued on 01-21-2011 (evaluation date) and expiring on 05-21-2011. The initial stock level is $S_0 =$ \$ 36.51 and the risk-free rate is the one-month Libor rate, $r = 0.48\%$ (per year) on 01-21-2011. We use daily stock returns from 01-13-2009 to 01-21-2011 as an estimation time window. We use the test of randomness of Genest and Rémillard (2004) with four consecutive values and we cannot reject the i.i.d structure of the random sample $R_1,\ldots,R_M$ at the level of 5%.

Figure 1 reports the time series of daily stock returns. It shows frequent peaks, which suggest a jump-diffusion dynamics for the underlying asset. We use maximum likelihood to estimate pure diffusion and conventional jump-diffusion models, while we use maximum likelihood coupled with calibration to estimate their extended versions. We start with maximum likelihood estimates in the conventional jump-diffusion setting, select some liquid option contracts, use calibration to ap-

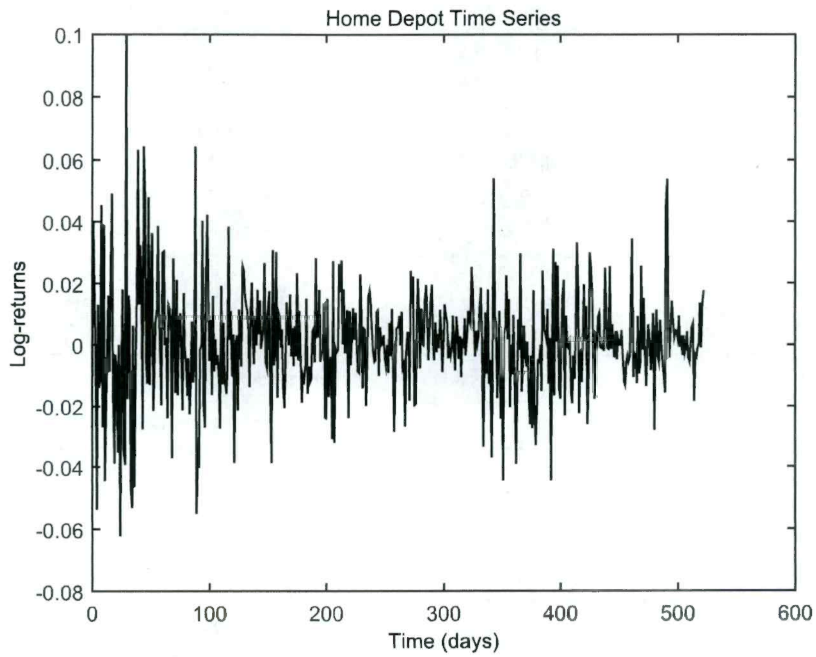

Figure 1.1: Home Depot log-returns

proximate α and β, and revise the maximum likelihood estimates, as explained in Section 2. Table 1.8 provides maximum-likelihood estimates. Their standard errors are indicated in brackets. Calibration results in $(\alpha,\beta) = (1, 0.228)$ for Merton and $(\alpha,\beta) = (0.9074, -0.8234)$ for Kou.

The estimated models present a key difference, that is, a low intensity and a high volatility for Merton's (1976) model and a high intensity and a low volatility for Kou's (2002) model. This illustrates the complexity to discern in between the pure component and the jump component of a jump-diffusion dynamics. Kou's model shows a small asymmetry between upward and downward jump amplitudes $(1/\eta_1 \simeq 1/\eta_2)$.

Tables 1.9-1.11 show that conventional jump-diffusion models outperform their pure diffusion counterparts. Moreover, extended jump-diffusion models further improve the final results.

Model	μ	σ	λ	γ	δ	η_1	η_2	p_1
BS	–	0.3472 (0.0209)	–	–	–	–	–	–
Merton	0.2720 (0.0191)	0.2554 (0.0098)	7.5066 (3.2955)	0.0474 (0.0068)	3.10^{-5} (0.0053)	–	–	–
Kou	-0.4901 (0.0568)	0.0430 (0.0118)	641.8961 (93.6113)	–	–	132.0057 (15.0636)	119.1263 (12.5161)	0.5904 (0.0467)

Table 1.8: Maximum likelihood estimates

	DP with a grid size p				Market	Relative
K	50	100	200	400	price	error
32	1.0182	1.0270	1.0273	1.0273	0.630	0.6306
33	1.3477	1.3393	1.3392	1.3391	0.830	0.6134
34	1.7187	1.7063	1.7065	1.7065	1.075	0.5874
35	2.1218	2.1300	2.1304	2.1304	1.395	0.5272
36	2.5987	2.6113	2.6109	2.6109	1.790	0.4586
37	3.1605	3.1465	3.1466	3.1468	2.265	0.3893

Table 1.9: American put options - BS (1973)

	DP with a grid size p				Market	Relative
K	50	100	200	400	price	error
Merton (1976)						
32	0.6446	0.6495	0.6496	0.6496	0.6300	0.0311
33	0.9085	0.9138	0.9138	0.9138	0.8300	0.1010
34	1.2367	1.2422	1.2424	1.2424	1.0750	0.1557
35	1.6429	1.6383	1.6384	1.6384	1.3950	0.1745
36	2.1092	2.1029	2.1030	2.1030	1.7900	0.1749
37	2.6419	2.6350	2.6350	2.6351	2.2650	0.1634
Extended Merton						
32	0.6372	0.6421	0.6422	0.6422	0.6300	0.0194
33	0.8997	0.9052	0.9051	0.9051	0.8300	0.0905
34	1.2268	1.2323	1.2325	1.2325	1.0750	0.1465
35	1.6322	1.6274	1.6276	1.6276	1.3950	0.1667
36	2.0979	2.0914	2.0915	2.0915	1.7900	0.1685
37	2.6303	2.6232	2.6233	2.6233	2.2650	0.1582

Table 1.10: American put options - Merton (1976)

	DP with a grid size p				Market	Relative
K	50	100	200	400	price	error
Kou (2002)						
32	0.6661	0.6605	0.6608	0.6608	0.6300	0.0490
33	0.9319	0.9229	0.9234	0.9235	0.8300	0.1126
34	1.2590	1.2499	1.2494	1.2494	1.0750	0.1622
35	1.6259	1.6427	1.6423	1.6424	1.3950	0.1773
36	2.1211	2.1019	2.1034	2.1035	1.7900	0.1751
37	2.6526	2.6303	2.6319	2.6320	2.2650	0.1620
Extended Kou						
32	0.5808	0.5881	0.5878	0.5878	0.6300	-0.0669
33	0.8284	0.8385	0.8380	0.8381	0.8300	0.0098
34	1.1651	1.1539	1.1535	1.1535	1.0750	0.0731
35	1.5545	1.5373	1.5385	1.5385	1.3950	0.1029
36	1.9749	1.9962	1.9948	1.9947	1.7900	0.1143
37	2.5435	2.5193	2.5216	2.5216	2.2650	0.1133

Table 1.11: American put options - Kou (2002)

1.6 Conclusion

We build a tractable framework for options pricing, where the underlying asset price follows a jump-diffusion process à la Merton (1976) and Kou (2002). Our construction is based on dynamic programming combined with finite elements. Our numerical investigation shows convergence and efficiency, and competes well against alternative methodologies. For the model estimation, we experiment with the maximum-likelihood approach alone and maximum likelihood combined with calibration. The latter outperforms the former.

Dynamic programming coupled with finite elements can be extended to handle two-dimensional jump-diffusion models. This is feasible as long as the transition tables can be computed in closed form and stored in the computer. For higher dimensional jump-diffusion models, dynamic programming coupled with Monte Carlo simulation is recommended. The state space becomes too large for a systematic split through regular hyper rectangles; rather, its is visited through random samples. Examples include the least squares Monte Carlo of Longstaff and Schwartz (2001) and Bally et al. (2005).

APPENDIX

1.A Change of measure

Recall that $h(x) = \alpha e^{\beta x}$, where $\alpha \in \mathbb{R}_+^*$ and $\beta \in \mathbb{R}$. For Merton's (1976), $a = \mathbb{E}[h(x)] = \alpha e^{\beta\gamma + \beta^2\delta^2/2}$, and the density $\tilde{f}(x)$ of $\tilde{v}$, under $\mathbb{Q}$, is

$$\begin{aligned}
\tilde{f}(x) &= h(x)f(x) \\
&= \alpha e^{\beta x}\lambda e^{-\frac{(x-\gamma)^2}{2\delta^2}}/(\sqrt{2\pi}\delta), \\
&= \alpha\lambda e^{-\frac{(x-\gamma)^2-2\beta\delta^2 x}{2\delta^2}}/(\sqrt{2\pi}\delta), \\
&= \alpha\lambda e^{\beta\gamma+\beta^2\delta^2/2} e^{-\frac{(x-\gamma-\beta\delta^2)^2}{2\delta^2}}, \\
&= \lambda e^{-\frac{(x-\gamma-\beta\delta^2)^2}{2\delta^2}}/(\sqrt{2\pi}\delta) \quad \text{for } x \in \mathbb{R},
\end{aligned}$$

where, under $\mathbb{Q}$, $\tilde{\lambda} = a\lambda$, $\tilde{\gamma} = \gamma + \beta\delta^2$, $\tilde{\delta} = \delta$, and $\tilde{\kappa} = e^{\tilde{\gamma}+\tilde{\delta}^2/2} - 1$.

For Kou's (1976) model, $\beta \in]-\eta_2, \eta_1[$, $a = \mathbb{E}[h(x)] = \alpha p_1\eta_1/(\eta_1 - \beta) + \alpha p_2\eta_2/(\eta_2 + \beta)$, and the density $\tilde{f}(x)$ of $\tilde{v}$, under $\mathbb{Q}$, is

$$\begin{aligned}
\tilde{f}(x) &= h(x)f(x) \\
&= \frac{\alpha\lambda(\eta_1-\beta)p_1\eta_1}{a(\eta_1-\beta)} e^{-(\eta_1-\beta)x}\mathbb{I}_{\{x\geq 0\}} + \frac{\alpha\lambda(\eta_2+\beta)p_2\eta_2}{a(\eta_2+\beta)} e^{(\eta_2+\beta)x}\mathbb{I}_{\{x<0\}}, \\
&= \frac{\lambda p_1\eta_1(\eta_1-\beta)}{(\eta_1-\beta)(\frac{p_1\eta_1}{\eta_1-\beta}+\frac{p_2\eta_2}{\eta_2+\beta})} e^{-(\eta_1-\beta)x}\mathbb{I}_{\{x\geq 0\}} + \\
&\qquad \frac{\lambda p_2\eta_2(\eta_2+\beta)}{(\eta_2+\beta)(\frac{p_1\eta_1}{\eta_1-\beta}+\frac{p_2\eta_2}{\eta_2+\beta})} e^{-(\eta_2+\beta)}\mathbb{I}_{\{x<0\}}, \\
&= \frac{p_1\eta_1(\eta_2+\beta)}{p_1\eta_1(\eta_2+\beta)+p_2\eta_2(\eta_1-\beta)}(\eta_1-\beta)e^{-(\eta_1-\beta)x}\mathbb{I}_{\{x\geq 0\}} + \\
&\qquad \frac{p_1\eta_2(\eta_1-\beta)}{p_2\eta_2(\eta_1-\beta)+p_1\eta_1(\eta_2+\beta)}(\eta_2+\beta)e^{-(\eta_2+\beta)}\mathbb{I}_{\{x<0\}}, \\
&= \tilde{p}_1\tilde{\eta}_1 e^{-\tilde{\eta}_1 x}\mathbb{I}_{\{x\geq 0\}} + \tilde{p}_2\tilde{\eta}_2 e^{\tilde{\eta}_2 x}\mathbb{I}_{\{x<0\}}, \quad \text{for } x \in \mathbb{R},
\end{aligned}$$

where, under $\mathbb{Q}$, $\tilde{\lambda} = a\lambda$, $\tilde{\eta}_1 = \eta_1 - \beta$, $\tilde{\eta}_2 = \eta_2 + \beta$, $\tilde{p}_1 = p_1\eta_1\tilde{\eta}_2/(p_1\eta_1\tilde{\eta}_2 + p_2\tilde{\eta}_1\eta_2)$, $\tilde{p}_2 = 1 - \tilde{p}_1$, and $\tilde{\kappa} = \tilde{p}_1\tilde{\eta}_1/(\tilde{\eta}_1 - 1) + \tilde{p}_2\tilde{\eta}_2/(\tilde{\eta}_2 + 1) - 1$.

1.B Transition tables - Merton (1976)

The transition parameters $T^{\nu}_{k,i}$, for $\nu \in \{0,1,2\}$, $k \in \{1,\ldots,p\}$, and $i \in \{0,\ldots,p\}$ are

$$T^{\nu}_{k,i} = \sum_{n=0}^{\infty} \mathbb{Q}(N_{\Delta t} = n)\ \eta^{\nu}_{k}(n) e^{c(n)^2/2} \left[\Phi\left(c_{k,i+1}(n) - c(n)\right) - \Phi\left(c_{k,i}(n) - c(n)\right)\right],$$

where $N_{\Delta t}$ is the number of jumps over $[t_m, t_{m+1}]$, $c(n) = \nu\sigma_n\sqrt{\Delta t}$, and

$$\mathbb{Q}(N_{\Delta t} = n) = e^{-\lambda\Delta t}\frac{(\lambda\Delta t)^n}{n!},$$

$$\sigma_n^2 = \sigma^2 + \frac{n}{\Delta t}\delta^2,$$

$$\eta_k(n) = a_k e^{\left(r-\bar{d}-\lambda\kappa-\sigma_n^2/2\right)\Delta t + n(\gamma+\delta^2/2)},$$

$$c_{k,i}(n) = \frac{\log\left(a_i/a_k\right) - \left(r-\bar{d}-\lambda\kappa-\sigma_n^2/2\right)\Delta t - n(\gamma+\delta^2/2)}{\sigma_n},$$

and $\Phi(\cdot)$ is the standard normal distribution function.

1.C Transition tables - Kou (2002)

The transition parameters $T^{\nu}_{k,i}$, for $\nu \in \{0,1,2\}$, $k \in \{1,\ldots,p\}$, and $i \in \{0,\ldots,p\}$ are

$$T^{0}_{k,i} = \Upsilon(\mu_0,\sigma,\lambda,p_1,\eta_1,\eta_2,x_{i+1},\Delta t) - \Upsilon(\mu_0,\sigma,\lambda,p_1,\eta_1,\eta_2,x_i,\Delta t),$$

$$T^{1}_{k,i} = \rho^{-1}a_k[\Upsilon(\mu_1,\sigma,\tilde{\lambda},\tilde{p}_1,\tilde{\eta}_1,\tilde{\eta}_2,x_{i+1},\Delta t) - \Upsilon(\mu_1,\sigma,\tilde{\lambda},\tilde{p}_1,\tilde{\eta}_1,\tilde{\eta}_2,x_i,\Delta t)],$$

$$T^{2}_{k,i} = b\rho^{-2}a_k^2[\Upsilon(\mu_2,\bar{\sigma},\bar{\lambda},\bar{p}_1,\bar{\eta}_1,\bar{\eta}_2,\bar{x}_{i+1},\Delta t) - \Upsilon(\mu_2,2\sigma,\bar{\lambda},\bar{p}_1,\bar{\eta}_1,\bar{\eta}_2,\bar{x}_i,\Delta t)],$$

where $\mu_0 = r - \frac{1}{2}\sigma^2 - \lambda\kappa, x_i = \log(a_i/a_k), \rho = \exp(-(r-\bar{d})\Delta t), \mu_1 = r + \frac{1}{2}\sigma^2 - \lambda\kappa, \tilde{\lambda} = \lambda(1+\kappa), \tilde{p}_1 = p\eta_1/(1+\kappa)(\eta_1-1), \tilde{\eta}_1 = \eta_1 - 1, \tilde{\eta}_2 = \eta_2 + 1, \bar{\sigma} = 2\sigma, \bar{\kappa} = p_1(\eta_1/2\bar{\eta}_1) + (1-p_1)(\eta_2/2\bar{\eta}_2) - 1, \mu_2 = 2r + \frac{1}{2}\bar{\sigma}^2 - \lambda\bar{\kappa}, \bar{\lambda} = \lambda(1+\bar{\kappa}), \bar{\eta}_1 = \eta_1/2 - 1, \bar{\eta}_2 = \eta_2/2 + 1, b = \exp(\sigma^2 + \lambda(\bar{\kappa} - 2\kappa)\Delta t)$, and $\bar{x}_i = x_i - \log(b)$. The function $\Upsilon(.)$ is defined by

$$
\begin{aligned}
\Upsilon(\mu,\sigma,\lambda,\eta_1,\eta_2,p_1,x_i,\Delta t) &= \frac{e^{(\sigma\eta_1)^2\Delta t/2}}{\sigma\sqrt{2\pi\Delta t}}\sum_{n=1}^{\infty}\pi_n\sum_{k=1}^{n}P_{n,k}\left(\sigma\sqrt{\Delta t}\eta_1\right)^k \\
&\times\ I_{k-1}\left(x_i-\mu\Delta t;-\eta_1,-\frac{1}{\sigma\sqrt{\Delta t}},-\sigma\eta_1\sqrt{\Delta t}\right) \\
&+\ \frac{e^{(\sigma\eta_2)^2\Delta t/2}}{\sigma\sqrt{2\pi\Delta t}}\sum_{n=1}^{\infty}\pi_n\sum_{k=1}^{n}Q_{n,k}\left(\sigma\sqrt{\Delta t}\eta_2\right)^k \\
&\times\ I_{k-1}\left(x_i-\mu\Delta t;\eta_2,\frac{1}{\sigma\sqrt{\Delta t}},-\sigma\eta_2\sqrt{\Delta t}\right) \\
&+\ \pi_0\Phi\left(-\frac{x_i-\mu\Delta t}{\sigma\sqrt{\Delta t}}\right),
\end{aligned}
$$

and

$$P_{n,k}=\Sigma_{i=k}^{n-1}\begin{pmatrix} n-k-1 \\ i-k \end{pmatrix}\begin{pmatrix} n \\ i \end{pmatrix}\cdot\left(\frac{\eta_1}{\eta_1+\eta_2}\right)^{i-k}\left(\frac{\eta_2}{\eta_1+\eta_2}\right)^{n-i}p_1^i p_2^{n-i},$$

$$Q_{n,k}=\Sigma_{i=k}^{n-1}\begin{pmatrix} n-k-1 \\ i-k \end{pmatrix}\begin{pmatrix} n \\ i \end{pmatrix}\cdot\left(\frac{\eta_1}{\eta_1+\eta_2}\right)^{n-i}\left(\frac{\eta_2}{\eta_1+\eta_2}\right)^{i-k}p_1^{n-i} p_2^{i},$$

$$I_n(c;\alpha,\beta,\delta)=\int_c^{\infty}e^{\alpha x}Hh_n(\beta x-\delta)\,dx,$$

for arbitrary constants α, c, $\beta\in\mathbb{R}$, and $n\in\mathbb{N}$.

1.D Validity of Assumption $\mathscr{D}$

Here we show that Assumption $\mathscr{D}$ holds true for linear and quadratic interpolations. One can guess that higher order polynomial interpolations also satisfy the assumption.

1.D.1 Linear interpolation

Suppose f is linearly interpolated on $I=[a,c]$ at points a,c by

$$\hat{f}(x)=f(a)+(x-a)\left\{\frac{f(c)-f(a)}{c-a}\right\}.$$

Then $\left\|f-\hat{f}\right\|_I \leq \frac{|f|_{Lip}}{2}(c-a)$. In fact,

$$f(x)-\hat{f}(x) = \frac{(c-x)(x-a)}{c-a}\left\{\frac{f(x)-f(a)}{x-a} - \frac{f(x)-f(c)}{x-c}\right\}$$

As a result, if $K = |f|_{Lip}$, then for any $x \in [a,c]$,

$$\left|f(x)-\hat{f}(x)\right| = 2K\frac{(c-x)(x-a)}{c-a} \leq \frac{K}{2}(c-a),$$

since $\sup_{x\in[a,c]}(c-x)(x-a) = \frac{(c-a)^2}{4}$. As a by-product, one obtains that if f is linearly interpolated on $I = [x_1,x_2],\ldots,[x_{n-1},x_n]$ at points $(x_1,x_2),\ldots,(x_{n-1},x_n)$, and the length of each interval is less or equal than h, then

$$\left\|f-\hat{f}\right\|_I = \sup_{x\in I}\left|f(x)-\hat{f}(x)\right| \leq \frac{|f|_{Lip,I}}{2}h,$$

where $|f|_{Lip,I}$ is the Lipschitz norm of f on $I = [x_1,x_n]$.

1.D.2 Quadratic interpolation

Suppose f is interpolated by a quadratic polynomial on $[a,c]$ at points a,b,c by

$$\hat{f}(x) = f(a) + (x-a)\left\{\frac{f(b)-f(a)}{b-a}\right\} + \frac{(x-a)(x-b)}{c-a}\left\{\frac{f(c)-f(b)}{c-b} - \frac{f(b)-f(a)}{b-a}\right\}.$$

It is then easy to check that $\left\|f-\hat{f}\right\| \leq 4|f|_{Lip}(c-a)$, since $|b-x| \leq (c-a)$, for any $x \in [a,c]$.

As a by-product, one obtains that if f is interpolated by a quadratic polynomial on $I = [x_1,x_n]$ at points $(x_1,x_2,x_3),(x_3,x_4,x_5),\ldots,(x_{n-2},x_{n-1},x_n)$, with n odd, and the length of each interval is less or equal than h, then

$$\left\|f-\hat{f}\right\|_I \leq 4|f|_{Lip,I}h.$$

The constant 4 is a little crude but it suffices for our purposes. Note that the smoother the function, the faster the convergence. In fact if f is differentiable and $|f'|_{Lip} < \infty$, then there is a constant C independent of f, such that $\left\|f-\hat{f}\right\|_I \leq C|f|_{Lip,I}h^2$.

1.E Proof of Theorem 1.3.2

For a given filtration $\mathbb{F}$ and a Markov process S adapted to the filtration, the value $v_0(x)$ at period 0 of the "Bermudan" option, with exercise values v_k^e and $S_0 = x$, is given by

$$v_0(x) = \sup_{\tau \in \mathscr{T}_{0,n}} E(\beta_\tau v_\tau^e(S_\tau)|S_0 = x),$$

where $\mathscr{T}_{k,n}$ stands for the set of all $\mathbb{F}$ stopping times with values in $\{k,\ldots,n\}$, $0 \leq k \leq n$. Here β_k is the discounted factor at k-th period, $\beta_0 = 1$, and $\beta_k/\beta_{k-1} = e^{-r_k}$, $k \in \{1,\ldots,n\}$. The sequence of functions

$$v_k(x) = \sup_{\tau \in \mathscr{T}_{k,n}} E(\beta_\tau v_\tau^e(S_\tau)|S_k = x)/\beta_k, \qquad k \in \{0,\ldots,n\},$$

is called the Snell envelope. It is well-known, e.g. Neveu (1975), that the Snell envelope can be computed in a recursive way. In fact, $v_n \equiv v_n^e$, and for $k = n-1,\ldots,0$,

$$v_k^h(x) = e^{-r_{k+1}} E\{v_{k+1}(S_{k+1})|S_k = x\}, \tag{1.10}$$

$$v_k(x) = \max\left\{v_k^e(x), v_k^h(x)\right\}. \tag{1.11}$$

In our model, $S_k = S_{k-1}Z_k = \pi_k(S_{k-1}, Z_k)$, where the $\log(Z_k) = X_{t_k} - X_{t_{k-1}}$ are increments of a Lévy process. Set

$$P_k g(x) = E\left[g\{\pi_k(x, Z_k)\}\right].$$

The approximation algorithm can be summarized as follows:

$$\hat{v}_n = v_n^e \text{ and } \begin{cases} \tilde{v}_{k-1}^h &= P_k \hat{v}_k \text{ on } \mathcal{G}_{k-1}, \\ \check{v}_{k-1} &= \max\left\{v_{k-1}^e, \tilde{v}_{k-1}^h\right\} \text{ on } \mathcal{G}_{k-1}, \quad k = n, \ldots, 1, \\ \hat{v}_{k-1} &= \mathcal{I}_{\mathcal{G}_{k-1}} \check{v}_{k-1}, \end{cases}$$

where $\mathcal{I}_{\mathcal{G}} g$ denotes the interpolation of g over the interval R associated with grid $\mathcal{G}$. The proof of the following proposition is easy. This result also appears in Bally et al. (2005).

Proposition 1.E.1. *For any real numbers x, y, z, w,*

$$|\max(x,z) - \max(y,w)| \leq \max(|x-y|, |z-w|).$$

In particular, $|\max(x,z) - \max(x,w)| \leq |z-w|$.

From now on, $S_k = S_{k-1} Z_k$, with $E(Z_k) = e^{-\delta_k} \leq 1$, where δ_k is the dividend factor period k.

We are now in a position to complete the proof.

Proof: First, note that since each v_k^e is Lipschitz over $[0, \infty)$, it follows that each $v_k^e(x) \leq v_k^e(0) + x \left|v_k^e\right|_{Lip}$ is bounded by a linear function in x. It also follows easily that v_k and v_k^h are Lipschitz over $[0, \infty)$, and $|v_k|_{Lip} \leq \max_{k \leq j \leq n} \left|v_j^e\right|_{Lip}$. In fact, by Proposition 1.E.1, for any $x \neq y$,

$$\begin{aligned} \frac{|v_k(x) - v_k(y)|}{|x-y|} &\leq \max\left\{\left|v_k^e(x) - v_k^e(y)\right| / |x-y|, \left|v_k^h(x) - v_k^h(y)\right| / |x-y|\right\} \\ &\leq \max\left\{\left|v_k^e\right|_{Lip}, \left|v_k^h\right|_{Lip}\right\}. \end{aligned}$$

Next,

$$\begin{aligned}\frac{\left|v_k^h(x)-v_k^h(y)\right|}{|x-y|} &\le \frac{1}{|x-y|}E\left\{\left|v_{k+1}(xZ_{k+1})-v_{k+1}(yZ_{k+1})\right|\right\}\\ &\le |v_{k+1}|_{Lip}E(Z_{k+1}) \le |v_{k+1}|_{Lip}.\end{aligned}$$

The, for any $k\in\{0,\ldots,n-1\}$,

$$|v_k|_{Lip} \le \max\left\{\left|v_k^e\right|_{Lip}, |v_{k+1}|_{Lip}\right\}.$$

As a result, for any $k\in\{0,\ldots,n-1\}$, $|v_k|_{Lip} \le \max_{k\le j\le n}\left|v_j^e\right|_{Lip}$.
For simplicity set $\mathscr{I}_{k-1}=\mathscr{I}_{\mathscr{G}_{k-1}}$ and $m_{k-1}=\mathrm{mesh}(\mathscr{G}_{k-1})$. To begin, one has, for any $k=1,\ldots,n$,

$$v_{k-1}-\hat{v}_{k-1} = (v_{k-1}-\mathscr{I}_{k-1}v_{k-1}) + (\mathscr{I}_{k-1}v_{k-1}-\mathscr{I}_{k-1}\check{v}_{k-1}).$$

Then, using Proposition 1.E.1, one obtains

$$\begin{aligned}\|v_{k-1}-\hat{v}_{k-1}\|_{\mathscr{R}_{k-1}} &\le Dm_{k-1}|v_{k-1}|_{Lip} + \|v_{k-1}-\check{v}_{k-1}\|_{\mathscr{G}_{k-1}}\\ &\le Dm_{k-1}|v_{k-1}|_{Lip} + \left\|v_{k-1}^h-\tilde{v}_{k-1}^h\right\|_{\mathscr{G}_{k-1}}\\ &\le Dm_{k-1}|v_{k-1}|_{Lip} + \|P_kv_k-P_k\hat{v}_k\|_{\mathscr{G}_{k-1}}\\ &\le Dm_{k-1}|v_{k-1}|_{Lip} + \|v_k-\hat{v}_k\|_{\mathscr{R}_k}\\ &\qquad + \left\|P_k\{(v_k-\hat{v}_k)1_{\mathscr{R}_k^c}\}\right\|_{\mathscr{G}_{k-1}}\\ &\le Dm_{k-1}|v_{k-1}|_{Lip} + \|v_k-\hat{v}_k\|_{\mathscr{R}_k}\\ &\qquad + \left\|P_k(v_k1_{\mathscr{R}_k^c})\right\|_{\mathscr{G}_{k-1}} + \left\|P_k\left(\hat{v}_k1_{R_k^c}\right)\right\|_{\mathscr{G}_{k-1}}.\end{aligned}$$

$R_k=\lambda_kR_{k-1}$, The aim is to show that if $\lambda_1,\ldots,\lambda_n$ are large enough, then the terms $\left\|P_k(v_k1_{\mathscr{R}_k^c})\right\|_{\mathscr{G}_{k-1}}$ and $\left\|P_k(\hat{v}_k1_{\mathscr{R}_k^c})\right\|_{\mathscr{G}_{k-1}}$ can be arbitrarily small, for all $1\le k\le n$. For a put, one can take $\lambda_k\equiv 1$ and $\mathscr{G}_k=\mathscr{G}$, with $m_k=m$. In this case, by construction, $\left\|P_k(\hat{v}_k1_{\mathscr{R}_k^c})\right\|_{\mathscr{G}_{k-1}}=0$. Also, it follows from Del Moral et al. (2012)

that v_k is decreasing and convex. Therefore, $\left\| P_k(v_k 1_{\mathscr{R}_k^c}) \right\|_{\mathscr{G}_{k-1}} \le v_k(R_0)$. Next, choose $R_0 > K$. Then for any $x \ge R_0$, $v_{n-1}(x) = v^h_{n-1}(x) \le KF_n(K/x)$, where F_k is the distribution function of Z_k, $k \in \{1,\ldots,n\}$. Now, suppose that if $x \ge K$, then $v_{n-k}(x) \le K\sum_{j=n-k+1}^n F_j\left\{(K/x)^{1/k}\right\}$. This is obviously true for $k = 1$. Then

$$
\begin{aligned}
v_{n-k-1}(x) &= v^h_{n-k-1}(x) + P_{n-k}v_{n-k}(x) = E\{v_{n-k}(xZ_{n-k})1_{\{Z_{n-k}\le y\}}\} \\
&\quad + E\{v_{n-k}(xZ_{n-k})1_{\{Z_{n-k}>y\}}\} \\
&\le KF_{n-k}(y) + v_{n-k}(xy) \\
&\le KF_{n-k}(y) + K\sum_{j=n-k+1}^{n} F_j\left\{(K/xy)^{1/k}\right\}.
\end{aligned}
$$

Taking $y = (K/x)^{1/(k+1)}$, one gets $xy \ge K$, so

$$
v_{n-k-1}(x) \le K\sum_{j=n-k}^{n} F_j\left\{(K/x)^{1/(k+1)}\right\}.
$$

Hence, for $k \in \{0,\ldots,n-1\}$, and for any $x \ge K$,

$$
v_k(x) \le K\sum_{j=k+1}^{n} F_j\left\{(K/x)^{1/(n-k)}\right\} \le K\sum_{j=1}^{n} F_j\left(K/x\right),
$$

which tends to 0 as $x \to \infty$. As a result,

$$
\|v_k - \hat{v}_k\|_{\mathscr{R}_0} \le (n-k)Dm + \sum_{j=k+1}^{n-1} v_j(R_0) \le (n-k)Dm + (n-k-1)K\sum_{j=1}^{n} F_j\left(K/R_0\right).
$$

It then suffices to take R_0 large enough so that $(n-1)K\sum_{j=1}^n F_j\left(K/R_0\right) < \varepsilon/2$, and then choose a grid $\mathscr{G}$ of $\mathscr{R}_0 = [0,R_0]$ with $m = \mathrm{mesh}(\mathscr{G}) < \varepsilon/(2nD)$.

For a call, note first that since v_k is Lipschitz with $|v_k|_{Lip} \le 1$, and since $v_k(0) = 0$, one gets $v_k(x) \le x$, for any $x \ge 0$. Also, using results in Del Moral et al. (2012), v_k

is convex and increasing. It then follows that

$$\left\| P_k\left(v_k 1_{\mathscr{R}_k^c}\right)\right\|_{\mathscr{G}_{k-1}} \le E\left\{v_k(R_{k-1}Z_k)1_{Z_k>\lambda_k}\right\} \le R_{k-1}E\left\{Z_k 1_{\{Z_k>\lambda_k\}}\right\}.$$

Since, by construction, $\left\| P_k(\hat{v}_k 1_{\mathscr{R}_k^c})\right\|_{\mathscr{G}_{k-1}} = 0$, one ends up with

$$\|v_k - \hat{v}_k\|_{\mathscr{R}_k} \le D\sum_{j=k}^{n-1} m_j + \sum_{j=k+1}^{n} R_{j-1}E\left\{Z_j 1_{\{Z_j>\lambda_j\}}\right\} \le D\sum_{j=0}^{n-1} m_j + \sum_{j=1}^{n} R_{j-1}E\left\{Z_j 1_{\{Z_j>\lambda_j\}}\right\}.$$

In order to complete the proof, first choose successively $\lambda_1,\ldots,\lambda_n$ so that R_{j-1} $E\left\{Z_j 1_{\{Z_j>\lambda_j\}}\right\} < \varepsilon/(2n)$. This is possible since each Z_k is positive and integrable. Next, recall that $R_k = R_{k-1}\lambda_k$, $k \in \{1,\ldots,n\}$. Then choose a grid $\mathscr{G} = \mathscr{G}_n$ of $\mathscr{R}_n$ with mesh $m \le \varepsilon/(2nD)$. Finally, set $\mathscr{G}_k = \mathscr{G} \cap \mathscr{R}_k$. It then follows that $m_k \le m$ and $\|v_k - \hat{v}_k\|_{\mathscr{R}_k} < \varepsilon$, for all $k \in \{0,\ldots,n\}$. Q.E.D.

In Bally et al. (2005), only the point convergence of $\hat{v}_0(S_0)$ to $v_0(S_0)$ is shown. Here we can prove uniform convergence on every finite interval $\mathscr{R}_0 = [0, R_0]$.

Remark 1.E.1. *The condition on the interpolation method is verified for the linear interpolation with $D = 1/2$ and the quadratic interpolation, for which one can take $D = 4$, even if it is a crude estimate. We needed that bound only to prove the convergence. The fact that we find a smaller D for the linear interpolation does not imply that its convergence is faster. In fact, it is quite the contrary, as shown in the numerical experiments.*

BIBLIOGRAPHY

Abramowitz, M. and Stegun, I. A. (1972). *Handbook of mathematical functions, 10th Printing.* U.S. National Bureau of Standards, Washington, D.C.

Ait-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics*, 74:487–528.

Almendral, A. and Oosterlee, C. W. (2005). Numerical valuation of options with jumps in the underlying. *Applied Numerical Mathematics*, 53:1–18.

Amin, K. (1993). Jump diffusion option valuation in discrete time. *Journal of Finance*, 48:1833–1863.

Andersen, L. and Andreasen, J. (2000). Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4:231–262.

Bally, V., Pagès, G., and Printems, J. (2005). A quantization tree method for pricing and hedging multidimensional American options. *Mathematical Finance*, 15:119–168.

Barone-Adesi, G. and Whaley, R. E. (1987). Efficient analytic approximation of American option values. *Journal of Finance*, 42:301–320.

Bates, D. S. (1991). The crash of '87: Was it expected? The evidence from options markets. *Journal of Finance*, 46:1009–1044.

Ben-Ameur, H., Breton, M., and L'Écuyer, P. (2002). A dynamic programming procedure for pricing american-style asian options. *Management Science*, 48:625–643.

Black, F. and Scholes, M. S. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–54.

Chan, R. and Hubbert, S. (2014). Options pricing under the one-dimensional jump-diffusion model using the radial basis function interpolation scheme. *Review of Derivatives Research*, 17:161–189.

Chiarella, C. and Ziogas, A. (2009). American call options under jump-diffusion processes:A Fourier transform approach. *Applied Mathematical Finance*, 16:37–79.

Chockalingam, A. and Muthuraman, K. (2010). Pricing American options when asset prices jump. *Operations Research Letters*, 38:82–86.

Cont, R. and Tankov, P. (2004a). *Financial Modelling with Jump Processes.* Chapman & Hall/CRC, Boca Raton, FL.

Cont, R. and Tankov, P. (2004b). Non-parametric calibration of jump-diffusion option pricing models. *Journal of Computational Finance*, 7:1–49.

Cont, R. and Voltchkova, E. (2005). Integro-differential equations for option prices in exponential Lévy models. *Finance and Stochastics*, 9:299–325.

Del Moral, P., Rémillard, B., and Rubenthaler, S. (2012). Monte Carlo approximations of American options that preserve monotonicity and convexity. In Carmona, R. A., Del Moral, P., Hu, P., and Oudjane, N., editors, *Numerical Methods in Finance*, volume 12 of *Springer Proceedings in Mathematics*, pages 115–143. Springer, Berlin Heidelberg.

DiCesare, J. and Mcleish, D. (2008). Simulation of jump diffusions and the pricing of options. *Insurance: Mathematics and Economics*, 43:316–326.

Feng, L. and Linetsky, V. (2008). Pricing options in jump-diffusion models: an extrapolation approach. *Operations Research*, 56:304–325.

Genest, C. and Rémillard, B. (2004). Test of independence and randomness based on the empirical copula process. *Test*, 13:335–369.

Gukhal, C. R. (2001). Analytical valuation of American options on jump-diffusion processes. *Mathematical Finance*, 11:97–115.

Gukhal, C. R. (2004). The compound option approach to American options on jump-diffusions. *Journal of Economic Dynamics and Control*, 28:2055–2074.

Hanson, F. and Zhu, Z. (2004). Comparison of market parameters for jump-diffusion distributions using multinomial maximum likelihood estimation. *Decision and Control, 43rd IEEE Conference*, 4:3919–3924.

He, C., Kennedy, J., Coleman, T., Forsyth, P., Li, Y., and Vetzal, K. (2006). Calibration and hedging under jump diffusion. *Review of Derivatives Research*, 9:1–35.

Hilliard, J. E. and Schwartz, A. (2005). Pricing European and American derivatives under a jump-diffusion process: A bivariate tree approach. *Journal of Financial & Quantitative Analysis*, 40:671–691.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48:1086–1101.

Kou, S. G. and Wang, H. (2004). Option pricing under a double exponential jump diffusion model. *Management Science*, 50:1178–1192.

Levendorskii, S. Z. (2004). Pricing of the American put under Lévy processes. *International Journal of Theoretical and Applied Finance*, 7:303–335.

Longstaff, F. A. and Schwartz, E. S. (2001). Valuing American options by simulation: A simple least-squares approach. *Review of Financial Studies*, 14:113–147.

Matache, A., Schwab, C., and Wihler, T. (2005). Fast numerical solution of parabolic integro-differential equations with applications in finance. *SIAM Journal on Scientific Computing*, 27:369–393.

Matache, A.-M., Nitsche, P.-A., and Schwab, C. (2003). Wavelet Galerkin pricing of American options on Lévy-driven assets. *Quantitative Finance*, 5:403–424.

Mayo, A. (2008). Methods for the rapid solution of the pricing {PIDEs} in exponential and merton models. *Journal of Computational and Applied Mathematics*, 222:128–143.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

Neveu, J. (1975). *Discrete-parameter Martingales.* North Holland, Amsterdam.

Press, S. J. (1967). A compound events model for security prices. *Journal of Business*, 40:317–335.

Rémillard, B. (2013). *Statistical Methods for Financial Engineering.* Chapman & Hall/CRC, Boca Raton, FL.

Salmi, S. and Toivanen, J. (2014). Imex schemes for pricing options under jump-diffusion models. *Applied Numerical Mathematics*, 84:33–45.

Simonato, J.-G. (2011). Computing American option prices in the lognormal jump diffusion framework with a Markov chain. *Finance Research Letters*, 8:220–226.

Toivanen, J. (2008). Numerical valuation of european and american options under kou's jump-diffusion model. *SIAM Journal on Scientific Computing*, 30:1949–1970.

Zhang, X. L. (1997). Numerical analysis of American option pricing in a jump-diffusion model. *Mathematics of Operations Research*, 22:668–690.

CHAPTER 2

DYNAMIC PROGRAMMING FOR VALUING AMERICAN OPTIONS UNDER THE VARIANCE-GAMMA PROCESS

Abstract

Hatem Ben Ameur, [1] Rim, Chérif, [2] Bruno Rémillard [3]

Lévy processes provide a solution to overcome the shortcomings of the lognormal hypothesis. A growing literature proposes the use of pure-jump Lévy processes such as the variance-gamma model. In this setting, explicit solutions for derivative prices are unavailable, for instance for the valuation of American options. We propose a dynamic programming approach coupled with finite elements for valuing American-style options under an extended variance-gamma model. Our numerical experiments confirm the convergence and show the efficiency of the proposed methodology. We also conduct a numerical investigation that focuses on American options on the S&P 500 futures contracts.

Keywords: American options; Variance Gamma process; Dynamic programming; finite elements, Calibration; Maximum likelihood.

2.1 Introduction

A special class of diffusion processes, specifically the generalized hyperbolic distributions, have straightforward parallels with Lévy processes. They have become extremely relevant in mathematical finance as they provide tools to accurately consider enough of the desired properties of asset returns, both in the real and the risk-neutral worlds. One member of the generalized hyperbolic family, and one of

---

1. Hatem Ben-Ameur is a Professor at HEC Montréal in the Department of Decision Sciences, and member of the GERAD.

2. Rim Chérif is a Ph.D. student at HEC Montréal.

3. Bruno Rémillard is a Professor at HEC Montréal in the Department of Decision Sciences, and member of the CRM and GERAD.

the most popular Lévy processes used in financial modeling, is the variance-gamma process that, in itself, synthesizes several desired movement types for the asset price. On the one hand, it allows an infinite number of low-amplitude jumps that behave like diffusion and, on the other, permits a finite number of high-amplitude jumps whose intensity decreases as amplitude increases.

The variance-gamma model is unique in that Brownian motion varies according to a stochastic time scale given by the gamma process. With this connection, price fluctuations are expressed according to a business time scale rather than calendar time. Thus, a stochastic time change can have two effects: it can speed up calendar time and subject the market to turbulence, or slow down calendar time and maintain an unperturbed market. The variance-gamma process enables accurate financial applications in, for example, modeling oil price dynamics (Askari and Krichene, 2008), credit risk (Fiorani et al., 2010), options on stocks, energy and currency prices (Pinho and Madaleno, 2011), (Daal and Madan, 2005). The variance-gamma process' flexibility is confirmed by Stein et al. (2007) and leveraged by the Bloomberg system through the implementation of the SKEW function, which compares option pricing under Black and Scholes (1973) and variance-gamma settings. The SKEW function inadvertently led to an even greater popularity for this process within the financial industry. All of this suggests the practical and theoretical relevance for further exploring the variance-gamma model to develop efficient numerical methods for financial applications.

The symmetrical version of variance-gamma was introduced by Madan and Seneta (1990) and Madan and Milne (1991), and then extended to the asymmetrical case by Madan et al. (1998). These authors have proposed a formula to value European options. However, given that solutions are only available for European-style options, a great deal of research is being dedicated to implementing efficient numerical methods to evaluate other types of options, such as Bermudan options.

Several algorithms have been developed to assess Bermudan options under the variance-gamma model. The most common are based on Monte Carlo methods in plain or enhanced form by using control and antithetic variates, quasi-Monte

Carlo methods, and importance sampling (Avramidis and L'Écuyer, 2006, Kaishev and Dimitrova, 2009, Becker, 2010, Dingeç and Hörmann, 2012). Other methods include the lattice models (Këllezi and Webber, 2004, Maller et al., 2006), partial-integro-differential equations (Fiorani, 1999, Hirsa and Madan, 2004, Almendral and Oosterlee, 2007a,b, Mayo, 2009), Fourier transforms (Carr and Madan, 1999, Wong and Guan, 2011, Haslip and Kaishev, 2014), and Markov chains (Konikov and Madan, 2002).

We evaluate options under the variance-gamma model in line with Ben-Ameur et al. (2016). Our approach is based on dynamic programming coupled with finite elements. The value function under consideration is approximated by a piecewise polynomial at each decision date. High-order polynomials are accurate but are time consuming. We are therefore looking for a compromise between accuracy and computing time. We compare piecewise-constant, -linear, and -quadratic approximations, and conclude this last one to be most efficient.

The remainder of this paper is organized as follows: Section 2.2 presents an extension of the variance-gamma model by Madan et al. (1998), Section 2.3 describes a stochastic dynamic program that solves the optimal Markov decision process embedded in the American-option valuation problem, Sections 2.4 and 2.5 present our numerical investigation and empirical study on S&P 500 futures contracts, and finally, Section 2.6 concludes our analysis.

2.2 The variance-gamma process

Let $\mathscr{M}$ be a frictionless market with a constant instantaneous interest rate r and risky asset S experiencing jumps at random times, and let $(\Omega, \mathscr{F}_t, \mathbb{P})$ be a complete probability space. Let the variance-gamma process be defined as a process obtained by evaluating a Brownian motion, with a drift at a random time given by a gamma process. Let $B = \{B(t;\theta,\sigma) = \theta t + \sigma W_t, t \geq 0\}$ be the Brownian motion with drift parameter $\theta \in \mathbb{R}$ and volatility parameter $\sigma > 0$, where W is standard Brownian motion and $B_0 = W_0 = 0$, for all $t \geq 0$. Let $G = \{G_t, t \geq 0\}$ be a Gamma process with unit mean rate and variance parameter $\nu > 0$, indicated by $G \sim GAMMA(\nu)$ and independent of B with independent increments, where $G_0 = 0$. For $t > 0$, the density of G_t is given by

$$f_{t,\nu}(x) = \frac{x^{\frac{t}{\nu}-1} \exp\left(-\frac{x}{\nu}\right)}{\nu^{\frac{t}{\nu}} \Gamma\left(\frac{t}{\nu}\right)}, \quad \text{for } x \geq 0,$$

where $\Gamma(\cdot)$ is the gamma function. The parameter $1/\nu$ controls the intensity of jumps of all sizes simultaneously and captures the decay rate of large jumps. We provide the variance-gamma process moments in Appendix 2.A.

More precisely, a variance-gamma process X with parameters (θ, σ, ν), indicated by $X \sim VG(\theta, \sigma, \nu)$, is a Lévy process defined by $X(t) = B(G_t)$, where $t \geq 0$, and where B and G were previously defined. By examining the four moments of the variance-gamma process (see Appendix 2.A), the parameters θ and ν can be interpreted as indirectly capturing the skewness and the kurtosis of stock returns.

The procedure for evaluating Brownian motion at a positive non-decreasing Lévy process is called subordination. Hence, the variance-gamma process is a time-changed (subordinated) Brownian motion, where the subordinator is a gamma process. This subordinated time process may be interpreted as market activity time, reflecting the random speedups and slowdowns of real-time business activity.

The Lévy process can be viewed as a continuous-time analog of a sequence of partial sums of independently and identically distributed random variables; its distribution is characterized by the Lévy-Khintchine representation given by Theorem 2.2.1.

Theorem 2.2.1. *Every Lévy process* (X_t) *is characterized by a triplet* (a, ϑ^2, k) *where* $a \in \mathbb{R}$, $\vartheta \in \mathbb{R}^+$, *and* k *is a measure satisfying* $k(\{0\}) = 0$ *and* $\int_{\mathbb{R}} \left(1 \wedge |x|^2\right) k(dx) \leq \infty$, *determining its characteristic exponent as*

$$\psi(u) = iua - \frac{u^2 \vartheta^2}{2} + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\mathbb{I}_{|x|<1}\right) k(dx),$$

i.e., $\mathbb{E}\left[e^{iuX_t}\right] = e^{t\psi(u)}$, *for all* $u \in \mathbb{R}$.

Thus, Lévy processes can be parameterized using the triplet (a, ϑ^2, k), where a is the drift, ϑ^2 is the variance of Brownian motion, and k is the so-called Lévy measure or jump measure, which characterizes the size and frequency of the jumps. In fact, if A is a closed interval not containing zero, then $N_A(t) = \sum_{0<s\leq t} \mathbb{I}(\Delta X_s \in A)$ is a Poisson process with intensity $k(A)$. As a result, if this measure is infinite, then the process has an infinite number of jumps of very small sizes in any small interval. Unlike Brownian motion B with characteristics $\left(0, \sigma^2, 0\right)$, the variance-gamma process is a Lévy process with characteristics $(a, 0, k)$ and with a Lévy measure associated with X having a density ℓ

$$\ell(x) = \frac{\exp(\theta x/\sigma^2)}{\nu|x|} \exp\left(-\frac{\sqrt{\frac{2}{\nu}+\frac{\theta^2}{\sigma^2}}}{\sigma}|x|\right), \quad (2.1)$$

where $a = \int_{|x|<1} x k(dx)$ and where the infinite activity of the process is observed in the Lévy density, which has the behavior of $\left(1/|x|\right)$ near zero. The integral of the Lévy density is therefore infinite. The special case of $\theta = 0$ in eq. (2.1) yields a Lévy measure that is symmetrical around zero, a case initially used by Madan and Seneta (1990) and Madan and Milne (1991). Furthermore, the density $X_t \sim (\theta, \sigma, \nu)$ of the variance-gamma model characterized by (θ, σ, ν) is given by

$$\begin{aligned} f_t^{VG}(x; \sigma, \theta, \nu) = & \frac{2\exp\left(\theta x/\sigma^2\right)}{\nu^{t/\nu}\sqrt{2\pi}\sigma\Gamma\left(\frac{t}{\nu}\right)} \left(\frac{x^2}{2\sigma^2/\nu+\theta^2}\right)^{\frac{t}{2\nu}-\frac{1}{4}} \times \\ & K_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{x^2\left(2\sigma^2/\nu+\theta^2\right)}\right), \end{aligned} \quad (2.2)$$

where K_z is the modified Bessel function of the second kind of index z.

Interestingly, the variance-gamma process may also be expressed as the difference of two independent increasing gamma processes Y_1, Y_2, viz., $X_t = \beta_1 Y_{1,t} - \beta_2 Y_{2,t}$, with $Y_i \sim GAMMA(\nu)$, $i \in \{1,2\}$, where β_1 and β_2 are defined by

$$\beta_1 = \frac{1}{2}\sqrt{2\sigma^2\nu + \theta^2\nu^2} + \frac{\theta\nu}{2} \quad \text{and} \quad \beta_2 = \frac{1}{2}\sqrt{2\sigma^2\nu + \theta^2\nu^2} - \frac{\theta\nu}{2}.$$

Using these results, one may conclude that the density of the Lévy measure associated with X can also be written as

$$\ell(x) = \frac{e^{-x/\beta_1}}{\nu|x|}\mathbb{I}_{\{x>0\}} + \frac{e^{x/\beta_2}}{\nu|x|}\mathbb{I}_{\{x<0\}}, \quad \text{for } x \in \mathbb{R}. \tag{2.3}$$

It is thus evident that expressions (2.1) and (2.3) are indeed the same.

In general, when the log-price is modeled by a Lévy process, the market is arbitrage free but incomplete, that is, there is an infinite number of equivalent martingale measures. This is mainly due to the presence of jumps, and is certainly true for the pure-jump model considered here. Based on Cont and Tankov (2004), we present a general change of measure for Lévy processes. To this end, let ϕ be a positive function satisfying the integrability condition

$$K_\phi = \int \{e^{\phi(x)/2} - 1\}^2 k(dx) < \infty. \tag{2.4}$$

Theorem 2.2.2. *Let X be a Lévy process with characteristics $\{a, \vartheta^2, k\}$. For every $b \in \mathbb{R}$ and ϕ satisfying (2.4), there exists an equivalent probability measure $\mathbb{Q}^{b,\phi}$ so that, under $\mathbb{Q}^{b,\phi}$, X is a Lévy process with characteristics $\{\tilde{a}, \tilde{\vartheta}^2, \tilde{k}\}$, with*

$$\tilde{a} = a + b\vartheta + \int_{-1}^{1} x\{e^{\phi(x)/2} - 1\}^2 k(dx), \tag{2.5}$$

$\tilde{\vartheta} = \vartheta$, and the Lévy measure $\tilde{k}$ having density e^ϕ with respect to k.

In particular, its cumulant function, provided it exists in the neighborhood of 0, is given by

$$\tilde{c}(u) = \tilde{a} + \frac{a^2}{2}\tilde{\vartheta} + \int_{\mathbb{R}} \left(e^{ux} - 1 - ux\mathbb{I}_{|x|<1}\right)\tilde{k}(dx). \tag{2.6}$$

Remark 2.2.1. *Given Theorem 2.2.2, it is easy to verify that the discounted value $e^{-(r-q)t}S_t$ is a martingale under $\mathbb{Q}^{b,\phi}$ if and only if $\tilde{c}$ given by (2.6) is finite for all*

$u \in [0,1]$, and $\tilde{c}(1) = r - q$. Here, q is the continuously compounded dividend yield.

Even if Theorem 2.2.2 gives a general change of measure for Lévy processes, it does not guarantee that under the change of measure, the process stays in the same family, which is a desirable property in practice. For example, in our case, we would like to concentrate on changes of measure determined by ϕ so that the process remains a variance gamma, possibly with new parameters $\left(\tilde{\theta}, \tilde{\sigma}, \tilde{v}\right)$. To this end, we propose to use the function (Rémillard, 2013, Remark.6.5.2)

$$\phi(x) = \xi_1 x \mathbb{I}(x > 0) + \xi_2 x \mathbb{I}(x \le 0), \quad \text{for } x \in \mathbb{R}, \tag{2.7}$$

where $\xi_1 < 1/\beta_1$ and $\xi_2 > -1/\beta_2$, so that the integrability condition (2.4) is met. In this case, the density $\tilde{\ell}$ of the associated Lévy measure is given by

$$\begin{aligned}\tilde{\ell}(x) = e^{\phi(x)}\ell(x) &= \frac{e^{-x(1-\xi_1\beta_1)/\beta_1}}{v|x|}\mathbb{I}_{\{x>0\}} + \frac{e^{x(1+\xi_2\beta_2)/\beta_2}}{v|x|}\mathbb{I}_{\{x<0\}} \\ &= \frac{e^{-x/\tilde{\beta}_1}}{v|x|}\mathbb{I}_{\{x>0\}} + \frac{e^{x/\tilde{\beta}_2}}{v|x|}\mathbb{I}_{\{x<0\}}, \quad \text{for } x \in \mathbb{R},\end{aligned}$$

where $\tilde{\beta}_1 = \beta_1/(1-\xi_1\beta_1)$, $\tilde{\beta}_2 = \beta_2/(1+\xi_2\beta_2)$, with $\tilde{v} = v$, $\tilde{\theta} = (\tilde{\beta}_1 - \tilde{\beta}_2)/\tilde{v}$, and $\tilde{\sigma} = \sqrt{2\tilde{\beta}_1\tilde{\beta}_2/\tilde{v}}$. This is our extended variance-gamma model.

It follows from Remark 2.2.1 that under this change of measure,

$$\tilde{S}_t = S_0 \exp\{(r - q + \tilde{w})t + \tilde{X}_t\}, \quad \text{for } t \ge 0, \tag{2.8}$$

where $r > 0$ is the risk-free rate of interest, q is the asset's continuously compounded dividend yield, $\tilde{w} = \ln(1 - \tilde{\theta}\tilde{v} - \tilde{\sigma}^2\tilde{v}/2)/\tilde{v}$, and $\tilde{X}$ is still a Lévy process represented by $\tilde{X}_t \sim VG\left(\tilde{\sigma}, \tilde{\theta}, \tilde{v}\right)$.

Madan et al. (1998) developed a risk-neutral valuation setting based on a general-equilibrium economy where the representative agent has a constant-relative risk-aversion utility function. Their equilibrium model corresponds to the simple case $\phi \equiv 0$, i.e, $\xi_1 = \xi_2 = 0$.

2.3 Valuing American-style options

2.3.1 The option contract

An American option on a stock with maturity T is characterized by its known exercise value $v_t^e(s)$, where $s = S_t$ is the stock price at time $t \in [0,T]$. For example, an American vanilla option is defined by

$$v_t^e(s) = \begin{cases} (s-K)^+ & \text{for a call option} \\ (K-s)^+ & \text{for a put option} \end{cases},$$

where K is the option's exercise price and $(x)^+ \equiv \max(0,x)$. We consider its Bermudan version which admits a finite number of exercise opportunities given by $t_0 = 0, \ldots, t_m, \ldots, t_M = T$. For simplicity, assume that $t_{m+1} - t_m = \Delta t$ is a fixed constant. By the no-arbitrage evaluation principle, the option's holding value at t_m is

$$\begin{aligned} v_m^h(s) &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r\Delta t} v_{m+1}(S_{t_{m+1}}) \mid S_{t_m} = s\right], \\ &\text{for } m = 0, \ldots, M, \end{aligned} \tag{2.9}$$

where $\mathbb{E}_{\mathbb{Q}}[\cdot | S_{t_m} = s]$ is the conditional expectation under $\mathbb{Q}$. With the convention that $v_M^h(\cdot) = 0$, the option must be exercised at maturity $t_M = T$. The option's overall value at time t_m is

$$v_m(s) = \max\left(v_m^e(s), v_m^h(s)\right), \quad \text{for } s > 0. \tag{2.10}$$

The associated European version is characterized by $v_m^e(\cdot) = 0$, for $m = 0, \ldots, M-1$.

2.3.2 The dynamic program

Let $\mathscr{G} = \{a_1, \ldots, a_p\}$ be the grid points, where $0 = a_0 < a_1 < \ldots < a_p < +\infty$ and $\Delta a_i = a_i - a_{i-1}$, $i = 1, \ldots, p$. The grid $\mathscr{G}$ must be selected so that $\text{mesh}(\mathscr{G}) = \max_{1 \le i \le p} \Delta a_i$, $\mathbb{Q}(S_t < a_1)$ and $\mathbb{Q}(S_t > a_p)$ all converge to 0 as $p \to \infty$, for $t \in$

$\{t_1,\ldots,t_M\}$. We select the grid points to be the quantiles of S_T.

Suppose that an approximation $\tilde{v}_{m+1}(\cdot)$ of the value function $v_{m+1}(\cdot)$ is available at a given future date t_{m+1} on the grid $\mathcal{G}$. This is not a strong assumption since the value function $v_M(\cdot) = v_M^e(\cdot)$ is known in closed form.

We use a piecewise polynomial and extend the approximation $\tilde{v}_{m+1}(\cdot)$ from $\mathcal{G}$ to the overall state space $\mathbb{R}_+^*$, that is, for $s > 0$,

$$\hat{v}_{m+1}(s) = \sum_{i=0}^{p} \left(\beta_i^0 + \beta_i^1\, s + \beta_i^2\, s^2\right) \mathbb{I}(a_i \leq s < a_{i+1}), \tag{2.11}$$

where the local coefficients depend on the time step $m+1$. Next, (2.9) and (2.11) yield

$$\begin{aligned} \tilde{v}_m^h(a_k) &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r\Delta t}\hat{v}_{m+1}\left(S_{t_{m+1}}\right)|S_{t_m} = a_k\right] \\ &= e^{-r\Delta t}\sum_{i=0}^{p}\left(\beta_i^0\, T_{ki}^0 + \beta_i^1\, T_{ki}^1 + \beta_i^2\, T_{ki}^2\right), \end{aligned} \tag{2.12}$$

where for $l \in \{0,1,2\}$

$$T_{ki}^l = \mathbb{E}_{\mathbb{Q}}\left[S_{t_{m+1}}^l \mathbb{I}\left(a_i \leq S_{t_{m+1}} < a_{i+1}\right)|S_{t_m} = a_k\right].$$

These transition matrix are derived and computed in Appendix 2.B. As is evident, we end up with several components: a number of explicit calculations and few valuations of integrals using Gauss-Legendre approximations.

From (2.10), the approximate value function at t_m is

$$\tilde{v}_m(a_k) = \max\left(v_m^e(a_k), \tilde{v}_m^h(a_k)\right), \quad \text{for } a_k \in \mathcal{G}, \tag{2.13}$$

and the approximate exercise policy at t_m is characterized by

$$\tilde{v}_m^h(a_k) < v_m^e(a_k), \quad \text{for } a_k \in \mathcal{G}.$$

The local coefficients of the piecewise-quadratic interpolation $\hat{v}_{m+1}(\cdot)$ in eq. (2.11) verify $\hat{v}_{m+1}(\cdot) = \tilde{v}_{m+1}(\cdot)$ on $\mathscr{G}$. Following Ben-Ameur et al. (2016), the dynamic program works as follows:

1. For $m = M - 1$, set $\tilde{v}_{m+1}(\cdot) = v_{m+1}(\cdot) = v^e_{m+1}(.)$ on $\mathscr{G}$;
2. Use a piecewise-quadratic interpolation as in eq. (2.11), and extend $\tilde{v}_{m+1}(\cdot)$, defined on $\mathscr{G}$, to $\hat{v}_{m+1}(\cdot)$, defined on the overall state space $\mathbb{R}^*_+$;
3. By eq. (2.12), compute $\tilde{v}^h_m(\cdot)$, defined on $\mathscr{G}$;
4. By eq. (2.13), compute $\tilde{v}_m(\cdot)$, defined on $\mathscr{G}$;
5. Exercise the option at (t_m, a_k) if $\tilde{v}^h_m(a_k) < v^e_m(a_k)$;
6. If $m = 0$, stop; else set $m = m - 1$, and go to step 2.

Ben-Ameur et al. (2016) show that the proposed methodology provides uniformly convergent approximations of the real prices over any given compact set at each time step.

2.4 Numerical investigation

The code lines is written in C, compiled with GCC, and executed using a standard laptop computer running Windows 7. For each set of results, we report the highest total CPU time in seconds. Our CPU times cannot be directly compared to the CPU times reported in the literature because they are obtained using different hardware characteristics, operation systems, and programming languages; regardless, we report them for information purposes.

To begin with, Tables 2.1 and 2.2 compare dynamic programming (DP) to the lattice method of Këllezi and Webber (2004) for European call options and Bermudan put options, respectively, both with 10 time steps. Table 2.1 also reports true values computed with the formulas proposed by Madan et al. (1998). Set $S_0 =$ \$100, $q = 0$, $\tilde{\sigma} = 0.12$, $T = 1$ (year), $r = 0.1$, $\tilde{\theta} = -0.14$, and $\tilde{\nu} = 0.2$. Dynamic programming is capable of reproducing benchmark call prices and shows convergence and efficiency. It ensures accuracy to the fourth digit within a few seconds. As we investigate the prices of Bermudan put options, dynamic programming appears to be accurate to the third digit at the finest grid size. It is also consistent with the prices proposed by Këllezi and Webber (2004), but with a greater competitive CPU time.

K	DP with a grid size p 75	150	300	Madan et al. (1998)	Lévy lattice
90	19.0996	19.0996	19.0994	19.0994	19.0994
95	15.0694	15.0706	15.0705	15.0705	15.0705
100	11.3669	11.3701	11.3700	11.3700	11.3700
105	8.1181	8.1190	8.1198	8.1198	8.1198
110	5.4276	5.4296	5.4296	5.4296	5.4296
115	3.3615	3.3648	3.3653	3.3654	3.3654
120	1.9203	1.9188	1.9210	1.9211	1.9211
CPU	(4.20)	(17.13)	(21.37)		(204.20)

Table 2.1: European call options – $M = 10$
DP versus Këllezi and Webber (2004)

K	DP with a grid size p 75	150	300	Lévy lattice
90	0.7625	0.7613	0.7613	0.7612
95	1.5221	1.5258	1.5259	1.5257
100	2.8774	2.8814	2.8816	2.8815
105	5.1585	5.1690	5.1705	5.1704
110	9.0417	9.0411	9.0406	9.0406
115	13.8754	13.8762	13.8762	13.8762
120	18.8097	18.8097	18.8097	18.8097
CPU	(4.29)	(18.05)	(20.01)	(188.80)

Table 2.2: Bermudan put options – $M = 10$
DP versus Këllezi and Webber (2004)

Tables 2.3 and 2.4 compare dynamic programming to the lattice method of Këllezi and Webber (2004) for European call options and Bermudan put options, with the same set of parameters, but with 20 time steps over time to maturity of

K	DP with a grid size p			Madan et al. (1998)	Lévy lattice
	75	150	300		
90	11.9719	11.9715	11.9717	11.9716	11.9715
95	7.4154	7.4205	7.4210	7.4210	7.4209
100	3.4379	3.4354	3.4381	3.4380	3.4379
105	0.8375	0.8320	0.8293	0.8292	0.8292
110	0.1464	0.1497	0.1496	0.1496	0.1496
115	0.0289	0.0291	0.0291	0.0291	0.0291
120	0.0060	0.0061	0.0061	0.0061	0.0061
CPU	(4.17)	(16.74)	(28.07)		(425.50)

Table 2.3: European call options – $M = 20$
DP versus Këllezi and Webber (2004)

0.2 years to consider a small time step of 0.01. Again, dynamic programming shows consistency with the prices computed by Këllezi and Webber (2004).

K	DP with a grid size p			Lévy lattice
	75	150	300	
90	0.2110	0.2126	0.2128	0.2125
95	0.6064	0.6121	0.6129	0.6128
100	1.6829	1.6861	1.6901	1.6899
105	4.9026	4.9156	4.9159	4.9159
110	9.8920	9.8921	9.8921	9.8921
115	14.8853	14.8853	14.8853	14.8853
120	19.8801	19.8801	19.8801	19.8801
CPU	(5.12)	(15.87)	(33.98)	(508.90)

Table 2.4: Bermudan put options – $M = 20$
DP versus Këllezi and Webber (2004)

Furthermore, Table 2.5 reports option pricing with the above set of parameters, maturing in $T = 0.2$ (years) and computed with 40 time steps. Option values computed with dynamic programming and with $M = 40$ time steps are consistent with those computed with $M = 20$ time steps. It is worth noting that Këllezi and Webber (2004) experienced problems in increasing the number of time steps higher than 40. In fact, under their setting, option values with 40 time steps are lower than those of 20 time steps, especially for the out-of-the-money options.

	DP with a grid size p			
K	75	150	300	Lévy lattice
90	0.2113	0.2133	0.2133	0.2112
95	0.6070	0.6139	0.6149	0.6128
100	1.6857	1.6906	1.6960	1.6940
105	4.9454	4.9574	4.9576	4.9575
110	9.9459	9.9460	9.9460	9.9459
115	14.9426	14.9426	14.9426	14.9426
120	19.9400	19.9400	19.9400	19.9400
CPU	(5.24)	(16.68)	(38.91)	(650.20)

Table 2.5: Bermudan put options – $M = 40$
DP versus Këllezi and Webber (2004)

Tables 2.6 and 2.7 compare dynamic programming to the multinomial model of Maller et al. (2006) for European (E-MN) and Bermudan put options (A-MN). Their procedure can be viewed as generalizing the binomial tree.

	DP with a grid size p				
K	75	150	300	E-FFT	E-MN
90	0.2264	0.2307	0.2310	0.2304	0.2280
95	0.6098	0.6210	0.6221	0.6218	0.6212
100	1.5445	1.5716	1.5710	1.5708	1.5720
105	3.6924	3.6929	3.6922	3.6925	3.6974
110	7.5567	7.5589	7.5578	7.5572	7.5594
CPU	(4.18)	(23.75)	(53.02)		

Table 2.6: Bermudan put option –
DP versus Maller et al. (2006)

European option prices obtained under a multinomial scheme and under dynamic programming are compared to those obtained from the fast Fourier transform solution (E-FFT) reported in Maller et al. (2006). Set $S_0 = \$100$, $q = 0$, $\tilde{\sigma} = 0.12$, $T = 0.25$ (years), $r = 0.1$, $\tilde{\theta} = -0.14$, and $\tilde{\nu} = 0.2$. Both methods run with 800 time steps. Here again, dynamic programming shows consistency and better convergence to the European prices, even with an increasing number of time

	DP with a grid size p			
K	75	150	300	A-MN
90	0.2620	0.2659	0.2667	0.2651
95	0.7156	0.7263	0.7284	0.7270
100	1.8651	1.8806	1.8832	1.8836
105	4.9918	4.9976	4.9975	5.0000
110	9.9966	9.9966	9.9966	10.0000
CPU	(3.04)	(19.75)	(54.12)	

Table 2.7: Bermudan put options – DP versus Maller et al. (2006)

steps. Maller et al. (2006) do not report their computation time.

Next, Tables 2.8 and 2.9 compare dynamic programming to the fast Fourier transform (FFT) network of Wong and Guan (2011) where both schemes run with 100 time steps. Set $S_0 = \$100$, $K = 100$, $q = 0$, $\tilde{\sigma} = 0.3$, $T = 1$ (year), $r = 0.05$, $\tilde{\theta} = 0.01$, and $\tilde{v} = 0.01$. Both Tables report option prices under Gaussian-Brownian motion (GBM) and variance-gamma settings, and the Monte Carlo (MC) prices are reported from Wong and Guan (2011) as a benchmark. Keeping in mind that the value of a Bermudan call option which pays no dividend is equal to the value of its European counterpart; we show that FFT method does not converge (unless a typo in the dividend parameter q), while DP shows a clear convergence.

	FFT			DP	
N	price	CPU	p	price	CPU
2^8	14.2059	(0.19)	75	14.2382	(0.01)
2^9	14.2076	(1.04)	150	14.2311	(0.03)
2^{10}	14.2082	(7.08)	300	14.2313	(0.19)
MC	14.2	(25.07)	Exact	14.2313	(0.01)

Table 2.8: Bermudan call options (GBM) – DP versus Wong and Guan (2011)

Rambeerich et al. (2011) examine the pricing of Bermudan options under an infinite-activity CGMY model. They use exponential time integration (ETI) to solve the system of ordinary differential equations resulting from spatial discretiza-

FFT			DP		
N	price	CPU	p	price	CPU
2^8	14.0763	(0.19)	75	14.2972	(1.57)
2^9	14.2354	(1.07)	150	14.2226	(6.16)
2^{10}	14.1364	(7.38)	300	14.2187	(16.82)
MC	14.1391	(38.8)	Exact	14.2188	(0.03)

Table 2.9: Bermudan call options (VG) – DP versus Wong and Guan (2011)

tion of PIDE. The variance-gamma model is a special case of the CGMY process, but with finite variations. Thus, in Table 2.10, we compare dynamic programming to the ETI-based method for pricing Bermudan put options. Set S_0 = \$100, K = \$100, $q=0$, $\tilde{\sigma}=0.2041$, $T=1$ (year), $r=0.06$, $\tilde{\theta}=0.0417$, and $\tilde{\nu}=1$. Rambeerich et al. (2011) use p spatial nodes, which settles the number of temporal steps to be set at $p/4$. We run our dynamic program with the same parameters p and $M=p/4$, since the grid size and the number of type steps are flexible.

The error term in Table 2.10 refers to the benchmark value taken from Rambeerich et al. (2011) and were computed using extreme size parameters. The authors report an average CPU time of 30 seconds for grid size of $p=2^{10}$. Given their CPU time of 30 seconds and their benchmark value, DP cuts their error by a factor of 2. It is important to note that the values for American options computed with

EIT		DP		
p	price	p	price	CPU
$2^8=256$	0.49527	75	0.49577	(7.08)
$2^9=512$	0.49573	150	0.49572	(19.44)
$2^{10}=1024$	0.49584	300	0.49589	(30.76)
Error	7.058E-05	Error	3.696E-05	

Table 2.10: Bermudan put options – Rambeerich et al. (2011)

dynamic programming, coupled with finite-element approximations, are highly accurate for $p=300$. In addition, the computational time is only about 30 seconds for generating 300 option prices.

2.5 Empirical Investigation

2.5.1 Case study: American options on futures contracts on S&P 500

Futures on stock indices have become an important and growing part of most financial markets. Today, you can buy or sell futures on Dow Jones, S&P 500, NASDAQ, and Value Line indices. An index futures entitles the buyer to any appreciation in the index over and above the index futures' price, and entitles the seller to any depreciation in the index from the same benchmark.

We use dynamic programming to value American options on the S&P 500 futures which were traded on the Chicago Mercantile Exchange (CME) in 2014. The futures contracts are cash settled and listed on a quarterly expiration cycle.

In addition to using dynamic programming, our empirical investigation adopts the following procedure: we gather the daily closing prices of the options and their underlying contracts along with the issue and maturity dates from Datastream. Moreover, we use the discount yields on Treasury bills provided by the Federal Reserve historical data as a proxy for the annual risk-free rate r. We compare dynamic programming prices to the corresponding closing prices quoted on CME, and consider futures options contracts that were issued on 22-04-2014 (evaluation date) and expired 14-08-2014. Hence, the initial futures level is $F_0 = \$1873.9$ and $r = 0.11\%$.

Parameters	θ	σ	ν
VG estimates	0.1169	0.1422	0.0023
Standard errors	0.1637	0.009	0.0074

Table 2.1: Maximum likelihood estimates

To estimate the variance-gamma parameters, we collected the daily futures returns from 03-01-2011 to 22-04-2014, and used a mixed moment-maximum likelihood method, as proposed in Rémillard (2013). Given these estimated parameters, we then select a number of liquid European option contracts on E-mini S&P 500 futures traded in 2014, and apply calibration on the quoted implied volatilities obtained from the Bloomberg system to approximate ξ_1 and ξ_2, as explained in Section 2.2.

K	DP with a grid size p 75	150	300	Market price	Relative error
1865	60.4838	60.4613	60.4623	54.6	-10.74%
1870	62.9371	62.9694	62.9698	56.7	-11.06%
1875	65.5184	65.5350	65.5373	58.8	-11.46%
1880	68.1994	68.1683	68.1682	60.9	-11.93%
1885	70.8845	70.8584	70.8578	63.1	-12.29%
1890	73.5765	73.6069	73.6104	65.4	-12.55%
1895	76.3994	76.4209	76.4216	67.8	-12.72%
CPU	(0.01)	(0.06)	(0.19)		

Table 2.2: American put options on the S&P 500 futures – BS (1973)

K	DP with a grid size p 75	150	300	Market price	Relative error
1865	56.5412	55.0200	54.8698	54.6	-0.49%
1870	56.2779	57.2015	57.3711	56.7	-1.18%
1875	58.6764	59.7638	59.9146	58.8	-1.90%
1880	61.2098	62.4193	62.5411	60.9	-2.69%
1885	63.8789	65.3506	65.2431	63.1	-3.40%
1890	66.6856	68.1456	67.9854	65.4	-3.95%
1895	69.6302	68.1456	70.8109	67.8	-4.44%
Extended Variance Gamma					
1865	56.2536	54.7404	54.5905	54.6	0.02%
1870	56.0030	56.9217	57.0903	56.7	-0.69%
1875	58.4005	59.4830	59.6332	58.8	-1.42%
1880	60.9337	62.1379	62.2590	60.9	-2.23%
1885	63.6031	65.0677	64.9607	63.1	-2.95%
1890	66.4108	67.8622	67.7031	65.4	-3.52%
1895	69.3566	70.6533	70.5295	67.8	-4.03%
CPU	(3.78)	(10.07)	(35.83)		

Table 2.3: American put options on the S&P 500 futures – Variance Gamma

Table 2.1 provides the mixed-moment-maximum likelihood estimates and their associated standard errors. Our calibration results in $(\xi_1, \xi_2) = (-0.2723, 1.8293)$, thus $\tilde{\theta} = 0.1313$ and $\tilde{\sigma} = 0.1415$.

In Tables 2.2 and 2.3, the options' values are computed under the Black and Scholes (1973) model with $\tilde{\sigma}_{BS} = 0.1556$, the variance-gamma model, and the extended variance-gamma model, which are then compared to market prices. Tables 2.2 and 2.3 show that the variance-gamma model outperforms the conventional diffusion model; it significantly decreases the absolute value of the relative error computed on behalf of the observed market price. Moreover, the extended variance-gamma model further improves the final results.

2.6 Conclusion

We build a tractable framework for options pricing where the underlying process is of a variance-gamma type. Our numerical method is mainly based on dynamic programming combined with finite elements. Our numerical investigation shows convergence and efficiency and competes well against alternative methodologies. We also conduct a numerical investigation that focuses on comparing the pricing performance of our extended variance-gamma model on American options on the S&P 500 futures contracts with the geometric-Brownian motion model. We find that our extended model better represents the options market.

Dynamic programming coupled with finite elements can accommodate larger families of derivative contracts, such as Asian options and barrier options, and extend to multidimensional Lévy-like models, such as two-dimensional jump diffusions and GARCH processes, as long as the transition parameters can be computed efficiently.

APPENDIX

2.A Variance-Gamma moments

Given a variance-gamma process X with parameters (θ, σ, ν), and using the moment of a gamma distribution with mean t and variance νt, one has $\mu_t = \mathbb{E}[X(t)] = \theta t$, $\mathbb{E}\left[X(t) - \mu_t^2\right] = \left(\theta^2 \nu + \sigma^2\right)t$, $\mathbb{E}\left[X(t) - \mu_t^3\right] = \left(2\theta^3\nu^2 + 3\sigma^2\theta\nu\right)t$, and $\mathbb{E}\left[X(t) - \mu_t^4\right] = \left(3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3\right)t + \left(3\sigma^4 + 6\sigma^2\theta^2\nu + 3\theta^4\nu^2\right)t^2$.

2.B Transition tables – Variance Gamma - Madan et al. (1998)

From Madan et al. (1998), we define the degenerate hypergeometric function of two variables $\Psi(a,b,\gamma)$ in terms of the modified Bessel function of the second kind $K(.)$ as

$$
\begin{aligned}
\Psi(a,b,\gamma) = {} & \frac{c^{\gamma+\frac{1}{2}} \exp(\mathrm{sign}(a)c)(1+u)^{\gamma}}{\sqrt{(2\pi)}\Gamma(\gamma)\gamma} K_{\gamma+\frac{1}{2}}(c) \times \\
& \Phi(\gamma, 1-\gamma, 1+\gamma; \frac{1+u}{2}, -\mathrm{sign}(a)c(1+u)) - \\
& \mathrm{sign}(a) \frac{c^{\gamma+\frac{1}{2}} \exp(\mathrm{sign}(a)c)(1+u)^{1+\gamma}}{\sqrt{(2\pi)}\Gamma(\gamma)(1+\gamma)} K_{\gamma-\frac{1}{2}}(c) \times \\
& \Phi(1+\gamma, 1-\gamma, 2+\gamma; \frac{1+u}{2}, -\mathrm{sign}(a)c(1+u)) + \\
& \mathrm{sign}(a) \frac{c^{\gamma+\frac{1}{2}} \exp(\mathrm{sign}(a)c)(1+u)^{\gamma}}{\sqrt{(2\pi)}\Gamma(\gamma)\gamma} K_{\gamma-\frac{1}{2}}(c) \times \\
& \Phi(\gamma, 1-\gamma, 1+\gamma; \frac{1+u}{2}, -\mathrm{sign}(a)c(1+u)),
\end{aligned}
$$

where $c = |a|\sqrt{2+b^2}$, $u = b/\sqrt{2+b^2}$, and where the degenerate hypergeometric function of two variables Φ has the integral representation

$$
\Phi(\alpha, \beta, \gamma; x, y) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \int_0^1 u^{\alpha-1}(1-u)^{\gamma-\alpha-1}(1-ux)^{-\beta} e^{uy} du.
$$

Let $x=\frac{1+u}{2}$, $\lambda=2\operatorname{sign}(a)c$, so that $c=|\lambda|/2$, and set

$$\Psi_1(x,\lambda,\gamma)=\frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma}}{2\sqrt{\pi}\Gamma(\gamma)\gamma}K_{\gamma+\frac{1}{2}}(c)\Phi(\gamma,1-\gamma,1+\gamma;x,-\lambda x)$$
$$=\frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma}}{2\sqrt{\pi}\Gamma(\gamma)}K_{\gamma+\frac{1}{2}}(c)\int_0^1 z^{\gamma-1}(1-zx)^{\gamma-1}e^{-\lambda zx}dz,$$

$$\Psi_2(x,\lambda,\gamma)=\frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma+1}}{\sqrt{\pi}\Gamma(\gamma)(\gamma+1)}K_{\gamma-\frac{1}{2}}(c)\Phi(1+\gamma,1-\gamma,2+\gamma;x,-\lambda x)$$
$$=\frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma+1}}{\sqrt{\pi}\Gamma(\gamma)}K_{\gamma-\frac{1}{2}}(c)\int_0^1 z^{\gamma}(1-zx)^{\gamma-1}e^{-\lambda zx}dz,$$

$$\Psi_3(x,\lambda,\gamma)=\frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma}}{2\sqrt{\pi}\Gamma(\gamma)}K_{\gamma+\frac{1}{2}}(c)\int_0^1 z^{\gamma-1}(1-zx)^{\gamma-1}e^{-\lambda zx}dz.$$

If $\lambda>0$, let $t=\lambda xz$, set

$$\Psi_1(x,\lambda,\gamma)=\frac{\sqrt{\lambda}e^{\lambda/2}}{2\sqrt{\pi}\Gamma(\gamma)}K_{\gamma+\frac{1}{2}}(\lambda/2)\int_0^{\lambda x}t^{\gamma-1}\left(1-\frac{t}{\lambda}\right)^{\gamma-1}e^{-t}dt,$$

and set $I_1(x,\lambda,\gamma)=\int_0^{\lambda x}t^{\gamma-1}\left(1-\frac{t}{\lambda}\right)^{\gamma-1}e^{-t}dt$. Integrating by parts,

$$\begin{aligned}I_1(x,\lambda,\gamma) &= \frac{(\lambda x)^{\gamma}}{\gamma}(1-x)^{\gamma-1}e^{-\lambda x}+\frac{1}{\gamma}\int_0^{\lambda x}t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-1}e^{-t}dt\\ &\quad+\frac{\gamma-1}{\lambda\gamma}\int_0^{\lambda x}t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-2}e^{-t}dt.\end{aligned}$$

Set $h_1(x,\lambda,\gamma)=\int_0^{\lambda x}t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-1}e^{-t}dt$ and $h_2(x,\lambda,\gamma)=\int_0^{\lambda x}t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-2}e^{-t}dt$, where h_1 and h_2 are evaluated by Gauss-Legendre quadrature. Then

$$\Psi_2(x,\lambda,\gamma) = \frac{e^{\lambda/2}}{\sqrt{\pi\lambda}} K_{\gamma-\frac{1}{2}}(\lambda/2) \int_0^{\lambda x} t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-1} e^{-t}dt,$$
$$= \frac{e^{\lambda/2}}{\sqrt{\pi\lambda}} K_{\gamma-\frac{1}{2}}(\lambda/2)\, h_1(x,\lambda,\gamma),$$

and $\Psi_3(x,\lambda,\gamma) = \frac{\sqrt{\lambda}e^{\lambda/2}}{2\sqrt{\pi}\Gamma(\gamma)} K_{\gamma-\frac{1}{2}}(\lambda/2)\, I_1(x,\lambda,\gamma)$, hence

$$\Psi(x,\lambda,\gamma) = \Psi_1(x,\lambda,\gamma) - \text{sign}(a)\,\Psi_2(x,\lambda,\gamma) + \text{sign}(a)\,\Psi_3(x,\lambda,\gamma).$$

If $\lambda < 0$, let $t = -\lambda x z$, let

$$\Psi_1(x,\lambda,\gamma) = \frac{\sqrt{-\lambda}e^{\lambda/2}}{2\sqrt{\pi}\Gamma(\gamma)} K_{\gamma+\frac{1}{2}}(-\lambda/2) \int_0^{-\lambda x} t^{\gamma-1}\left(1+\frac{t}{\lambda}\right)^{\gamma-1} e^{t}dt,$$

and set $I_2(x,\lambda,\gamma) = \int_0^{-\lambda x} t^{\gamma-1}\left(1+\frac{t}{\lambda}\right)^{\gamma-1} e^{t}dt$. Integrating by parts,

$$I_2(x,\lambda,\gamma) = \frac{(-\lambda x)^{\gamma}}{\gamma}(1-x)^{\gamma-1}e^{-\lambda x} - \frac{1}{\gamma}\int_0^{-\lambda x} t^{\gamma}\left(1+\frac{t}{\lambda}\right)^{\gamma-1} e^{t}dt$$
$$- \frac{\gamma-1}{\lambda\gamma}\int_0^{-\lambda x} t^{\gamma}\left(1+\frac{t}{\lambda}\right)^{\gamma-2} e^{t}dt.$$

Set $h_3(x,\lambda,\gamma) = \int_0^{-\lambda x} t^{\gamma}\left(1+\frac{t}{\lambda}\right)^{\gamma-1} e^{t}dt$ and $h_4(x,\lambda,\gamma) = \int_0^{-\lambda x} t^{\gamma}\left(1+\frac{t}{\lambda}\right)^{\gamma-2} e^{t}dt$, where h_3 and h_4 are evaluated by Gauss-Legendre quadrature. Then

$$\Psi_2(x,\lambda,\gamma) = \frac{\sqrt{(-\lambda)}e^{\lambda/2}}{\lambda\sqrt{\pi}\Gamma(\gamma)} K_{\gamma-\frac{1}{2}}(-\lambda/2)\, h_3(x,\lambda,\gamma),$$
$$\Psi_3(x,\lambda,\gamma) = \frac{\sqrt{(-\lambda)}e^{\lambda/2}}{2\sqrt{\pi}\Gamma(\gamma)} K_{\gamma-\frac{1}{2}}(-\lambda/2)\, I_2(x,\lambda,\gamma).$$

As a result,

$$\Psi(x,\lambda,\gamma) = \Psi_1(x,\lambda,\gamma) - \text{sign}(a)\,\Psi_2(x,\lambda,\gamma) + \text{sign}(a)\,\Psi_3(x,\lambda,\gamma).$$

The transition parameters $T_{k,i}^{j}$, for $j \in \{0,1,2\}$, $k \in \{1,\dots,p\}$, and $i \in \{0,\dots,p\}$ are

$$T_{k,i}^{0} = \Psi(x_0, \lambda_i^{(0)}, dt/v) - \Psi(x_0, \lambda_{i+1}^{(0)}, dt/v),$$

$$T_{k,i}^{1} = \rho^{-1} a_k \Big[\Psi(x_1, \lambda_i^{(1)}, dt/v) - \Psi(x_1, \lambda_{i+1}^{(1)}, dt/v) \Big],$$

$$T_{k,i}^{2} = e^{\eta_2} \rho^{-2} a_k^2 \Big[\Psi(x_2, \lambda_i^{(2)}, dt/v) - \Psi(x_2, \lambda_{i+1}^{(2)}, dt/v) \Big],$$

where $\rho = \exp\{-(r-q)dt\}$, $x_0 = \frac{1+u_0}{2}$, $u_0 = \frac{b_0}{\sqrt{2+b_0^2}}$, $b_0 = \alpha \sqrt{\frac{v}{1-\xi_2}}$, $\xi_2 = \frac{v\alpha^2}{2}$, $\alpha = \zeta s$, with $\zeta = \frac{\theta}{\sigma^2}$ and $s = \frac{\sigma}{\sqrt{1+\left(\frac{\theta}{\sigma}\right)^2 \frac{v}{2}}}$. Thus $\lambda_i^{(0)} = 2\,\text{sign}\,(a_0)c_0$, with $c_0 = |a_0| \sqrt{2+b_0^2}$, $a_0 = d_i \sqrt{\frac{1-\xi_2}{v}}$, $\xi_1 = \frac{v(\alpha+s)^2}{2}$, and

$$d_i^{(0)} = \frac{1}{s} \left[\ln\left(\frac{a_k}{a_i}\right) + r dt + \frac{dt}{v} \ln\left(\frac{1-\xi_1}{1-\xi_2}\right) \right].$$

Next, $x_1 = \frac{1+u_1}{2}$, $u_1 = \frac{b_1}{\sqrt{2+b_1^2}}$, $b_1 = (\alpha+s)\sqrt{\frac{v}{1-\xi_1}}$, $\lambda_i^{(1)} = 2\,\text{sign}\,(a_1)c_1$, with $c_1 = |a_1| \sqrt{2+b_1^2}$, $a_1 = d_i^{(0)} \sqrt{\frac{1-\xi_1}{v}}$.

Further, let $r_2 = 2r$, $\sigma_2 = 2\sigma$, $\theta_2 = 2\theta$, $q_2 = 2q$, $\alpha_2 = \zeta_2 s_2$, with $\zeta_2 = \frac{\theta_2}{\sigma_2^2}$ and $s_2 = \frac{\sigma_2}{\sqrt{1+\left(\frac{\theta_2}{\sigma_2}\right)^2 \frac{v}{2}}}$. Then $x_2 = \frac{1+u_2}{2}$, $u_2 = \frac{b_2}{\sqrt{2+b_2^2}}$, $b_2 = (\alpha_2+s_2)\sqrt{\frac{v}{1-\xi_1^{(2)}}}$, with $\xi_1^{(2)} = \frac{v(\alpha_2+s_2)^2}{2}$. Thus, $\lambda_i^{(2)} = 2\,\text{sign}\,(a_2)c_2$, with $c_2 = |a_2| \sqrt{2+b_2^2}$, $a_2 = d_i^{(2)} \sqrt{\frac{1-\xi_2^{(2)}}{v}}$, $\xi_2^{(2)} = \frac{v\alpha_2^2}{2}$ and

$$d_i^{(2)} = \frac{1}{s_2} \left[2\ln\left(\frac{a_k}{a_i}\right) + r_2 dt + \eta_2 + \frac{dt}{v} \ln\left(\frac{1-\xi_1^{(2)}}{1-\xi_2^{(2)}}\right) \right],$$

where $\eta_2 = 2w_1 - w_2$, $w_1 = \frac{dt}{v} \ln\left(\frac{1-\xi_1}{1-\xi_2}\right)$, and $w_2 = \frac{dt}{v} \ln\left(\frac{1-\xi_1^{(2)}}{1-\xi_2^{(2)}}\right)$.

Thus, for $j \geq 3$

$$T_{k,i}^{j} = e^{\eta_j} \rho^{-j} a_k^j \left[\Psi(x_j, \lambda_i^{(j)}, dt/\nu) - \Psi(x_j, \lambda_{i+1}^{(j)}, dt/\nu) \right],$$

where $r_j = jr$, $\sigma_j = j\sigma$, $\theta_j = j\theta$, $q_j = jq$, $\alpha_j = \zeta_j s_j$, with $\zeta_j = \frac{\theta_j}{\sigma_j^2}$ and $s_j = \frac{\sigma_j}{\sqrt{1+\left(\frac{\theta_j}{\sigma_j}\right)^2 \frac{\nu}{2}}}$.

Hence, $x_j = \frac{1+u_j}{2}$, $u_j = \frac{b_j}{\sqrt{2+b_j^2}}$, $b_j = (\alpha_j + s_j)\sqrt{\frac{\nu}{1-\xi_1^{(j)}}}$, with $\xi_1^{(j)} = \frac{\nu(\alpha_j+s_j)^2}{2}$, $\xi_2^{(j)} = \frac{\nu\alpha_j^2}{2}$, and $\lambda_i^{(j)} = 2\,\mathrm{sign}\,(a_j) c_j$, with $c_j = |a_j|\sqrt{2+b_j^2}$, $a_j = d_i^{(j)}\sqrt{\frac{1-\xi_2^{(j)}}{\nu}}$, and

$$d_i^{(j)} = \frac{1}{s_j}\left[j \ln\left(\frac{a_k}{a_i}\right) + r_j dt + \eta_j + \frac{dt}{\nu} \ln\left(\frac{1-\xi_1^{(j)}}{1-\xi_2^{(j)}}\right)\right],$$

where $\eta_j = jw_1 - w_j$, $w_1 = \frac{dt}{\nu}\ln\left(\frac{1-\xi_1}{1-\xi_2}\right)$, and $w_j = \frac{dt}{\nu}\ln\left(\frac{1-\xi_1^{(j)}}{1-\xi_2^{(j)}}\right)$.

BIBLIOGRAPHY

Almendral, A. and Oosterlee, C. W. (2007a). Accurate evaluation of European and American options under the CGMY process. *SIAM Journal on Scientific Computing*, 29(1):93–117.

Almendral, A. and Oosterlee, C. W. (2007b). On American options under the variance gamma process. *Applied Mathematical Finance*, 14(2):131–152.

Askari, H. and Krichene, N. (2008). Oil price dynamics (2002–2006). *Energy Economics*, 30(5):2134–2153.

Avramidis, A. N. and L'Écuyer, P. (2006). Efficient monte carlo and quasi-monte carlo option pricing under the variance gamma model. *Management Science*, 52(12):1930–1944.

Becker, M. (2010). Unbiased monte carlo valuation of lookback, swing and barrier options with continuous monitoring under variance gamma models. *The Journal of Computational Finance*, 13(4):35–61.

Ben-Ameur, H., Chérif, R., and Rémillard, B. (2016). American-style options in jump-diffusion models: estimation and evaluation. *Quantitative Finance*, 16(8):1313–1324.

Black, F. and Scholes, M. S. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–54.

Carr, P. and Madan, D. (1999). Option valuation using the fast fourier transform. *Journal of computational finance*, 2(4):61–73.

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, Boca Raton, FL.

Daal, E. A. and Madan, D. B. (2005). An empirical examination of the variance-gamma model for foreign currency options. *The Journal of Business*, 78(6):2121–2152.

Dingeç, K. D. and Hörmann, W. (2012). A general control variate method for option pricing under Lévy processes. *European Journal of Operational Research*, 221(2):368–377.

Fiorani, F. (1999). The Variance-Gamma process for option pricing. *Journal of Computational Finance*, 2(4):61–73.

Fiorani, F., Luciano, E., and Semeraro, P. (2010). Single and joint default in a structural model with purely discontinuous asset prices. *Quantitative Finance*, 10(3):249–263.

Haslip, G. G. and Kaishev, V. K. (2014). Lookback option pricing using the fourier transform b-spline method. *Quantitative Finance*, 14(5):789–803.

Hirsa, A. and Madan, D. B. (2004). Pricing American options under variance gamma. *Journal of Computational Finance*, 7(2):63–80.

Kaishev, V. K. and Dimitrova, D. S. (2009). Dirichlet bridge sampling for the variance gamma process: pricing path-dependent options. *Management Science*, 55(3):483–496.

Këllezi, E. and Webber, N. (2004). Valuing Bermudan options when asset returns are Lévy processes. *Quantitative Finance*, 4(1):87–100.

Konikov, M. and Madan, D. B. (2002). Option pricing using variance gamma markov chains. *Review of Derivatives Research*, 5(1):81–115.

Madan, D. B., Carr, P. P., and Chang, E. C. (1998). The variance gamma process and option pricing. *European Finance Review*, 2(1):79–105.

Madan, D. B. and Milne, F. (1991). Option pricing with v. g. martingale components1. *Mathematical Finance*, 1(4):39–55.

Madan, D. B. and Seneta, E. (1990). The Variance Gamma (V.G.) model for share market returns. *The Journal of Business*, 63(4):511–524.

Maller, R. A., Solomon, D. H., and Szimayer, A. (2006). A multinomial approximation for American option prices in Lévy process models. *Mathematical Finance*, 16(4):613–633.

Mayo, A. (2009). On the numerical evaluation of option prices in the variance gamma model. *International Journal of Computer Mathematics*, 86(2):251–260.

Pinho, C. and Madaleno, M. (2011). Using exponential lévy models to study implied volatility patterns for electricity options. *World Academy of Science, Engineering and Technology, International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, 5(11):1512–1523.

Rambeerich, N., Tangman, D. Y., and Bhuruth, M. (2011). Numerical pricing of American options under infinite activity Lévy processes. *Journal of Futures Markets*, 31(9):809–829.

Rémillard, B. (2013). *Statistical Methods for Financial Engineering.* Chapman & Hall/CRC, Boca Raton, FL.

Stein, H. J., Hogan, A., and Carr, P. (2007). Time for a change: The Variance Gamma Model and option pricing.

Wong, H. Y. and Guan, P. (2011). An FFT-network for Lévy option pricing. *Journal of Banking & Finance*, 35(4):988–999.

CHAPTER 3

A DYNAMIC PROGRAM UNDER LÉVY PROCESSES FOR VALUING CORPORATE SECURITIES

Hatem Ben Ameur,[1] Rim, Chérif,[2] Bruno Rémillard[3]

Abstract

Most structural models for valuing corporate securities assume a geometric-Brownian motion to describe the firm's assets value. However, this does not reflect market-stylized features; the default is more often conducted by sudden informations and shocks, which are not captured by the Gaussian model assumption. To remedy this, we propose a dynamic program for valuing corporate securities under various Lévy processes. Specifically, we study two jump diffusions and a pure-jump process. Under these settings, we build and experiment with a flexible framework, which accommodates the balance-sheet equality, arbitrary corporate debts, multiple seniority classes, tax benefits, and bankruptcy costs. While our approach applies to several Lévy processes, we compute and detail the equity's, debt's, and firm's total values, as well as the debt's credit-spreads under Gaussian, double exponential, and variance-gamma-jump models.

Keywords: Lévy processes; Dynamic programming; Finite elements; Credit risk; Corporate securities; Equity options.

3.1 Introduction

According to the Securities Industry and Financial Markets Association (SIFMA), the size of the U.S. corporate debt issuance reached $1.25 trillion at the end of 2016.

---

1. Hatem Ben-Ameur is a Professor at HEC Montréal in the Department of Decision Sciences, and member of the GERAD.

2. Rim Chérif is a Ph.D. student at HEC Montréal.

3. Bruno Rémillard is a Professor at HEC Montréal in the Department of Decision Sciences, and member of the CRM and GERAD.

Public corporate debt is traded through a dealer-based system. Market signals are partially inaccurate since part of public corporate debt is not heavily traded, and there is a complete lack of information for private corporate debt. Hence, there is a great need for a rational modeling framework for valuing corporate securities.

The structural approach enables the valuation of various corporate securities, including corporate debt portfolio, related by the balance-sheet equality. Corporate securities are seen as derivatives on the firm's assets value and interpreted as the state process. Default is triggered when that state variable crosses a default barrier. Despite their great success, several empirical studies have been critical of the structural approach developed under geometric-Brownian motion. The major drawbacks of such an approach are that it does not reflect data from the bond market; instead, it over-evaluates the debt, and thus, under-evaluates credit spreads. It generates a zero credit spread when the maturity of the debt approaches zero. However, several money-market debt instruments, such as commercial paper, trade at significant spreads above zero. Moreover, an ongoing issue has been to explain the credit spreads of bonds belonging to high investment-grade firms. Huang and Huang (2012) state that even firms with low default risk still have credit spreads that are sizable and positive. Tauchen and Zhou (2011) conclude that firms are exposed to large sudden movements and unexpected information, which confirms the existence of jumps in financial markets. Geske and Delianedis (2001) have also confirmed this claim and sate that jumps are one of the most important components of credit risk.

Our aim in this paper is to build and experiment with a flexible structural model where the value of the firm's assets is described by a Lévy process. We propose this as a solution to some of the major drawbacks under the geometric-Brownian motion model. Specifically, our model allows for significant credit spreads over short maturities. Lévy processes effectively model the firm's assets value, offering realistic financial features and enriching the structural modeling framework. While it initially allows for the leptokurtic feature observed in the financial market, it also introduces unexpected components that allow for low-leverage companies to default, even over short time intervals, without compromising economic intuition. In fact, we combine the unpredictability of the default event generated by reduced-form models with the economic background of structural models.

Merton (1974) pioneered the structural approach by considering a firm that has a trivial capital structure (zero coupon) with a possibility of defaulting only upon maturity of the debt. The proposed model is based on options theory and links to the evaluation of corporate debt when the firm's assets follow geometric-Brownian motion. Many other researchers have worked on extending Merton's (1974) model to achieve higher levels of realism, either by proposing a more complex capital structure, by incorporating frictions, or by experimenting with alternative Markov stochastic processes. Under geometric-Brownian motion modeling, some researchers have proposed exogenous-barrier models (Black and Cox, 1976, Ericsson and Reneby, 1998, Collin-Dufresne et al., 2001), while others have opted for endogenous barriers (Geske, 1977, Leland, 1994, Leland and Toft, 1996, Anderson and Sundaresan, 1996, François and Morellec, 2004, Nivorozhkin, 2005).

Several researchers have developed work on credit modeling under Lévy processes; however, some shortcomings still remain. Zhou (1997) generalizes Merton's (1974) work by considering Gaussian jumps for the firm's assets. Hilberink and Rogers (2002) work under a perpetual debt structure and use a negative spectral Lévy process that only allows for negative jumps. Cariboni and Schoutens (2007) present a structural model and price credit-default-swap under the variance-gamma process using integro-differential equations. Le Courtois and Quittard-Pinon (2008) propose an extension of the Hilberink and Rogers (2002) model using specific stable Lévy processes that also consider upward and downward jumps. Chen and Kou (2009) use a jump-diffusion model with double-exponential-type jumps à la Kou (2002), and also generalize the models of Leland (1994) and Leland and Toft (1996). These methodologies make it possible to solve some of the structural models' problems, but they all work under restrictive hypotheses: a simple debt structure, a predefined, often exponential, debt-maturity structure, and fixed coupons. In addition, they do not allow for seniority classes.

We extend the flexible framework of Altieri and Vargiolu (2000) and Ayadi et al. (2016) to Lévy processes for valuing corporate securities. Our setting is based on dynamic programming coupled with finite elements and accommodates arbitrary corporate debts, market frictions, and multiple seniority classes. Moreover, our research considers three types of Lévy processes: symmetric Gaussian jumps, asymmetric double-exponential jumps, and a pure-jump process, specifically, the variance-gamma process.

Our paper is organized as follows: Section 3.2 provides a theoretical background on Lévy processes and presents two finite-activity Lévy models (Merton, 1976, Kou, 2002) and an infinite-activity Lévy process, namely Madan et al.'s (1998) variance-gamma model. Section 3.3 defines the optimal stopping problem and presents the structural model under Lévy processes. Section 3.4 describes our dynamic programming, Section 3.5 presents our numerical investigation which discusses the impact of jumps in the credit structure, and finally, Section 3.6 concludes the paper.

3.2 The Lévy process

We assume that the firm's assets value follows an exponential Lévy process. On the one hand, our attention is focused on finite-activity-finite-variation Lévy processes (i.e. Lévy jump-diffusions), specifically Merton's (1976) model where the jump size has normal distribution, and the double exponential jump model of Kou (2002). On the other hand, we work with a special case of infinite-activity-finite-variation Lévy process to describe the firm's assets value, which is the variance-gamma process.

We describe the asset value of the firm by a stochastic process $V = \{V_t, 0 \leq t \geq T\}$ of the form

$$V_t = V_0 \exp(L_t),$$

where $L = \{L_t, t \geq 0\}$ is an exponential Lévy process.

3.2.1 Finite-activity Lévy processes

Pure-jump Lévy processes of finite activity are characterized as compound Poisson processes and referred to as jump-diffusion processes. We provide some theoretical background to define finite-activity-Lévy processes in Appendix 3.A. For our default model, we consider Merton's (1976) and Kou's (2002) models, where jumps are considered rare events so that there are a finite number of jumps in any given finite interval. Under this setting, the firm's assets value described with Lévy process L_t is

$$L_t = at + \sigma W_t + \sum_{n=1}^{N_t} J_n - t\lambda\kappa,$$

where $a \in \mathbb{R}$ is the drift , $\sigma \geq 0$ is the volatility, $W = \{W_t, t \geq 0\}$ is a standard Brownian motion, $N = \{N_t, t \geq 0\}$ is a Poisson process with parameter λ, and $J = \{J_n, n \in \mathbb{N}^*\}$ is an i.i.d sequence of random variables with probability distribution F. It describes the jumps that arrive according to a Poisson process, where $\mathbb{E}[e^{J_n} - 1] = \kappa < \infty$. The processes W, N, and J are assumed to be independent under $\mathbb{P}$. Under Merton's (1976) model, jumps are supposed to be normally distributed and thus symmetric, whereas under Kou's (2002) setting, the distribution of jumps are assumed to be double exponential and, as a result, asymmetric.

3.2.2 Infinite-activity Lévy processes

For infinite-activity Lévy processes, the Lévy measure has infinite mass, implying that there are an infinite number of jumps in a finite time interval. We provide some theoretical background to define infinite-activity Lévy processes in Appendix 3.A. Many of these models can be obtained with Brownian subordination $W = \{W_{G_t}, t \geq 0\}$ where G is called a subordinator. The subordinator is an increasing Lévy process that has no diffusion component. For instance, if the subordinator G is a gamma process, then W_{G_t} leads to the variance-gamma model upon which we base our default model. The firm's assets value at time t, under the risk-neutral measure $\mathbb{Q}$, is assumed to follow a Lévy process of the variance-gamma type L_t

$$L_t = (r - q + w)t + X(t; \sigma, \nu, \theta),$$

where r is the constant riskless rate, q is the dividend rate, and w is a compensator that makes the risk neutral return on V equal to $r - q$. Define $\mathbb{E}\left[\exp(X_t)\right] = \left(1 - \theta\nu - \frac{1}{2}\sigma^2\nu\right)^{-t/\nu}$, we set

$$w = \nu^{-1}\log\left(1 - \theta\nu - \frac{1}{2}\sigma^2\nu\right),$$

and $X(t; \sigma, \nu, \theta)$ is a variance-gamma process with parameters σ, ν and θ. The advantage behind the variance-gamma-structural-default model is that the firm's value process becomes asymmetric and that the jump structure allows for random default times.

3.3 Problem Formulation

We consider a market with risky asset V which represents the total asset value of the firm. Let $(\Omega, \mathscr{F}_t, \mathbb{P})$ be a complete probability space and $\mathscr{P} = \{t_0, t_1, \ldots t_n, \ldots, t_N = T\}$ be a set of payment dates. We assume that the interest rate r_n, for $n \in \{1, \ldots, N+1\}$, are known from the beginning. For each $k \in \{0, \ldots, N+1\}$, where k is the bankruptcy time, set the discount factor $\beta_0 = 1$ and $\beta_k = e^{-r_k}\beta_{k-1}$. The case $k = N+1$ means that the firm survives until time t_N. We assume that the stochastic process $V = (V_t)_{t\in[0,T]}$ describing the firm's asset value is modeled as a Lévy process with $V_0 > 0$. In this setting, the firm is financed by debt and equity. More precisely, the firm's capital structure is characterized by senior and junior bonds and a common stock. The firm distributes a coupon payment by which the firm receives tax benefits. The model assumes that shareholders determine the time to default by maximizing the firm's equity value which is subject to limited liability constraints. We assume herein the strict priority rule. The value functions, in terms of the bankruptcy time k, are expressed as follows:

Bankruptcy costs

We assume that on default, a fraction $w \in [0,1]$ of the firm's asset value is due to cover the bankruptcy costs, at time t_n, defined by

$$BC_k^{(n)} = \begin{cases} 0, & k < n \text{ or } k = N+1, \\ w\frac{\beta_k}{\beta_n} V_k, & n \leq k \leq N. \end{cases}$$

Debt

At each date t_n, the firm undertakes to pay a total outflow indicated by $d_n = d_n^{(jun)} + d_n^{(sen)}$ to its bondholders, where $d_n^{(sen)}$ and $d_n^{(jun)}$ refers to the senior and junior bondholders respectively. These payments include principal as well as interest payments. The latter are indicated by d_n^{int}. At the beginning of the contracts, the payment structure are known to all investors. Define the total debt by $D_k^{(n)} = DS_k^{(n)} + DJ_k^{(n)}$, where $DS_k^{(n)}$ and $DJ_k^{(n)}$ are referring to senior and junior debt evaluated at date t_n. The last payment dates for senior and junior debts are indicated by $T^{(sen)}$ and $T^{(jun)}$, with $0 \leq T^{(sen)} \leq T^{(jun)} = T$. For each $n \in \{0,\ldots,N\}$ and $k \in \{0,\ldots,N+1\}$, the senior, junior and total debts are defined as follows:

$$DS_k^{(n)} = \begin{cases} 0, & k < n, \\ \frac{\beta_k}{\beta_n} \min\left\{(1-w)V_k, d_k^{(sen)}\right\} + \frac{1}{\beta_n}\sum_{j=n}^{k-1} \beta_j d_j^{(sen)}, & n \leq k \leq N, \\ \frac{1}{\beta_n}\sum_{j=n}^{N} \beta_j d_j^{(sen)}, & k = N+1, \end{cases}$$

$$DJ_k^{(n)} = \begin{cases} 0, & k < n, \\ \frac{\beta_k}{\beta_n} \max\left\{(1-w)V_k - d_k^{(sen)}, 0\right\} + \frac{1}{\beta_n}\sum_{j=n}^{k-1} \beta_j d_j^{(jun)}, & n \leq k \leq N, \\ \frac{1}{\beta_n}\sum_{j=n}^{N} \beta_j d_j^{(jun)}, & k = N+1. \end{cases}$$

If the firm defaults at time t_n, the senior bondholders receive $\min\left\{(1-w)V_n, d_n^{(sen)}\right\}$, while the junior bondholders receive $\max\left\{(1-w)V_n - d_n^{(sen)}, 0\right\}$, so that the bondholders receive, in total, $(1-w)V_n$. If the firm survives at time t_n, then the senior bondholders receive $d_n^{(sen)}$ while the junior bondholders receive $d_n^{(jun)}$, for a total of $d_n^{(sen)} + d_n^{(jun)} = d_n$. and

$$D_k^{(n)} = \begin{cases} 0, & k < n, \\ \frac{1}{\beta_n} \sum_{j=n}^{N} \beta_j d_j, & k = N+1, \end{cases}$$

where $\sum_{j=n}^{n-1} = 0$ by convention.

Tax benefits

The firm benefits from the tax shield from debt financing as long as it survives. In case of default, tax benefits cannot be claimed. Let $r_n^c \in [0,1]$ be the periodic corporate tax rate over time $[t_n, t_{n+1}]$ that is considered as a known constant. The tax benefit is seen as a security that pays a coupon $tb_n = r_n^c d_n^{int}$. The tax benefit is indicated by $TB_k^{(n)}$ and given by

$$TB_k^{(n)} = \begin{cases} 0, & k < n, \\ \frac{1}{\beta_n} \sum_{j=n}^{k-1} \beta_j tb_j, & n \leq k \leq N+1. \end{cases}$$

The total value of the firm

The total value of the firm $W_k^{(n)}$ is the sum of the firm's assets V_n, and the tax shield of interest payment $TB_k^{(n)}$, minus the bankruptcy costs $BC_k^{(n)}$. That is

$$W_k^{(n)} = V_n + TB_k^{(n)} - BC_k^{(n)} = \begin{cases} 0, & k < n, \\ V_n + \frac{1}{\beta_n} \sum_{j=n}^{k-1} \beta_j tb_j - w \frac{\beta_k}{\beta_n} V_k, & n \leq k \leq N, \\ V_n + \frac{1}{\beta_n} \sum_{j=n}^{N} \beta_j tb_j, & k = N+1. \end{cases}$$

Equity

In the case that the firm survives survival to date t_n, the stockholders receive the total value of the firm minus the total debt defined by

$$\mathscr{E}_k^{(n)} = W_k^{(n)} - D_k^{(n)}.$$

Now, let $\mathscr{T}$ be the set of stopping times with values in $\{0,\ldots,N+1\}$. As a result, for any stopping time $\tau \in \mathscr{T}$ with $\tau \geq n$, we obtain

$$E\left(\mathscr{E}_\tau^{(n)}|\mathscr{F}_n\right) = \mathscr{B}_n^{(\tau)}\mathbf{1}(\tau > n),$$

where $\mathscr{B}_N^{(\tau\vee N)} = \mathscr{B}_N = V_N + tb_N - d_N$ and

$$\mathscr{B}_n^{(\tau)} = V_n + tb_n - d_n - E\left(e^{-r_{n+1}}V_{n+1}|\mathscr{F}_n\right) + E\left(e^{-r_{n+1}}\mathscr{E}_{\tau\vee(n+1)}^{(n+1)}|\mathscr{F}_n\right),$$

for all $n \in \{0,\ldots,N-1\}$.

Since $\mathscr{B}_n^{(\tau)}$ only depends on $\tau\vee(n+1)$, it follows that the definition is sound.

Definition 1.

$$\mathscr{T}_n = \left\{\tau \in \mathscr{T}; \tau \geq n,\ \{\tau > k\} \subset \left\{E\left(\mathscr{E}_{\tau\vee k}^{(k)}|\mathscr{F}_k\right) > 0\right\},\ \text{ for } k \geq n\right\}.$$

Finally, define $J_\tau^{(n)} = TB_\tau^{(n)} - BC_\tau^{(n)}$, and set

$$\bar{J}_n = \sup_{\tau\in\mathscr{T}_n} E\left(J_\tau^{(n)}\middle|\mathscr{F}_n\right),$$

for all $n \in \{0,\ldots,N\}$. Note that $\sup_{\tau\in\mathscr{T}_n} E\left\{W_\tau^{(n)}\middle|\mathscr{F}_n\right\} = V_n + \bar{J}_n$.

The main aim is to find a sequence of stopping times $\tau_n^\star \in \mathscr{T}_n$, corresponding to optimal bankruptcy times, so that the total expected wealth at time n is maximized; that is $V_n + \bar{J}_n = E\left\{W_{\tau_n^\star}^{(n)}|\mathscr{F}_n\right\}$. The solution follows.

Theorem 1. Set

$$\mathscr{E}_N = \max\left(V_N + tb_N - d_N, 0\right), \tag{3.1}$$

and for any $k \in \{0, \ldots, N-1\}$, set

$$\mathscr{E}_k = \max\left\{V_k + tb_k - d_k - E\left(e^{-r_{k+1}} V_{k+1} | \mathscr{F}_k\right) + E\left(e^{-r_{k+1}} \mathscr{E}_{k+1} | \mathscr{F}_k\right), 0\right\}. \tag{3.2}$$

Next, define

$$\tau_k^\star = \begin{cases} N+1, & \text{if } \mathscr{E}_j > 0 \text{ for all } j \in \{k, \ldots, N\}, \\ \min\{k \leq j \leq N; \mathscr{E}_j = 0\}, & \text{otherwise.} \end{cases}$$

Then

$$\bar{J}_N = E\left(J_{\tau_N^\star}^{(N)} | \mathscr{F}_N\right) = -wV_N \mathbf{1}(\mathscr{E}_N = 0) + tb_N \mathbf{1}(\mathscr{E}_N > 0),$$

and for all $k \in \{0, \ldots, N-1\}$,

$$\begin{aligned} \bar{J}_k &= E\left(J_{\tau_k^\star}^{(k)} | \mathscr{F}_k\right) \\ &= -wV_k \mathbf{1}(\mathscr{E}_k = 0) + \left\{tb_k + E\left(e^{-r_{k+1}} \bar{J}_{k+1} | \mathscr{F}_k\right)\right\} \mathbf{1}(\mathscr{E}_k > 0). \end{aligned}$$

The proof of Theorem 1 is provided in Ben Abdellatif et al. (2016).

Remark 1. If $(\beta_k V_k)_{k=0}^N$ is a martingale, then

$$\mathscr{E}_n = \max\left\{E\left(e^{-r_{n+1}} \mathscr{E}_{n+1} | \mathscr{F}_n\right) + tb_n - d_n, 0\right\}. \tag{3.3}$$

3.3.1 Expressions for the debts and equity

Using Theorem 1 from Ben Abdellatif et al. (2016), we have the following expressions for the debt: for any $n \in \{0, \ldots, N\}$, set $D_n = E\left(D_{\tau_n^\star}^{(n)} | \mathscr{F}_n\right)$, $DS_n =$

$E\left(DS^{(n)}_{\tau^\star_n}|\mathscr{F}_n\right)$, $DJ_n = E\left(DJ^{(n)}_{\tau^\star_n}|\mathscr{F}_n\right)$. Then

$$\begin{aligned}
\mathscr{E}_N &= \max\left(V_N + tb_N - d_N, 0\right), &(3.4)\\
D_N &= (1-w)V_N\mathbf{1}(\mathscr{E}_N = 0) + d_N\mathbf{1}(\mathscr{E}_N > 0), &(3.5)\\
DS_N &= \min\left\{(1-w)v, d_N^{(sen)}\right\}\mathbf{1}(\mathscr{E}_N = 0) + d_N\mathbf{1}(\mathscr{E}_N > 0), &(3.6)\\
DJ_N &= \max\left\{(1-w)V_N - d_N^{(sen)}, 0\right\}\mathbf{1}(\mathscr{E}_N = 0) + d_N^{(jun)}\mathbf{1}(\mathscr{E}_N > 0), &(3.7)
\end{aligned}$$

and for any $n \in \{0,\dots,N-1\}$,

$$\begin{aligned}
\mathscr{E}_n &= \max\left\{E\left(e^{-r_{n+1}}\mathscr{E}_{n+1}|\mathscr{F}_n\right) + tb_n - d_n, 0\right\}. &(3.8)\\
D_n &= (1-w)V_n\mathbf{1}(\mathscr{E}_n = 0) + \left\{d_n + \left(e^{-r_{n+1}}D_{n+1}|\mathscr{F}_n\right)\right\}\mathbf{1}(\mathscr{E}_n > 0), &(3.9)\\
DS_n &= \min\left\{(1-w)V_n, d_n^{(sen)}\right\}\mathbf{1}(\mathscr{E}_n = 0)\\
&\quad + \left\{d_n^{(sen)} + E\left(e^{-r_{n+1}}DS_{n+1}|\mathscr{F}_n\right)\right\}\mathbf{1}(\mathscr{E}_n > 0), &(3.10)\\
DJ_n &= \max\left\{(1-w)V_n - d_n^{(sen)}, 0\right\}\mathbf{1}(\mathscr{E}_n = 0)\\
&\quad + \left\{d_n^{(jun)} + E\left(e^{-r_{n+1}}DJ_{n+1}|\mathscr{F}_n\right)\right\}\mathbf{1}(\mathscr{E}_n > 0). &(3.11)
\end{aligned}$$

3.3.2 Expressions for tax benefits and bankruptcy costs

Using Theorem 1 from Ben Abdellatif et al. (2016), we have the following expressions for the tax benefits and bankruptcy costs: for any $n \in \{0,\dots,N\}$, set $TB_n = E\left(TB^{(n)}_{\tau^\star_n}|\mathscr{F}_n\right)$ and set $BC_n = E\left(BC^{(n)}_{\tau^\star_n}|\mathscr{F}_n\right)$, and $v = V_n$. Then

$$\begin{aligned}
TB_N &= tb_N\mathbf{1}(\mathscr{E}_N > 0), &(3.12)\\
BC_N &= wV_N\mathbf{1}(\mathscr{E}_N = 0), &(3.13)
\end{aligned}$$

and for any $n \in \{0, \ldots, N-1\}$,

$$TB_n = \left\{tb_n + E\left(e^{-r_{n+1}} TB_{n+1} | \mathscr{F}_n\right)\right\} \mathbf{1}(\mathscr{E}_n > 0), \tag{3.14}$$

$$BC_n = wV_n \mathbf{1}(\mathscr{E}_n = 0) + E\left(e^{-r_{n+1}} BC_{n+1} | \mathscr{F}_n\right) \mathbf{1}(\mathscr{E}_n > 0). \tag{3.15}$$

Note that under our Markovian case and for any integrable function Ψ on $\mathbb{R} \times [0, \infty)$, one has

$$E\left[e^{-r_{n+1}} \Psi\{V(t_{n+1})\} | \mathscr{F}_n\right] = \Psi\{V(t_n)\},$$

where $V_n = V(t_n)$ and $\mathscr{F}_n = \sigma\{V(u); 0 \leq u \leq t_n\}$.

3.4 Dynamic programming approach

The implementation of the optimal stopping time problem presented in Section 3.3 is performed by using dynamic programming coupled with finite elements. Let $\mathscr{G} = \{v_1, \ldots, v_p\}$ be a mesh of grid points for the firm's assets value, where $0 = v_0 < v_1 < \ldots < v_p < +\infty$ and $\Delta v_i = v_i - v_{i-1}$, $i = 1, \ldots, p$. The grid $\mathscr{G}$ must be selected so that $\text{mesh}(\mathscr{G}) = \max_{1 \leq i \leq p} \Delta v_i$, $\mathbb{Q}(V_t < v_1)$ and $\mathbb{Q}(V_t > v_p)$ all converge to 0 as $p \to \infty$, for $t \in \{t_1, \ldots, t_N\}$. We use the quantile of the state process $\{V\}$ at time $t_N = T$ for grid construction; however, the optimal choice of $\mathscr{G}$ is not addressed in this paper. For simplicity, we assume constant annual interest rate r and $dt = t_{n+1} - t_n$ a positive constant. The dynamic program works as follows:

1. At maturity the value functions are known in closed form and computed using Eq. (3.4), Eq. (3.5), Eq. (3.6), Eq. (3.7), Eq. (3.12), and Eq. (3.13).
2. Suppose that an approximation of the value functions are available at a given future decision date t_{n+1} on the grid $\mathscr{G}$, indicated by $\widehat{\Psi}_{n+1}(v_k)$, for $k = 1, \ldots, p$. This is not a strong assumption since the value functions are known in closed form at maturity. We use a piecewise polynomial interpolation for each value

function $\widehat{\Psi}_{n+1}$ at t_{n+1} from $\mathscr{G}$ to the overall state space by setting

$$\widehat{\Psi}_{n+1}(v) = \sum_{i=0}^{p} \left(\beta_i^0 + \beta_i^1 \; v + \cdots + \beta_i^d \; v^d\right) \mathbb{I}(v_i \leq s < v_{i+1}),$$
$$\text{for } v > 0,$$

where d is the degree of the piecewise polynomial, whose local coefficients depend on the time step t_{n+1}.

3. Approximate every expected discounted value function on $\mathscr{G}$ by

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-rdt}\widehat{\Psi}_{n+1}\left(V_{t_{n+1}}\right)|V_{t_n} = v_k\right] = e^{-rdt}\sum_{i=0}^{p}\left(\beta_i^0 \; T_{ki}^0 + \cdots + \beta_i^d \; T_{ki}^d\right), \quad (3.16)$$

where the transition tables of order j are defined as

$$T_{ki}^j = \mathbb{E}_{\mathbb{Q}}\left[V_{t_{m+1}}^j \mathbb{I}\left(v_i \leq V_{t_{m+1}} < v_{i+1}\right)|V_{t_m} = v_k\right],$$
$$\text{for } j = 0, \ldots, d.$$

For example, T_{ki}^0 is the conditional probability that the firm's assets value at t_{n+1} falls in the interval $[v_i, v_{i+1})$, given that the firm's assets value at time t_n is v_k. We present the computation of the transition tables under Merton's (1976), Kou's (2002), and Madan et al.'s (1998) models in Appendices 3.B.1, 3.B.2, and 3.B.3.

4. Compute the value functions at t_n on $\mathscr{G}$ following Eq. (3.8), Eq. (3.9), Eq. (3.10), Eq. (3.11), Eq. (3.14), and Eq. (3.15), using Eq. (3.16).

5. Go to step 2 and repeat until $n = 0$.

3.5 Numerical investigation

In this section, we study the overall impact of introducing jumps in the dynamic of the firm's assets value. For our numerical investigation, we first focus on Kou's (2002) model and then report some results for the pure-jump credit model of Madan et al. (1998). These settings lead to several scenarios for the credit structure. Under the double-exponential jump model, we concentrate on the scenario of infrequent large jumps. We omit the case of a moderate number of small jumps since it is akin to the pure-diffusion case, especially for short-maturity bonds.

Figure 3.1 examines the impact of jump volatility versus diffusion volatility on credit spreads when fixing the total volatility for both asset processes. We found that the large infrequent jumps scenario reduces credit spreads for long-maturity bonds while increasing credit spreads for short-maturity bonds. Thus, we overcome the major drawbacks of structural models based on pure diffusion which are incapable of producing significant credit spreads for short-maturity bonds. Hence, by including jumps, our results remain consistent with the empirical work of Sarig and Warga (1989) and Fons (1994), and are also consistent with Chen and Kou (2009). But, unlike Chen and Kou (2009), our framework does not suppose an exponential maturity profile for the debt, rather, it enables arbitrary debt structure, accommodating multiple seniority classes, for instance, senior and junior debt.

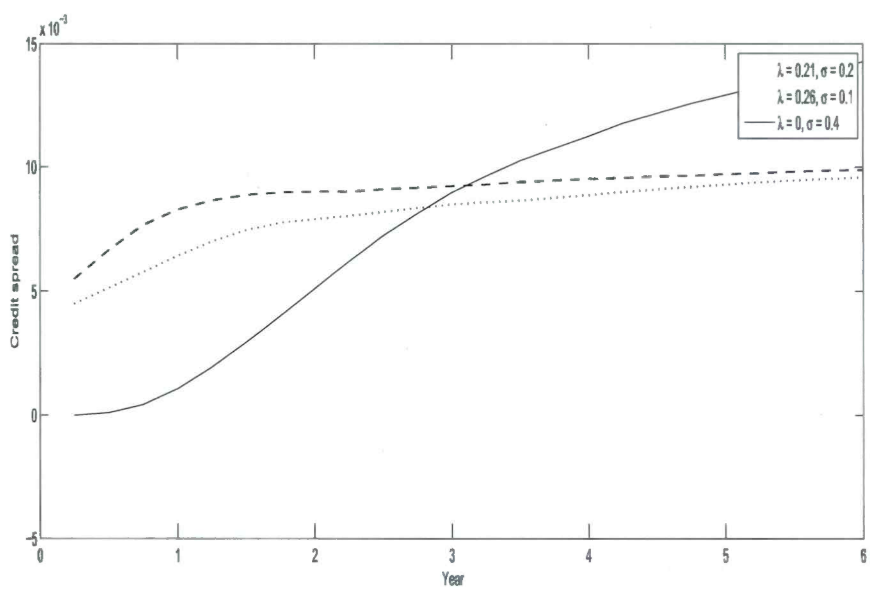

Figure 3.1: The impact of jump volatility versus diffusion volatility on credit spreads. The debt is a 10% bond with a maturity of 6 years. We use a leverage ratio (debt principal over firm's assets) of 30%. Set $r^c = 35\%$ (per year) and $w = 0.25$. The diffusion parameters are given by $r = 8\%$ (per year), and the jump parameters by $\eta_1 = 3$, $\eta_2 = 2$, and $p_u = 0.5$.

Figure 3.2 shows that our model can produce upward, humped and downward shapes for term structures of credit spreads depending on the financial situation of the firm. These credit spreads are observable in bond markets and documented in the empirical work of Sarig and Warga (1989). In fact, the authors find a downward slope for bonds rated B or C, a hump shape for rating class BB, and an upward slope for the investment grade classes. Chen and Kou (2009) argue that the credit spread is normally upward sloping and, as the firm's financial situation worsens, it becomes humped and even downward. Moreover, Fons'(1994) empirical paper suggests that speculative bonds may have humped or downward credit spread curves.

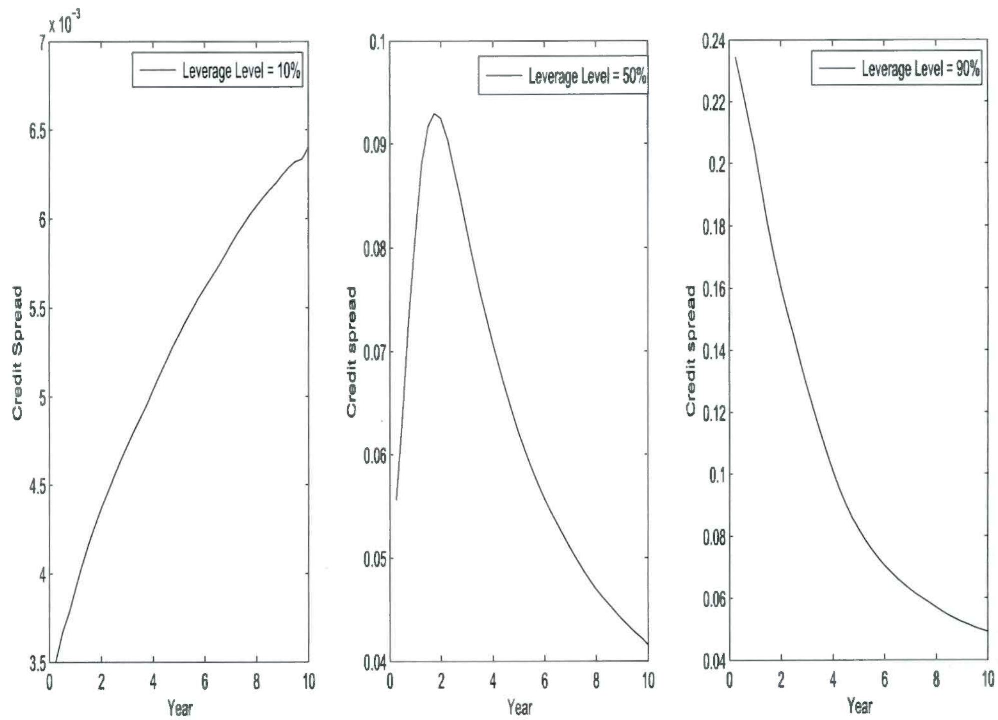

Figure 3.2: Several shapes of yield spread for a 20% bond for different leverage ratio. Set $V_0 = 100$, $r^c = 35\%$ (per year), and $w = 0.5$. The diffusion parameters are given by $r = 8\%$ (per year) and $\sigma = 0.2$, and the jump parameters by $\eta_1 = 3$, $\eta_2 = 2$, $\lambda = 0.2$ (per year) and with probability of upward jumps $p_u = 0.75$.

Figure 3.3 shows that our credit-Lévy model is able to generate an upward shape for speculative bonds with a leverage level of 50%. This feature is pointed out in the empirical work of Helwege and Turner (1999) who claim that speculative bonds can have upward-shaped. Figure 3.3 also shows that the pure-diffusion model is unable to produce a significant credit spread for very short maturity and tends to approach zero as maturity reaches zero. Collin-Dufresne et al. (2001) and Chen and Kou (2009) generate a similar upward credit curve for speculative bonds. With greater flexibility, our model agrees with Chen and Kou's (2009) model. However, Collin-Dufresne et al. (2001) work with diffusion models and are unable to generate non-zero credit spreads as the maturity approaches zero.

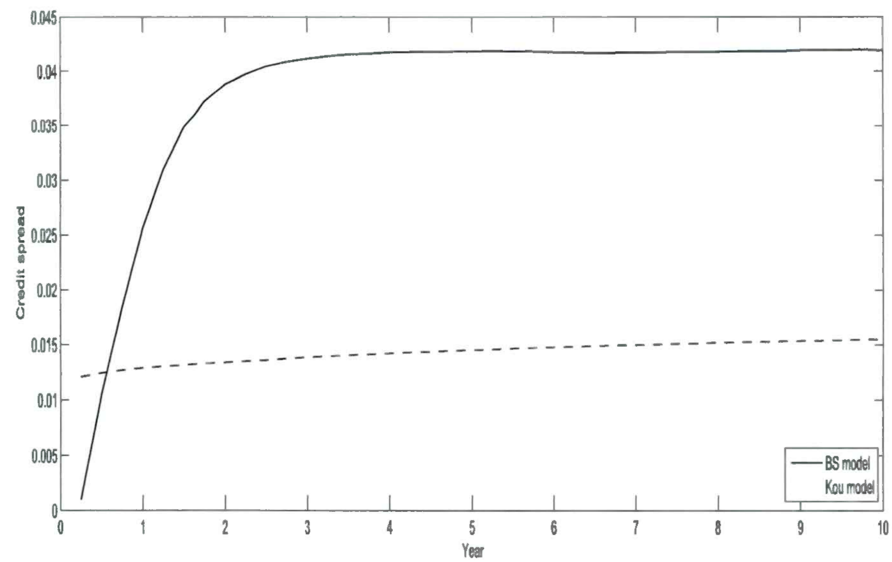

Figure 3.3: Upward credit spread for high-risk bonds. The debt is a 20% bond with a maturity of 10 years and a principal amount of 50\$. Set $V_0 = 100$, $r^c = 35\%$ (per year), and $w = 0.25$. The jump parameters are $\eta_1 = 3$, $\eta_2 = 2$, $p_u = 0.5$, $\lambda = 0.2062$ (per year), and with $\sigma = 0.1$, for a total volatility of 0.4. Under the pure-diffusion case, the volatility $\sigma = 0.4$.

Figure 3.4 shows that our setting reproduces the negative relation between the risk-free rate and the credit spreads discussed in empirical papers. Hence, it shows that as the interest rate increases, credit spreads decrease. This feature is discussed in Longstaff and Schwartz (1995) who point out that a higher spot rate increases the risk-neutral drift of the Lévy-value process. Consequently, this reduces the probability of default and further decreases the credit spread.

Figures 3.5 and 3.6 highlight the effect of diffusion volatility and of the jump's frequency on credit spreads. In fact, credit spreads increase with σ and with λ.

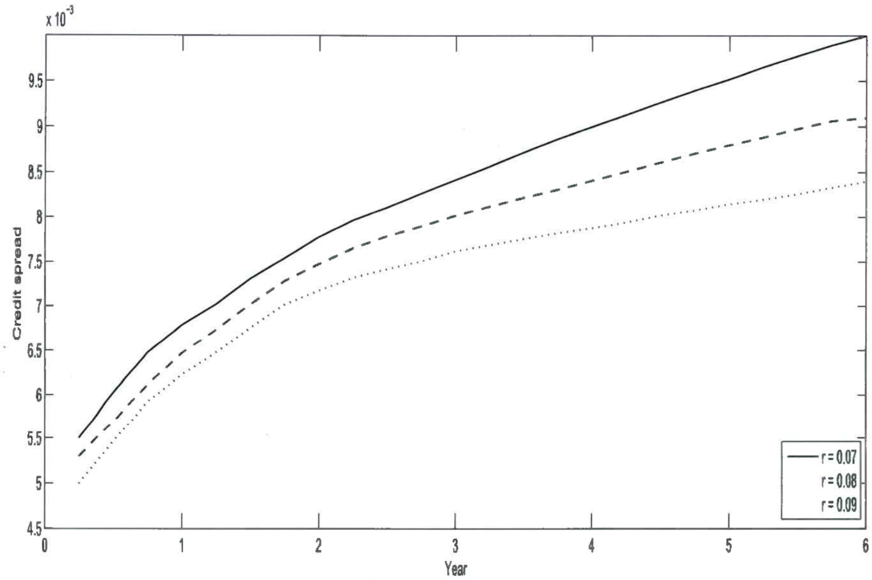

Figure 3.4: The effect of the risk-free rate on credit spreads. The debt is a 10% bond with a maturity of 6 years, and with a principal amount of 50\$. Set $V_0 = 100$, and $r^c = 35\%$ (per year), and $w = 0.25$. The jump parameters are given by $\eta_1 = 3$, $\eta_2 = 2$, $p_u = 0.5$, $\lambda = 0.2$ (per year), and with $\sigma = 0.2$.

Figure 3.5 shows that the effect of diffusion volatility on credit spreads increases with maturity. In other words, diffusion volatility σ has a significant impact on default for long-maturity bonds. Our findings remain consistent with Chen and Kou (2009) results on medium to long maturity bonds.

We still work under the exponential-jump diffusion model and also compute a call-equity option and obtain the implied volatility by inverting the Black-Scholes formula. While Chen and Kou (2009) evaluates equity option by Monte Carlo simulation, we consider equity options as an additional derivative on the Lévy's assets value and use dynamic programming for valuing all these contingent claims. Thus, we suppose a numerical error, but not a statistical one. Figures 3.7 and 3.8 show how default and jumps can, together, generate significant volatility smile, and suggest that the implied volatility and credit spreads tend to be positively correlated. We illustrate the effect of diffusion volatility σ and jump frequency λ on the implied volatility of 1 year maturity call options. Essentially, the implied volatility seems to be increasing in σ and λ. These findings are in line with Chen

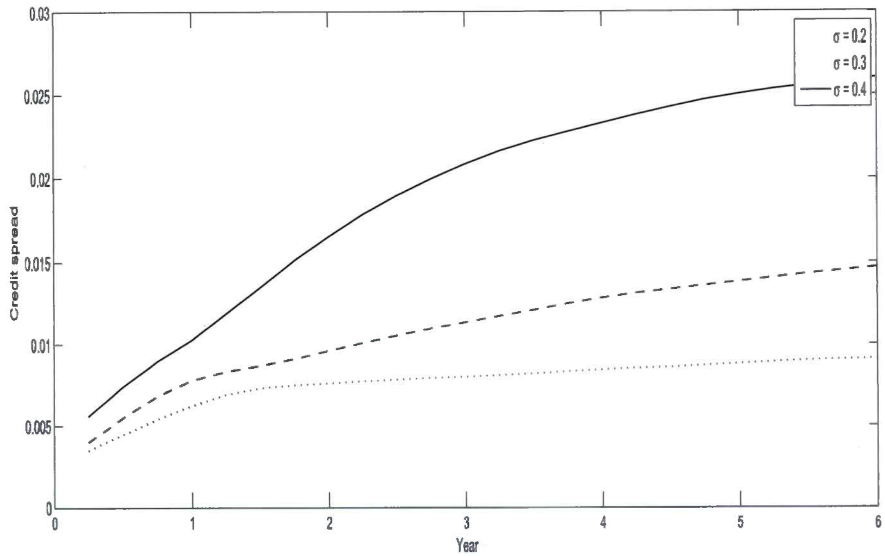

Figure 3.5: The effect of diffusion volatility on credit spreads. The debt is a 10% bond with a maturity of 6 years, and a principal amount of 50$. Set $V_0 = 100$, $r = 8\%$ (per year), $r^c = 35\%$ (per year), and $w = 0.25$. The jump parameters are given by $\eta_1 = 3$, $\eta_2 = 2$, $p_u = 0.5$, and $\lambda = 0.2$ (per year).

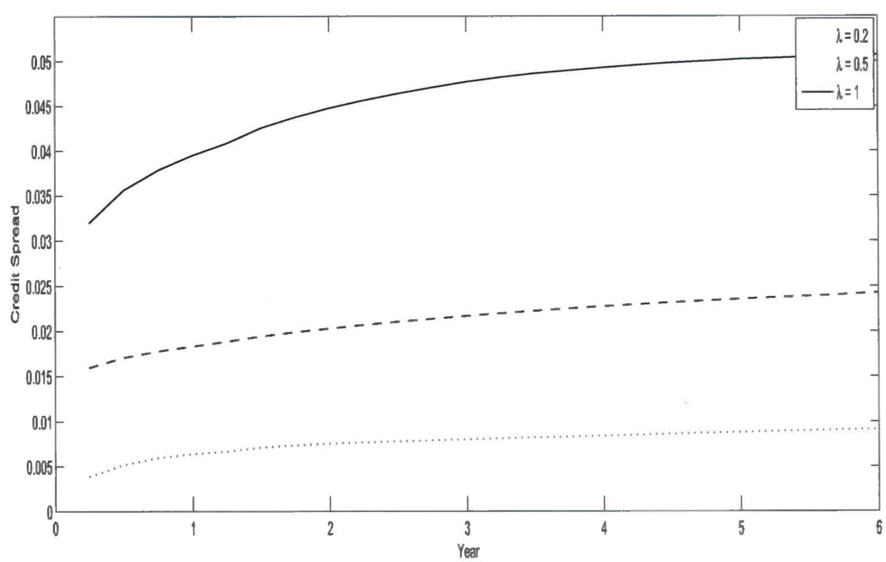

Figure 3.6: The effect of jump frequency on credit spreads. The debt is a 10% bond with a maturity of 6 years, a principal amount of 50$. Set $V_0 = 100$, $r = 8\%$ (per year), $\sigma = 0.2$, $r^c = 35\%$ (per year), and $w = 0.25$. The jump parameters are given by $\eta_1 = 3$, $\eta_2 = 2$, and $p_u = 0.5$.

and Kou (2009).

As a final investigation, we examine the impact of jumps on credit spreads when

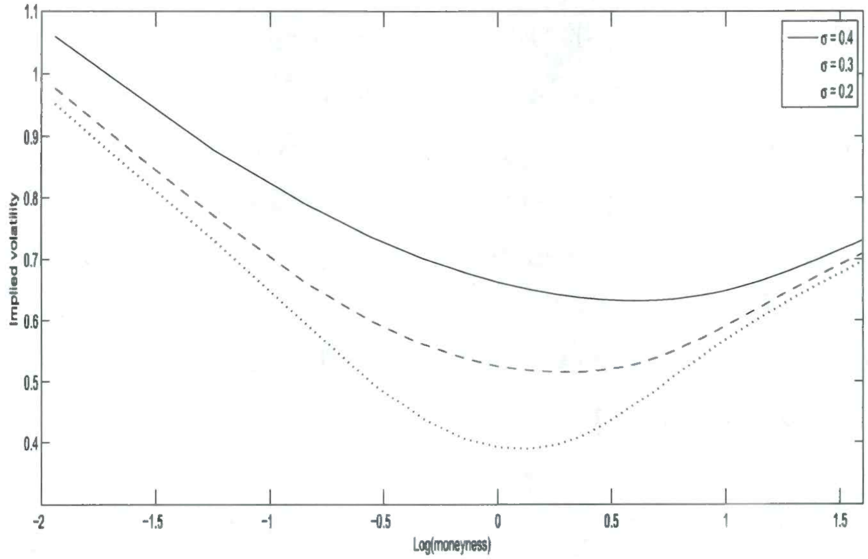

Figure 3.7: Implied volatility versus credit spreads. The senior debt is a 10% bond with a maturity of 1 year. We use a leverage ratio of 30%, $r = 8\%$ (per year), $r^c = 35\%$ (per year), and $w = 0.25$. The jump parameters are given by $\eta_1 = 3$, $\eta_2 = 2$, $p_u = 0.5$, and $\lambda = 0.2$.

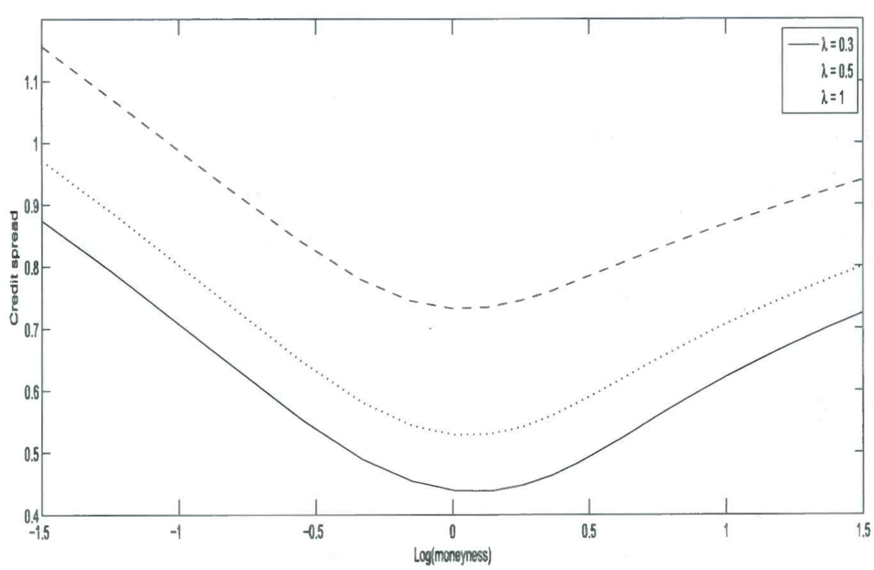

Figure 3.8: The effect of jump's frequency on implied volatility. The senior debt is a 10% bond with a maturity of 1 year. We use a leverage ratio of 30%, $r = 8\%$ (per year), $\sigma = 0.2$, $r^c = 35\%$ (per year), and $w = 0.25$. The jump parameters are given by $\eta_1 = 3$, $\eta_2 = 2$, and $p_u = 0.5$.

modeling the Lévy's assets value with the variance-gamma process. Cariboni and Schoutens (2007) calibrate the variance-gamma model to the credit-default-swaps term structure and found a set of estimated parameters for several companies.

We selected a Lévy process with the following estimated parameters: $\sigma = 0.3553$, $\nu = 2.8132$, and $\theta = -0.0824$, for a total variance $\sigma_{total} = 0.3812$. Under this setting, we can see that the variance-gamma model generates a positive short-term credit spread. Conversely, for the same total volatility, the geometric-Brownian-motion model fails to capture this main feature observed in actual bond markets. Furthermore, as shown in Figure 3.9, the impact of jump volatility is limited on longer-term spreads.

Figure 3.10 reports a credit spread under the variance-gamma process with several total volatility. The applied parameters are also from Cariboni and Schoutens (2007) for two different companies with varying credit-risk ratings. Set the first Lévy's parameters as $\sigma = 0.2041$, $\nu = 0.9644$, and $\theta = -0.0851$, for a total variance of $\sigma_{total} = 0.2205$ and with rating Baa3, relying on Moody's database. Set the second estimated Lévy's parameters as $\sigma = 0.3553$, $\nu = 2.8132$, and $\theta = -0.0824$, for a total variance of $\sigma_{total} = 0.3812$, and with rating A3. From Figure 3.10, we see that this rating is consistent with the credit spreads generated by our structural Lévy model under the variance-gamma process.

For non-redundancy, we do not report all credit spread figures based on the variance-gamma process. We emphasize that several features discussed under Kou's (2002) model are also confirmed under the variance-gamma model. As we provide all the detailed calculations under the three-Lévy-type models, one can undertake an empirical study and work with the most suitable process by conducting a goodness-of-fit test.

3.6 Conclusion

We propose a dynamic program that introduces Lévy-structural-credit models. Our setting accommodates arbitrary corporate debt, several classes of seniority, tax benefits, and bankruptcy costs. We focus on the contribution of jumps by correcting the shortcomings of modeling the firm's assets under geometric-Brownian motion. In fact, Lévy processes are important in financial modeling as they can mimic the

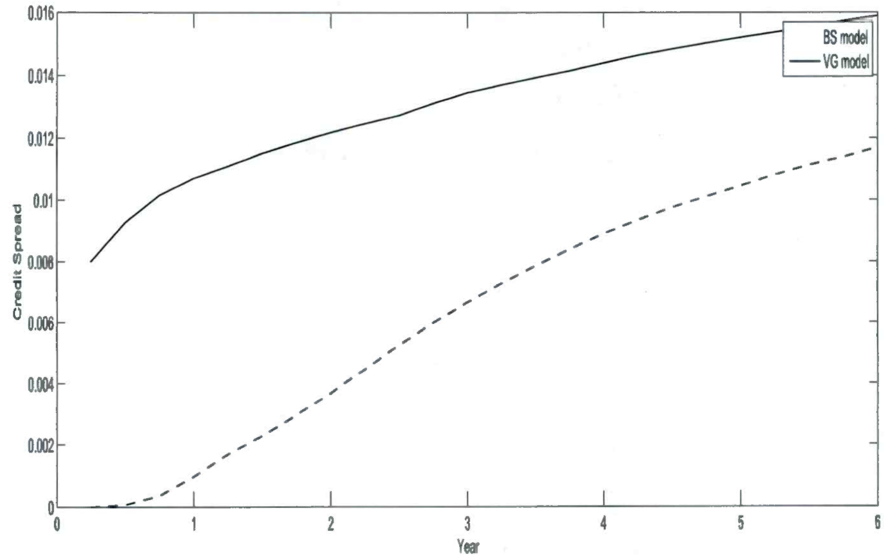

Figure 3.9: The impact of jump volatility versus diffusion volatility on credit spreads. The senior debt is a 10% bond with a maturity 6 years. We use a leverage ratio of 30%. Set $r = 8\%$ (per year), $r^c = 35\%$ (per year), and $w = 0.25$, and with a total variance $\sigma_{total} = 0.3812$.

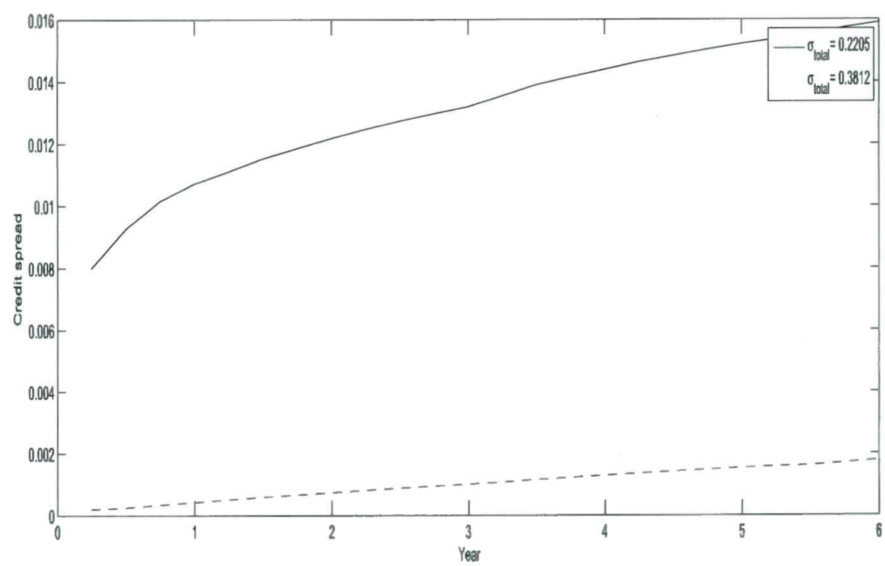

Figure 3.10: Credit spreads under the variance-gamma process. The debt is a 10% bond with a maturity of 6 years. We use a leverage ratio of 30%. Set $r = 8\%$ (per year), $r^c = 35\%$ (per year), and $w = 0.25$.

stylized feature observed in bond markets discussed in empirical papers such as in Sarig and Warga (1989). By adding jumps in the firm's assets dynamic, we capture the impact of unexpected components.

Future research avenues that could be explored consist in: working under a reorganization process, valuing bonds with embedded options under Lévy processes, handling structural Lévy frameworks for multidimensional corporate securities, and finally, introducing a default framework under non-Markovian state process.

APPENDIX

3.A Finite versus infinite-activity-Lévy-processes

Proposition 3.A.1. *From Rémillard (2013), let L be a Lévy process with characteristics (a,b,k), where k is the Lévy measure defined on $\mathbb{R}$ such that $k(\{0\}) = 0$ and $\int_{\mathbb{R}\backslash 0}\left(1 \wedge |x|^2\right) k(dx) \leq \infty$. In a finite time interval, the number of jumps of Lévy process can be finite or infinite, according as $k(\mathbb{R}) < \infty$ or $k(\mathbb{R}) = \infty$.*

Proposition 3.A.2. *From Rémillard (2013), let L be a Lévy process with characteristics (a,b,k), where k is the Lévy measure defined on $\mathbb{R}$ such that $k(\{0\}) = 0$ and $\int_{\mathbb{R}\backslash 0}\left(1 \wedge |x|^2\right) k(dx) \leq \infty$. A Lévy process has jumps of finite variation if and only if $b = 0$ and $\int_{|x|<1} |x| k(dx) < \infty$.*

3.B Transition tables

3.B.1 Transition tables - Merton (1976)

The transition parameters $T_{k,i}^{\nu}$, for $\nu \in \{0,1,2\}$, $k \in \{1,\ldots,p\}$, and $i \in \{0,\ldots,p\}$ are

$$T_{k,i}^{\nu} = \sum_{n=0}^{\infty} \mathbb{Q}(N_{\Delta t} = n)\ \eta_k^{\nu}(n) e^{c(n)^2/2} \left[\Phi\left(c_{k,i+1}(n) - c(n)\right) - \Phi\left(c_{k,i}(n) - c(n)\right)\right],$$

where $N_{\Delta t}$ is the number of jumps over $[t_m, t_{m+1}]$, $c(n) = \nu\sigma_n\sqrt{\Delta t}$, and

$$\begin{aligned}
\mathbb{Q}(N_{\Delta t} = n) &= e^{-\lambda\Delta t}\frac{(\lambda\Delta t)^n}{n!},\\
\sigma_n^2 &= \sigma^2 + \frac{n}{\Delta t}\delta^2,\\
\eta_k(n) &= a_k e^{\left(r-\bar{d}-\lambda\kappa-\sigma_n^2/2\right)\Delta t + n(\gamma+\delta^2/2)},\\
c_{k,i}(n) &= \frac{\log\left(a_i/a_k\right) - \left(r - \bar{d} - \lambda\kappa - \sigma_n^2/2\right)\Delta t - n(\gamma + \delta^2/2)}{\sigma_n},
\end{aligned}$$

and $\Phi(\cdot)$ is the standard normal distribution function.

3.B.2 Transition tables - Kou (2002)

The transition parameters $T^{\nu}_{k,i}$, for $\nu \in \{0,1,2\}$, $k \in \{1,\dots,p\}$, and $i \in \{0,\dots,p\}$ are

$$T^0_{k,i} = \Upsilon(\mu_0,\sigma,\lambda,p_1,\eta_1,\eta_2,x_{i+1},\Delta t) - \Upsilon(\mu_0,\sigma,\lambda,p_1,\eta_1,\eta_2,x_i,\Delta t),$$

$$T^1_{k,i} = \rho^{-1}a_k[\Upsilon(\mu_1,\sigma,\tilde{\lambda},\tilde{p}_1,\tilde{\eta}_1,\tilde{\eta}_2,x_{i+1},\Delta t) - \Upsilon(\mu_1,\sigma,\tilde{\lambda},\tilde{p}_1,\tilde{\eta}_1,\tilde{\eta}_2,x_i,\Delta t)],$$

$$T^2_{k,i} = b\rho^{-2}a_k^2[\Upsilon(\mu_2,\bar{\sigma},\bar{\lambda},\bar{p}_1,\bar{\eta}_1,\bar{\eta}_2,\bar{x}_{i+1},\Delta t) - \Upsilon(\mu_2,2\sigma,\bar{\lambda},\bar{p}_1,\bar{\eta}_1,\bar{\eta}_2,\bar{x}_i,\Delta t)],$$

where $\mu_0 = r - \frac{1}{2}\sigma^2 - \lambda\kappa, x_i = \log(a_i/a_k), \rho = \exp(-(r-\bar{d})\Delta t), \mu_1 = r + \frac{1}{2}\sigma^2 - \lambda\kappa, \tilde{\lambda} = \lambda(1+\kappa), \tilde{p}_1 = p\eta_1/(1+\kappa)(\eta_1 - 1), \tilde{\eta}_1 = \eta_1 - 1, \tilde{\eta}_2 = \eta_2 + 1, \bar{\sigma} = 2\sigma, \bar{\kappa} = p_1(\eta_1/2\bar{\eta}_1) + (1-p_1)(\eta_2/2\bar{\eta}_2) - 1, \mu_2 = 2r + \frac{1}{2}\bar{\sigma}^2 - \lambda\bar{\kappa}, \bar{\lambda} = \lambda(1+\bar{\kappa}), \bar{\eta}_1 = \eta_1/2 - 1, \bar{\eta}_2 = \eta_2/2 + 1, b = \exp(\sigma^2 + \lambda(\bar{\kappa} - 2\kappa)\Delta t)$, and $\bar{x}_i = x_i - \log(b)$. The function $\Upsilon(.)$ is defined by

$$\begin{aligned}
\Upsilon(\mu,\sigma,\lambda,\eta_1,\eta_2,p_1,x_i,\Delta t) &= \frac{e^{(\sigma\eta_1)^2\Delta t/2}}{\sigma\sqrt{2\pi\Delta t}} \sum_{n=1}^{\infty} \pi_n \sum_{k=1}^{n} P_{n,k}\left(\sigma\sqrt{\Delta t}\eta_1\right)^k \\
&\times\ I_{k-1}\left(x_i - \mu\Delta t; -\eta_1, -\frac{1}{\sigma\sqrt{\Delta t}}, -\sigma\eta_1\sqrt{\Delta t}\right) \\
&+\ \frac{e^{(\sigma\eta_2)^2\Delta t/2}}{\sigma\sqrt{2\pi\Delta t}} \sum_{n=1}^{\infty} \pi_n \sum_{k=1}^{n} Q_{n,k}\left(\sigma\sqrt{\Delta t}\eta_2\right)^k \\
&\times\ I_{k-1}\left(x_i - \mu\Delta t; \eta_2, \frac{1}{\sigma\sqrt{\Delta t}}, -\sigma\eta_2\sqrt{\Delta t}\right) \\
&+\ \pi_0\Phi\left(-\frac{x_i - \mu\Delta t}{\sigma\sqrt{\Delta t}}\right),
\end{aligned}$$

and by

$$P_{n,k} = \Sigma_{i=k}^{n-1} \begin{pmatrix} n-k-1 \\ i-k \end{pmatrix} \begin{pmatrix} n \\ i \end{pmatrix} \cdot \left(\frac{\eta_1}{\eta_1+\eta_2}\right)^{i-k} \left(\frac{\eta_2}{\eta_1+\eta_2}\right)^{n-i} p_1^i p_2^{n-i},$$

$$Q_{n,k} = \Sigma_{i=k}^{n-1} \begin{pmatrix} n-k-1 \\ i-k \end{pmatrix} \begin{pmatrix} n \\ i \end{pmatrix} \cdot \left(\frac{\eta_1}{\eta_1+\eta_2}\right)^{n-i} \left(\frac{\eta_2}{\eta_1+\eta_2}\right)^{i-k} p_1^{n-i} p_2^{i},$$

$$I_n(c;\alpha,\beta,\delta) = \int_c^\infty e^{\alpha x} Hh_n(\beta x - \delta)\, dx,$$

for arbitrary constants α, c, $\beta \in \mathbb{R}$, and $n \in \mathbb{N}$.

3.B.3 Transition tables – Variance Gamma - Madan et al. (1998)

From Madan et al. (1998), we define the degenerate hypergeometric function of two variables $\Psi(a,b,\gamma)$ in terms of the modified Bessel function of the second kind $K(.)$ as

$$\begin{aligned}\Psi(a,b,\gamma) = {} & \frac{c^{\gamma+\frac{1}{2}} \exp(\text{sign}(a)c)(1+u)^{\gamma}}{\sqrt{(2\pi)}\Gamma(\gamma)\gamma} K_{\gamma+\frac{1}{2}}(c) \times \\ & \Phi(\gamma, 1-\gamma, 1+\gamma; \frac{1+u}{2}, -\text{sign}(a)c(1+u)) - \\ & \text{sign}(a) \frac{c^{\gamma+\frac{1}{2}} \exp(\text{sign}(a)c)(1+u)^{1+\gamma}}{\sqrt{(2\pi)}\Gamma(\gamma)(1+\gamma)} K_{\gamma-\frac{1}{2}}(c) \times \\ & \Phi(1+\gamma, 1-\gamma, 2+\gamma; \frac{1+u}{2}, -\text{sign}(a)c(1+u)) + \\ & \text{sign}(a) \frac{c^{\gamma+\frac{1}{2}} \exp(\text{sign}(a)c)(1+u)^{\gamma}}{\sqrt{(2\pi)}\Gamma(\gamma)\gamma} K_{\gamma-\frac{1}{2}}(c) \times \\ & \Phi(\gamma, 1-\gamma, 1+\gamma; \frac{1+u}{2}, -\text{sign}(a)c(1+u)),\end{aligned}$$

where $c = |a|\sqrt{2+b^2}$, $u = b/\sqrt{2+b^2}$, and where the degenerate hypergeometric function of two variables Φ has the integral representation

$$\Phi(\alpha,\beta,\gamma;x,y) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \int_0^1 u^{\alpha-1}(1-u)^{\gamma-\alpha-1}(1-ux)^{-\beta} e^{uy} du.$$

Let $x = \frac{1+u}{2}$, $\lambda = 2\,\mathrm{sign}(a)c$, so that $c = |\lambda|/2$, and set

$$\begin{aligned}\Psi_1(x,\lambda,\gamma) &= \frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma}}{2\sqrt{\pi}\Gamma(\gamma)\gamma}K_{\gamma+\frac{1}{2}}(c)\,\Phi(\gamma,1-\gamma,1+\gamma;x,-\lambda x)\\ &= \frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma}}{2\sqrt{\pi}\Gamma(\gamma)}K_{\gamma+\frac{1}{2}}(c)\int_0^1 z^{\gamma-1}(1-zx)^{\gamma-1}e^{-\lambda zx}dz\end{aligned}$$

$$\begin{aligned}\Psi_2(x,\lambda,\gamma) &= \frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma+1}}{\sqrt{\pi}\Gamma(\gamma)(\gamma+1)}K_{\gamma-\frac{1}{2}}(c)\,\Phi(1+\gamma,1-\gamma,2+\gamma;x,-\lambda x)\\ &= \frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma+1}}{\sqrt{\pi}\Gamma(\gamma)}K_{\gamma-\frac{1}{2}}(c)\int_0^1 z^{\gamma}(1-zx)^{\gamma-1}e^{-\lambda zx}dz\end{aligned}$$

$$\Psi_3(x,\lambda,\gamma) = \frac{|\lambda|^{\gamma+\frac{1}{2}}e^{\lambda/2}x^{\gamma}}{2\sqrt{\pi}\Gamma(\gamma)}K_{\gamma+\frac{1}{2}}(c)\int_0^1 z^{\gamma-1}(1-zx)^{\gamma-1}e^{-\lambda zx}dz.$$

If $\lambda > 0$, let $t = \lambda xz$, set

$$\Psi_1(x,\lambda,\gamma) = \frac{\sqrt{\lambda}e^{\lambda/2}}{2\sqrt{\pi}\Gamma(\gamma)}K_{\gamma+\frac{1}{2}}(\lambda/2)\int_0^{\lambda x} t^{\gamma-1}\left(1-\frac{t}{\lambda}\right)^{\gamma-1}e^{-t}dt,$$

and set $I_1(x,\lambda,\gamma) = \int_0^{\lambda x} t^{\gamma-1}\left(1-\frac{t}{\lambda}\right)^{\gamma-1}e^{-t}dt$. Integrating by parts,

$$\begin{aligned}I_1(x,\lambda,\gamma) &= \frac{(\lambda x)^{\gamma}}{\gamma}(1-x)^{\gamma-1}e^{-\lambda x}+\frac{1}{\gamma}\int_0^{\lambda x} t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-1}e^{-t}dt\\ &\quad+\frac{\gamma-1}{\lambda\gamma}\int_0^{\lambda x} t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-2}e^{-t}dt.\end{aligned}$$

Set $h_1(x,\lambda,\gamma) = \int_0^{\lambda x} t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-1}e^{-t}dt$ and $h_2(x,\lambda,\gamma) = \int_0^{\lambda x} t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-2}e^{-t}dt$, where h_1 and h_2 are evaluated by Gauss-Legendre quadrature. Then

$$\begin{aligned}\Psi_2(x,\lambda,\gamma) &= \frac{e^{\lambda/2}}{\sqrt{\pi\lambda}} K_{\gamma-\frac{1}{2}}(\lambda/2) \int_0^{\lambda x} t^{\gamma}\left(1-\frac{t}{\lambda}\right)^{\gamma-1} e^{-t} dt \\ &= \frac{e^{\lambda/2}}{\sqrt{\pi\lambda}} K_{\gamma-\frac{1}{2}}(\lambda/2) h_1(x,\lambda,\gamma),\end{aligned}$$

and $\Psi_3(x,\lambda,\gamma) = \frac{\sqrt{\lambda}e^{\lambda/2}}{2\sqrt{\pi}\Gamma(\gamma)} K_{\gamma-\frac{1}{2}}(\lambda/2) I_1(x,\lambda,\gamma)$, hence

$$\Psi(x,\lambda,\gamma) = \Psi_1(x,\lambda,\gamma) - \text{sign}(a)\Psi_2(x,\lambda,\gamma) + \text{sign}(a)\Psi_3(x,\lambda,\gamma).$$

If $\lambda < 0$, let $t = -\lambda xz$, let

$$\Psi_1(x,\lambda,\gamma) = \frac{\sqrt{-\lambda}e^{\lambda/2}}{2\sqrt{\pi}\Gamma(\gamma)} K_{\gamma+\frac{1}{2}}(-\lambda/2) \int_0^{-\lambda x} t^{\gamma-1}\left(1+\frac{t}{\lambda}\right)^{\gamma-1} e^{t} dt,$$

and set $I_2(x,\lambda,\gamma) = \int_0^{-\lambda x} t^{\gamma-1}\left(1+\frac{t}{\lambda}\right)^{\gamma-1} e^{t} dt$. Integrating by parts,

$$\begin{aligned}I_2(x,\lambda,\gamma) = &\frac{(-\lambda x)^{\gamma}}{\gamma}(1-x)^{\gamma-1}e^{-\lambda x} - \frac{1}{\gamma}\int_0^{-\lambda x} t^{\gamma}\left(1+\frac{t}{\lambda}\right)^{\gamma-1} e^{t} dt \\ &- \frac{\gamma-1}{\lambda\gamma}\int_0^{-\lambda x} t^{\gamma}\left(1+\frac{t}{\lambda}\right)^{\gamma-2} e^{t} dt.\end{aligned}$$

Set $h_3(x,\lambda,\gamma) = \int_0^{-\lambda x} t^{\gamma}\left(1+\frac{t}{\lambda}\right)^{\gamma-1} e^{t} dt$ and $h_4(x,\lambda,\gamma) = \int_0^{-\lambda x} t^{\gamma}\left(1+\frac{t}{\lambda}\right)^{\gamma-2} e^{t} dt$, where h_3 and h_4 are evaluated by Gauss-Legendre quadrature. Then

$$\begin{aligned}\Psi_2(x,\lambda,\gamma) &= \frac{\sqrt{(-\lambda)}e^{\lambda/2}}{\lambda\sqrt{\pi}\Gamma(\gamma)} K_{\gamma-\frac{1}{2}}(-\lambda/2) h_3(x,\lambda,\gamma), \\ \Psi_3(x,\lambda,\gamma) &= \frac{\sqrt{(-\lambda)}e^{\lambda/2}}{2\sqrt{\pi}\Gamma(\gamma)} K_{\gamma-\frac{1}{2}}(-\lambda/2) I_2(x,\lambda,\gamma).\end{aligned}$$

As a result,

$$\Psi(x,\lambda,\gamma) = \Psi_1(x,\lambda,\gamma) - \text{sign}(a)\Psi_2(x,\lambda,\gamma) + \text{sign}(a)\Psi_3(x,\lambda,\gamma).$$

The transition parameters $T^j_{k,i}$, for $j \in \{0,1,2\}$, $k \in \{1,\dots,p\}$, and $i \in \{0,\dots,p\}$ are

$$T^0_{k,i} = \Psi(x_0, \lambda^{(0)}_i, dt/\nu) - \Psi(x_0, \lambda^{(0)}_{i+1}, dt/\nu),$$

$$T^1_{k,i} = \rho^{-1} a_k \Big[\Psi(x_1, \lambda^{(1)}_i, dt/\nu) - \Psi(x_1, \lambda^{(1)}_{i+1}, dt/\nu)\Big],$$

$$T^2_{k,i} = e^{\eta_2}\rho^{-2} a_k^2 \Big[\Psi(x_2, \lambda^{(2)}_i, dt/\nu) - \Psi(x_2, \lambda^{(2)}_{i+1}, dt/\nu)\Big],$$

where $\rho = \exp\{-(r-q)dt\}$, $x_0 = \frac{1+u_0}{2}$, $u_0 = \frac{b_0}{\sqrt{2+b_0^2}}$, $b_0 = \alpha\sqrt{\frac{\nu}{1-\xi_2}}$, $\xi_2 = \frac{\nu\alpha^2}{2}$, $\alpha = \zeta s$, with $\zeta = \frac{\theta}{\sigma^2}$ and $s = \frac{\sigma}{\sqrt{1+\left(\frac{\theta}{\sigma}\right)^2\frac{\nu}{2}}}$. Thus $\lambda^{(0)}_i = 2\,\mathrm{sign}(a_0)c_0$, with $c_0 = |a_0|\sqrt{2+b_0^2}$, $a_0 = d_i\sqrt{\frac{1-\xi_2}{\nu}}$, $\xi_1 = \frac{\nu(\alpha+s)^2}{2}$, and

$$d^{(0)}_i = \frac{1}{s}\left[\ln\left(\frac{a_k}{a_i}\right) + r\,dt + \frac{dt}{\nu}\ln\left(\frac{1-\xi_1}{1-\xi_2}\right)\right].$$

Next, $x_1 = \frac{1+u_1}{2}$, $u_1 = \frac{b_1}{\sqrt{2+b_1^2}}$, $b_1 = (\alpha+s)\sqrt{\frac{\nu}{1-\xi_1}}$, $\lambda^{(1)}_i = 2\,\mathrm{sign}(a_1)c_1$, with $c_1 = |a_1|\sqrt{2+b_1^2}$, $a_1 = d^{(0)}_i\sqrt{\frac{1-\xi_1}{\nu}}$.

Further, let $r_2 = 2r$, $\sigma_2 = 2\sigma$, $\theta_2 = 2\theta$, $q_2 = 2q$, $\alpha_2 = \zeta_2 s_2$, with $\zeta_2 = \frac{\theta_2}{\sigma_2^2}$ and $s_2 = \frac{\sigma_2}{\sqrt{1+\left(\frac{\theta_2}{\sigma_2}\right)^2\frac{\nu}{2}}}$. Then $x_2 = \frac{1+u_2}{2}$, $u_2 = \frac{b_2}{\sqrt{2+b_2^2}}$, $b_2 = (\alpha_2+s_2)\sqrt{\frac{\nu}{1-\xi^{(2)}_1}}$, with $\xi^{(2)}_1 = \frac{\nu(\alpha_2+s_2)^2}{2}$. Thus, $\lambda^{(2)}_i = 2\,\mathrm{sign}(a_2)c_2$, with $c_2 = |a_2|\sqrt{2+b_2^2}$, $a_2 = d^{(2)}_i\sqrt{\frac{1-\xi^{(2)}_2}{\nu}}$, $\xi^{(2)}_2 = \frac{\nu\alpha_2^2}{2}$ and

$$d^{(2)}_i = \frac{1}{s_2}\left[2\ln\left(\frac{a_k}{a_i}\right) + r_2 dt + \eta_2 + \frac{dt}{\nu}\ln\left(\frac{1-\xi^{(2)}_1}{1-\xi^{(2)}_2}\right)\right],$$

where $\eta_2 = 2w_1 - w_2$, $w_1 = \frac{dt}{\nu}\ln\left(\frac{1-\xi_1}{1-\xi_2}\right)$, and $w_2 = \frac{dt}{\nu}\ln\left(\frac{1-\xi^{(2)}_1}{1-\xi^{(2)}_2}\right)$.

Thus, for $j \geq 3$

$$T_{k,i}^{j} = e^{\eta_j} \rho^{-j} a_k^j \left[\Psi(x_j, \lambda_i^{(j)}, dt/\nu) - \Psi(x_j, \lambda_{i+1}^{(j)}, dt/\nu) \right],$$

where $r_j = jr$, $\sigma_j = j\sigma$, $\theta_j = j\theta$, $q_j = jq$, $\alpha_j = \zeta_j s_j$, with $\zeta_j = \frac{\theta_j}{\sigma_j^2}$ and $s_j = \frac{\sigma_j}{\sqrt{1+\left(\frac{\theta_j}{\sigma_j}\right)^2 \frac{\nu}{2}}}$.

Hence, $x_j = \frac{1+u_j}{2}$, $u_j = \frac{b_j}{\sqrt{2+b_j^2}}$, $b_j = (\alpha_j + s_j)\sqrt{\frac{\nu}{1-\xi_1^{(j)}}}$, with $\xi_1^{(j)} = \frac{\nu(\alpha_j+s_j)^2}{2}$, $\xi_2^{(j)} = \frac{\nu\alpha_j^2}{2}$, and $\lambda_i^{(j)} = 2\,\mathrm{sign}(a_j)c_j$, with $c_j = |a_j|\sqrt{2+b_j^2}$, $a_j = d_i^{(j)}\sqrt{\frac{1-\xi_2^{(j)}}{\nu}}$, and

$$d_i^{(j)} = \frac{1}{s_j}\left[j \ln\left(\frac{a_k}{a_i}\right) + r_j dt + \eta_j + \frac{dt}{\nu} \ln\left(\frac{1-\xi_1^{(j)}}{1-\xi_2^{(j)}}\right) \right],$$

where $\eta_j = jw_1 - w_j$, $w_1 = \frac{dt}{\nu} \ln\left(\frac{1-\xi_1}{1-\xi_2}\right)$, and $w_j = \frac{dt}{\nu} \ln\left(\frac{1-\xi_1^{(j)}}{1-\xi_2^{(j)}}\right)$.

BIBLIOGRAPHY

Altieri, A. and Vargiolu, T. (2000). Optimal default boundary in discrete time models. *Rendiconti per gli Studi Economici Quantitativi*, 1(3):1–20.

Anderson, R. W. and Sundaresan, S. (1996). Design and valuation of debt contracts. *Review of Financial Studies*, 9(1):37–68.

Ayadi, M. A., Ben-Ameur, H., and Fakhfakh, T. (2016). A dynamic program for valuing corporate securities. *European Journal of Operational Research*, 249(2):751–770.

Ben Abdellatif, M., Chérif, R., and Rémillard, B. (2016). Valuing corporate securities. Technical report, HEC Montréal.

Black, F. and Cox, J. C. (1976). Valuing corporate securities: some effects of bond indenture provisions. *The Journal of Finance*, 31(2):pp. 351–367.

Cariboni, J. and Schoutens, W. (2007). Pricing credit default swaps under Lévy models. *Journal of Computational Finance*, 10(4):71.

Chen, N. and Kou, S. (2009). Credit spreads, optimal capital structure, and implied volatility with endogenous default and jump risk. *Mathematical Finance*, 19(3):343–378.

Collin-Dufresne, P., Goldstein, R. S., and Martin, J. S. (2001). The determinants of credit spread changes. *The Journal of Finance*, 56(6):2177–2207.

Ericsson, J. and Reneby, J. (1998). A framework for valuing corporate securities. *Applied Mathematical Finance*, 5(3-4):143–163.

François, P. and Morellec, E. (2004). Capital structure and asset prices: some effects of bankruptcy procedures. *The Journal of Business*, 77(2):387–411.

Geske, R. (1977). The valuation of corporate liabilities as compound cptions. *Journal of Financial and Quantitative Analysis*, 12(04):541–552.

Geske, R. L. and Delianedis, G. (2001). The components of corporate credit spreads: default, recovery, taxes, jumps, liquidity, and market factors. Working paper, UCLA.

Helwege, J. and Turner, C. M. (1999). The slope of the credit yield curve for speculative-grade issuers. *The Journal of Finance*, 54(5):1869–1884.

Hilberink, B. and Rogers, L. C. (2002). Optimal capital structure and endogenous default. *Finance and Stochastics*, 6(2):237–263.

Huang, J. Z. and Huang, M. (2012). How much of the corporate-treasury yield spread is due to credit risk? *Review of Asset Pricing Studies*, 2(2):153–202.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48:1086–1101.

Le Courtois, O. and Quittard-Pinon, F. (2008). The optimal capital structure of the firm with stable Lévy assets returns. *Decisions in Economics and Finance*, 31(1):51–72.

Leland, H. E. (1994). Corporate debt value, bond covenants, and optimal capital structure. *The Journal of Finance*, 49(4):1213–1252.

Leland, H. E. and Toft, K. B. (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance*, 51(3):987–1019.

Longstaff, F. A. and Schwartz, E. S. (1995). A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3):789–819.

Madan, D. B., Carr, P. P., and Chang, E. C. (1998). The variance gamma process and option pricing. *European Finance Review*, 2(1):79–105.

Merton, R. C. (1974). On the pricing of corporate debt: the risk structure of interest rates. *The Journal of Finance*, 29(2):449–470.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

Nivorozhkin, E. (2005). Market discipline of subordinated debt in banking: the case of costly bankruptcy. *European Journal of Operational Research*, 161(2):364–376.

Rémillard, B. (2013). *Statistical Methods for Financial Engineering.* Chapman & Hall/CRC, Boca Raton, FL.

Sarig, O. and Warga, A. (1989). Some empirical estimates of the risk structure of interest rates. *The Journal of Finance*, 44(5):1351–1360.

Tauchen, G. and Zhou, H. (2011). Realized jumps on financial markets and predicting credit spreads. *Journal of Econometrics*, 160(1):102–118.

Zhou, C. (1997). A jump-diffusion approach to modeling credit risk and valuing defaultable securities. Working paper, Board of Governors of the Federal Reserve System (U.S.).

CONCLUSION

In financial mathematics, Lévy processes have gained increasing popularity because of their ability to replicate the observed empirical features of the financial market. The aim of this thesis has been to present a framework for the evaluation of financial products whose underlying assets are described according to exponential Lévy processes. For this purpose, we propose a numerical methodology based on dynamic programming coupled with finite elements approximations.

In the first two experiments, we developed a framework for evaluating derivative products for different types of Lévy processes, namely under two jump-diffusion models belonging to finite activity Lévy process and then under infinite-activity pure jump model. Our methodology may be applicable to other types of exotic options and generalizable to multidimensional settings. We can also consider other schemes of interpolations, such as spectral methods, and study their efficiencies. Apart from evaluating complex derivative contracts, dynamic programming also provides us hedging parameters. We do not report these results in this thesis, but they will be considered in the future.

In the last essay, we propose a structural model under Lévy process. The algorithm is based on dynamic programming and incorporates a flexible debt structure, several classes of seniority, tax savings, and bankruptcy costs. The aim was to analyze the ability to reproduce several observed features on the bond market when describing the firm's asset value by Lévy processes. In future research, it would be interesting to study a reorganization process when a firm is unable to honour its commitments and is placed under chapter 11 bankruptcy protection. Other avenues of research can be considered, namely the evaluation of corporate bonds with embedded options, the development of a multidimensional credit model under a Lévy process, and the extension of our structural model under a non-Markov state process.