Pricing Ratchet EIA under Heston's Stochastic Volatility with Deterministic Interest

Dezhao Han

A Thesis

in

The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science (Mathematics) at

Concordia University

Montreal, Quebec, Canada

August 2012

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Dezhao Han

Entitled: Pricing Ratchet EIA under Heston's Stochastic Volatility with Deterministic Interest

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Examiner

Dr. J. Garrido

_____ Examiner

Dr. C. Hyndman

_____ Thesis Supervisor

Dr. P. Gaillardetz

Approved by _____

Chair of Department or Graduate Program Director

_____2012 _____

Dean of Faculty

ABSTRACT

Pricing Ratchet EIA under Heston's Stochastic Volatility
with Deterministic Interest

Dezhao Han

Since its introduction in 1995, equity-indexed annuities (EIAs) received increasing attention from investors. Most of the pricing and hedging for different types of EIAs have been obtained in the Black-Scholes (BS) framework. In this framework the underlying asset is assumed to follow a geometric Brownian motion. However, the BS model is plagued by its assumption of constant volatility, while stochastic volatility models have become increasingly popular. In this paper we assume that the asset price follows Heston's stochastic volatility model with deterministic interest, and introduce two methods to price the ratchet EIA.

The first method is called the joint transition probability density function (JTPDF) method. Given the JTPDF of the asset price and variance, pricing ratchet EIAs boils down to a question of solving multiple integrals. Here, the multiple integral is solved using Quasi Monte Carlo methods and the importance sampling technique. The other method used to evaluate EIAs prices is called the conditional expectation (CE) approach. Conditioning on the volatility path, we first price the rachet EIA analytically in a BS framework. Then the price in Heston framework can be obtained by simulating the volatility path. Greeks for the ratchet EIA can also be calculated by the JTPDF and CE methods. At the end, we carry out some sensitivity analyses for ratchet EIAs' prices and Greeks.

Acknowledgements

I would like to express my special thanks and gratitude to my supervisor Dr. Patrice Gaillardetz, who chose me such an interesting topic and provided me with constant encouragement and inspiration throughout my academic research. I would also like to thank my co-supervisor, Professor José Garrido, for all his help throughout my studies at Concordia. Many thanks to Dr. Xiaowen Zhou, who makes me feel at home.

Finally, I would like to thank my families, specially my parents, and also my friends. All of you make me brave for life.

To my grandmother: Wang Songmei.

Contents

List of Figures

List of Tables

Introduction

An equity-indexed annuity (EIA) is an innovative hybrid insurance product which provides participation in the financial market. The EIA's return is linked to the performance of an equity index, typically the S&P 500 index, while a minimal return is guaranteed on its initial investment. Investors sacrifice some potential upside return for the downside protection, which means that the policy earns a non-negative return even when the market performs poorly. EIAs are increasingly popular since their introduction in 1995 by Keyport Life Insurance, and according to Marrion et al. (2010), EIAs' sales have grown dramatically from $3.00 billion in 1997 to $30.2 billion in 2009. In 2008 the sales of fixed-indexed annuities, the generalization of EIAs, represented 42% of the annuities sold by agents, but in 2010 their market share shrank to 25% (see soa.org). This could be explained by the fact that the companies failed to hedge EIAs during the financial crisis and they became reluctant to issue such contracts. Thus pricing and hedging EIAs are interesting topics.

Designs of EIAs vary according to the companies that sell them. The simplest EIA is the point-to-point design where the policy earns the realized return on the index over a certain period of time at a prescribed participation rate, but with a minimal guarantee. The most popular EIA is the ratchet EIA, which represents 85% of the current market (see annuityadvantage.com). The return of ratchet EIAs is reset annually, and in each year it is the maximum of the prescribed portion of the index return and the minimal guarantee.

Because of their popularity, EIAs have received considerable attention in the actuarial literature. In a Black-Scholes framework, Tiong (2000) derived explicit prices for some popular EIAs using the Esscher transform. Lee (2003) extended Tiong's method to some other path-dependent EIAs. Lin and Tan (2003) and Kijima and Wong (2007) argued that the effects of stochastic interest rates are crucial in pricing EIAs due to their long-term maturity. Assuming stochastic interest and mortality rates, Qian et al. (2010) priced

the ratchet EIA analytically. For most designs it is possible to price EIAs analytically in the Black-Scholes framework because of the Markovian property of the return. Little variations on the financial model have been considered except for Cheung and Yang (2005) and Lin et al. (2009) in which a Markov regime-switching model was applied to evaluate an optimal surrender strategy and price.

In our paper, we use Heston's stochastic volatility model which is a popular method to generalize the Black-Scholes model. Since there is a closed-form for the price of European call option in Heston's model, MacKay (2011) priced the point-to-point EIA analytically by transforming its payoff to a European call's payoff. However, there is no closed-formula for the price of a ratchet EIA due to its complexity. Lin and Tan (2003) and Lin et al. (2009) valued ratchet EIAs by simulation, which makes it difficult to evaluate the Greeks. In this paper, we introduce two approaches to evaluate the price and Greeks of the ratchet EIA.

The structure of this thesis is as follows: In Chapter 1, we describe the Heston Model. First, we give a review of financial frameworks. Since Heston's assumption of a constant interest rate does not hold for long-term investments, we generalize Heston's model to a case of deterministic interest and give a semi-closed expression for the price of European call options. In the end, using a global optimization algorithm, named differential evolution, we calibrate Heston's model using observed European call option prices.

Chapters 2 and 3 introduce two methods to evaluate the prices of ratchet EIAs. The first method is called the joint transition probability density function (JTPDF) approach. Lipton (2001) and Lamoureux and Paseka (2009) derived different explicit formulas for the JTPDF of the Heston process. We generalize this formula to the case of deterministic interest. Though the formula is analytic, it is hard to calculate the oscillatory integral in it. Following Ballestra et al. (2007), we solve the oscillatory integral by the Filon-type quadrature. Given the JTPDF, pricing ratchet EIAs boils down to a question of solving multiple integrals. Again, following Ballestra et al. (2007) the multiple integrals are solved by the so-called importance sampling technique. However, our samples are generated by

Quasi-Monte Carlo methods.

We call the other method the conditional expectation (CE) approach. Conditioning on the volatility path, we first price the ratchet EIA in a Black-Scholes framework. Since there are explicit formulas for the prices of ratchet EIAs in the Black-Schloles model, we can evaluate its price in Heston framework by simulating the stochastic volatility. The idea of taking conditional expectation can be dated back to Hull and White (1987) in which stochastic volatility was first introduced in Finance. Broadie and Kaya (2004) adopted the same idea to evaluate the Greeks for European call and Asian options. In terms of the CE method, the key point is how to simulate the integrated variance. Broadie and Kaya (2006) and Glasserman and Kim (2008) introduced exact simulation schemes, but according to Tse and Wan (2010) and Bégin et al. (2012) both exact schemes are time-consuming. In our paper, we approximate the integrated variance by a summation of gamma distributions, which is faster and with acceptable errors.

Chapter 4 discusses the equity-indexed annuities, especially the ratchet EIA. We applied the JTPDF and CE methods to evaluate ratchet guarantees. The Greeks are also derived using the JTPDF or CE methods. Numerical results are presented in the last chapter. Some sensitivity tests are also conducted in Chapter 5.

1 Economic Model

1.1 From Black-Scholes to Stochastic Volatility

Since the introduction of Black-Scholes model in 1973, there have been a lot of empirical examples showing that the assumption of constant volatility does not correctly describe stock returns. Firstly, Blattberg and Gonedes (1974) show that the log-return of the stock does not follow the normal distribution but a leptokurtic density, which has heavier tails and a higher peak. However, this phenomenon can be explained by time dependent stochastic volatilities and people realized that the volatility also has a mean-reverting property. Furthermore, Beckers (1980) presents statistical evidence of a negative relationship between the level of stock price and its volatility. Nandi (1998) evaluate the correlation between the stochastic return and the volatility and its importance. The volatility smile obtained from the implied volatility casts more doubts on the Black-Scholes model.

In order to describe these behaviors such as a leptokurtic density, mean-reverting property, negative correlation and volatility smile, a proper model is needed to describe the variability of the volatility. Cox (1975) develops a constant elasticity of variance (CEV) diffusion model. It assumes that the volatility is a decreasing function of the stock price. Derman and Kani (1994), Dupire (1994) and Rubinstein (1994) suggest to use the so-called local volatility model, *i.e.* the volatility should be a deterministic function of the stock price and time. They also develop appropriate binomial or trinomial option pricing procedures. However, Dumas et al. (1998) prove that although the deterministic volatility (DV) model beats constant volatility models, the DV model's predictions get worse with the complexity of the assumptions on the volatility and the hedge ratios are not as reliable as those in Black-Scholes model.

Stochastic volatility models describe the volatility using a diffusion process. Denote by V_t the stochastic volatility, then the general process is given by

$$dV_t = p\left(S_t, V_t, t\right) dt + q\left(S_t, V_t, t\right) dW\left(t\right),$$

where $p(S_t, V_t, t)$ and $q(S_t, V_t, t)$ are functions of S_t, V_t and t, and $W(t)$ is a standard Brownian motion. Different $p(S_t, V_t, t)$ and $q(S_t, V_t, t)$ lead to different models (here, we do not consider cases with jumps). Research works on stochastic volatility can be found in Johnson and Shanno (1987), Wiggins (1987), Scott (1987), Hull and White (1987), Stein and Stein (1991), Heston (1993), Schöbel and Zhu (1999), Lewis (2000), and Zhu (2010). Table 1, which is from Zhu (2010), gives an overview of some representative stochastic volatility models. In Table 1 V_t means the stochastic volatility and v_t stands

Table 1: Stochastic volatility models

Johnsom and Shanno (1987)	(1): $dV_t = \kappa V_t dt + \sigma V_t dW(t)$
Wiggins (1987)	(2): $dlnV_t = \kappa\left(\theta - lnV_t\right) dt + \sigma dW(t)$
Hull and White (1987)	(3): $dV_t^2 = \kappa V_t^2 dt + \sigma V_t^2 dW(t)$
Hull and White (1987)	(4): $dV_t^2 = \kappa V_t^2\left(\theta - V_t\right) dt + \sigma V_t^2 dW(t)$
Stein and Stein (1991)	(5): $dV_t = \kappa\left(\theta - V_t\right) dt + \sigma dW(t)$
Heston (1993)	(6): $dv_t = \kappa\left(\theta - v_t\right) dt + \sigma\sqrt{v_t} dW(t)$
Lewis (2000)	(7): $dv_t = \kappa\left(\theta - v_t^2\right) dt + \sigma v_t^{3/2} dW(t)$
Zhu (2010)	(8): $dv(t) = \kappa\left(\theta - \sqrt{v_t} - \lambda v_t\right) dt + \sigma\sqrt{v_t} dW(t)$

for the stochastic variance so that $V_t = \sqrt{v_t}$. Among the models given in Table 1, only (5) (6), (8) have analytic option prices in the case of non-zero correlation between the stock returns and volatility. Models (1) and (3) can not produce the mean-reverting property. Zhu (2010) shows that models (1), (2), (3), (4), (7) are not stationary processes, so they violate the feature of stationarity of volatility or variance. Hence models (5), (6), (8) are worth studying in details.

In fact, each stochastic volatility model in Table 1 has a corresponding discrete model. Discretizing them leads to autoregressive random variance models. The generalized autoregressive conditional heteroscedasticity (GARCH) models also play important roles in studying volatility, but since we only focus on continuous models this kind of approach

will not be considered here.

1.2 Heston's Stochastic Volatility

In this paper, we assume that the asset price S_t follows the Heston model which, under the real probability measure $\mathbb{P}$, is described as follows:

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dW_s(t), \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dW_v(t), \\ d\langle W_s(\cdot), W_v(\cdot)\rangle_t &= \rho dt, \end{aligned} \tag{1.1}$$

where $W_s(t)$ and $W_v(t)$ are two standard Brownian motions with negative correlation ρ, μ is the drift and v_0 and S_0 are known. In general, the initial asset price S_0 is observable from the market and v_0 can be calibrated.

In this model, the asset price S_t still follows a geometric Brownian motion, but with volatility $\sqrt{v_t}$. Instead of modeling the stochastic volatility directly, Heston described the variance v_t by the mean-reverting square root process with long-run mean θ, rate of reversion κ, volatility of volatility σ. The mean-reversion is a desired property for stochastic volatility or variance and is well documented by many empirical studies. The process for v_t is the square root process, which is called the Cox-Ingersoll-Ross (CIR) process. Cox et al. (1985) apply this process to model interest rates. It is also called the Feller process because of William Feller's early work on this process. The CIR process has the following properties.

Proposition 1.1 *For the CIR process*

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dW_v(t).$$

1. *The conditional probability density function of $v_{t+\tau}$ given its current value v_t can be expressed as*

$$p\left(v_{t+\tau}|v_t\right) = ce^{-\xi-\nu}\left(\frac{\nu}{\xi}\right)^{q/2} I_q\left[2\sqrt{\xi\nu}\right], \quad (v_t, v_{t+\tau} \geq 0). \tag{1.2}$$

where

$$\begin{aligned} c &= \frac{2\kappa}{\sigma^2\left(1-e^{-\kappa\tau}\right)}, \\ \xi &= c\cdot v_t\cdot e^{-\kappa\tau}, \\ \nu &= c\cdot v_{t+\tau}, \\ q &= \frac{2\kappa\theta}{\sigma^2}-1. \end{aligned}$$

and $I_q[\cdot]$ is the modified Bessel function[1] *of the first kind of order q.*

2. *The conditional expectation of $v_{t+\tau}$ given v_t is*

$$E\left[v_{t+\tau}|v_t\right] = \theta + (v_t-\theta)\,e^{-\kappa\tau}. \tag{1.3}$$

Proof.

See Cox et al. (1985).□

For a given t, the random variable $2cv_t$ follows a noncentral chi-square with $2q+2$ degrees of freedom and parameter of noncentrality 2ξ. When the Feller condition ($2\kappa\theta > \sigma^2$) is satisfied, the process v_t is strictly positive.

Heston's stochastic volatility model is one of the most popular stochastic models for equities because of the following reasons. Firstly, a suitable set of the Heston parameters $\{\kappa,\theta,\rho,\sigma\}$ can produce a leptokurtic distribution (high peak and heavy tails) of asset returns and the negative correlation between the asset price and volatility. Examples are available in Moodley (2005). Secondly, the volatility has the mean-reverting property. Thirdly, it captures the volatility smile. Finally, a closed-formula for the European call option is available so that it is possible to calibrate the Heston model.

1.3 Deterministic Interest Rate

Heston (1993) assumes that the interest rate is a constant. It is acceptable for short-

[1]The definition of the modified Bessel function of first kind can be found in Bowman (1958) and it is evaluated numerically with the method given by Amos (1985) in this paper.

term investment products since their maturities usually are less than 2 years. In such a relatively low interest rate market, a constant interest rate could give an acceptable approximation. However, for long-term investment products whose maturities range from 5 to 15 years, the risk-free interest rate is more volatile. Table 2 in Section 1.5 illustrates the observed yield rates of T-bills on August 9th, 2011, which are regarded as risk-free investment products. The maturities of the T-bills in Table 2 range from 1 month to 30 years, and the table shows that after 2 years the yield rate varies a lot. Hence, it is unreasonable to assume a constant risk-free interest rate when we are dealing with long-term investment products.

The constant assumption could be overcome by assuming deterministic or stochastic interest rate models. A deterministic risk-free rate is described by a function of time t, while a stochastic process is used to represent a stochastic risk-free rate. Though a deterministic interest rate model is not as accurate as a stochastic model, some deterministic models still fit the data correctly in the long run. Lin and Tan (2003) apply stochastic interest rates to equity-indexed annuities (EIAs), and show how the interest rate affects participation rates. However, their results show that randomness just introduce little impact on the participation rates. Moreover, the problem gets more complex when stochastic interest rate models are introduced in stochastic volatility models. Hence, we use deterministic interest in this paper.

Instead of modeling the risk-free interest rate r_t directly, we turn to the yield rates of T-bills y_t. Present value factor can be obtained from yield rates using

$$e^{-\int_0^t r_s\, ds} = e^{-ty_t}.$$

1.4 Non-Arbitrage Assumption and Heston Framework

Generally speaking, the non-arbitrage assumption says that it is impossible to make money out of nothing. This concept is important in mathematical finance.

Black and Scholes (1973) derive a PDE for option pricing under the non-arbitrage assumption and solve the PDE, leading to a closed-form equation for the price of European call options. Harrison and Pliska (1981) prove that under the non-arbitrage assumption, there is a risk neutral measure $\mathbb{Q}$ under which the asset price earns a risk-free interest rate and introduced the popular risk-neutral pricing formula. Under the non-arbitrage assumption, Heston (1993) also derives a PDE for prices of European call options, and also gives the asset prices dynamic under $\mathbb{Q}$.

Adopting the non-arbitrage assumption, the Heston model with deterministic interest under $\mathbb{Q}$ is given by:

Proposition 1.2 *Heston's process under the risk-neutral measure $\mathbb{Q}$ is given by*

$$\begin{aligned} dS_t &= r_t S_t dt + \sqrt{v_t} S_t dW_s^{\mathbb{Q}}(t), \\ dv_t &= \kappa^*(\theta^* - v_t)dt + \sigma\sqrt{v_t} dW_v^{\mathbb{Q}}(t), \\ d\langle W_s^{\mathbb{Q}}(\cdot), W_v^{\mathbb{Q}}(\cdot)\rangle_t &= \rho dt, \end{aligned} \tag{1.4}$$

where $W_s^{\mathbb{Q}}(t)$ and $W_v^{\mathbb{Q}}(t)$ are standard Brownian motions under $\mathbb{Q}$; r_t is the risk-free rate. $\kappa^* = \kappa + \lambda_0$ and $\theta^* = \frac{\kappa\theta}{\kappa+\lambda_0}$. Here, $\{\kappa, \theta, \sigma, \rho\}$ are the same as the parameters in (1.1). Finally, λ_0, a constant, is related to the market price of the volatility risk.

Proof. For the value of an option Π, a PDE can be derived according to (A.11) in Appendix A. That is

$$\begin{aligned} &\frac{1}{2}v_t S_t^2 \frac{\partial^2 \Pi}{\partial S_t^2} + \frac{1}{2}\sigma^2 v_t \frac{\partial^2 \Pi}{\partial v_t^2} + \rho\sigma S_t v_t \frac{\partial^2 \Pi}{\partial v_t \partial S_t} + \frac{\partial \Pi}{\partial t} - r_t \Pi + r_t S_t \frac{\partial \Pi}{\partial S_t} \\ &\qquad + \left(\kappa(\theta - v_t) - \lambda(s, v, t)\sigma\sqrt{v_t}\right)\frac{\partial \Pi}{\partial v_t} = 0. \end{aligned} \tag{1.5}$$

In Heston's model, $\lambda(S_t, v_t, t)$ is assumed to be $\frac{\lambda_0\sqrt{v_t}}{\sigma}$, that is the market price of volatility risk.

Similar to (A.11) in Appendix A, the drift terms of dS_t and dv_t under $\mathbb{Q}$ should be

$r_t S_t$ and $\kappa(\theta - v_t) - \lambda_0 v_t$ respectively. Hence, (1.1) can be rewritten as follows,

$$\frac{dS_t}{S_t} = r_t dt + \sqrt{v_t}\left(\frac{\mu - r_t}{\sqrt{v_t}}dt + dW_s(t)\right),$$
$$dv_t = (\kappa(\theta - v_t) - \lambda_0 v_t)\,dt + \sigma\sqrt{v_t}\left(\frac{\lambda_0\sqrt{v_t}}{\sigma}dt + dW_v(t)\right),$$

$$d\langle W_s(\cdot), W_v(\cdot)\rangle_t = \rho dt,$$

$\mathbb{Q}$ is just the measure under which

$$\begin{aligned}\frac{\mu - r_t}{\sqrt{v_t}}dt + dW_s(t) &= dW_s^{\mathbb{Q}}(t),\\ \frac{\lambda\sqrt{v_t}}{\sigma}dt + dW_v(t) &= dW_v^{\mathbb{Q}}(t).\end{aligned}$$

□

One term which calls for attention is the market price of volatility risk. Though it does not appear in (1.1), the PDE in (1.5) shows that it has an impact on the option prices. However, the market price of volatility risk is hard to estimate since the volatility is not traded in the financial market. In Heston (1993) it is assumed to be $\frac{\lambda_0\sqrt{v_t}}{\sigma}$. This assumption at least fits common sense, as the market price of volatility should be higher when the volatility is large and lower when v_t is small. Moreover, under this assumption the asset prices keep the same dynamic under both measures $\mathbb{P}$ and $\mathbb{Q}$.

1.5 European Call Price

In order to price and hedge derivatives in the Heston framework, we need to identify the parameters in (1.4), these are $\{\kappa^*, \theta^*, \sigma, \rho, v_0\}$. The parameters are calibrated by comparing the observed European call prices and the prices obtained using the Heston model. The closed-form expression for European call prices is derived in Heston's original paper, but as we pointed out in the previous section, the assumption of constant interest does not hold for pricing EIAs. In this section, following Heston (1993) and by studying the characteristic function of the log-return, we give an analytical formula for the price of European calls under Heston model with deterministic interest.

1.5.1 Characteristic Functions

The payoff of European calls is $\max\{S_T - K, 0\}$, where S_T is the asset price at maturity T and K is the strike price. It means that if the asset price at maturity S_T is larger than the strike K then the payoff of one unit of European call is $S_T - K$. Otherwise, the payoff is zero. In order to price European call options, we use the risk-neutral valuation formula which was introduced in Harrison and Pliska (1981) as follows:

Proposition 1.3 *Denote by $Y(T)$ the valuation of an attainable[2] claim exercised at time T, then its price at time t is given by*

$$Y(t) = B(t)E^{\mathbb{Q}}\left[\frac{Y(T)}{B(T)}\middle|\mathcal{F}_t\right], \tag{1.6}$$

where

$$B(t) = e^{\int_0^t r_w\, dw}$$

and $\mathcal{F}_t$ is the filtration (or the information) up to time t.

Proof. See Harrison and Pliska (1981).□

Remark 1.4 *$B(t)$ is also called the numeraire[3] under $\mathbb{Q}$ such that $\frac{Y(t)}{B(t)}$ is a $\mathbb{Q}$-martingale.*

According to Proposition 1.3, the price of a European call option, Π_C, is given by

$$\begin{aligned}\Pi_C(t,T,S_t,v_t,K) &= E^{\mathbb{Q}}\left[e^{-\int_t^T r_w\, dw}\max\{S_T-K,0\}\middle|\mathcal{F}_t\right]\\ &= e^{-\int_t^T r_w\, dw}E^{\mathbb{Q}}\left[S_T I\{S_T>K\}|\mathcal{F}_t\right]\\ &\quad -Ke^{-\int_t^T r_w\, dw}\mathbb{Q}\left\{S_T>K|\mathcal{F}_t\right\},\end{aligned} \tag{1.7}$$

where $\mathbb{Q}\{S_T > K|\mathcal{F}_t\}$ is the probability of the event $\{S_T > K|\mathcal{F}_t\}$ under $\mathbb{Q}$ and $I\{\cdot\}$ stands for the indicator function.

In order to evaluate $E^{\mathbb{Q}}[S_T I\{S_T > K\}|\mathcal{F}_t]$ we change $\mathbb{Q}$ to another measure $\mathbb{Q}_s$ by the technique of change of numeraire. Before doing that, it is necessary to introduce two lemmas.

[2] A contingent claim attainable if it can be replicated by a self-financing portfolio.

[3] A numeraire is a price process $X(t)$, almost surely strictly possitive for each $t \in [0,T]$. For details of numeraire please see Geman et al. (1995).

Lemma 1.5 *Assume that $B(t)$, $\mathbb{Q}$, $\mathcal{F}_t$ are as in Proposition 1.3 and let $X(t)$ be a non-dividend paying numeraire such that $X(t)/B(t)$ is a $\mathbb{Q}$-martingale. Then there exists a probability measure $\mathbb{Q}_X$ defined by its Radon-Nikodym derivative with respect to $\mathbb{Q}$, that is*

$$\frac{d\mathbb{Q}_X}{d\mathbb{Q}}|\mathcal{F}_T = \frac{B(0)/B(T)}{X(0)/X(T)},$$

such that the basic security prices discounted with respect to X are $\mathbb{Q}_X$ martingales.

Proof. See Geman et al. (1995). □

Lemma 1.6 (Bayes Formula) *Assume that $\tilde{\mathbb{P}}$ is absolutely continuous[4] with respect to $\mathbb{P}$ and $\mathbb{Z}$ is its Radon-Nikodym derivative with respect to $\mathbb{P}$. If Y is bounded (or $\tilde{\mathbb{P}}$-integrable) and $\mathcal{F}_T$ measurable, then*

$$E^{\tilde{\mathbb{P}}}[Y|\mathcal{F}_t] = \frac{1}{\mathbb{Z}(t)} E^{\mathbb{P}}[Y\mathbb{Z}(T)|\mathcal{F}_t], \qquad a.s. \text{ for } t \leq T.$$

Proof. See Exercises 5.1 in Bingham and Kiesel (2004). □

Proposition 1.7 *Given the Heston model (1.4), let $B(t)$ be the numeraire under $\mathbb{Q}$. In other words, $B(t)$ stands for the money market account $e^{\int_0^t r_w\, dw}$. Then*

1. *$S_t/B(t)$ is a $\mathbb{Q}$-martingale.*

2. *There exist a measure $\mathbb{Q}_S$ such that the basic security prices discounted with respect to S_t are $\mathbb{Q}_S$ martingales.*

3. *Heston's model under the measure $\mathbb{Q}_S$ is given by*

$$\begin{aligned} dS_t &= (r_t + v_t) S_t dt + \sqrt{v_t} S_t dW_s^{\mathbb{Q}_S}(t), \\ dv_t &= [\kappa^*(\theta^* - v_t) + \rho\sigma v_t] dt + \sigma\sqrt{v_t} dW_v^{\mathbb{Q}_S}(t), \\ d\langle W_s^{\mathbb{Q}_S}(\cdot), W_v^{\mathbb{Q}_S}(\cdot)\rangle &= \rho dt, \end{aligned} \tag{1.8}$$

where $W_s^{\mathbb{Q}_S}(t)$ and $W_v^{\mathbb{Q}_S}(t)$ are two standard Brownian motions under $\mathbb{Q}_S$.

[4] Measure μ is absolutely continuous with respect to ν is there exists a function g such that $\mu(A) = \int_A g\, d\nu$.

Proof.

1. Apply Itô's formula, we have

$$\begin{aligned} d\left[\frac{S_t}{B(t)}\right] &= \frac{dS_t}{B(t)} + S_t dB(t) + d\langle S., \frac{1}{B(\cdot)}\rangle_t \\ &= \frac{1}{B(t)}\left[r_t S_t dt + \sqrt{v_t} S_t dW_s^{\mathbb{Q}}(t)\right] + S_t\left[-\frac{1}{B(t)^2}\right] B(t) r_t dt \\ &= \frac{1}{B(t)}\left[r_t S_t dt + \sqrt{v_t} S_t dW_s^{\mathbb{Q}}(t) - r_t S_t\right] \\ &= \frac{1}{B(t)}\sqrt{v_t} S_t dW_s^{\mathbb{Q}}(t). \end{aligned}$$

2. This is true because of Lemma 1.5 and the previous result.

3. Let $Y_t = \ln S_t$, applying Itô's formula to Y_t leads to

$$\begin{aligned} dY_t &= \frac{\partial Y_t}{\partial S_t} + \frac{1}{2}\frac{\partial^2 Y_t}{\partial S_t^2} d\langle S., S.\rangle_t + \frac{\partial Y_t}{\partial t} dt \\ &= \frac{1}{S_t} dS_t + \frac{1}{2}\left(-\frac{1}{S_t^2}\right) S_t^2 v_t dt \\ &= \left(r_t - \frac{1}{2} v_t\right) dt + \sqrt{v_t} dW_s^{\mathbb{Q}}(t). \end{aligned}$$

Hence,

$$\ln \frac{S_t}{S_0} = Y_t - Y_0 = \int_0^t r_s - \frac{1}{2} v_s \, ds + \int_0^t \sqrt{v_s} dW_s^{\mathbb{Q}}(s) \, ds.$$

Then, according to Lemma 1.5,

$$\begin{aligned}\frac{d\mathbb{Q}_S}{d\mathbb{Q}} &= \frac{B(0)/B(t)}{S_0/S_t} \\ &= \frac{1}{e^{\int_0^t r_s\,ds}}\exp\left\{\int_0^t r_s - \frac{1}{2}v_s\,ds + \int_0^t \sqrt{v_s}dW_s^{\mathbb{Q}}(s)\,ds\right\} \\ &= \exp\left\{\int_0^t \sqrt{v_s}dW_s^{\mathbb{Q}}(s)ds - \int_0^t \frac{1}{2}v_s\,ds\right\}. \qquad (1.9)\end{aligned}$$

By Girsanov's theorem

$$dW_s^{\mathbb{Q}_S}(t) = dW_s^{\mathbb{Q}}(t) - \sqrt{v_t}dt. \qquad (1.10)$$

Since $dW_v^{\mathbb{Q}_S}(t)$ can be written as $\rho dW_s^{\mathbb{Q}_S}(t) + \sqrt{1-\rho^2}dZ^{\mathbb{Q}_S}(t)$, where $Z^{\mathbb{Q}_S}(t)$ is a standard Brownian Motion under $\mathbb{Q}_S$ which is independent of $W_s^{\mathbb{Q}_S}(t)$, then

$$\begin{aligned}dW_v^{\mathbb{Q}_S}(t) &= \rho dW_s^{\mathbb{Q}_S}(t) + \sqrt{1-\rho^2}dZ^{\mathbb{Q}_S}(t) \\ &= \rho dW_s^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}dZ^{\mathbb{Q}_S}(t) - \sqrt{v_t}dt \\ &= \rho dW_s^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}dZ^{\mathbb{Q}}(t) - \sqrt{v_t}dt \qquad (1.11) \\ &= \rho dW_v^{\mathbb{Q}}(t) - \sqrt{v_t}dt. \qquad (1.12)\end{aligned}$$

Equation (1.11) holds since while we change the measure by (1.9), $Z^{\mathbb{Q}}(t)$ is still a standard Brownian motion under the new measure, that is $Z^{\mathbb{Q}_S}(t) = Z^{\mathbb{Q}}(t)$.

Hence, (1.8) is true because of (1.10) and (1.12). □

Remark 1.8 *Wong and Heyde (2006) prove that* $\frac{S_t}{B(t)}$ *is a true martingale only if* $\kappa^* \geq \sigma\rho$. *In Part* 1 *of Proposition 1.7, the condition* $\kappa^* \geq \sigma\rho$ *is not necessary since by definition* κ^* *and* σ *must be larger or equal to* 0, *and in* (1.1) *we assume that* $\rho < 0$.

By the definition of $\mathbb{Q}_S$, both $\mathbb{Q}_S$ and $\mathbb{Q}$ are absolutely continuous to each other. Denote the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{Q}_S$ by $\mathbb{Z}_t$, s.t.

$$\mathbb{Z}_t = \frac{d\mathbb{Q}}{d\mathbb{Q}_S}\bigg|\mathcal{F}_t = \frac{S_0/S_t}{B(0)/B(t)},$$

by Lemma 1.6 we have

$$E^{\mathbb{Q}}\left[S_T I\{S_T > K\}|\mathcal{F}_t\right] = \frac{1}{\mathbb{Z}_t} E^{\mathbb{Q}_S}\left[\mathbb{Z}_T S_T I\{S_T > K\}|\mathcal{F}_t\right].$$

Hence,

$$\begin{aligned}
e^{-\int_t^T r(w)\,dw} E^{\mathbb{Q}}\left[S_T I\{S_T > K\}|\mathcal{F}_t\right] &= \frac{B(t)}{B(T)} E^{\mathbb{Q}}\left[S_T I\{S_T > K\}|\mathcal{F}_t\right] \\
&= \frac{B(t)}{B(T)} \frac{1}{\mathbb{Z}_t} E^{\mathbb{Q}_S}\left[\mathbb{Z}_T S_T I\{S_T > K\}|\mathcal{F}_t\right] \\
&= E^{\mathbb{Q}_S}\left[\frac{B(t)}{B(T)} \frac{\mathbb{Z}_T}{\mathbb{Z}_t} S_T I\{S_T > K\}|\mathcal{F}_t\right] \\
&= E^{\mathbb{Q}_S}\left[S_t I\left\{S_T > K\right\}|\,\mathcal{F}_t\right] \\
&= S_t \mathbb{Q}_S\left\{S_T > K|\mathcal{F}_t\right\}.
\end{aligned}$$

where $\mathbb{Q}_S\left\{S_T > K|\mathcal{F}_t\right\}$ stands for the probability of the event $\left\{S_T > K|\mathcal{F}_t\right\}$ under measure $\mathbb{Q}_S$. Then the call price (1.7) can be written as

$$\begin{aligned}
\Pi_C(t, T, S_t, v_t, K) &= S_t \mathbb{Q}_S\left\{S_T > K|\mathcal{F}_t\right\} - K e^{-\int_t^T r_w\,dw} \mathbb{Q}\left\{S_T > K|\mathcal{F}_t\right\} \\
&= S_t P_1 - K e^{-\int_t^T r_w\,dw} P_2 \\
&\triangleq e^{x_t} P_1 - K e^{-\int_t^T r_w\,dw} P_2. \qquad (1.13)
\end{aligned}$$

where $P_1 \triangleq \mathbb{Q}_S\{S_T > K|\mathcal{F}_t\}$, $P_2 \triangleq \mathbb{Q}\{S_T > K|\mathcal{F}_t\}$ and $x_t = \ln S_t$.

We turn to the log-price x_t in order to simplify the calculations. Applying Itô's formula to x_t, we obtain the SDE for the log-price as follows,

$$\begin{aligned}
dx_t &= \left(r_t - \frac{1}{2} v_t\right) dt + \sqrt{v_t} dW_s^{\mathbb{Q}}(t), \\
dv_t &= \kappa^*\left(\theta^* - v_t\right) dt + \sigma \sqrt{v_t} dW_v^{\mathbb{Q}}(t), \qquad (1.14) \\
d\langle W_s^{\mathbb{Q}}(\cdot), W_v^{\mathbb{Q}}(\cdot)\rangle_t &= \rho dt.
\end{aligned}$$

Hence, to obtain a formula for $\Pi(t, T, S_t, v_t, K)$ we just need to seek for P_1 and P_2. Whereas, Rollin et al. (2010) reported that the the probability density function of the log-return or log-price is still not well known, and maybe do not have a closed-form. However, the characteristic function of the log-price has nice properties. Hence the valuation of options

via characteristic function (CF) is popular: in fact in Heston (1993), the original closed-form for call price is derived using the CF; Bakshi and Madan (2000) shows that CF plays a significant role in pricing options and simplify the problem; Drăgulescu and Yakovenko (2002) derived the PDF of the "log-return" in Heston's framework via CF, assuming the initial stochastic variance follows a stationary distribution. Zhu (2010) gave a thorough description of applying the CF to Heston's model. The following results explain how to derive the call price via the characteristic function of the log-price.

Proposition 1.9 *Given the characteristic function $\phi(s)$ of a random variable X, the CDF of X, $F_X(x)$, is given by*

$$F_X(x) = \frac{1}{2} - \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-isx}\phi(s)}{is}\, ds. \tag{1.15}$$

Proof. See Gil-Pelaez (1951). □

Corollary 1.10 *Assume that $\phi_j(s)$ is the characteristic function corresponding to P_j, for $j = 1, 2$ in (1.13), then we have the following relationship:*

$$P_j = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} Re\left[\frac{e^{-islnK}\phi_j(s)}{is}\right] ds, \tag{1.16}$$

where $Re[z]$ is the real part of z.

Proof. According to Proposition 1.9, for $j = 1, 2$,

$$\begin{aligned}
P_j &= 1 - F(\ln K) \\
&= \frac{1}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-is\ln K}\phi_j(s)}{is}\, ds \\
&= \frac{1}{2} + \frac{1}{2\pi}\left[-\int_{+\infty}^{0} \frac{e^{-i(-\tilde{s})\ln K}\phi_j(-\tilde{s})}{-i\tilde{s}}\, d\tilde{s} + \int_0^{\infty} \frac{e^{-is\ln K}\phi_j(s)}{is}\, ds\right] \\
&= \frac{1}{2} + \frac{1}{2\pi}\left[\int_0^{+\infty} \frac{e^{i\tilde{s}\ln K}\overline{\phi_j(\tilde{s})}}{-i\tilde{s}}\, d\tilde{s} + \int_0^{\infty} \frac{e^{-is\ln K}\phi_j(s)}{is}\, ds\right] \qquad (1.17) \\
&= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} Re\left[\phi_j(s)\frac{e^{-is\ln K}}{is}\right] ds, \qquad (1.18)
\end{aligned}$$

where $\overline{a+ib} = a - ib$, $a, b \in \mathbf{R}$, $Im[z]$ and $Re[z]$ denote the imaginary and real part of z.

Equation (1.17) is true since for a characteristic function $\phi(-s) = \overline{\phi(s)}$ and (1.18) holds since the probability density P_j must be real, so that the integrals w.r.t the imaginary part can be eliminated. $\square$

Note that Corollary 1.10 is model free, so that the problem of pricing a European call option can be converted to a problem of identifying the characteristic functions of the asset price (or log-price) under $\mathbb{Q}$ and $\mathbb{Q}_S$. In Heston's model, ϕ_j, $j = 1, 2$, are obtained through the Heston PDE.

1.5.2 Heston PDE

In order to derive the Heston PDE, we need to introduce the following theorem.

Theorem 1.11 (Feynman-Kac Theorem)

Suppose that

1. $\mathbf{x}_t$ *follows the stochastic process in* n *dimensions*

$$d\mathbf{x}_t = \boldsymbol{\mu}(\mathbf{x}_t, t) + \boldsymbol{\sigma}(\mathbf{x}_t, t)\, d\boldsymbol{W}^{\mathbb{Q}}(t), \tag{1.19}$$

where $\mathbf{x}_t$ *and* $\boldsymbol{\mu}(\mathbf{x}_t, t)$ *are* n*-dimensional column vectors,* $\boldsymbol{\sigma}(\mathbf{x}_t, t)$ *is a* $n \times m$ *matrix and* $\boldsymbol{W}_t^{\mathbb{Q}}$ *is a* m*-dimensional* $\mathbb{Q}-$*Brownian motion. That is*

$$d\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} \mu_1(\mathbf{x_t}, t) \\ \vdots \\ \mu_n(\mathbf{x_t}, t)) \end{pmatrix} dt + \begin{pmatrix} \sigma_{11}(\mathbf{x_t}, t) & \cdots & \sigma_{1m}(\mathbf{x_t}, t) \\ \vdots & \ddots & \vdots \\ \sigma_{n1}(\mathbf{x_t}, t) & \cdots & \sigma_{nm}(\mathbf{x_t}, t) \end{pmatrix} \begin{pmatrix} dW_1^{\mathbb{Q}}(t) \\ \vdots \\ dW_m^{\mathbb{Q}}(t) \end{pmatrix},$$

where $\mu_i(\mathbf{x}_t, t)$ *and* $\sigma_{ij}(\mathbf{x}_t, t)$ *are functions from* $\mathbf{R}^{n+1}$ *to* $\mathbf{R}$.

2. *By definition, the generator of the process* $\mathbf{x}_t$ *is*

$$\mathcal{A} = \sum_{i=1}^{n} \mu_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right)_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}, \tag{1.20}$$

where for convenience $\mu_i = \mu_i(\mathbf{x}_t, t), \boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{x}_t, t)$ *and* $\left(\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right)_{ij}$ *is the element* (i, j) *of the matrix* $\boldsymbol{\sigma}\boldsymbol{\sigma}^T$.

Now let $f \in C_0^2(\mathbf{R}^n)$, $q \in C(\mathbf{R}^n)$ and is lower bounded, we have

1. *Put*

$$Y(t,\mathbf{x}) = E\left[exp\left(-\int_0^t q(\mathbf{x}_s)\, ds\right) f(\mathbf{x}_t) \middle| \mathbf{x}\right], \tag{1.21}$$

then

$$\frac{\partial Y}{\partial t} = \mathcal{A}Y - qY; \quad t > 0, \ \mathbf{x} \in \mathbf{R}^n, \tag{1.22}$$

$$Y(0,\mathbf{x}) = f(\mathbf{x}); \quad \mathbf{x} \in \mathbf{R}^n. \tag{1.23}$$

2. *Moreover, if $w(t,\mathbf{x}) \in C^{1,2}(\mathbf{R} \times \mathbf{R}^n)$ is bounded on $K \times \mathbf{R}^n$ for each compact $K \subset \mathbf{R}$ and w solves (1.22), and (1.23), then $w(t,\mathbf{x}) = Y(t,\mathbf{x})$, given by (1.21).*

Proof. See Oksendal (2002).□

The Heston PDE can be derived directly from Theorem 1.11. Recall that in the Heston framework the asset price dynamics under the risk-neutral measure are described by (1.14). The process for $\mathbf{x_t} = (x_t, v_t)$ can be written in terms of two independent Brownian motions Z_1 and Z_2 as

$$d\begin{pmatrix} x_t \\ v_t \end{pmatrix} = \begin{pmatrix} r_t - \frac{1}{2}v_t \\ \kappa^*(\theta^* - v_t) \end{pmatrix} dt + \begin{pmatrix} \sqrt{v_t} & 0 \\ \sigma\rho\sqrt{v_t} & \sigma\sqrt{v_t(1-\rho^2)} \end{pmatrix} \begin{pmatrix} dZ_1(t) \\ dZ_2(t) \end{pmatrix}. \tag{1.24}$$

To apply the Feymann-Kac Theorem, we need the generator given in (1.20). Since

$$\begin{aligned} \boldsymbol{\sigma\sigma}^T &= \begin{pmatrix} \sqrt{v_t} & 0 \\ \sigma\rho\sqrt{v_t} & \sigma\sqrt{v_t(1-\rho^2)} \end{pmatrix} \begin{pmatrix} \sqrt{v_t} & 0 \\ \sigma\rho\sqrt{v_t} & \sigma\sqrt{v_t(1-\rho^2)} \end{pmatrix}^T \\ &= \begin{pmatrix} v_t & \rho\sigma v_t \\ \rho\sigma v_t & \sigma^2 v_t \end{pmatrix}, \end{aligned}$$

the generator in (1.20) becomes

$$\mathcal{A} = \left(r_t - \frac{1}{2}v_t\right)\frac{\partial}{\partial x_t} + \kappa^*(\theta^* - v_t)\frac{\partial}{\partial v_t} + \frac{1}{2}\left[v_t\frac{\partial^2}{\partial x_t^2} + 2\sigma\rho v_t\frac{\partial^2}{\partial x_t \partial v_t} + \sigma^2 v_t\frac{\partial^2}{\partial v_t^2}\right]. \tag{1.25}$$

Then (1.22) can be written as

$$\frac{\partial Y}{\partial t}+\left(r_t-\frac{1}{2}v_t\right)\frac{\partial Y}{\partial x_t}+\kappa^*\left(\theta^*-v_t\right)\frac{\partial Y}{\partial v_t}+\frac{1}{2}v_t\frac{\partial^2 Y}{\partial x_t^2}+\sigma\rho v_t\frac{\partial^2 Y}{\partial x_t\partial v_t}+\frac{1}{2}\sigma^2 v_t\frac{\partial^2 Y}{\partial v_t^2}-r_t Y=0. \tag{1.26}$$

Equation (1.26) is called the Heston PDE which is used to solve for ϕ_j and $j=1,2$.

1.5.3 Solving the Characteristic Functions

Note that the definition of Y in (1.26) is given by (1.21), we can replace Y by Π_C in (1.26) and get the following PDE by setting $\tau=T-t$,

$$\begin{aligned}-\frac{\partial \Pi_C}{\partial \tau}+\frac{1}{2}v_t\frac{\partial^2\Pi_C}{\partial x_t^2}+\left(r_{T-\tau}-\frac{1}{2}v_t\right)\frac{\partial\Pi_C}{\partial x_t}+\rho\sigma v_t\frac{\partial^2\Pi_C}{\partial v_t\partial x_t}\\ +\frac{1}{2}\sigma^2 v_t\frac{\partial^2\Pi_C}{\partial v_t^2}-r_{T-\tau}\Pi_C+\left[\kappa^*(\theta^*-v_t)\right]\frac{\partial\Pi_C}{\partial v_t}=0.\end{aligned} \tag{1.27}$$

Inserting (1.13) into (1.27) leads to

$$\begin{aligned}&e_t^x\left[-\frac{\partial P_1}{\partial\tau}+\left(r_{T-\tau}+\frac{1}{2}v_t\right)\frac{\partial P_1}{\partial x_t}+\frac{1}{2}v_t\frac{\partial^2 P_1}{\partial x_t^2}+\rho\sigma v_t\frac{\partial^2 P_1}{\partial v_t\partial x_t}+\left[\rho\sigma v_t+\kappa^*\left(\theta^*-v_t\right)\right]\frac{\partial P_1}{\partial v_t}+\right.\\ &\left.\frac{1}{2}\sigma^2 v_t\frac{\partial^2 P_1}{\partial v^2}\right]-Ke^{-\int_{T-\tau}^T r(w)\,dw}\left[-\frac{\partial P_2}{\partial\tau}+\frac{1}{2}v_t\frac{\partial^2 P_2}{\partial x_t^2}+\left(r_{T-\tau}-\frac{1}{2}\right)\frac{\partial P_2}{\partial x_t}+\rho\sigma v_t\frac{\partial^2 P_1}{\partial x_t\partial v_t}+\right.\\ &\left.\kappa^*\left(\theta^*-v_t\right)\frac{\partial P_2}{\partial v_t}+\frac{1}{2}\sigma^2 v_t\frac{\partial^2 P_2}{\partial v_t^2}\right]=0.\end{aligned} \tag{1.28}$$

Since (1.28) must be true for all strike values, the term multiplying K and the term independent of K should each be zero. Hence, (1.28) can be rewritten into two PDEs, that are

$$\begin{aligned}-\frac{\partial P_1}{\partial\tau}+\left(\frac{1}{2}v_t+r_{T-\tau}\right)\frac{\partial P_1}{\partial x_t}+\frac{1}{2}v_t\frac{\partial^2 P_1}{\partial x_t^2}+\rho\sigma v_t\frac{\partial^2 P_1}{\partial v_t\partial x_t}\\ +\left[\rho\sigma v_t+\kappa^*(\theta^*-v_t)\right]\frac{\partial P_1}{\partial v_t}+\frac{1}{2}\sigma^2 v_t\frac{\partial^2 P_1}{\partial v_t^2}=0,\end{aligned} \tag{1.29}$$

and

$$\begin{aligned}-\frac{\partial P_2}{\partial\tau}+\rho\sigma v_t\frac{\partial^2 P_2}{\partial v_t\partial x_t}+\frac{1}{2}\frac{\partial^2 P_2}{\partial x_t^2}+\frac{1}{2}v_t\sigma^2\frac{\partial^2 P_2}{\partial v_t^2}\\ +\left(r_{T-\tau}-\frac{1}{2}v_t\right)\frac{\partial P_2}{\partial x_t}+\left[\kappa^*(\theta^*-v_t)\right]\frac{\partial P_2}{\partial v_t}=0.\end{aligned} \tag{1.30}$$

We can rewrite (1.29) and (1.30) by

$$-\frac{\partial P_j}{\partial \tau}+\rho\sigma v_t\frac{\partial^2 P_j}{\partial x_t\partial v_t}+\frac{1}{2}v_t\frac{\partial^2 P_j}{\partial x_t^2}+\frac{1}{2}v_t\sigma^2\frac{\partial^2 P_j}{\partial v_t^2}$$
$$+\left(r_{T-\tau}+u_j v_t\right)\frac{\partial P_j}{\partial x_t}+\left(a-b_j v_t\right)\frac{\partial P_j}{\partial v_t}=0, \quad (1.31)$$

where

$$u_1=\frac{1}{2}, u_2=-\frac{1}{2}, a=\kappa^*\theta^*, b_1=\kappa^*-\rho\sigma, b_2=\kappa^*.$$

In order to obtain PDEs of ϕ_j, for $j=1,2$, we need to introduce Kolmogorov's backward equation given in the following theorem.

Theorem 1.12 (Kolmogorov's backward equation)
Let $f\in C_0^2\left(\mathbf{R}^n\right)$ and $\mathbf{x}_t$ be and Itô diffusion in $\mathbf{R}_n$ with generator $\mathcal{A}$.

1. *Define*

$$u\left(t,\mathbf{x}\right)=E\left[f\left(\mathbf{x}_t\right)\left|\mathbf{x}\right.\right]. \quad (1.32)$$

Then $u\left(t,\cdot\right)\in\mathcal{D}_{\mathcal{A}}$ for each t and

$$\frac{\partial u}{\partial t}=\mathcal{A}u, \quad t>0,\ \mathbf{x}\in\mathbf{R}^n, \quad (1.33)$$
$$u\left(0,\mathbf{x}\right)=f\left(\mathbf{x}\right); \quad \mathbf{x}\in\mathbf{R}^n, \quad (1.34)$$

where the right hand side is to be interpreted as $\mathcal{A}$ applied to the function $\mathbf{x}\rightarrow u\left(t,\mathbf{x}\right)$.

2. *Moreover, if $w\left(t,\mathbf{x}\right)\in C^{1,2}\left(\mathcal{R}\times\mathcal{R}^n\right)$ is a bounded function satisfying (1.33), (1.34) then $w\left(t,\mathbf{x}\right)=u\left(t,\mathbf{x}\right)$, given by (1.32).*

Proof. See Oksendal (2002). □

According to Theorem 1.12, the characteristic functions ϕ_j must satisfy the PDEs of P_j but with different boundary conditions, which are

$$\phi_j\big(x_0,v_0,\tau=0\big)=e^{isx_0}, \quad j=1,2.$$

In fact, by definition when $\tau = 0$

$$\phi_1(s) = E^{\mathbb{Q}_S}\left[e^{isx_0}\right] = e^{isx_0},$$
$$\phi_2(s) = E^{\mathbb{Q}}\left[e^{isx_0}\right] = e^{isx_0}.$$

Following Heston (1993) we assume the characteristic function is the form of:

$$\phi_j(s) = \exp\{C_j(\tau, s) + D_j(\tau, s)v_t + isx_t\}. \tag{1.35}$$

Hence, $D_j(0, s) = 0$ and $C_j(0, s) = 0$. Inserting (1.35) into (1.31) leads to

$$-\left(\frac{\partial C_j}{\partial \tau} + \frac{\partial D_j}{\partial \tau}v_t\right) + \rho\sigma v_t i s D_j - \frac{1}{2}v_t s^2 + \frac{1}{2}v_t\sigma^2 D_j^2 + (r_{T-\tau} + u_j v_t)\, is + (a - b_j v_t)\, D_j = 0.$$

Grouping the terms including v_t, we have

$$v_t\left(-\frac{\partial D_j}{\partial \tau} + \rho\sigma i s D_j - \frac{1}{2}s^2 + \frac{1}{2}\sigma^2 D_j^2 + u_j i s - b_j D_j\right) - \frac{\partial C_j}{\partial \tau} + r_{T-\tau} i s + a D_j = 0. \tag{1.36}$$

Note that (1.36) must be true for all values of v_t, so that in (1.36) the term with v_t and the term free from v_t are both supposed to be zero. Hence, (1.36) can be rewritten using the following two equations:

$$\frac{\partial D_j}{\partial \tau} = \rho\sigma i s D_j - \frac{1}{2}s^2 + \frac{1}{2}\sigma^2 D_j^2 + u_j i s - b_j D_j \tag{1.37}$$
$$\frac{\partial C_j}{\partial \tau} = r_{T-\tau} i s + a D_j. \tag{1.38}$$

Since $D(0, s) = 0$, (1.37) is solvable according to Appendix C. Setting

$$L_j = u_j i s - \frac{1}{2}s^2, Q_j = \rho\sigma i s - b_j \quad \text{and} \quad R = \frac{1}{2}\sigma^2,$$

(1.37) can be written as

$$\frac{\partial D_j}{\partial \tau} = L_j + Q_j D_j + R D_j^2.$$

According to Appendix C, denote by

$$B_0 = \frac{-Q_j + \sqrt{Q_j^2 - 4RL_j}}{2R},$$
$$A = 2RB_0 + Q_j,$$

the unique solution to (1.37) with $D(0, s) = 0$ is given by

$$
\begin{aligned}
D_j &= \frac{B_0\left(A-RB_0\right)e^{-At}-\left(A-RB_0\right)B_0}{(A-RB_0)e^{-At}+RB_0} \\
&= \frac{A-RB_0}{R}\frac{e^{-A\tau}-1}{\frac{A-RB_0}{RB_0}e^{-A\tau}+1} \\
&= \frac{\rho\sigma is-b_j-d_j}{\sigma^2}\frac{e^{d_j\tau}-1}{\frac{\rho\sigma is-b_j-d_j}{-\rho\sigma is+b_j-d_j}e^{d_j\tau}+1} \\
&= \frac{b_j-\rho\sigma is+d_j}{\sigma^2}\frac{1-e^{d_j\tau}}{1-g_je^{d_j\tau}},
\end{aligned}
\tag{1.39}
$$

where

$$
\begin{aligned}
d_j &= -\sqrt{\left(\rho\sigma is-b_j\right)^2-\sigma^2\left(2u_jis-s^2\right)}, \\
g_j &= \frac{b_j-\rho\sigma is+d_j}{b_j-\rho\sigma is-d_j}.
\end{aligned}
$$

Given D_j, the solution for C_j with boundary conditions $C_j\left(0,s\right)=0$ is given by integrating (1.38),

$$
\begin{aligned}
C_j &= \int_0^\tau r(T-\tau)is\,d\tau+a\left(\frac{b_j-\rho\sigma is+d_j}{\sigma^2}\right)\int_0^\tau\frac{1-e^{d_j\tau}}{1-g_je^{d_j\tau}}\,d\tau \\
&= \int_0^\tau r\left(T-\tau\right)is\,d\tau+\frac{a}{\sigma^2}\left[\left(b_j-\rho\sigma is+d_j\right)\tau-2ln\left(\frac{1-g_je^{d_j\tau}}{1-g_j}\right)\right] \\
&= is\int_t^T r(w)\,dw+\frac{a}{\sigma^2}\left[\left(b_j-\rho\sigma is+d_j\right)\tau-2ln\left(\frac{1-g_je^{d_j\tau}}{1-g_j}\right)\right].
\end{aligned}
\tag{1.40}
$$

Obtaining C_j and D_j, (1.35),(1.39) and (1.40) give us the formula for ϕ_j, $j=1,2$. Hence, the formula of call price could be obtained by combining (1.35), (1.17) and (1.13).

Though a closed (or semi-closed) formula for the European call price is derived, there are still some numerical problems which make it unpractical. Albrecher et al. (2007) pointed out that due to the discontinuity in the branch cut of the complex logarithm in C_1, $\phi_1\left(s\right)$ is unstable for the long term maturity or for certain parameters, while $\phi_2\left(s\right)$ does not suffer this problem. This numerical problem was solved by Bakshi and Madan (2000) in which the relationship between $\phi_1\left(s\right)$ and $\phi_2\left(s\right)$ was found as follows:

$$
\phi_1\left(s\right)=\frac{\phi_2\left(s-i\right)}{\phi_2\left(-i\right)}.
\tag{1.41}
$$

Since ϕ_1 in our paper suffers the same problem as Heston's original formula, we use $\phi_2(s)$ and the relationship between the two characteristic functions given in (1.41) to evaluate P_1 and P_2.

Finally, we give the European call price under Heston model with deterministic interest as follows: Under Heston's model with deterministic interest, the price for a European call option with strike K and maturity T at time $t\,(\leq T)$ given current asset price S_t and current stochastic variance v_t can be expressed as

$$\Pi_C(t,T,S_t,v_t,K) = e^{x_t}P_1 - Ke^{-\int_t^T r_w\,dw}P_2, \tag{1.42}$$

where

$$\begin{aligned}
x_t &= \ln S_t,\\
P_1 &= \frac{1}{2}+\frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-is\ln K}\phi_2(s-i)}{is\phi_2(-i)}\right]ds,\\
P_2 &= \frac{1}{2}+\frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-is\ln K}\phi_2(s)}{is}\right]ds,\\
\phi_2(s) &= exp\{C(\tau,s)+D(\tau,s)\,v_t+ix_ts\},\\
C(\tau,s) &= is\int_t^T r(w)\,dw+\frac{\kappa^*\theta^*}{\sigma^2}\left[(\kappa^*-\rho\sigma is+d)\,\tau-2\ln\left(\frac{1-ge^{d\tau}}{1-g}\right)\right],\\
D(\tau,s) &= \frac{\kappa^*-\rho\sigma is+d}{\sigma^2}\left(\frac{1-e^{d\tau}}{1-ge^{d\tau}}\right),\\
g &= \frac{\kappa^*-\rho\sigma is+d}{\kappa^*-\rho\sigma is-d},\\
d &= -\sqrt{(\rho\sigma is-\kappa^*)^2+\sigma^2(is+s^2)},\\
\tau &= T-t.
\end{aligned}$$

Please note that "d" in our paper is the opposite to the "d" in Heston's original paper. In fact, there could be two values for "d" because it is the square root of a complex number and they lead to the same result. This is because of the uniqueness of the solution to the Riccati function (1.37). (See details in Appendix C). However, compared with the call prices obtained by simulation and the numerical methods that we introduced in Chapter 3, Heston's choice generates a little error in practice while the other value of "d" leads

to consistent results. This is due to the branch cut of the square root function of a complex number. Another explanation is that when d is positive the computation error in evaluating $e^{d\tau}$ is larger than $e^{-d\tau}$.

1.6 Calibrating the Heston Model

The parameters of the asset price dynamic are calibrated using observed European call prices. Since most of our formulas are derived under the risk neutral measure, we pay more attention to the calibration of the Heston parameters $G_2 = \{\kappa^*, \theta^*, \sigma\rho, v_0\}$.

In order to calibrate the Heston model with deterministic interest, we first identify and calibrate the model for yield rates. In this thesis, we assume that

$$y_t = \beta_1 + \beta_2 \left[\frac{1 - e^{-t/\lambda_1}}{t/\lambda_1}\right] + \beta_3 \left[\frac{1 - e^{-t/\lambda_1}}{t/\lambda_1} - e^{-t/\lambda_1}\right] + \beta_4 \left[\frac{1 - e^{-t/\lambda_2}}{t/\lambda_2} - e^{-t/\lambda_2}\right], \quad (1.43)$$

where t is the time to maturity. $\beta_1, \beta_2, \beta_3, \beta_4, \lambda_1$ and λ_2 are parameters to be estimated. This model, which is called Nelson-Siegel-Svensson (NSS) model, is introduced in Svensson (1994). Although we specify a certain formula for the deterministic yield rate, our financial model does not limited to the particular one.

We calibrate $G_1 = \{\beta_1, \beta_2, \beta_3, \beta_4, \lambda_1, \lambda_2\}$ according to the observed rates from the market. Table 2 illustrates the yield rates of the U.S. Treasury bills (T-bills)[5] which were observed on August 9th, 2011 and the maturities of corresponding T-bills range from 1 month to 30 years. Following Gilli et al. (2010), we use DE method to calibrated G_1 according to the yield rates given in Table 2 (brief introduction to the DE algorithm is given in Appendix B). The objective function is

$$G_1 = \arg\min_{\Omega_1} \sum_i \left(y_i^{Obs} - y_i^{NSS}\right)^2,$$

where Ω_1 is the space of parameters, y_i^{Obs} stands for the observed yield rates given by the U.S. Department of Treasury and y_i^{NSS} is for the yield rates obtained from (1.43). The calibration results are given in Table 3.

[5]Data was obtained from http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yield

Table 2: Yield curve in percentage

Time	1 Month	3 Month	6 Month	1 Year	2 Year	3 year
Yield Rate	0.02	0.03	0.06	0.11	0.19	0.33
Time	5 Year	7 Year	10 Year	20 Year	30 Year	
Yield Rate	0.91	1.53	2.20	3.17	3.56	

Table 3: NSS model calibration

Parameter	Calibration Results
β_1	4.233068
β_2	-4.233048
β_3	-25.918993
β_4	19.522368
λ_1	1.572826
λ_2	1.367069

The error is estimated by the root of mean squared error which is,

$$\sqrt{\frac{1}{N}\sum_i \left(y_i^{Obs} - y_i^{NSS}\right)^2} = 0.004768\%.$$

Furthermore, Figure 1 shows the calibration results as well as the NSS model fits the observed data well.

After calibrating the NSS model, the Heston parameters can be calibrated according to the observed data. In this paper we calibrate $G_2 = \{\kappa^*, \theta^*, \sigma\rho, v_0\}$ based on observed European call options. In other words, calibrating the Heston parameters is an optimization problem as follows,

$$G_2 = \arg\min_{\Omega_2} \sum_{i=1}^{N} w_i \left[\Pi_C^{Obs}(0, T_i, S_0, v_0, K_i) - \Pi_C^{Heston}(0, T_i, S_0, v_0, K_i)\right]^2, \tag{1.44}$$

where Ω_2 is the space for G_2, $\Pi_C^{Obs}(0, T_i, S_0, v_0, K_i)$ is the observed European call option price, $\Pi_C^{Heston}(0, T_i, S_0, v_0, K_i)$ is the European call option price obtained by (1.42), N is

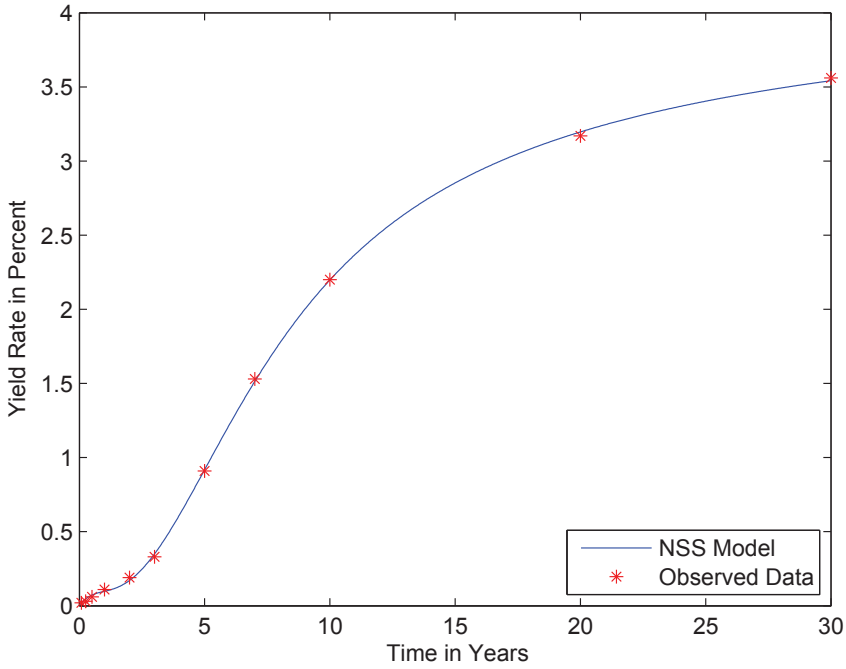


Figure 1: Calibration result for NSS model.

the number of observed prices and $\{w_i\}$ are the weights. Because the market prices are given as a bid-ask spread[6], it is impossible to determine the real price. Hence, the average of the bid and asked prices is used to define the observed European call prices. Following Moodley (2005), we define the weights as $\frac{1}{|\text{bid}_i - \text{ask}_i|}$, where bid_i is the bid price and ask_i is the asked price. This weight makes sense: the larger the difference between the asked and bid price, the harder it is to determine the real price, hence we are supposed to assign less weights to such prices.

In our paper, forty European call options on S&P 500 observed on 9th August, 2011[7] are used to calibrate Heston's parameters. On that day the S&P 500 index closed at $1,172.53$ that is used as the initial asset price S_0. The maturities of the observed call options range from 2 month to 14 month and the strikes are between 800 and $1,400$.

[6]Bid prices are the prices at which the agents sell the calls and ask prices are the prices at which the agents buy calls.

[7]Data from marketwatch.com

As we said before, the Heston model under the risk-neutral measure has five parameters that need to be estimated. Minimizing the objective function is a nonlinear programming problem. The objective function is far from being convex and Mikhailov and Nögel (2003) pointed out that there exist many local extrema so that global optimizers should be applied here. In our paper we use the Differential Evolution (DE) method introduced by Storn and Price (1997) to minimize the objective function. The DE method is widely and successfully used when calibrating Heston model and other financial models, see Gilli and Schumann (2010), Schoutens et al. (2004), and Vollrath and Wendland (2009).

The parameters estimation is given in Table 4: Note that this set of parameters

Table 4: Heston parameters under risk neutral measure

Parameter	Calibration Results
κ^*	4.1
θ^*	0.046
σ	0.605
ρ	-0.7736
v_0	0.077931

satisfies the Feller condition, that is, $2\kappa^*\theta^* - \sigma^2 > 0$. This result is obtained using the DE optimization algorithm with parameters $n_p = 50, F = 0.5, CR = 0.9, N_G = 200$.

The error is evaluated by the root of the weighted mean squared error, which is

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N} w_i \left[\Pi_i^{Obs} - \Pi_i^{Heston}\right]^2} = 1.4700.$$

Compared with the weighted average call option, that is

$$\frac{1}{N}\sum_{i=1}^{N} w_i \Pi_i^{model} = 31.39$$

the calibration error seems acceptable. Furthermore, in Figure 2 the black stars stand for observed European call option prices and the surface is generated by (1.42) with the

parameters given in Table 4. From Figure 2 we can see the calibrated prices fit the observed prices well.

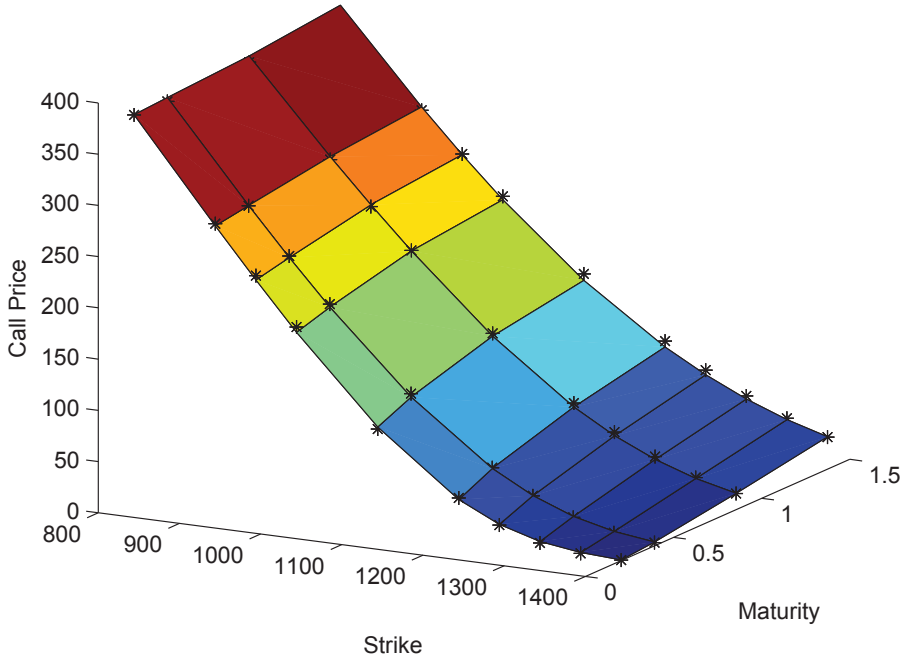


Figure 2: Calibration result for Heston parameters.

2 Joint Transition PDF of Heston Model

In this section, we introduce the conditional probability density functions of the asset dynamics and volatility.

As we presented in Section 1.2, the volatility follows a noncentral chi-square distribution. In terms of the marginal PDF of the asset price, there is still no closed formula though a number of research works have tried to address this problem. One interesting result is given by Drăgulescu and Yakovenko (2002). They first obtain $p\left(\ln(S_t/S_0), v_t|v_0\right)$ by inverting its characteristic function, then a formula for the marginal PDF of the asset price, $p\left(\ln\left(S_t/S_0\right)\right)$, is derived by integrating v_t and v_0 out. They call their result the DY formula and show that the DY formula fits the data of the Dow-Jones index from 1982 to 2002. Silva and Yakovenko (2003) continue to illustrate that the DY formula successfully fits the data of S&P 500 and Nasdaq indices from 1980s to 2000s. However, in integrating v_0 out they assumed it follows the stationary distribution of $p\left(v_{t+\tau}|v_t\right)$. This makes the DY formula an approximation. Though this assumption makes sense for long-time periods, it is not reasonable for short terms. Rollin et al. (2010) proved that the density of the log-return has a C^∞(or smooth) density and can be written as an infinite convolution of Bessel type densities.

Though the marginal PDF of asset prices is unknown, Lipton (2001) as well as Lamoureux and Paseka (2009) both derived closed formulas for the joint transition density probability function (JTPDF) of the asset and volatility. The Heston process is described by the stochastic processes S_t (or x_t) and v_t, so that it is characterized by the JTPDF $p\left(x_{t+\tau}, v_{t+\tau}|x_t, v_t\right)$. The JTPDF is advantageous in dealing with the path-dependent derivatives because of the Markovian property of (x_t, v_t). This enables us to address the path-dependent derivatives period by period, rather than dealing with the whole path.

2.1 The Joint Transition Probability Density Function

In this section, we generalize formulas in Lamoureux and Paseka (2009) to the case of deterministic interest.

Proposition 2.1 *Given the Heston process with deterministic interest rate under $\mathbb{Q}$, that is (1.4), the JTPDF of the process is given by*

$$p\left(x_{t+\tau}, v_{t+\tau} | x_t, v_t\right) = \frac{1}{\pi} e^{\frac{\nu+1}{2}\kappa^*\tau} \left(\frac{v_{t+\tau}}{v_t}\right)^{\nu/2} \int_0^{+\infty} e^{imk} h\left(k\right)\, dk, \tag{2.1}$$

where

$$\begin{aligned}
m &= -\left(x_{t+\tau} - x_t - \int_t^T r_w\, dw + \frac{\nu+1}{2}\rho\sigma\tau\right), \\
h\left(k\right) &= \psi I_\nu\left[2\sqrt{\varphi v_t v_{t+\tau}}\right] e^{\frac{\tau d}{2} + (A_2 - \psi)v_t - (A_1 + \psi)v_{t+\tau}}, \\
\nu &= \frac{2\kappa^*\theta^*}{\sigma^2} - 1, \\
\psi &= \frac{d}{R_1\left(e^{d\tau} - 1\right)}, \\
\varphi &= \psi^2 e^{d\tau}, \\
A_1 &= \frac{1}{2R_1}\left[-R_2 + d\right], \\
A_2 &= \frac{1}{2R_1}\left[-R_2 - d\right], \\
R_1 &= \frac{\sigma^2}{2}, \\
R_2 &= ik\rho\sigma - \kappa^*, \\
d &= \sqrt{(\kappa^* - ik\rho\sigma)^2 + \sigma^2\left(k^2 + ik\right)},
\end{aligned}$$

and $I_\nu\left[\cdot\right]$ is the modified Bessel function of the first kind of order ν.

Proof. Lamoureux and Paseka (2009) give a proof when the interest rate is constant. This idea can be applied to the case of deterministic interest rates. See Ballestra et al. (2007) on how to change the bounds of the interval. □

Figure 3 consists of the graphs related to the JTPDF. The plot on the left is the figure of the JTPDF, the plot on the top right is the projection on the surface of JTPDF and x_t,

and the third plot at the bottom right is the projection on the surface of JTPDF and v_t. The parameters used for plotting are obtained from the calibration. Figure 3 shows how the log-price and the stochastic variance of S&P 500 would change after August 9th, 2011 based on the market of European call prices that we observed. Firstly, the support of the JTPDF is rather small compared to its domain which is $[-\infty, +\infty] \times [0, +\infty]$. Secondly, the log-price tends to increase with a large probability and has a heavier tail on the left side. Thirdly, the volatility tends to decrease with large probabilities.

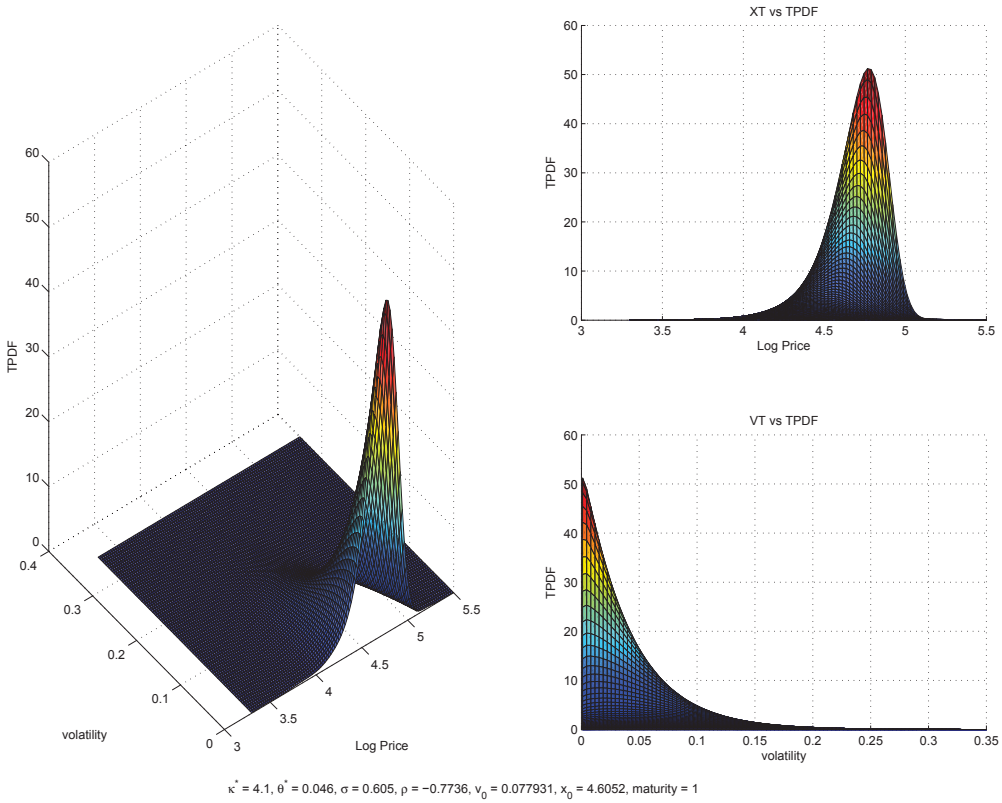


Figure 3: Figure of the transition PDF

2.2 The Quadrature of The TPDF Integral

The integral in (2.1) is an oscillatory integral due to the term e^{imk}. When traditional quadrature is applied to evaluate such integrals, the result is not reliable. In fact, to evaluate an integral, for example $I = \int_a^b f(x)\,dx$, the traditional quadrature first generates N points $\{x_i\}$, $i = 1, 2, \ldots, N$ according to a certain rule from the interval $[a, b]$, then the following summation is applied to approximate the integral

$$I = \sum_{i=1}^{N} \omega_i f(x_i),$$

where ω_i is the weight assigned to $f(x_i)$. Setting $\omega_i = \frac{1}{N}$ leads to the Monte-Carlo method. When $f(x)$ is difficult to evaluate, $f(x)$ is approximated by an interpolation $g(x)$ according to $(x_i, f(x_i))$, $i = 1, 2, \ldots N$. The function $g(x)$ is selected such that it is always easy to evaluate. Hence, the basic idea of typical quadrature is

$$\begin{aligned} I &= \int_a^b f(x)dx \\ &\approx \sum_{i=1}^{N} \omega_i g(x_i). \end{aligned}$$

However, when $f(x)$ is an oscillatory function the integral of $f(x)$ is referred to an oscillatory integral and the points might not be representative of the behavior of the function. Figure 4 shows how these interpolation points could miss their target. In Figure 4 the blue line is the plot for function $x^2\cos(60x)$, which is a typical oscillatory function. To approximate this function we randomly generate 20 points that are marked by red stars. These interpolation points are fooled by the behavior of $x^2\cos(60x)$, and the typical quadrature gives the integral of the function presented by the red line. Some Gaussian quadratures such as the adaptive Gauss-Kronrod or Lobatto quadratures works at evaluating the oscillatory integral in (2.1). However, this is because the oscillatory factor m in (2.1) is too small for our problem. When the oscillatory factor is large, like the function in Figure 4, the results obtained by a typical quadrature are not reliable. Hence, the Gaussian-type quadrature is not recommended to evaluate oscillatory integrals. Firstly, it

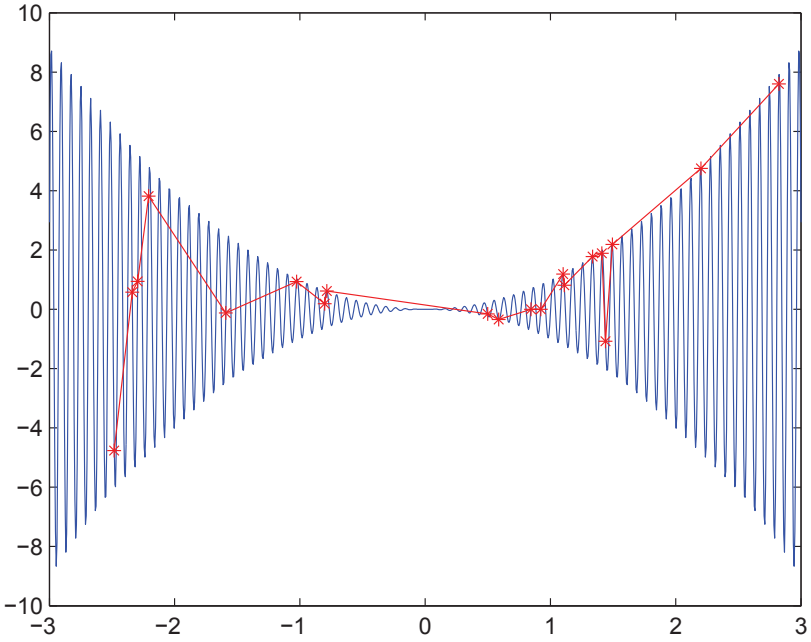


Figure 4: $f(x) = x^2 \cos(60x), \quad x \in [-3, 3]$

is time-consuming since the method requires a lot of interpolation points when evaluating the oscillatory integral. Moreover, the Gaussian-type quadrature gives modest accuracy. However, to evaluate an oscillatory integral there are many other numerical methods. Olver (2008) gives a good review of the literature on oscillatory integrals. In our paper we use the Filon-type quadrature which was introduced in Filon (1928) to solve the integral

$$I = \int_a^b e^{imk} h(k)\, dk. \tag{2.2}$$

Though it was introduced almost one hundred years ago, it is fairly good in solving oscillatory integrals when the oscillatory factor m is given. So far there have been just limited improvements to Filon's idea.

Filon's method approximates the function $h(k)$ by polynomials instead of interpolating the whole integrand in (2.2). When $h(k)$ is continuous the accuracy of the approximation is guaranteed by the Weierstrass approximation theorem (see Jeffreys and Jeffreys (1988)). Assume $h(k)$ could be approximated by a polynomial $P(k)$ such that $\max_k |h(k)-$

$P(k)| < \epsilon$, then

$$\begin{aligned} I - Q &= \int_a^b e^{imk} h(k)\, dk - \int_a^b e^{imk} P(k)\, dk \\ &= \int_a^b e^{imk} [h(k) - P(k)]\, dk \\ &\leq \int_a^b |e^{imk} [h(k) - P(k)]|\, dk \\ &\leq \int_a^b |h(k) - P(k)|\, dk \\ &\leq |b-a| \max_k |h(k) - P(k)|. \end{aligned}$$

The last two lines hold since $|e^{imk}|$ is always less or equal to 1. Hence the error of the Filon quadrature is bounded as follows:

$$error(I - Q) \leq |b-a| \max_k |h(k) - P(k)|,$$

which goes to zero if $P(k)$ is an excellent approximation to $h(k)$. From the above error estimation, the accuracy of Filon quadrature is free from the oscillatory factor.

Following Ballestra et al. (2007) we apply a Filon-type quadrature to our problem. Before doing that we change the integral in (2.1) as follows:

$$\begin{aligned} \int_0^{+\infty} e^{imk} h(k)\, dk &= \int_0^{+\infty} [cos(mk) + isin(mk)] \times [Re\{h(k)\} + iIm\{h(k)\}]\, dk \\ &= \int_0^{+\infty} cos(mk)\, Re\{h(k)\} - sin(mk)\, Im\{h(k)\}\, dk \\ &\quad + \int_0^{+\infty} i\, [cos(mk)\, Im\{h(k)\} + sin(mk)\, Re\{h(k)\}]\, dk \\ &= \int_0^{+\infty} cos(mk)\, Re\{h(k)\} - sin(mk)\, Im\{h(k)\}\, dk. \end{aligned}$$

The last equation holds since the integral is a probability density function which must be real. The upper bound of the integral $+\infty$ can be replaced by a constant k_{max} since the absolute value of the integrand goes to zero as k increases. Hence,

$$\int_0^{+\infty} e^{imk} h(k)\, dk \approx \int_0^{k_{max}} e^{imk} h(k)\, dk.$$

It is easy to choose a k_{max} such that $|e^{imk}h(k)| < \epsilon$ for all possible x_t and v_t, given the Heston parameters, x_0 and v_0. In this paper, we choose $k_{max} = 60$ such that

$$|e^{imk}h(k)| < \epsilon = 10^{-5},$$

when $k > k_{max}$.

In what follows, we illustrate how we can interpolate $h(k)$. Our idea is to use piece-wise interpolation. We first divide the interval $[0, k_{max}]$ into M subintervals at points $x_0, x_1, \ldots, x_M$, then we interpolate $h(k)$ on each interval. In order to have efficient computation time, we use 6th-order Lagrange interpolation in each subinterval, and the interpolation points in each subinterval are a linear transformation of the roots of the Legendre polynomial of degree 7. In the ith subinterval, the interpolation points are denoted by $x_0^i, x_1^i, \ldots, x_6^i$. In Filon (1928) the interpolation was a quadratic spline, the spline could be more accurate but it costs more time since the derivatives of $h(k)$ and the interpolation function at boundaries points of each subinterval are set to be the same for spline. We denote the piece-wise Lagrange interpolation as $L(k)$. Using this method the integral (2.2) becomes

$$\int_0^{k_{max}} e^{imk} L(k)\, dk \tag{2.3}$$

and it can be solved analytically since $L(k)$ is a polynomial. The quadrature error depends on the Lagrange interpolation, that is, for any k in $[a, b]$

$$error\left(h(k) - L(k)\right) \leq \max_i \max_{x_i^*} \left| \frac{f^{(n+1)}(\xi_i)}{(n+1)!} \omega_i(x_i^*) \right|,$$

where $\xi_i, x_i^* \in [x_{i-1}, x_i]$ and

$$\omega_i(x) = \left(x - x_0^i\right)\left(x - x_1^i\right)\ldots\left(x - x_6^i\right).$$

Hence, the quadrature error is bounded by

$$k_{max} \max_i \max_{x_i^*} \left| \frac{f^{(n+1)}(\xi_i)}{(n+1)!} \omega_i(x_i^*) \right|.$$

Here, we set $n = 6$ which is large enough. The Filon-type quadrature is faster since all the computations are evaluated analytically except for the modified Bessel function in (2.2).

2.3 Pricing European Path-Dependent Derivatives using TPDF

It is challenging to price path-dependent derivatives in the Heston framework. Although they can be priced by simulating the asset price, simulation methods have some drawbacks. It is hard to evaluate and eliminate the discretization error. Another shortcoming is that it is hard to simulate the Greeks directly, which represent the sensitivities of the price with respect to financial market changes. In this subsection we show how to price European pathdependent derivatives by JTPDFs, a method which overcomes the simulation's shortcoming.

Since Heston's process is a Markovian process of x_t and v_t, given the initial log-price x_0 and stochastic variance v_0, the joint PDF of $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ can be written as:

$$p\left(\mathbf{x}, \mathbf{v} | x_0, v_0\right) = \prod_{t=1}^{n} p\left(x_t, v_t | x_{t-1}, v_{t-1}\right). \tag{2.4}$$

Using (2.4) it is possible to price European path-dependent derivatives analytically.

Let $\mathbf{S} = (S_0, S_1, \ldots, S_n)$, we denote the payoff of a European path-dependent derivative by $C_{\text{path}}(\mathbf{S}, n)$. Note that $S_i = S_i\left(x_i\right) = e^{x_i}$, the payoff is rewritten as $C_{\text{path}}(\mathbf{S}\left(\mathbf{x}\right), n)$, where $\mathbf{S}\left(\mathbf{x}\right) = \left(S_0\left(x_0\right), S_1\left(x_1\right), \ldots, S_n\left(x_n\right)\right)$. Note that the time unit represents one period, so that we let the initial time t_0 be 0 but it can represent any time before maturity n. Then, according to Proposition 1.3, the price of the European path-dependent derivative, $\Pi_{\text{path}}\left(0, n, x_0, v_0\right)$, is the expectation of its discounted payoff, that is,

$$\begin{aligned}
\Pi_{\text{path}}\left(0, n, x_0, v_0\right) &= E^{\mathbb{Q}}\left[e^{-\int_0^n r_w\, dw} C_{\text{path}}(\mathbf{S}\left(\mathbf{x}\right), n) \middle| x_0, v_0\right] \\
&= e^{-\int_0^n r_w\, dw} E^{\mathbb{Q}}\left[C_{\text{path}}(\mathbf{S}\left(\mathbf{x}\right), n) \middle| x_0, v_0\right] \\
&= e^{-\int_0^n r_w\, dw} \underbrace{\int \cdots \int}_{(\mathbf{x},\mathbf{v})\in\Omega} C_{\text{path}}(\mathbf{S}\left(\mathbf{x}\right), n) \prod_{j=1}^{n} p\left(x_j, v_j | x_{j-1}, v_{j-1}\right)\, d\mathbf{x} d\mathbf{v},
\end{aligned} \tag{2.5}$$

where $\Omega = [-\infty, +\infty]^n \times [0, +\infty]^n$.

Note that (2.5) could be solved recursively by tacking

$$\iint\limits_{(x_n,v_n)\in[-\infty,+\infty]\times[0,+\infty]} C_{\text{path}}(\mathbf{S}(\mathbf{x}),n)p(x_n,v_n|x_{n-1},v_{n-1})\,dx_n dv_n,$$

for given x_{n-1} and v_{n-1}. Hence, it becomes possible to solve (2.5). Lipton (2001) first used this idea to price the forward starting option. However, while n gets large it is rather challenging to solve the multidimensional integral. Ballestra et al. (2007) introduced a numerical method to price the path-dependent derivatives and they succeed to price an one-year arithmetic Asian option. Because the main purpose of our paper is to price a 7-year EIA contract, which means that (2.5) is a 20-dimensional integral, we solve the multiple integral by a numerical method.

2.4 Importance Sampling

2.4.1 Importance Sampling

Following Ballestra et al. (2007), we solve the multiple integral in (2.5) using the method of importance sampling. In this case the standard Monte Carlo method is inefficient to evaluate the expectation since it is difficult to generate samples from the density $p(x_j,v_j|x_{j-1},v_{j-1})$. Thus the importance sampling method is introduced to overcome this problem.

The idea of importance sampling is as follows. Instead of generate samples from $p(x_j,v_j|x_{j-1},v_{j-1})$, we first approximate $p(x_j,v_j|x_{j-1},v_{j-1})$ by another similar function $\tilde{p}(x_j,v_j|x_{j-1},v_{j-1})$ which must be 0 where $p(x_j,v_j|x_{j-1},v_{j-1})=0$. Then (2.5) becomes:

$$\begin{aligned}
&\Pi_{\text{path}}(0,n,x_0,v_0)\\
=\;& e^{-\int_0^n r_w\,dw}\int\limits_{(\mathbf{x},\mathbf{v})\in\Omega}\cdots\int C_{\text{path}}(\mathbf{S}(\mathbf{x}),n)\times\\
&\frac{\prod_{j=1}^n p(x_j,v_j|x_{j-1},v_{j-1})}{\prod_{j=1}^n \tilde{p}(x_j,v_j|x_{j-1},v_{j-1})}\prod_{j=1}^n \tilde{p}(x_j,v_j|x_{j-1},v_{j-1})\,d\mathbf{x}d\mathbf{v}\\
=\;& e^{-\int_0^n r_w\,dw}E_{\tilde{p}}^{\mathbb{Q}}\left[C_{\text{path}}(\mathbf{S}(\mathbf{x}),n)\frac{\prod_{j=1}^n p(x_j,v_j|x_{j-1},v_{j-1})}{\prod_{j=1}^n \tilde{p}(x_j,v_j|x_{j-1},v_{j-1})}\middle| x_0,v_0\right],
\end{aligned}\tag{2.6}$$

where $E_{\tilde{p}}^{\mathbb{Q}}[\cdot]$ is the expectation under $\mathbb{Q}$ w.r.t $\prod_{j=1}^{n} \tilde{p}(x_j, v_j | x_{j-1}, v_{j-1})$.

Hence, to evaluate the multiple integral in (2.5), we just need to generate samples from $\tilde{p}(x_j, v_j | x_{j-1}, v_{j-1})$. Hence, what is left is to find an appropriation $\tilde{p}$ whose samples are easy to generate.

2.4.2 JTPDF Approximation

Again following Ballestra et al. (2007), $\tilde{p}(x_j, v_j | x_{j-1}, v_{j-1})$ is given by bilinear interpolation. This is because samples of a density described by a bilinear interpolation function are easy to generate. An introduction to bilinear interpolation is given in Appendix D.

In terms of approximating the JTPDF $p(x_{t+\tau}, v_{t+\tau} | x_t, v_t)$, note that in (2.1) both $x_{t+\tau}$ and x_t always appear together in the term of $x_{t+\tau} - x_t$, so that (2.1) is actually a function of $v_{t+\tau}$, v_t and $\Delta x_t = x_{t+\tau} - x_t$. Hence, the JTPDF is a function of three variables,

$$p(x_{t+\tau}, v_{t+\tau} | x_t, v_t) = p(\Delta x_t, v_{t+\tau} | 0, v_t). \tag{2.7}$$

In fact, (2.7) shows that, given v_t, the process for x_t is time homogeneous. Since the support is rather small compared to its domain, we focus on approximating p on its support.

Assume that v_t is given, we can find an upper bound v_{max} for $v_{t+\tau}$ using the transition PDF in (1.2), such that $p(v_{t+\tau} | v_t) < \varepsilon_v$ when $v_{t+\tau}$ is outside the interval $[0, v_{max}]$.

In terms of Δx_t, note that the process is described as

$$dx_t = \left(r_t - \frac{1}{2} v_t\right) dt + \sqrt{v_t} dW_s^{\mathbb{Q}}(t).$$

We discretize the previous process as

$$\Delta x_t \approx \int_t^{t+\tau} r(w)\, dw - \frac{1}{2} (\sigma^*)^2 \tau + \sigma^* Z,$$

where $Z \sim N(0, \tau)$. In terms of σ^*, for v_t is a mean-reverting process it makes sense to set $\sigma^* = \sqrt{E[v_{t+\tau} | v_t]}$. Then the PDF of Δx_t is estimated by a normal distribution $N\left(\int_t^{t+\tau} r(w)\, dw - \frac{1}{2}(\sigma^*)^2 \tau, (\sigma^* \sqrt{\tau})^2\right)$.

Thus, the support of Δx_t is estimated by

$$
\begin{aligned}
& [\Delta_{min}, \Delta_{max}] \\
\triangleq & \left[\int_t^{t+\tau} r(w)\, dw - \frac{1}{2}(\sigma^*)^2 \tau - \alpha_x \sigma \sqrt{\tau}, \quad \int_t^{t+\tau} r(w)\, dw - \frac{1}{2}(\sigma^*)^2 \tau - \alpha_x \sigma \sqrt{\tau} \right]
\end{aligned}
$$

where α_x is a number such that the probability for Δx_t to be outside the above interval is smaller than ε_x. Hence $[0, v_{max}] \times [\Delta_{min}, \Delta_{max}]$ is used as the support of the TPDF, $p\{\Delta x_t, v_{t+\tau}|0, v_t\}$.

Divide the above support into $N_v \times N_x$ grids. Then the look-up table can be generated using the Filon quadrature. Finally, $p(\Delta x_t, v_{t+\tau}|0, v_t)$, given v_t, can be approximated by a bilinear interpolation $\tilde{p}(\Delta x_t, v_{t+\tau}|0, v_t)$.

In the above, v_t is assumed to be known while in practice it is changing from time to time. Hence, $p(\Delta x_t, v_{t+\tau}|0, v_t)$ is also a function of v_t. Suppose that $v_{t+\tau}$ and Δx_t are given, we use piece-wise interpolation to approximate JTPDF as a function of v_t. Since the transition density function is given in (1.2) and the process v_t has a mean-reverting property, we can find an interval $[0, \mu_{max}]$ for v_t, given any possible $v_{t-\tau}$, such that $p(v_t|v_{t-\tau}) < \varepsilon_v$ when v_t is outside $[0, \mu_{max}]$. Hence, v_t must be in the interval $[0, \mu_{max}]$. Divide the interval $[0, \mu_{max}]$ into M subintervals and denote the break-points by $\{\mu_1, \mu_2, \ldots, \mu_M, \mu_{M+1}\}$, then v_t must fall into one of the subintervals. Say v_t falls into the interval $[\mu_{i-1}, \mu_i]$, given Δx_t and $v_{t+\tau}$, then the approximation of $p\{\Delta x_t, v_{t+\tau}|0, v_t\}$ is given by

$$
\tilde{p}(\Delta x_t, v_{t+\tau}|0, v_t) \approx \frac{\mu_i - v_t}{\mu_i - \mu_{i-1}} \tilde{p}(\Delta x_t, v_{t+\tau}|0, \mu_{i-1}) + \frac{v_t - \mu_{i-1}}{\mu_i - \mu_{i-1}} \tilde{p}(\Delta x_t, v_{t+\tau}|0, \mu_i), \quad (2.8)
$$

where u_i, $i = 1, \ldots, M+1$, are already known.

This leads to a good approximation because v_{max} cannot be too large since the initial volatility is usually smaller than 0.1. Also, the JTPDF is a continuous function of v_t and if M is big enough the error remains acceptable.

In summary, the algorithm for approximating the JTPDF is given in Algorithm 1 below.

Algorithm 1 Approximating $p\left(\Delta x_t, v_{t+\tau}|0, v_t\right)$ by Bilinear Interpolation

Given $v_t, v_{t+\tau}, x_t, x_{t+\tau}$, initialize $N_x, N_v, M, \Delta x_t$ and $[0, \mu_{max}]$.

Divide $[0, \mu_{max}]$ into M subintervals. Determine $v_t \in [\mu_{i-1}, \mu_i]$.

Generate look-up tables according to μ_{i-1} and μ_i respectively.

Compute $\tilde{p}\left(\Delta x_t, v_{t+\tau}|0, \mu_{i-1}\right)$ and $\tilde{p}\left(\Delta x_t, v_{t+\tau}|0, \mu_{i-1}\right)$ by bilinear interpolation.

Evaluate the approximation $\tilde{p}$ according to (2.8).

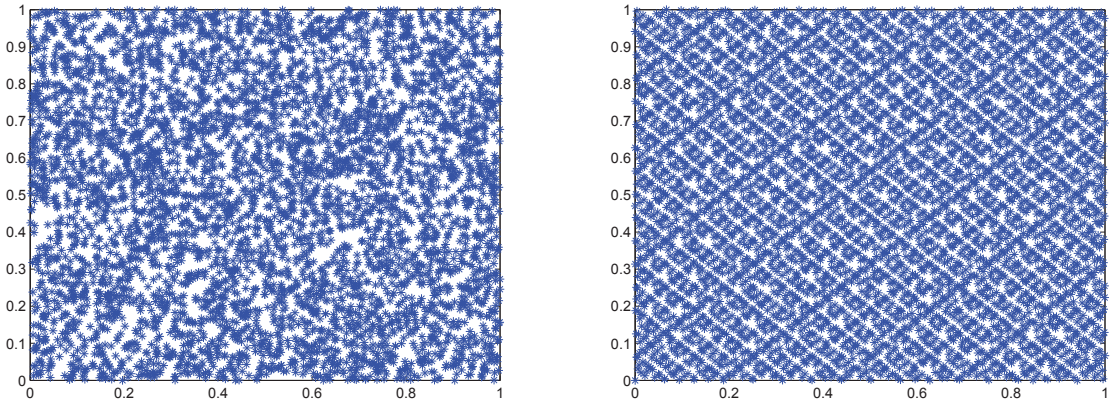
2.5 Generate $\tilde{p}$ Samples using Quasi-Monte Carlo Method

As we said before, we need samples from $\tilde{p}$ to compute (2.6). Here, samples from $\tilde{p}$ are obtained by a Quasi-Monte Carlo method.

2.5.1 Quasi-Monte Carlo Method

The Quasi-Monte Carlo (QMC) method is a modification of the original Monte-Carlo (MC) method. It is well documented in solving multiple integrals and details are given in Schürer (2003), Paskov and Traub (1995) and Joy et al. (1996). Here we use QMC to generate samples from $\tilde{p}$.

In practice, the randomnes in the MC method is always generated by pseudo-random numbers, while QMC uses low-discrepancy sequences. A low-discrepancy sequence is a set of well-chosen deterministic points that are defined from some results in number theory. Although it is a deterministic set, Figure 5 shows that the low-discrepancy sequence fills the space more efficiently than the pseudo-random numbers. There are different kinds of low-discrepancy sequences, some popular ones are Halton's, Faure's, and Sobol's sequences. Krykova (2003) compares them and suggests that Sobol's sequence works best. Here, we use Sobol's sequence in all our Quasi Monte Carlo methods. For details of low-discrepancy sequences, please refer to Krykova (2003).



(a) pseudo-random numbers

(b) Sobol Sequence

Figure 5: Pseudo-random numbers vs. Sobol's sequence, 5000-by-5000

2.5.2 Generating Samples From $\tilde{p}$

Note that

$$\tilde{p}\left(\Delta_t, v_{t+\tau}|0, v_t\right) = \tilde{p}\left(v_{t+\tau}|0, v_t\right)\tilde{p}\left(\Delta x_t|0, v_t, v_{t+\tau}\right),$$

where we can generate Δx_t and $v_{t+\tau}$ individually. Since $\tilde{p}\left(\Delta x_t, v_{t+\tau}|0, v_t\right)$ is an approximation of a transition probability density function the integral of $\tilde{p}\left(\Delta x_t, v_{t+\tau}|0, v_t\right)$ over its whole support (or space) is close to 1. Hence, before generating samples we need to normalize $\tilde{p}$ by setting

$$\tilde{p}_n\left(\Delta x_t, v_{t+\tau}|0, v_t\right) = \frac{\tilde{p}\left(\Delta x_t, v_{t+\tau}|0, v_t\right)}{\iint\limits_{S_{xv}} \tilde{p}\left(\Delta x_t, v_{t+\tau}|0, v_t\right)\, dx_{t+\tau}dv_{t+\tau}},$$

where $S_{xv} = [x_{min}, x_{max}] \times [v_{min}, v_{max}]$ is the support for $\tilde{p}\left(\Delta x_t, v_{t+\tau}|0, v_t\right)$. In fact, we are generating samples from the normalized density $\tilde{p}_n$.

First, generate $v_{t+\tau}$ from $\tilde{p}_n\left(v_{t+\tau}|0, v_t\right)$, which is obtained by first integrating Δx_t. Using the inverse transform algorithm, given u_1 is a sample from a uniform distribution on $[0, 1]$, the sample of $v_{t+\tau}$ is obtained by solving

$$\int_{v_{min}}^{v_{t+\tau}} \tilde{p}_n\left(v_{t+\tau}|0, v_t\right)\, dv_{t+\tau} = u_1, \tag{2.9}$$

where v_{min} stands for the lower bound of the support of $v_{t+\tau}$.

Note that the sample of $v_{t+\tau}$ has been generated from (2.9) and the following equation holds

$$\tilde{p}\left(\Delta x_t|0, v_t, v_{t+\tau}\right) = \frac{\tilde{p}\left(\Delta x_t, v_{t+\tau}|0, v_t\right)}{\tilde{p}\left(v_{t+\tau}|0, v_t\right)},$$

then the sample of Δx_t can be generated by solving the equation

$$\int_{\delta_{min}}^{\Delta x_t} \tilde{p}_n\left(\Delta x_t|0, v_t, v_{t+\tau}\right) d\Delta x_t = \frac{\int_{\delta_{min}}^{\Delta x_t} \tilde{p}_n\left(\Delta x_t, v_{t+\tau}|0, v_t\right) d\Delta x_t}{\tilde{p}_n\left(v_{t+\tau}|0, v_t\right)} = u_2, \tag{2.10}$$

where u_2 is another sample that follows a uniform distribution on $[0, 1]$. In summary, the samples of $(v_{t+\tau}, \Delta x_t)$ from $p\left(\Delta x_t, v_{t+\tau}|0, v_t\right)$ are generated by (2.9) and (2.10). Then it is trivial to generate samples of $(x_{t+\tau}, v_{t+\tau})$ from $\tilde{p}\left(x_{t+\tau}, v_{t+\tau}|x_t, v_t\right)$.

In the end, the price of the European path-dependent derivative is evaluated using the JTPDF approach by the following algorithm.

Algorithm 2 Evaluate $E^{\mathbb{Q}}\left[e^{-\int_0^T r_w\, dw} C_{\text{path}}\left(\mathbf{S}\left(\mathbf{x}\right), n\right) \middle| x_0, v_0\right]$

Generate look-up tables.

for $i = 1 \to N_{QMC}$ **do**

for $t = 1 \to n$ **do**

Generate $u_1^{(i)}$ and $u_2^{(i)}$ from Sobol's sequence.

Generate sample of $v_t^{(i)}$ based on $x_{t-1}^{(i)}, v_{t-1}^{(i)}$ according to (2.9).

Generate sample of $x_t^{(i)}$ based on $x_{t-1}^{(i)}, v_{t-1}^{(i)}$ and $v_t^{(i)}$ according to (2.10).

Compute $p\left(x_t^{(i)}, v_t^{(i)} \middle| x_{t-1}^{(i)}, v_{t-1}^{(i)}\right)$ and $\tilde{p}\left(x_t^{(i)}, v_t^{(i)} \middle| x_{t-1}^{(i)}, v_{t-1}^{(i)}\right)$.

end for

$p\left(\mathbf{x}_i, \mathbf{v}_i \middle| x_0, v_0\right) = \prod_{t=1}^n p\left(x_t^{(i)}, v_t^{(i)} \middle| x_{t-1}^{(i)}, v_{t-1}^{(i)}\right)$

$\tilde{p}\left(\mathbf{x}_i, \mathbf{v}_i \middle| x_0, v_0\right) = \prod_{t=1}^n \tilde{p}\left(x_t^{(i)}, v_t^{(i)} \middle| x_{t-1}^{(i)}, v_{t-1}^{(i)}\right)$.

Compute $C_{\text{path}}\left(\mathbf{S}\left(\mathbf{x}_i\right), n\right) \frac{p(\mathbf{x}_i, \mathbf{v}_i|x_0, v_0)}{\tilde{p}(\mathbf{x}_i, \mathbf{v}_i|x_0, v_0)}$.

end for

Compute $e^{-\int_0^n r(w)\, dw} \frac{1}{N_{QMC}} \sum_{i=1}^{N_{QMC}} C_{\text{path}}\left(\mathbf{S}\left(\mathbf{x}_i\right), n\right) \frac{p(\mathbf{x}_i, \mathbf{v}_i|x_0, v_0)}{\tilde{p}(\mathbf{x}_i, \mathbf{v}_i|x_0, v_0)}$.

---

3 The Conditional Probability Density Function of Asset Prices

This chapter introduces another method to price derivatives in Heston framework.

3.1 Pricing Derivatives via Conditional Expectation

According to Proposition 1.3, given the payoff $C(\mathbf{S}, T)$ at maturity T the price of the derivative at time t is the expectation of the discounted payoff under $\mathbb{Q}$. Conditioning on the path of the stochastic variance, that is $\mathbf{v}^* = \{v_s | s \in [t, T]\}$, the price can be rewritten as

$$\begin{aligned} \Pi &= E^{\mathbb{Q}}\left[e^{\int_t^T r_s\, ds} C(\mathbf{S}, n) \middle| \mathcal{F}_t\right] \\ &= E^{\mathbb{Q}}\left[E^{\mathbb{Q}}\left[e^{\int_t^T r_s\, ds} C(\mathbf{S}, n) \middle| \mathbf{v}^*\right] \middle| \mathcal{F}_t\right]. \end{aligned} \tag{3.1}$$

In Hull and White (1987), the paper in which the stochastic volatility model was introduced, they show that the inner conditional expectation in (3.1) is similar to the case under Black-Scholes' assumption. Thus, what matters is the outer expectation. In Broadie and Kaya (2004) this idea was also used to derived the statistical estimators for the Greeks.

In the Black-Scholes framework, the inner expectation in (3.1) varies a lot due to different payoffs, but many of them can be evaluated analytically. In terms of the outer expectation, it is estimated by simulating the path of the stochastic variance.

3.2 Conditional Density of the Asset Price

In order to evaluate (3.1), we need to know the conditional distribution of S_t. Note that $x_t = \ln S_t$, given (1.4). Then applying Itô's formula to x_t leads to

$$
\begin{aligned}
dx_t &= d\ln S_t \\
&= \frac{\partial x_t}{\partial S_t} dS_t + \frac{1}{2}\frac{\partial^2 x_t}{\partial S_t^2} d\langle S., S.\rangle + \frac{\partial x_t}{\partial t} dt \\
&= \frac{1}{S_t} dS_t - \frac{1}{2}\left(\rho^2 v_t + \left(1-\rho^2\right) v_t\right) dt \\
&= \left(r_t - \frac{1}{2} v_t\right) dt + \rho\sqrt{v_t} dW_v^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}\sqrt{v_t} dZ^{\mathbb{Q}}(t),
\end{aligned}
$$

where $Z^{\mathbb{Q}}(t)$ is a standard Brownian motion under $\mathbb{Q}$ which is independent of $W_v^{\mathbb{Q}}(t)$. Hence,

$$
\begin{aligned}
S_T &= S_t \exp\left\{\int_t^T r_s\, ds - \frac{1}{2}\int_t^T v_s\, ds + \rho\int_t^T \sqrt{v_s}\, dW_v^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}\int_t^T \sqrt{v_t} dZ^{\mathbb{Q}}(t)\right\} \\
&= S_t \exp\left\{\int_t^T r_s\, ds - \frac{1}{2}\int_t^T v_s\, ds + \right. \\
&\qquad \left. \rho\int_t^T \sqrt{v_s}\frac{dv_s - \kappa^*\left(\theta^* - v_s\right) ds}{\sigma\sqrt{v_s}} + \sqrt{1-\rho^2}\int_t^T \sqrt{v_t} dZ^{\mathbb{Q}}(t)\right\} \\
&= S_t \exp\left\{\int_t^T r_s\, ds + \frac{\rho}{\sigma}\left[v_T - v_t - \kappa^*\theta^*\left(T-t\right)\right] + \right. \\
&\qquad \left. \left(\frac{\rho\kappa^*}{\sigma} - \frac{1}{2}\right)\int_t^T v_s\, ds + \sqrt{1-\rho^2}\int_t^T \sqrt{v_t} dZ^{\mathbb{Q}}(t)\right\}.
\end{aligned}
$$

Note that $Z^{\mathbb{Q}}(t)$ and v_t are independent because of the independence between $Z^{\mathbb{Q}}(t)$ and $W_v^{\mathbb{Q}}(t)$. According to Andersen (2007), $\int_t^T \sqrt{v_t} dZ^{\mathbb{Q}}(t)$ follows a normal distribution with mean 0 and variance $\int_t^T v_s$, that is $\mathcal{N}\left(0, \int_t^T v_s\, ds\right)$.

Here $\mathbf{v}^*$ contains the information about v_t, v_T and $\int_t^T v_s\, ds$. Hence, conditional on the path of the stochastic volatility, the density of the asset's log-return under $\mathbb{Q}$ is given by

$$
\ln\left(\frac{S_T}{S_t}\right)\Big|_{\mathbf{v}^*} \sim \mathcal{N}\left(\mu_{\mathbb{Q}}, \sigma_{\mathbb{Q}}^2\right), \tag{3.2}
$$

where

$$
\begin{aligned}
\mu_{\mathbb{Q}} &= \int_t^T r_s\, ds + \frac{\rho}{\sigma}\left[v_T - v_t - \kappa^*\theta^*\left(T-t\right)\right] + \left(\frac{\rho\kappa^*}{\sigma} - \frac{1}{2}\right)\int_t^T v_s\, ds, \\
\sigma_{\mathbb{Q}}^2 &= \left(1-\rho^2\right)\int_t^T v_s ds.
\end{aligned}
$$

Similarly using (1.8), conditioning on the path of the stochastic variance, the density of the asset's log-return under $\mathbb{Q}_S$ is given by

$$\ln\left(\frac{S_T}{S_t}\right)\Big|_{\mathbf{v}^*} \sim \mathcal{N}\left(\mu_{\mathbb{Q}_S}, \sigma^2_{\mathbb{Q}_S}\right), \tag{3.3}$$

where

$$\begin{aligned}
\mu_{\mathbb{Q}_S} &= \int_t^T r_s\, ds + \frac{\rho}{\sigma}\left[v_T - v_t - \kappa^*\theta^*\left(T-t\right)\right] + \left(\frac{\rho\kappa^*}{\sigma} + \frac{1}{2}\right)\int_t^T v_s\, ds \\
&= \mu_{\mathbb{Q}} + \int_t^T v_s\, ds, \\
\sigma^2_{\mathbb{Q}_S} &= \left(1-\rho^2\right)\int_t^T v_s ds = \sigma^2_{\mathbb{Q}}.
\end{aligned}$$

3.3 Simulate the Path for the Stochastic Variance and the Integrated Variance

By Proposition 1.1, v_T, given v_t, follows a noncentral chi-squared distribution, thus the path of the stochastic variance can be simulated according to (1.2). Hence, the method of pricing via conditional expectation boils down to the question of simulating $\int_t^T v_s\, ds$.

Since $\mathbf{v}^*$ is a random process, $\int_t^T v_s\, ds$ is a random variable. Dufresne (2001) derived the moment generating function (MGF) of the integrated variance and discussed the properties of the MGF. Broadie and Kaya (2006) discussed how to simulate the integral of v_t exactly, they first obtained the characteristic function of the integrated variance and applied (1.15) to find the CDF of $\int_t^T v_s\, ds$. Then samples of the integrated variance can be generated by the inverse transform method in simulation. This method provides the exact simulation of the integrated variance but is plagued by its complex computations.

Glasserman and Kim (2008) point out that the integrated variance can be represented explicitly by a gamma expansion and their method is much faster than that of Broadie and Kaya (2006). Though both, the exact simulation and the gamma expansion, describe the exact probabilistic properties of the integrated variance, it is impossible to use them because each of them includes infinite summations. However, each truncated

error can be controlled. In order to avoid complex computations, Tse and Wan (2010) show that when $T-t$ goes to infinity $\int_t^T v_s\,ds$ converges to a inverse Gaussian distribution and they suggest using an inverse Gaussian to approximate the integrated variance. Although it is not as accurate as the exact simulation or the gamma expansion, according to Tse and Wan (2010) it is much faster and the accuracy is acceptable. Instead of using an inverse Gaussian, Bégin et al. (2012) showed that the gamma distribution is more efficient in approximating the integrated variance. Since our project is much more complicated than pricing European call options, it is necessary to sacrifice some accuracy in exchange of a faster processing time. Hence, we use a gamma distribution to approximate the integrated variance.

To approximate $\int_t^T v_s\,ds$ we let

$$\int_t^T v_s\,ds = \sum_{j=1}^{m} \int_{t_{j-1}}^{t_j} v_s\,ds, \tag{3.4}$$

where $t = t_0 < t_1 < \cdots < t_{m-1} = t_m = T$. Since the first and second moments of each $\int_{t_{j-1}}^{t_j} v_s\,ds$ can be obtained, given $\mathbf{v}^*$, by matching the first and second moments each integral on the right hand side to that of the approximating by a gamma distribution. Then the samples of the integral can be simulated.

3.4 Comparison of the JTPDF and CE Methods

Results for the JTPDF and CE methods are obtained through numerical and statistical methods. In the JTPDF method, importance sampling is used to evaluate a multiple integral up to $2n$ (n is the maturity) dimensions. Applying the CE method, the path of the stochastic variance and the integrated variance are simulated.

The JTPDF method requires an appropriate method to evaluate the multiple integrals and it is difficult to find one. The primary errors stem from (i) quadrature to compute the modified Bessel function, (ii) applying Filon's quadrature to evaluate the oscillatory integral, (iii) approximating the JTPDF by bilinear functions, (iv) applying Quasi-Monte Carlo to evaluate the expectation. However, except for the last one, the other numerical

errors can be controlled. The JTPDF method can be used to price any European option, even when it takes a long time to compute the result.

In terms of processing time, the CE method is faster since most of the work is derived analytically except for the outer expectation. But one prerequisite to use the CE is that the inner expectation can be priced explicitly. Otherwise, CE is exactly the same scheme as in Broadie and Kaya (2006). Errors of this method are (i) the discretization errors in simulating the path of the volatility. (ii) the error in approximating the integrated variance. (iii) the truncated error while applying the law of large numbers to evaluate the expectation.

4 Application to Ratchet EIA

4.1 Investment Guarantees

The objective of life insurance policies is to provide financial protection to policyholders and their beneficiaries in the case of death. Traditionally, a level premium is paid to purchase a whole life insurance policy which provides a fixed death benefit. The level premium means that policyholders pay the same amount at each contract anniversary. The fixed death benefit is the amount of money that is paid to beneficiaries when the insured dies.

As the financial markets changes, the policyholders demand extra investment opportunities outside the insurance sector. In 1970s, the high interest rate environment resulted in the introduction of the universal life insurance policy (UL). The difference between UL and traditional life insurance is that the premiums are deposited in an account which earns a return linked to the performance of the insurance company. Typically, a guaranteed return between 0% to 4% is provided and it could be higher if the company receives a higher return from the market. Each month spreads such as management fees are directly withdrew from the account. If the value of the account reaches zero, the account is closed. In general, in the case that the insured dies a specified benefit is paid plus whatever is left in the account. According to Klugman et al. (2012) UL represents around 40% of the market, which is twice the market of traditional life insurance. The rest consists of equity-linked products and term insurance.

Variable annuity(VA), which is more of an long-term investment product, is similar to UL. The primary difference is that the policyholder determines how the account is invested and how the return of the account is credited. VAs contracts are flexible, the account linked to VAs can be invested in bonds, a stock index, or commodities. The VA is one of the most important life insurance products because it enables policyholders to participate in financial markets. Suppose that the account is invested into the stock market, if the return on the stock is positive the account receives positive return while

the account loses money in the case of a stock market collapse. Guarantees of the account can be obtained by purchasing extra riders on the contract, such as guaranteed minimum income benefit (GMIB) and guaranteed minimum death benefit (GMDB).

Another kind of equity-linked products is called Equity-indexed annuity (EIA), which allows policyholders to invest in a financial index, typically the S&P 500, while the return of the linked account is guaranteed.

4.2 Introduction to EIAs

Here we focus on equity-indexed annuities (EIA). The EIA is one of the most innovative life insurance products and its market shares increase rapidly since it was introduced by Keypord Life in 1995. Generally speaking, the policyholder pays an initial premium to purchase a unit of EIA. The premium is deposited into an account whose return is credited according to the performance of an external index, typically S&P 500. A minimal guarantee, or a minimal interest rate, is provided in EIA contracts to protect the policyholders against possible losses. The account earns a portion of the index's return if it is larger than the minimal guarantee. Otherwise, the minimal guarantee is credited. Thus, even though the stock market collapses the account still earns a non-negative return. Hence, the EIA is attractive to people who seek market appreciations with downside protections. In practice, the upside return of EIAs is always "limited" by the participation rate, cap or spread.

- The participation rate is the portion at which the external index return is credited to the EIA's return. For example, if the participation rate is 90% and the return of the external index is 10%, the EIA's return can be as much as 9%.

- The cap is the maximum of the return for a specified period which could be smaller than the portion of the index's return. For instance, if the cap is 8% in the previous example the return is at most 8% rather than 9%, which is 90% of the index' return.

- The spread is a reduction from the fund's interest rate such as the management fee, etc.

Most EIAs have the characteristics we discussed before, but different designs lead to different EIAs. Here, we focus on the ratchet EIA which is the most popular design. The return from the ratchet EIA is reset annually, and according to annuityadvantage.com EIAs with reset designs represent 85% of the current market.

In each year, the ratchet EIA's return is linked to the performance of the external index. It is as least equal to the minimal guarantee, and its upside return is limited by the cap, spread, or participation rate. The value of the account of n-year ratchet EIA is given by

$$C_{Rat}(\mathbf{S}, n) = \prod_{t=1}^{n} \max\left\{\min\left[e^{\alpha(Y_t-\gamma)}, e^{\zeta}\right], e^{g}\right\}, \tag{4.1}$$

where α stands for the participation rate, γ is the annual yield spread, ζ is the cap, g stands for the minimal guarantee, $\mathbf{S} = \{S_t\}_{0\le t\le n}$, $Y_t = \ln\frac{S_t^*}{S_{t-1}}$ and S_t^* takes different forms according to the EIA's design. In practice S_t^* could be defined as

$$S_t^* = \begin{cases} S_t, & \text{term-end point design} \\ \frac{1}{12}\sum_{k=0}^{11} S_{t-\frac{k}{12}}, & \text{Asian-end design} \\ \max\limits_{1\le k\le 12t} S_{\frac{k}{12}}. & \text{high-water-mark design} \end{cases} \tag{4.2}$$

Note that in the term-end point design, Y_t only depends on S_t and S_0. The other two designs are path-dependent. The Asian-end design is less volatile so that it leads to a less expensive price. The high-water-mark design credits the appreciation at the highest index price during a whole period so that this kind of appreciation leads to an expensive product. According to annuityadvantage.com the term-end point design is the most popular one. Tiong (2000) priced it analytically in a Black-Scholes model, so that we focus on the first design.

There is another representation for the value of the fund at the end of year n, that is

$$\tilde{C}_{Rat}(\mathbf{S}, n) = \max\left\{\prod_{t=1}^{n} \max\left[\min\left(1+\alpha R_t - \tilde{\gamma}, 1+\zeta\right), 1\right], \beta\left(1+\tilde{g}\right)^n\right\}, \tag{4.3}$$

where R_t is defined as $\frac{S_t^*}{S_{t-1}} - 1$, $\tilde{\gamma}$ is the spread, β is the portion at which the minimal guarantee is credited, and $\tilde{g}$ is the minimal guarantee. The rates $\tilde{\gamma}$ and $\tilde{g}$ could be different with the spread and minimal guarantee in (4.1).

The returns produced by (4.1) and (4.3) are similar. The primary difference between them is that in the former the guarantee is compared annually while in the latter the guarantee is compared at the end of the period. The payoff expressed by (4.3) is closer to what most companies use, while (4.1) is a simplification of (4.3) and according to Tiong (2000) and Qian et al. (2010) it can be priced analytically in a Black-Scholes model. Hence, from now on we use (4.1), in which $Y_t = \frac{S_t}{S_{t-1}}$, as the ratchet EIA's survival benefit.

4.3 Pricing EIAs

Due to their complex payoffs, it is challenging to price EIAs. The risk involved in EIAs can be decomposed into the financial and mortality risks. For the financial risk it is important to find the present value of the account linked to EIAs. In terms of the mortality risk, we need to determine when the benefit is claimed.

Actually, in this thesis we do not consider the mortality risk but focus on the survival benefit. This is because mortality risk can be reduced by using a well diversified homogeneous portfolio. Under this diversification assumption, EIA prices become linear combinations of prices with different maturities. Hence, we focus on the financial risk that is the most important risk involved in EIAs.

4.3.1 Pricing Ratchet EIAs via the TPDF Method.

Denote the beginning of the contract as time 0 and the maturity as time n, let i represent the integers between 0 and n, at time t where $0 \leq i-1 \leq t < i \leq n$, the survival benefit is given by (4.1). Using (2.5), the price at time t can be written as

$$
\begin{aligned}
& \Pi_{Rat}\left(t, n, x_t, v_t\right) \\
= \; & e^{-\int_t^n r_w\, dw} \int_{(\mathbf{x},\mathbf{v})\in\Omega_{\mathbf{xv}}} \cdots \int C_{Rat}\left(\mathbf{S}\left(\mathbf{x}\right), n\right) p\left(\mathbf{x}, \mathbf{v} | x_t, v_t\right)\, d\mathbf{x}d\mathbf{v} \\
= \; & e^{-\int_t^n r_w\, dw} \int_{(\mathbf{x},\mathbf{v})\in\Omega_{\mathbf{xv}}} \cdots \int C_{Rat}\left(\mathbf{S}\left(\mathbf{x}\right), n\right) p\left(x_i, v_i | x_t, v_t\right) \prod_{j=i+1}^{n} p\left(x_j, v_j | x_{j-1}, v_{j-1}\right)\, d\mathbf{x}d\mathbf{v},
\end{aligned}
\tag{4.4}
$$

where $\mathbf{S}\left(\mathbf{x}\right) = \left(S_i\left(x_i\right), \ldots, S_n\left(x_n\right)\right)$, $\mathbf{x} = \left(x_i, x_{i+1}, \ldots, x_n\right)$, $\mathbf{v} = \left(v_i, v_{i+1}, \ldots, v_n\right)$, $\Omega_{\mathbf{xv}} = \left(-\infty, +\infty\right)^{n+1-i} \times \left(0, +\infty\right)^{n+1-i}$. Here $p\left(x_{t+\tau}, v_{t+\tau} | x_t, v_t\right)$ is the JTPDF given in (2.1).

Given (4.4), the price can be found according to Algorithm 2.

4.3.2 Pricing Ratchet EIAs via the CE Method

If a contingent claim can be priced analytically in the Black-Scholes framework, it is easy to price the same derivative in the stochastic volatility model using the CE approach. In terms of the ratchet EIA, Tiong (2000) derived its analytical price in a Black-Scholes model using Esscher's transform and proved the following result:

Proposition 4.1 *1. In Black-Scholes model, there exists a unique risk-neutral Esscher parameter h^* such that under the Esscher transform with parameter h^* the real probability measure is changed to the risk-neutral measure $\mathbb{Q}$.*

2. Suppose that Y follows $\mathcal{N}(\mu, \sigma^2)$, under the Esscher transform with parameter h

$$Y \sim \mathcal{N}(\mu + h\sigma^2, \sigma^2);$$

Under the Esscher's transform with parameter $h + \alpha$,

$$Y \sim \mathcal{N}(\mu + h\sigma^2 + \alpha\sigma^2, \sigma^2).$$

Proof. See Tiong (2000). □

Conditioning on the variance path and the integral of the volatility, the density of the log-return in $\mathbb{Q}$ is given by (3.2). Since in a Black-Scholes framework $\mathbb{Q}$ is unique, (3.2) must

represent the density of the log-return under Esscher's transform with the risk-neutral parameter h^*. By Proposition 4.1 under Esscher transform with parameter $h^*+\alpha$,

$$ln\left(\frac{S_T}{S_t}\right)\Big|_{h^*+\alpha} \sim \mathcal{N}\left(\mu_{\mathbb{Q}}+\alpha\sigma_{\mathbb{Q}}^2, \sigma_{\mathbb{Q}}^2\right). \tag{4.5}$$

Proposition 4.2 *In Black-Scholes framework with deterministic risk-free interest rate, the expectation of the discounted survival benefit (4.1) under $\mathbb{Q}$ is given by*

$$\prod_{t=1}^{n}\left[e^{-\int_{t-1}^{t} r_w\, dw+g} P\left\{Y_t \le \frac{g}{\alpha}+\gamma; h^*\right\} + e^{\zeta-\int_{t-1}^{t} r_w\, dw} P\left\{Y_t > \frac{\zeta}{\alpha}+\gamma; h^*\right\} + e^{-(1-\alpha)\int_{t-1}^{t} r_w\, dw+\frac{1}{2}\alpha(\alpha-1)\sigma_{\mathbb{Q}}^2} P\left\{\frac{g}{\alpha}+\gamma < Y_t \le \frac{\zeta}{\alpha}+\gamma; h^*+\alpha\right\}\right], \tag{4.6}$$

where under the Esscher's transform with parameter h^*

$$Y_t\big|_{h^*} \sim \mathcal{N}(\mu_{\mathbb{Q}}, \sigma_{\mathbb{Q}}^2),$$

and under the Esscher's transform with parameter $h^*+\alpha$

$$Y_t\big|_{h^*+\alpha} \sim \mathcal{N}(\mu_{\mathbb{Q}}+\alpha\sigma_{\mathbb{Q}}^2, \sigma_{\mathbb{Q}}^2).$$

Proof. See Tiong (2000). □

Before showing how to apply the CE method to price the ratchet design in Heston's framework, we rewrite the survival benefit (4.1) as follows:

$$\begin{aligned} & C_{Rat}(\mathbf{S}, n) \\ \triangleq\ & \prod_{j=1}^{i-1} \max\left\{\min\left[e^{\alpha(Y_j-\gamma)}, e^{\zeta}\right], e^{g}\right\} \prod_{j=i+1}^{n} \max\left\{\min\left[e^{\alpha(Y_j-\gamma)}, e^{\zeta}\right], e^{g}\right\} \times \\ & \left(\frac{S_t}{S_{i-1}}\right)^{\alpha} e^{-\alpha\gamma} \max\left\{\min\left[e^{\alpha Y^*}, e^{\zeta^*}\right], e^{g^*}\right\}, \end{aligned} \tag{4.7}$$

where

$$\begin{aligned} Y^* &= \ln\left(\frac{S_i}{S_t}\right), \\ \zeta^* &= \zeta+\alpha\gamma-\alpha\ln\left(\frac{S_t}{S_{i-1}}\right), \end{aligned}$$

$$g^* \quad = \quad g + \alpha\gamma - \alpha \ln\left(\frac{S_t}{S_{i-1}}\right).$$

Insert (4.7) into (3.1), the price of the ratchet EIA at time t, where $0 \leq i-1 \leq t < i \leq n$, is given by

$$\begin{aligned} & \Pi_{Rat}\left(t, n, x_t, v_t\right) \\ = \; & E^{\mathbb{Q}}\left[E^{\mathbb{Q}}\left[e^{-\int_t^n r_w\, dw} C_{Rat}\left(\mathbf{S}, n\right) \middle| \{v_s\}_{s\in[0,n]}\right] \middle| \mathcal{F}_t\right] \\ = \; & \left(\frac{S_t}{S_{i-1}}\right)^{\alpha} e^{-\alpha\gamma} \prod_{j=1}^{i-1} \max\left\{\min\left[e^{\alpha(Y_j-\gamma)}, e^{\zeta}\right], e^{g}\right\} E^{\mathbb{Q}}\left[\Pi_{\text{inner}} \middle| \mathcal{F}_t\right], \end{aligned} \tag{4.8}$$

where

$$\begin{aligned} \Pi_{\text{inner}} \quad = \quad & E^{\mathbb{Q}}\left[\max\left\{\min\left[e^{\alpha Y^*}, e^{\zeta^*}\right], e^{g^*}\right\} \prod_{j=i+1}^{n} \max\left\{\min\left[e^{\alpha(Y_j-\gamma)}, e^{\zeta}\right], e^{g}\right\} \middle| \mathbf{v}^*\right] \\ = \quad & E^{\mathbb{Q}}\left[\max\left\{\min\left[e^{\alpha Y^*}, e^{\zeta^*}\right], e^{g^*}\right\} \middle| \{v_s\}_{s\in[0,n]}\right] \times \\ & E^{\mathbb{Q}}\left[\prod_{j=i+1}^{n} \max\left\{\min\left[e^{\alpha(Y_j-\gamma)}, e^{\zeta}\right], e^{g}\right\} \middle| \mathbf{v}^*\right]. \end{aligned} \tag{4.9}$$

Applying Proposition 4.2 to (4.9) leads to

$$\begin{aligned} \Pi_{\text{inner}} \quad = \quad & \left[e^{-\int_t^i r_w\, dw + g} P\left\{Y^* \leq \frac{g^*}{\alpha}; h^*\right\} + e^{\zeta^* - \int_t^i r_w\, dw} P\left\{Y^* > \frac{\zeta^*}{\alpha}; h^*\right\} + \right. \\ & \left. e^{-(1-\alpha)\int_t^i r_w\, dw + \frac{1}{2}\alpha(\alpha-1)\sigma_{\mathbb{Q}}^2} P\left\{\frac{g^*}{\alpha} < Y^* \leq \frac{\zeta^*}{\alpha}; h^* + \alpha\right\}\right] \times \\ & \prod_{j=i+1}^{n} \left[e^{-\int_{j-1}^j r_w\, dw + g} P\left\{Y_j \leq \frac{g}{\alpha} + \gamma; h^*\right\} + e^{\zeta - \int_{j-1}^j r_w\, dw} P\left\{Y_j > \frac{\zeta}{\alpha} + \gamma; h^*\right\} + \right. \\ & \left. e^{-(1-\alpha)\int_{j-1}^j r_w\, dw + \frac{1}{2}\alpha(\alpha-1)\sigma_{\mathbb{Q}}^2} P\left\{\frac{g}{\alpha} + \gamma < Y_j \leq \frac{\zeta}{\alpha} + \gamma; h^* + \alpha\right\}\right]. \end{aligned}$$

In the end, the outer expectation in (4.8) is evaluated by simulating

$$\{v_s\}_{s\in[t,\ldots,n]} \quad \text{and} \quad \left\{\int_{t_{min}}^{t_{max}} v_s\right\}_{t\leq t_{min}\leq t_{max}\leq n}, \tag{4.10}$$

using the method introduced in Chapter 3.

4.4 Greeks of the Ratchet EIA

After obtaining the price of ratchet EIAs, it is natural to think about deriving its Greeks. Greeks are defined as the sensitivities of an option's price to its parameters. They are usually used to set up a replicating portfolio in order to reduce the financial risk. The Greeks that are most used are the Delta, Gamma, Vega[8], and Rho. Denote by Π_t the price of a derivative at time t, and S_t, v_t and r_t as in (1.4), then the Greeks mentioned above are defined for Π_t in the Table 5.

Table 5: Greeks

Greeks	Formula
Delta	$\Delta_t = \frac{\partial}{\partial S_t}\Pi_t$
Gamma	$\Gamma_t = \frac{\partial^2}{\partial S_t^2}\Pi_t$
Vega	$\Lambda_t = \frac{\partial}{\partial V_t}\Pi_t$
Rho	$\rho_t = \frac{\partial}{\partial r_t}\Pi_t$

The delta measures the sensitivity of the price to changes in the underlying asset price. In the replicating portfolio Delta represents the number of asset shares in order to reduce the risk induced by price movements in the underlying. Gamma is the second derivative of the price w.r.t. the underlying asset price, according to McDonald (2006) it is introduced since Delta hedging (replicating portfolio only consists of asset shares) fails when the change in asset price is large. Rho is used to measure the sensitivity to interest rate changes.

Vega measures the sensitivity of the price to the volatility of the underlying. It is undervalued in Black-Scholes models due to the assumption of constant volatility, but it becomes important in stochastic volatility models.

In terms of the Greeks for ratchet EIAs, it is easy to derive the Delta and Gamma

[8] "Vega" is not a Greek letter. Sometimes it is represented by "Kappa" or "Lambda". See McDonald (2006).

using (4.8), since the expression depends on S_t in the terms of $\left(\frac{S_t}{S_{i-1}}\right)^{\alpha}$. Hence, at time t $(0 \leq i-1 \leq t < i \leq n)$, we have

$$\begin{aligned}
\Delta_t &= \frac{\partial}{\partial S_t}\Pi_{Rat}\left(t, n, x_t, v_t\right) \qquad (4.11)\\
&= \frac{\alpha\left(S_t\right)^{\alpha-1}}{\left(S_{i-1}\right)^{\alpha}} e^{-\alpha\gamma} \prod_{j=1}^{i-1} \max\left\{\min\left[e^{\alpha(Y_j-\gamma)}, e^{\zeta}\right], e^{g}\right\} E^{\mathbb{Q}}\left[\Pi_{\text{inner}}\middle|\mathcal{F}_t\right],\\
\Gamma_t &= \frac{\partial^2}{\partial S_t^2}\Pi_{Rat}\left(t, n, x_t, v_t\right), \qquad (4.12)\\
&= \frac{\alpha\left(\alpha-1\right)\left(S_t\right)^{\alpha-2}}{\left(S_{i-1}\right)^{\alpha}} e^{-\alpha\gamma} \prod_{j=1}^{i-1} \max\left\{\min\left[e^{\alpha(Y_j-\gamma)}, e^{\zeta}\right], e^{g}\right\} E^{\mathbb{Q}}\left[\Pi_{\text{inner}}\middle|\mathcal{F}_t\right],
\end{aligned}$$

where Π_{inner} is given by (4.9). It is difficult to derive Λ_t, using the CE method, since the initial variance is required. However, Λ_t can be derived using the JTPDF method.

Recall that the JTPDF method evaluates the price by solving multiple integrals. If it is possible to interchange the differential operator with the integrals, the JTPDF method can be applied to evaluate the corresponding Greeks. Although it is hard to see whether $\frac{\partial}{\partial S_t}$ is interchangeable with the integrals, however, it is true for $\frac{\partial}{\partial v_t}$ (see a proof in Appendix E).

Note that we denote by V_t the stochastic volatility, in other words $v_t = V_t^2$. According to Appendix D, at time t $(0 \leq t < n)$

$$\begin{aligned}
\Lambda_t &= \frac{\partial \Pi}{\partial V_t} = \frac{\partial \Pi}{\partial v_t}\frac{dv_t}{dV_t}\\
&= 2\sqrt{v_t} e^{\int_t^n r(w)\,dw} E^{\mathbb{Q}}\left[C^*\left(\mathbf{X}, n\right)\frac{\frac{\partial}{\partial v_t}p\left(x_i, v_i|x_t, v_t\right)}{p\left(x_i, v_i|x_t, v_t\right)}\middle|\mathcal{F}_t\right]. \qquad (4.13)
\end{aligned}$$

In terms of the term $\frac{\partial}{\partial v_t}p\left(x_i, v_i|x_t, v_t\right)$, recall that in (2.1) only $\left(\frac{v_{t+\tau}}{v_t}\right)$ and $h\left(k\right)$ are functions of v_t, so that

$$\begin{aligned}
&\frac{\partial}{\partial v_t}p\left(x_{t+\tau}, v_{t+\tau}|x_t, v_t\right)\\
=\ &\frac{1}{\pi}e^{\frac{\nu+1}{2}\kappa^*\tau}\left\{\left[\frac{\partial}{\partial v_t}\left(\frac{v_{t+\tau}}{v_t}\right)^{\nu/2}\right]\int_0^{+\infty} e^{imk}h\left(k\right)\,dk + \left(\frac{v_{t+\tau}}{v_t}\right)^{\nu/2}\int_0^{+\infty}\frac{\partial}{\partial v_t}e^{imk}h\left(k\right)\,dk\right\}\\
=\ &\frac{1}{\pi}e^{\frac{\nu+1}{2}\kappa^*\tau}\left(\frac{v_{t+\tau}}{v_t}\right)^{\nu/2}\left\{-\frac{\nu}{2v_t}\int_0^{+\infty} e^{imk}h\left(k\right)\,dk + \int_0^{+\infty} e^{imk}\frac{\partial}{\partial v_t}h\left(k\right)\,dk\right\}, \qquad (4.14)
\end{aligned}$$

and

$$\frac{\partial}{\partial v_t} h(k) \tag{4.15}$$

$$= \psi e^{\frac{\tau d}{2} + (A_2 - \psi) v_t - (A_1 + \psi) v_{t+\tau}} \times \left\{ \frac{1}{2} \left(I_{\nu-1}\left[2\sqrt{\varphi v_t v_{t+\tau}}\right] + I_{\nu+1}\left[2\sqrt{\varphi v_t v_{t+\tau}}\right] \right) \sqrt{\frac{\varphi v_{t+\tau}}{v_t}} + I_\nu \left[2\sqrt{\varphi v_t v_{t+\tau}}\right] (A_2 - \psi) \right\}. \tag{4.16}$$

The previous equation holds since $\frac{\partial}{\partial x} I_\nu[x] = \frac{1}{2}\left(I_{\nu-1}[x] + I_{\nu+1}[x]\right)$, see Bowman (1958).

Given (4.14) and (4.16), Vega in (4.13) can be evaluated using Algorithm 2.

5 Numerical Results

This section shows numerical results from the implementation of the JTPDF and CE methods to evaluate prices and Greeks for ratchet EIAs. All the results obtained with the by JTPDF method use the following parameters

$$N_{QMC} = 10,000,\ N_x = 500,\ N_v = 300,\ M = 50,\ \varepsilon_v = 10^{-5},\ \alpha_x = 4.6031,\ \varepsilon_x = 10^{-6}.$$

In terms of the CE method, the number of samples in simulating the stochastic variance was $15,000$ and the time step 0.005. In (3.4), $t_j - t_{j-1} = 0.01$ for all j.

5.1 European Call Option

The analytical price of a European call option is given by (1.42), its explicit Greeks can be found in Reiss and Wystup (2001). These explicit formulas can be a benchmark to compare the performance of the JTPDF and CE methods.

Broadie and Kaya (2004) discussed the Greeks of the European call option, we adopt their parameters and use the JTPDF and CE methods to price an at-the-money European call option with strike 100. The prices are illustrated in Table 6, where the exact value is the number derived by the explicit formula. The results suggest that both the JTPDF and CE methods work well.

Table 6: European call prices

	Exact Value	JTPDF	CE
Price	6.8061	6.8097	6.7995

In carrying out the JTPDF method, we just simulate the price once since the variance of the Quasi-Monte Carlo (QMC) method is small. We also calculate the integral of

$$\iint\limits_{(x_1,v_1)\in[-\infty,+\infty]\times[0,+\infty]} p\left(x_1, v_1|x_0, v_0\right) \frac{\tilde{p}\left(x_1, v_1|x_0, v_0\right)}{p\left(x_1, v_1|x_0, v_0\right)}\, dx_1 dv_1.$$

The result is supposed to be 1, and our result is 0.999930. This number measures the error of the bilinear interpolation approximation and it suggests that the approximation error is negligible.

In carrying out the CE approach, we simulated the call price 30 times. The 5-th percentile is 6.7694, and the 95-th percentile is 6.8338. We also simulated the Greeks 30 times for the European call at time 0. The results are summarized in Table 7 where "Exact Value" stands for the number evaluated by explicit formulas.

Table 7: Applying the CE approach to European call options

	Exact Value	Mean	5% estimator	95% estimator
Δ_0^C	0.6958	0.6952	0.6918	0.7015
Γ_0^C	0.0265	0.0264	0.0263	0.0266

5.2 Ratchet EIAs

There is no closed-form expression for the prices of ratchet EIAs, but we can evaluate these using the JTPDF or CE methods. In order to check the accuracy of the two methods, we use the QE scheme introduced in Andersen (2007) as a benchmark. The QE method is recognized as a good scheme to simulate from the Heston model. In carrying out the QE method, we set the number of simulations to $20,000$ and the time step to 0.005.

5.2.1 Prices

We first price a 7-year ratchet EIA at time 0. Here, we use the Heston parameters in Table 4, and assume that at time 0 the T-bill yield rate follows the NSS model with parameters in Table 3. Given a minimal guarantee of 2%, a participation rate of 30%, a cap of 10% and a spread of 3%, Table 8 summaries the prices Π_0^{Rat}, evaluated individually by the QE, JTPDF, and CE methods. For the CE methods, we simulated the price 30 times, the mean is used as the price while the 5th percentile is 1.0864 and 95th percentile is 1.0872.

Table 8: Price of the 7-year ratchet EIA

	QE	JTPDF	CE
Π_0^{Rat}	1.1084	1.1094	1.0872

The JTPDF, CE, and QE methods are totally different approaches. However, in Table 8 the results are consistent with each other. Thus both the JTPDF and CE methods work well in dealing with the ratchet EIA. Although there is a deviation of around 0.02 for the CE method, it is still acceptable. The error of the CE approach is generated when approximating the integrated variance in (3.4).

5.2.2 Fair Values

In Table 8 prices of the ratchet EIA is above 1, this contract is unfair for companies since we assume that the investors just pays one dollar to obtain one unit of ratchet EIA. Fair contract parameters should be set at values at which $\Pi_0^{Rat} = 1$.

In this section, we focus on the fair cap and fair participation rate. In solving for the fair cap we first fix the participation rate at 100%, and then evaluated the fair cap using different spreads and minimal guarantees. In terms of fair participation rates, they are calculated without a cap, that is $\zeta = +\infty$, using different spreads and minimal guarantees. A bisection (or the binary search) method is applied here to solve for the fair values, and the tolerance is set to be 10^{-6}.

Table 9 illustrates the fair cap and participation rates evaluated using the QE scheme, and they are regarded as benchmarks to compare the fair values given by JTPDF and CE approaches. The left part of Table 9 summaries the fair caps with 100% participation rates and the right part illustrates the fair participation rates without a cap. Tables 10 and 11 present the errors of fair ζ and α values calculated by the JTPDF and CE methods respectively[9].

[9]The errors in Table 10 and 11 are defined as $\frac{\text{QE}-\text{Approximation}}{\text{QE}}$.

Table 9: Fair values (in percentage) given by the QE method

	ζ ($\alpha = 100\%$)			α ($\zeta = +\infty$)		
$\gamma \setminus g$	0.00%	0.30%	0.50%	0.00%	0.30%	0.50%
0.00%	2.91	2.64	2.46	19.70	17.21	16.66
0.50%	2.97	2.69	2.50	20.49	18.58	17.21
1.00%	3.04	2.74	2.54	21.26	19.24	17.79

Table 10: Errors given by the JTPDF approach

	Fair ζ ($\alpha = 100\%$)			Fair α ($\zeta = +\infty$)		
$\gamma \setminus g$	0.00%	0.30%	0.50%	0.00%	0.30%	0.50%
0.00%	-0.24%	-0.39%	-0.15%	0.34%	-3.92%	0.39%
0.50%	-0.20%	-0.07%	-0.12%	0.35%	0.39%	0.44%
1.00%	-0.39%	-0.40%	-0.43%	0.40%	0.44%	0.47%

Table 9 shows that when the participation rate is one, fair cap is an increasing function of the spread, and it decreases when the minimal guarantee gets larger. This is because a large spread reduces the EIA's gain from the external index, and a small minimal guarantee weakens the EIA's downside protection.

In this case if the cap keeps constant, a smaller potential return leads to the lower price. Thus a higher fair cap is expected to yield a price that is more stable.

In terms of the fair participation rate, it behaves the same as the fair cap when spreads and minimal guarantees change. Table 9 also suggests that fair values are more sensitive to changes in g. According to annuityadvantage.com, our fair caps are consistent with what companies are adopting. This also confirms our assumption that the financial risk is the primary risk involved in EIAs. Finally, even if our S_0 is not based on observed data, it does not influence the return.

Again, Table 10 shows that the results derived by the JTPDF method are close to

Table 11: Errors given by the CE approach

	Fair ζ ($\alpha = 100\%$)			Fair α ($\zeta = +\infty$)		
$\gamma \setminus g$	0.00%	0.30%	0.50%	0.00%	0.30%	0.50%
0.00%	-3.58%	-3.01%	-2.26%	-56.27%	-62.74%	-56.15%
0.50%	-2.64%	-2.77%	-1.33%	-49.63%	-50.23%	-50.31%
1.00%	-2.88%	-2.61%	-2.11%	-44.03%	-44.91%	-44.56%

those obtained by the QE scheme. The fair caps calculated in Table 11 are still acceptable, but the fair participation rates appear to suggest that the CE approach fails. However, although in the CE method we sacrifice accuracy for processing time, Table 8 shows that the result is still acceptable. What the errors tell us is that the fair participation rates are sensitive to the changes in prices.

5.2.3 Greeks

This section introduces the Greeks for the ratchet design. As we said before, the JTPDF method is good for evaluating Λ_t^{Rat}, while it is easy to derive Δ_t^{Rat} and Γ_t^{Rat} using the CE approach. All the Greeks are evaluated using fair contract parameters.

The Greeks of a 7-year ratchet design are evaluated at time 0. It is interesting that both Δ_0^{Rat} and Γ_0^{Rat} are zero no matter what the contract parameters are. This is because at time 0 the price in (4.8) is free from S_0. Zero Δ_0^{Rat} and Γ_0^{Rat} means that when the company sells the ratchet EIA policy, it is not necessary to hold any assets in the replicating portfolio at the beginning. In terms of Λ_0^{Rat}, they are evaluated by the JTPDF method using the parameters shown in Table 9, and the results are illustrated in Table 12.

The left part in Table 12 presents Λ_0^{Rat} with a 100% participation rate. Different spreads and minimal guarantees are presented in this table. The caps are the fair values obtained in Table 9. For instance, in Table 9 the fair cap corresponding to a 100% participation rate, 0 spread, and 0 minimal guarantee is 2.91%. Thus the left top Λ_0^{Rat} in

Table 12 is evaluated using $(\alpha, \gamma, \zeta, g) = (100\%, 0, 2.91\%, 0)$.

One interesting fact is that in Table 12 Λ_0^{Rat} values can be negative. It appears to violate common sense that larger volatilities lead to higher prices, but the negative Λ_0^{Rat} values result from the cap. A small cap limits the gains from the index, so that a larger volatility only contributes to a higher downside risk in which case the price should be lower. Our argument is verified by the positive Λ_0^{Rat} when $\zeta = +\infty$. Thus, the higher the cap, the larger Λ_0^{Rat}. Since Λ_0^{Rat} has different signs, holding a portfolio of ratchet EIAs can reduce the volatility risk. For instance, the Λ_0^{Rat} of a portfolio consists of one contract with $(\alpha, \gamma, \zeta, g) = (100\%, 0, 2.64\%, 0.30\%)$ and 41 contracts with $(\alpha, \gamma, \zeta, g) = (17.21\%, 0.50\%, +\infty, 0.50\%)$ is zero.

The results in Table 13 are used to investigate how the changes in α affect Λ_0^{Rat}. Those numbers are evaluated using $(\alpha, \gamma, \zeta) = (100\%, 0, +\infty)$ and different minimal guarantees. Comparing the numbers in Tables 13 and 12, we find that Λ_0^{Rat} is an increasing function of α. When $\zeta = +\infty$, larger participation rates leads to more volatile returns, in which case a change in volatility contributes to big changes in price.

All the effects of γ and g on Λ_0^{Rat} in the right part can be traced back to how they impact the behaviors of fair participation rates. When $\alpha = 100\%$ and for fix g, Λ_0^{Rat} is an increasing function of γ. This is because how γ affects the fair cap. When γ is fix, the higher the minimal guarantee, the lower the fair cap is. Thus the space for the gain from the index is narrow, which leads to the reduction in the absolute values in Λ_0^{Rat}.

Table 12: Λ_0^{Rat} of ratchet EIAs

	Fair ζ ($\alpha = 100\%$)			Fair α ($\zeta = +\infty$)		
$\gamma \setminus g$	0.00%	0.30%	0.50%	0.00%	0.30%	0.50%
0.00%	-0.0128	-0.0123	-0.0121	0.0008	-0.0004	-0.0006
0.50%	-0.0124	-0.0121	-0.0119	0.0014	0.0005	-0.0003
1.00%	-0.0121	-0.0118	-0.0116	0.0019	0.0009	0.0001

Table 13: Λ_0^{Rat} of ratchet EIA (II)

$\alpha = 100\%,\ \zeta = +\infty,\ \gamma = 0$			
g	0.00%	0.30%	0.50%
Λ_0^{Rat}	0.0834	0.0844	0.0852

At time 0, Δ_0^{Rat} and Γ_0^{Rat} are derived by theory, here $\Delta_{1.5}^{\text{Rat}}$ and $\Gamma_{1.5}^{\text{Rat}}$ are evaluated by the CE method. In doing so we assume that $S_0 = 100$, $S_1 = 90$, $S_{1.5} = 100$, and the applied contract parameters are given in Table 9.

Table 14 presents the results for $\Delta_{1.5}^{\text{Rat}}$. To investigate how ζ and α influence $\Delta_{1.5}^{\text{Rat}}$, we evaluate it using $(\alpha, \gamma, \zeta) = (100\%, 0, +\infty)$ and different minimal guarantees. The results are given in Table 16. These, together with the results in Table 14, suggest that $\Delta_{1.5}^{\text{Rat}}$ is an increasing function of α and ζ. This is because when α or ζ are large, the gain from the index is supposed to be more volatile. From the left part of Table 14, we can see that given a fixed g a higher γ leads to a larger $\Delta_{1.5}^{\text{Rat}}$, and this is the reason for the increase in fair caps. When $\zeta = +\infty$, $\Delta_{1.5}^{\text{Rat}}$ is an increasing function of γ, and it decreases when g gets larger. The effects of γ or g on $\Delta_{1.5}^{\text{Rat}}$ are determined by how they affect the fair participation rates.

Table 15 illustrates the results for $\Gamma_{1.5}^{\text{Rat}}$. All of the numbers are equal to or near zero, which means there are no big changes in asset prices. This is because we assume there are no jumps in our model for S_t. When $\alpha = 100\%$, according to (4.12), $\Gamma_{1.5}^{\text{Rat}} = 0$. When $\zeta = +\infty$, the changes in $\Gamma_{1.5}^{\text{Rat}}$ behave similarly with those of Λ_0^{Rat}. This is true since in the replicating portfolio both Γ_t and Λ_t are used as the number of similar derivatives.

Table 14: $\Delta_{1.5}^{\text{Rat}}$ of ratchet EIA

	Fair ζ ($\alpha = 100\%$)			Fair α ($\zeta = +\infty$)		
$\gamma \setminus g$	0.00%	0.30%	0.50%	0.00%	0.30%	0.50%
0.00%	0.0102	0.0104	0.0104	0.0020	0.0017	0.0017
0.50%	0.0103	0.0104	0.0104	0.0021	0.0019	0.0018
1.00%	0.0104	0.0104	0.0104	0.0022	0.0020	0.0018

Table 15: $\Gamma_{1.5}^{\text{Rat}}$ of ratchet EIA

	Fair ζ ($\alpha = 100\%$)			Fair α ($\zeta = +\infty$)		
$\gamma \setminus g$	0.00%	0.30%	0.50%	0.00%	0.30%	0.50%
0.00%	0	0	0	-0.0016%	-0.0014%	-0.0014%
0.50%	0	0	0	-0.0016%	-0.0015%	-0.0015%
1.00%	0	0	0	-0.0017%	-0.0016%	-0.0015%

Table 16: $\Delta_{1.5}^{\text{Rat}}$ of ratchet EIA (II)

$\alpha = 100\%,\ \zeta = +\infty,\ \gamma = 0$			
g	0.00%	0.30%	0.50%
$\Delta_{1.5}^{\text{Rat}}$	0.0134	0.0136	0.0137

6 Conclusions

In this thesis, we applied the Heston stochastic volatility model to analyze equity-indexed annuities (EIAs). Since the 1980s the volatility of stock market began to change markedly. Thus the Black-Scholes model has been plagued by the assumption of constant volatility, so stochastic volatility models became popular. The Heston model is one of the most used models. However, it assumes that the interest rate is constant, which makes sense for most short-term financial options. But for the EIA, whose maturity is between 10 and 15 years, the constant interest rate assumption does not hold. For long terms, the interest is volatile and a deterministic or stochastic model is expected. Here, we generalized Heston model to the case of a deterministic interest rate and give a semi-closed formula for the price of European call prices. We calibrate Heston model according to the observed European call prices, using a global optimization algorithm called differential evolution (DE).

The equity-indexed annuity is an innovative life insurance product. However, it is challenging to evaluate them properly. Hence, pricing and hedging of EIAs are interesting topics. In this thesis, we focused on ratchet EIAs, and two different methods are presented to evaluate them.

The first method is the joint transition probability density function (JTPDF) approach. We generalize the formula for the JTPDF of Heston's process, which can be found in Lipton (2001) and Lamoureux and Paseka (2009), to the case of deterministic interest. Following Ballestra et al. (2007), we applied a Filon-type quadrature to solve the oscillatory integral involved in the JTPDF. We find that the Filon-type quadrature is much faster and more accurate in solving oscillatory integrals than the traditional Gaussian quadratures. Given the JTPDF, the risk-neutral pricing formula reduced pricing ratchet EIAs to a problem on solving multiple integrals. The dimension of the multiple integrals is always above 10 and the integrand is too complex to be solved by basic cubature or Monte Carlo methods. Again, following Ballestra et al. (2007) we adopted the importance sampling technique to solve the multiple integral, but we generate samples by

a Quasi-Monte Carlo method.

Another method is the conditional expectation (CE) method. By conditioning on the volatility path, we can first price the ratchet EIA under Black-Scholes (BS) assumptions. Since a closed-form is known for the price of ratchet EIAs under BS, the price in the Heston framework is obtained by simulating the volatility path. The key point in the CE method is how to approach the integrated variance. Broadie and Kaya (2006) and Glasserman and Kim (2008) showed how to simulated exactly by different methods, but their method is time-consuming. Inspired by Bégin et al. (2012), we approximate the integrated variance by summations of gamma distributions. This approach is faster while the results are still acceptable. The CE method is much faster than the JPTDF approach, since CE does not require to calculate Bessel functions.

In the last chapter, we present numerical results for ratchet EIA prices as well as Greeks. We obtained the fair caps and fair participation rates. Our results are consistent with the parameters used by companies (see annuityadvantage.com). One interesting result is that the sign of Vegas changes for different contracts, thus the volatility risk could be reduced by holding a portfolio of ratchet EIAs with different contract parameters. At the end, sensitivity tests were carried out for the price and Greeks of ratchet EIAs.

Future work could focus on applying stochastic mortality rates to EIAs, since it is recognized that the mortality risk observed from life tables is a non-diversifiable risk. It is also interesting to consider inflation rates in dealing with EIAs, since it is always better to provide inflation protection for life insurance products.

References

H. Albrecher, P. Mayer, W. Schoutens, and Jurgen T. The little Heston trap. *Willmot Magazine*, pages 83–92, January 2007.

D.E. Amos. Algorithm 644: A portable package for Bessel functions of a complex argument and nonegative order. *ACM Transactions on Mathematical Software (TOMS)*, 12:265–273, 1985.

L. Andersen. Efficient simulation of the Heston stochastic volatility model. Working paper, Bank of America Securities, 2007.

G. Bakshi and D. Madan. Spanning and derivative-security valuation. *Journal of Fianancial Economics*, 55:205–238, 2000.

L.V. Ballestra, G. Pacelli, and F. Zirilli. A numerical method to price exotic path-dependent options on an underlying described by the Heston stochastic volatility model. *Journal of Banking and Finance*, 31:3420–3437, 2007.

S. Beckers. The constant elasticity of variance model and its implications for option pricing. *Journal of Finance*, 35:661–673, 1980.

J.F. Bégin, M. Bedard, and P. Gaillardetz. Efficient simulation scheme for the Heston model by gamma approximation. Unpublished, 2012.

N.H. Bingham and R. Kiesel. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives, 2nd.* Springer, 2004.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81:637–654, 1973.

R. Blattberg and N. Gonedes. A comparison of the stable and student distributions as statistical models for stock prices. *Journal of Business*, 47:244–280, 1974.

F. Bowman. *Introduction to Bessel Functions*. Dover, 1958.

M. Broadie and Ö. Kaya. Exact simulation of option Greeks under stochastic volatility and jump diffusion model. In *Proceedings of 2004 Winter Simulation Conference*, 2004.

M. Broadie and Ö. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54:217–231, 2006.

K.C. Cheung and H. Yang. Optimal stopping behavior of equity-linked investment products with regime switching. *Insurance: Mathematics and Economics*, 37:599–614, 2005.

J.C. Cox. Notes on option pricing i: Constant elasticity of variance diffusion. Working paper, 1975.

J.C. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

E. Derman and I. Kani. Riding on the smile. *Risk*, 7:32–39, 1994.

A.A. Drăgulescu and V.M. Yakovenko. Probability distribution of returns in the Heston model with stochastic volatility. *Quantitative Finance*, 2:443–453, 2002.

D. Dufresne. The integrated square-root process. University of Montreal, Research Paper Number 90, 2001.

B. Dumas, J. Fleming, and R. Whaley. Implied volatility functions: Empirical tests. *Journal of Finance*, 53:111–127, 1998.

B. Dupire. Pricing with a smile. *Risk*, 7:18–20, 1994.

L.N.G. Filon. On a quadrature formula for trigonometric integrals. In *Proc. Roy. Soc. Edinburgh*, 1928.

H. Geman, N. EL Karoui, and Rochet J.C. Changes of numeraire, changes of probability measure and option pricing. *Journal of Applied Probability*, 32:443–458, 1995.

J. Gil-Pelaez. Note on the inversion theorem. *Biometrika*, 38:481–482, 1951.

M. Gilli and E. Schumann. Calibrating option pricing models with heuristics. COMISEF Working Papers Series WPS-030, 2010.

M. Gilli, S. Grose, and E. Schumann. Calibrating the Nelson-Siegel-Svensson model. COMISEF Wording Papers Series WPS-031, 2010.

P. Glasserman and K.K. Kim. Gamma expansion of the Heston stochastic volatility model. Working paper, 2008.

J.M. Harrison and S.R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981.

S.L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6, 1993.

J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.

H. Jeffreys and B.S. Jeffreys. *Methods of Mathematical Physics, 3rd ed.* Cambridge University Press, 1988.

H Johnson and D. Shanno. Option pricing when the variance is changing. *Journal of Financial and Quantitative Analysis*, 22:143–151, 1987.

C. Joy, P.P. Boyle, and K.S. Tan. Quasi-Monte Carlo methods in numerical finance. *Management Science*, 42:926–938, 1996.

M. Kijima and T. Wong. Pricing of ratchet equity-indexed annuities under stochastic interest rates. *Insurance: Mathematics and Econommics*, 41:317–338, 2007.

S.A. Klugman, J.A. Beckley, P.L. Scahill, M.C. Varitek, and T.A. White. *Understanding Actuarial Practice*. Society of Actuaries, 2012.

I. Krykova. Evaluating of path-dependent securities with low discrepancy mehtods, 2003.

C.G. Lamoureux and A. Paseka. Information in option prices and the underlying asset dynamics. Working paper, 2009.

H. Lee. Pricing equity-indexed annuities with path-dependent options. *Insurance: Mathematics and Economics*, 33, 2003.

A. Lewis. *Option Valuation Under Stochastic Volatility*. Finance Press, 2000.

X.S. Lin and K.S. Tan. Valuation of equity-indexed annuities under stochastic interest rates. *North American Actuarial Journal*, 7, 2003.

X.S. Lin, K.S. Tan, and H. Yang. Pricing annuity guarantees under a regime-switching model. *North American Actuarial journal*, 13(3):316–338, 2009.

A. Lipton. *Mathematical methods for foreign exchange: a financial engineer's approach*. World Scientific, 2001.

A. MacKay. Pricing and hedging equity-linked products under stochastic volatility models, 2011.

J. Marrion, G. VanderPal, and D.F. Babbel. Real world index annuity returns. Working Paper: Wharton Financial Institutions Center, Personal Finance, 2010.

R.L. McDonald. *Derivatives Markets, 2nd Edition*. Prentice Hall, 2006.

S. Mikhailov and U. Nögel. Heston's stochastic volatility model implementation, calibration and some extensions. *Willmot Magazine*, 4:74–79, 2003.

N. Moodley. The Heston model: A practical approach, 2005.

S. Nandi. How important is the correlation between returns and volatility in a stochastic volatility model? Empirical evidence from pricing and hedging in the index options market. *Journal of Banking and Finance*, 22:589–610, 1998.

B.K. Oksendal. *Stochastic Differential Equations: An Introduction with Applications, 5th edition*. Springer, 2002.

S. Olver. *Numerical Approxmation of Highly Oscillatory Integrals*. PhD thesis, University of Cambridge, 2008.

S.H. Paskov and J.F. Traub. Faster evaluation of financial derivatives. *Journal of Portfolio Management*, 22:113–120, 1995.

L. Qian, Wang W., Wang R., and Tang Y. Valuation of equity-indexed annuity under stochastic mortality and interest rate. *Insurance: Mathematics and Economics*, 47: 123–129, 2010.

O. Reiss and U. Wystup. Efficient computation of option price sensitivities using homogeneity and other tricks. *The Journal of Derivatives*, 9:41–53, 2001.

S.B. Rollin, A. Ferreiro-Castilla, and F. Utzet. On the density of log-spot in the Heston volatility model. *Stochastic Processes and their Applications*, 120:2037–2063, 2010.

M. Rubinstein. Implied binomial trees. *Journal of Finance*, 49:771–818, 1994.

R. Schöbel and J.W. Zhu. Stochastic volatility with an ornsten-uhlenbeck process: an extension. *European Finance Review*, 3:23–46, 1999.

W. Schoutens, E. Simons, and J. Tistaert. A perfect calibration! now what? *Willmot Magazine*, pages 66–78, March 2004.

R. Schürer. A comparison between (quasi-) monte carlo and cubature rule based methods for solving high-dimensional integration problems. *Mathematics and Computers in Simulation*, 62:509–517, 2003.

L.O. Scott. Option pricing when the variance changes randomly: Theory, estimation and an application. *Journal of Financial and Quantitative Analysis*, 22:419–438, 1987.

A.C. Silva and V.M. Yakovenko. Comparison between the probability distribution of returns in the Heston model and empirical data for stock indexes. *Physica A*, 324: 303–310, 2003.

E.M. Stein and J.C. Stein. Stock price distribution with stochastic volatility: An analytical approach. *Review of Financial Studies*, 4:727–752, 1991.

R. Storn and K Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359, 1997.

R.M. Storn, K. Price, and J.A. Lampinen. *Differential Evolution: A Practical Approach to Global Optimization*. Springer, 2005.

L.E.O. Svensson. Estimating and interpreting forward interest tates: Sweden 1992-1994. IMF Working paper, 1994.

S. Tiong. Valuing equity-indexed annuities. *North American Actuarial Journal*, 4, 2000.

S.T. Tse and J. W. L. Wan. Low-bias simulation scheme for the Heston model by inverse gaussian approximation. SSRN eLibrary, 2010.

I.E. Vollrath and J. Wendland. Calibration of interest rate and option models using differential evolution. Working Paper, FINCAD Corporation, 2009.

J.B. Wiggins. Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics*, 19:351–372, 1987.

P. Wilmott. *Paul Wilmott on quantitative finance, 2nd ed, Vol.*3. Wiley, 2006.

B. Wong and C.C. Heyde. On changes of measure in stochastic volatility models. *Journal of Applied Mathematics and Stochastic Analysis*, pages 1–13, 2006.

J.W. Zhu. *Applications of Fourier Transform to Smile Modeling, 2nd Edition*. Springer, 2010.

A Appendix A

In this section, we present how the general stochastic volatility model is defined under the risk neutral measure $\mathbb{Q}$. This work follows from Wilmott (2006).

In general, the stochastic volatility model is given by

$$\frac{dS_t}{S_t} = \mu dt + V_t dW_s(t) \tag{A.1}$$

$$dV_t = p(S_t, V_t, t)\, dt + q(S_t, V_t, t)\, dW_v(t), \tag{A.2}$$

$$\langle dW_s(\cdot), dW_v(\cdot) \rangle = \rho dt, \tag{A.3}$$

where μ is the drift, V_t is the return volatility, $p(S_t, V_t, t)$ and $q(S_t, V_t, t)$ determine the model of stochastic volatility, $W_s(t)$ and $W_v(t)$ are two standard Brownian motion under measure $\mathbb{P}$ with correlation ρ.

Compared with the Black-Scholes framework in which the hedging portfolio consists of a risk-free bond and certain asset shares, there are two source of randomness in a stochastic volatility model that are the asset ($W_s(t)$) and the volatility ($W_v(t)$). Hence, we need another option to hedge the risk. Suppose that a portfolio consists of an option $\Pi(S_t, V_t, t)$, $-\Delta$ shares of the asset, and $-\Delta_1$ shares of another option $\Pi_1(S_t, V_t, t)$, the value of the portfolio at time t given by

$$O(S_t, V_t, t) = \Pi(S_t, V_t, t) - \Delta S_t - \Delta_1 \Pi_1(S_t, V_t, t). \tag{A.4}$$

For convenience we write $O(S_t, V_t, t)$ as O and similarly with Π, Π_1, p, and q. Applying Itô's formula to (A.4) leads to

$$
\begin{aligned}
dO &= \left(\frac{\partial \Pi}{\partial S_t}dS_t + \frac{\partial \Pi}{\partial V_t}dV_t + \frac{1}{2}V_t^2 S_t^2 \frac{\partial^2 \Pi}{\partial S_t^2}dt + \frac{1}{2}q^2\frac{\partial^2 \Pi}{\partial V_t^2}dt + \rho q V_t S_t \frac{\partial \Pi}{\partial t}dt\right) - \Delta dS_t \\
&\quad -\Delta_1\left(\frac{\partial \Pi_1}{\partial S_t}dS_t + \frac{\partial \Pi_1}{\partial V_t}dV_t + \frac{1}{2}V_t^2 S_t^2 \frac{\partial^2 \Pi_1}{\partial S_t^2}dt + \frac{1}{2}q^2\frac{\partial^2 \Pi_1}{\partial V_t^2}dt + \rho q S_t V_t \frac{\partial \Pi_1}{\partial t}dt\right) \\
&= \left(\frac{\partial \Pi}{\partial S_t} - \Delta - \Delta_1\frac{\partial \Pi_1}{\partial S_t}\right)dS_t + \left(\frac{\partial \Pi}{\partial V_t} - \Delta_1\frac{\partial \Pi_1}{\partial V_t}\right)dV_t \\
&\quad + \left[\left(\frac{\partial \Pi}{\partial t} + \frac{1}{2}S_t^2 V_t^2 \frac{\partial \Pi^2}{\partial^2 S_t} + \rho q S_t V_t \frac{\partial \Pi^2}{\partial V_t \partial S_t} + \frac{1}{2}q^2\frac{\partial^2 \Pi}{\partial V_t^2}\right)\right. \\
&\quad \left. - \Delta_1\left(\frac{\partial \Pi_1}{\partial t} + \frac{1}{2}S_t^2 V_t^2 \frac{\partial \Pi_1^2}{\partial^2 S_t} + \rho q S_t V_t \frac{\partial \Pi_1^2}{\partial V_t \partial S_t} + \frac{1}{2}q^2\frac{\partial^2 \Pi_1}{\partial V_t^2}\right)\right]dt. \qquad \text{(A.5)}
\end{aligned}
$$

Note that portfolio O is set up to hedge the option Π, so that O is free of asset and volatility risks. Hence, the terms multiplying dS_t and dV_t in (A.5) must be 0, that is

$$
\begin{cases}
\frac{\partial \Pi}{\partial S_t} - \Delta - \Delta_1\frac{\partial \Pi_1}{\partial S_t} &= 0, \\
\frac{\partial \Pi}{\partial V_t} - \Delta_1\frac{\partial \Pi_1}{\partial V_t} &= 0.
\end{cases} \qquad \text{(A.6)}
$$

This is equivalent to

$$
\begin{cases}
\Delta &= \frac{\partial \Pi}{\partial S_t} - \frac{\partial \Pi/\partial V_t}{\partial \Pi_1/\partial V_t}\frac{\partial \Pi_1}{\partial S_t}, \\
\Delta_1 &= \frac{\partial \Pi/\partial V_t}{\partial \Pi_1/\partial V_t}.
\end{cases} \qquad \text{(A.7)}
$$

Note that when the terms multiplying dS_t and dV_t in (A.5) are 0, (A.5) becomes

$$
\begin{aligned}
dO &= \left[\left(\frac{\partial \Pi}{\partial t} + \frac{1}{2}S_t^2 V_t^2 \frac{\partial \Pi^2}{\partial^2 S_t} + \rho q S_t V_t \frac{\partial \Pi^2}{\partial V_t \partial S_t} + \frac{1}{2}q^2\frac{\partial^2 \Pi}{\partial V_t^2}\right)\right. \\
&\quad \left. - \Delta_1\left(\frac{\partial \Pi_1}{\partial t} + \frac{1}{2}S_t^2 V_t^2 \frac{\partial \Pi_1^2}{\partial^2 S_t} + \rho q S_t V_t \frac{\partial \Pi_1^2}{\partial V_t \partial S_t} + \frac{1}{2}q^2\frac{\partial^2 \Pi_1}{\partial V_t^2}\right)\right]dt. \qquad \text{(A.8)}
\end{aligned}
$$

Given that O is free of asset and volatility risks, it earns risk-free interest rate. Thus

$$
dO = rOdt = r\left(\Pi - \Delta S_t - \Delta_1 \Pi_1\right)dt. \qquad \text{(A.9)}
$$

Combining and (A.8) and (A.9) leads to a PDE, that is

$$
\begin{aligned}
&\left(\frac{\partial \Pi}{\partial t}+\frac{1}{2} S_t^2 V_t^2 \frac{\partial \Pi^2}{\partial^2 S_t}+\rho q S_t V_t \frac{\partial \Pi^2}{\partial V_t \partial S_t}+\frac{1}{2} q^2 \frac{\partial^2 \Pi}{\partial V_t^2}\right) \\
&-\Delta_1\left(\frac{\partial \Pi_1}{\partial t}+\frac{1}{2} S_t^2 V_t^2 \frac{\partial \Pi_1^2}{\partial^2 S_t}+\rho q S_t V_t \frac{\partial \Pi_1^2}{\partial V_t \partial S_t}+\frac{1}{2} q^2 \frac{\partial^2 \Pi_1}{\partial V_t^2}\right) \\
= \quad & r\left(\Pi-\Delta S_t-\Delta_1 \Pi_1\right)
\end{aligned} \tag{A.10}
$$

Inserting (A.7) into (A.10) leads to

$$
\begin{aligned}
&\frac{\frac{\partial \Pi}{\partial t}+\frac{1}{2} S_t^2 V_t^2 \frac{\partial \Pi^2}{\partial^2 S_t}+\rho q S_t V_t \frac{\partial \Pi^2}{\partial V_t \partial S_t}+\frac{1}{2} q^2 \frac{\partial^2 \Pi}{\partial V^2}+r S_t \frac{\partial \Pi}{\partial S_t}-r \Pi}{\frac{\partial \Pi}{\partial V_t}} \\
= \quad & \frac{\frac{\partial \Pi_1}{\partial t}+\frac{1}{2} S_t^2 V_t^2 \frac{\partial \Pi_1^2}{\partial^2 S_t}+\rho q S_t V_t \frac{\partial \Pi_1^2}{\partial V_t \partial S_t}+\frac{1}{2} q^2 \frac{\partial^2 \Pi_1}{\partial V_t^2}+r S_t \frac{\partial \Pi_1}{\partial S_t}-r \Pi_1}{\frac{\partial \Pi_1}{\partial V_t}}
\end{aligned}
$$

This is an intimidating PDE, but fortunately the left hand side is a function w.r.t. Π and the other side is a function only w.r.t. Π_1. Since there are no constraints posed on the contract of Π and Π_1, the previous equation must be true for any Π and Π_1. This means that both sides must be equal to a quantity which only depends on S_t, V_t and t. We denote quantity $-\left(p\left(S_t, V_t, t\right)-\lambda\left(S_t, V_t, t\right) q\left(S_t, V_t, t\right)\right)$ or $-(p-\lambda q)$ for short. (Though we give the quantity a form, but it can still be any function of S_t, V_t and t since we do not propose any constraint on $\lambda\left(S_t, V_t, t\right)$. Later, the readers will find more information about $\lambda\left(S_t, V_t, t\right)$.) Then the following equation is true

$$
\frac{\partial \Pi}{\partial t}+\frac{1}{2} S_t^2 V_t^2 \frac{\partial^2 \Pi}{\partial S_t^2}+\rho q S_t V_t \frac{\partial^2 \Pi}{\partial V_t \partial S_t}+\frac{1}{2} q^2 \frac{\partial^2 \Pi}{\partial V_t^2}+r S_t \frac{\partial \Pi}{\partial S_t}-r \Pi=-(p-\lambda q) \frac{\partial \Pi}{\partial V_t},
$$

that is

$$
\frac{\partial \Pi}{\partial t}+\frac{1}{2} S_t^2 V_t^2 \frac{\partial^2 \Pi}{\partial S_t^2}+\rho q S_t V_t \frac{\partial \Pi^2}{\partial V_t \partial S_t}+\frac{1}{2} q^2 \frac{\partial^2 \Pi}{\partial V_t^2}+r S_t \frac{\partial \Pi}{\partial S_t}+(p-\lambda q) \frac{\partial \Pi}{\partial V_t}-r \Pi=0. \tag{A.11}
$$

Here (A.11) is important since solving it leads to the price Π. Similarly to Black-Scholes PDE, the drift of dS_t/S_t, that is μ, does not appear in (A.10), but the risk-free rate r does. Note the terms multiplying $\frac{\partial \Pi}{\partial S_t}$ and $\frac{\partial \Pi}{\partial V_t}$. We claim that under the risk-neutral measure $\mathbb{Q}$ the drift terms of dS_t and dV_t are supposed to be $rS_t dt$ and $(p-\lambda q)\, dt$. Hence

the general stochastic volatility model under $\mathbb{Q}$ is given by

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + V_t dW_s^{\mathbb{Q}}(t), \\ dV_t &= \left[p(S_t, V_t, t) - \lambda(S_t, V_t, t)\, q(S_t, V_t, t)\right] dt + q(S_t, V_t, t)\, dW_v^{\mathbb{Q}}(t), \end{aligned} \tag{A.12}$$

$$d\langle W_s(\cdot), W_v(\cdot)\rangle = \rho dt.$$

Here the Randon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ can also be derived.

The term $\lambda(S_t, V_t, t)$, λ for short, is of interest. Following Wilmott (2006) we discuss its properties. Suppose that we delta-hedge an option Π satisfying (A.11), then we have a portfolio $\tilde{O}$ given by

$$\tilde{O}(S_t, V_t, t) = \Pi(S_t, V_t, t) - \Delta S_t.$$

Again, for short we denote $\tilde{O}(S_t, V_t, t)$ and $\Pi(S_t, V_t, t)$ by $\tilde{O}$ and Π. Applying Itô's formula to $\tilde{O}$ leads to

$$\begin{aligned} d\tilde{O} &= d\Pi - \Delta dS_t \\ &= \left(\frac{\partial \Pi}{\partial S_t} - \Delta\right) dS_t + \frac{\partial \Pi}{\partial V_t} dV_t + \left(\frac{\partial \Pi}{\partial t} + \frac{1}{2} S_t^2 V_t^2 \frac{\partial^2 \Pi}{\partial S_t^2} + \frac{1}{2} q^2 \frac{\partial^2 \Pi}{\partial V_t^2} + \frac{1}{2} \rho q S_t V_t \frac{\partial^2 \Pi}{\partial V_t \partial S_t}\right) dt \\ &= \frac{\partial \Pi}{\partial V_t} dV_t + \left(\frac{\partial \Pi}{\partial t} + \frac{1}{2} S_t^2 V_t^2 \frac{\partial^2 \Pi}{\partial S_t^2} + \frac{1}{2} q^2 \frac{\partial^2 \Pi}{\partial V_t^2} + \frac{1}{2} \rho q S_t V_t \frac{\partial^2 \Pi}{\partial V_t \partial S_t}\right) dt. \end{aligned}$$

The last equation holds since we are delta-hedging Π, that is $\Delta = \frac{\partial \Pi}{\partial S_t}$. The difference between $d\tilde{O}$ and $r\tilde{O}dt$ is interesting. It is given by

$$\begin{aligned} d\tilde{O} - r\tilde{O}dt &= \frac{\partial \Pi}{\partial V_t} dV_t + \left(\frac{\partial \Pi}{\partial t} + \frac{1}{2} S_t^2 V_t^2 \frac{\partial^2 \Pi}{\partial S_t^2} + \frac{1}{2} q^2 \frac{\partial^2 \Pi}{\partial V_t^2} + \frac{1}{2} \rho q S_t V_t \frac{\partial^2 \Pi}{\partial V_t \partial S_t}\right) dt \\ &\quad -r(\Pi - \Delta S_t) dt \\ &= \frac{\partial \Pi}{\partial V_t} dV_t + \left(\frac{\partial \Pi}{\partial t} + \frac{1}{2} S_t^2 V_t^2 \frac{\partial^2 \Pi}{\partial S_t^2} + \frac{1}{2} q^2 \frac{\partial^2 \Pi}{\partial V_t^2} + \frac{1}{2} \rho q S_t V_t \frac{\partial^2 \Pi}{\partial V_t \partial S_t}\right. \\ &\quad \left. -r\Pi + r S_t \frac{\partial \Pi}{\partial S_t}\right) dt \\ &= \frac{\partial \Pi}{\partial V_t} dV_t - (p - \lambda q) \frac{\partial \Pi}{\partial V_t} dt \qquad \text{(A.13)} \\ &= \frac{\partial \Pi}{\partial V_t} \left(dV_t - (p - \lambda q) dt\right) \\ &= q \frac{\partial \Pi}{\partial V_t} \left(\lambda dt + dW_v(t)\right). \qquad \text{(A.14)} \end{aligned}$$

Here, (A.13) holds because of (A.11), while (A.14) holds from (A.2).

Observe that the error in delta-hedging, which is an instantaneous rate, can be described by the change rate in Π w.r.t. V_t and "for every unit of volatility risk, represented by" $dW_v(t)$, "there are λ units of extra return, represented by dt". Hence, λ or $\lambda(S_t, V_t, t)$, is called the market price of volatility risk.

B Appendix B

Here we give a brief introduction to DE's update procedure: DE deals with a population of n_p solutions stored in real-valued vectors $x_{i,G} = (x_{1,i,G}, ..., x_{D,i,G}), i = 1, 2, ..., n_p$, where G is the generation number and D is the number of parameters ($D = 5$ in the Heston model). Given the upper bounds and lower bounds for each parameter, s.t., $x_j^L \leq x_{j,i,1} \leq x_j^U$, the entry $x_{j,i,1}, (j = 1, ..., D, i = 1, ..., n_p)$ of the initial vectors $x_{i,1}$ is uniformly chosen from the interval $\left[x_j^L, x_j^U\right]$. Then the initial vector is updated as follows: For every vector $x_{i,G}$, randomly select three other vectors $x_{r_1,G}, x_{r_2,G}, x_{r_3,G}$ where i, r_1, r_2, r_3 are different. Define the donor vector as

$$v_{i,G+1} = x_{r_1,G} + F\left(x_{r_2,G} - x_{r_3,G}\right),$$

where F is a constant from $[0, 2]$. The element of the trial vector $u_{i,G+1}$ is determined by the target vector $x_{i,G}$ and the donor vector $v_{i,G+1}$. In detail, the entries of the donor vector are the entries the trial vector with probability CR, that is

$$u_{j,i,G+1} = \begin{cases} v_{j,i,G+1} & \text{if } u_{i,j} \leq CR \text{ or } j = I_{rand}, \\ x_{j,i,G} & \text{if } u_{i,j} \geq CR \text{ and } j \neq I_{rand}, \end{cases} \tag{B.1}$$

where $u_{i,j}$ is chosen uniformly from $[0, 1]$, and I_{rand} is a random integer from $\{1, 2, ..., D\}$. At last, vector $x_{j,i,G}$ is updated by the trial vector $u_{i,G+1}$ and itself, is:

$$x_{i,G+1} = \begin{cases} u_{i,G+1} & \text{if } f(u_{i,G+1} < f(x_{i,G})), \\ x_{i,G} & \text{otherwise.} \end{cases} \tag{B.2}$$

In the following the update procedure is repeated until the iteration number reaches the maximum limit N_G. The update procedure is summarized in Algorithm 1.

A series of articles have proved that DE is more accurate and efficient than several other optimization methods including four genetic algorithms, simulated annealing and evolutionary programming. For more details on the DE method please refer to Storn et al. (2005).

Algorithm 3 DE's update procedure

initialize parameters D, n_p, N_G, F, CR, the upper and lower bounds.

initialize $x_{i,1}, i = 1, 2, \cdots, n_p$.

for $G = 2 \rightarrow N_G - 1$ **do**

generate $x_{r_1,G}, x_{r_2,G}, x_{r_3,G}$ for each $x_{i,G}$

$v_{i,G+1} = x_{r_1,G} + F\left(x_{r_2,G} - x_{r_3,G}\right)$

for $j = 1 \rightarrow D$ **do**

define $u_{j,i,G+1}$ according to equation (B.1).

end for

update $x_{i,G+1}$ according to equation (B.2).

end for

C Appendix C

Proposition C.1 *Suppose a kind of Riccati function is given as follows:*

$$\frac{dY(t)}{dt} = \alpha Y(t)^2 + \beta Y(t) + \gamma, \tag{C.1}$$

and with boundary condition $Y(0) = 0$*, where* $\alpha, \beta, \gamma \in \mathbb{C}$*.*

This kind of Ricatti equation has a unique solution. And the solution is given by

$$Y(t) = \frac{Y_0 (A - \alpha Y_0) e^{-At} - (A - \alpha Y_0) Y_0}{(A - \alpha Y_0) e^{-At} + \alpha Y_0},$$

where

$$\begin{aligned} Y_0 &= \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}, \\ A &= 2\alpha Y_0 + \beta \end{aligned}$$

Remark C.2 *Thought there are two values for* Y_0*, they lead to the same value of* $Y(t)$ *that satisfies* (C.1) *and* $Y(0) = 0$*.*

Proof. First of all, it is possible to find a constant solution to $Y(t)$ by solving

$$0 = \alpha Y(t)^2 + \beta Y(t) + \gamma \tag{C.2}$$

and the solution is

$$Y_0 = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}.$$

Define $U(t) = Y(t) - Y_0$, that is $Y(t) = U(t) + Y_0$. It is easy to derive that

$$\begin{aligned} \frac{dY(t)}{dt} &= \frac{dU(t)}{dt}, \\ Y(t)^2 &= U(t)^2 + Y_0^2 + 2U(t) Y_0. \end{aligned}$$

Hence (C.1) can be written as

$$\begin{aligned}
\frac{dU(t)}{dt} &= \alpha\left[U(t)^2 + Y_0^2 + 2U(t)Y_0\right] + \beta\left[U(t) + Y_0\right] + \gamma \\
&= \alpha U(t)^2 + (2\alpha Y_0 + \beta)U(t) + \left(\alpha Y_0^2 + \beta Y_0 + \gamma\right) \qquad \text{(C.3)} \\
&= \alpha U(t)^2 + (2\alpha Y_0 + \beta)U(t). \qquad \text{(C.4)}
\end{aligned}$$

(C.3) holds since Y_0 is the solution to (C.2).

Let $W(t) = \frac{1}{U(t)}$, that is $U(t) = \frac{1}{W(t)}$. The substitution leads to

$$\begin{aligned}
\frac{dU(t)}{dt} &= -\frac{1}{W(t)^2}\frac{dW(t)}{dt}, \\
U(t)^2 &= \frac{1}{W(t)^2}.
\end{aligned}$$

Then (C.4) can be written as

$$-\frac{1}{W(t)^2}\frac{dW(t)}{dt} = \alpha\frac{1}{W(t)^2} + (2\alpha Y_0 + \beta)\frac{1}{W(t)}.$$

This is equivalent to

$$\frac{dW(t)}{dt} + (2\alpha Y_0 + \beta)W(t) = -\alpha.$$

Note that the previous ODE is a first order linear equation, so that the solution to $W(t)$ is given by

$$W(t) = \frac{C(2\alpha Y_0 + \beta)e^{-(2\alpha Y_0 + \beta)t} - \alpha}{2\alpha Y_0 + \beta},$$

where C is a constant.

Define $A = 2\alpha Y_0 + \beta$,

$$U(t) = \frac{1}{W(t)} = \frac{A}{CAe^{-At} - \alpha}$$

and

$$\begin{aligned}
Y(t) &= U(t) + Y_0 \\
&= \frac{A}{CAe^{-At} - \alpha} + Y_0 \\
&= \frac{CY_0Ae^{-At} + A - \alpha Y_0}{CAe^{-At} - \alpha}
\end{aligned}$$

In the previous equation $Y(t)$ must satisfy the boundary condition $Y(0) = 0$, this leads to

$$C = \frac{\alpha Y_0 - A}{AY_0}.$$

Finally, the solution to (C.1), given the boundary condition $Y(t) = 0$, is

$$Y(t) = \frac{Y_0 (A - \alpha Y_0) e^{-At} - (A - \alpha Y_0) Y_0}{(A - \alpha Y_0) e^{-At} + \alpha Y_0}.$$

The uniqueness is because the solution to a first order linear equation is unique. □

D Appendix D: Bilinear Interpolation

Assume that a smooth function f of x and v over $[x_{min}, x_{max}] \times [v_{min}, v_{max}]$ is given and its values at four given points (shown in Figure 6) are known as $f_{ij} = f(x_i, v_j)$ for $i, j = 1, 2$, where $[x_1, x_2] \times [v_1, v_2] \subset [x_{min}, x_{max}] \times [v_{min}, v_{max}]$. The value of $f(x, v)$ for any $(x, v) \in [x_1, x_2] \times [v_1, v_2]$ can be approximated using linear interpolation. First, the value of g at two points $R_1(x, v_1)$ and $R_2(x, v_2)$ (shown in Figure 6) is estimated using linear interpolation with (x_1, v_i) and (x_2, v_i), for $i = 1, 2$:

$$\begin{aligned} g(R_1) &= g(x, v_1) = \frac{x_2 - x}{x_2 - x_1} f(x_1, v_1) + \frac{x - x_1}{x_2 - x_1} f(x_2, v_1), \\ g(R_2) &= g(x, v_2) = \frac{x_2 - x}{x_2 - x_1} f(x_1, v_2) + \frac{x - x_1}{x_2 - x_1} f(x_2, v_2). \end{aligned}$$

Note that (x, y), $R_1(x, v_1)$ and $R_2(x, v_2)$ share the same x coordinate, then $f(x, y)$ can be approximated by linear interpolation using points R_1, R_2 and

$$\begin{aligned} g(x, y) &= \frac{v_2 - v}{v_2 - v_1} g(R_1) + \frac{v - v_1}{v_2 - v_1} g(R_2) \\ &= \frac{v_2 - v}{v_2 - v_1} \left\{ \frac{x_2 - x}{x_2 - x_1} f(x_1, v_1) + \frac{x - x_1}{x_2 - x_1} f(x_2, v_1) \right\} \\ &\quad + \frac{v - v_1}{v_2 - v_1} \left\{ \frac{x_2 - x}{x_2 - x_1} f(x_1, v_2) + \frac{x - x_1}{x_2 - x_1} f(x_2, v_2) \right\}. \end{aligned} \tag{D.1}$$

Using the above method, if we divide $[x_{min}, x_{max}] \times [v_{min}, v_{max}]$ into N_x-by-N_v grids, given the values of f at the intersection points we could estimate the value of f at any point in $[x_{min}, x_{max}] \times [v_{min}, v_{max}]$ by bilinear interpolation. This is because any point $(x, v) \in [x_{min}, x_{max}]$ must be located in one grid. The set of the values of f at the intersection points is usually called a look-up table. Note that

$$g \to f \quad \text{when} \quad N_x \to \infty \quad \text{and} \quad N_v \to \infty \tag{D.2}$$

holds because of (D.1) and the assumption that f is smooth over $[x_{min}, x_{max}] \times [v_{min}, v_{max}]$.

Moreover, some useful formulas are listed as follows.

1. The integral of $g(x, v)$ with respect to x from x_1 to x_2 is

$$\int_{x_1}^{x_2} g(x, v)\, dx = \left(1 - \frac{v - v_1}{v_2 - v_1}\right) \frac{(x_2 - x_1)(f_{11} + f_{21})}{2} + \frac{v - v_1}{v_2 - v_1} \frac{(x_2 - x_1)(f_{12} + f_{22})}{2}.$$

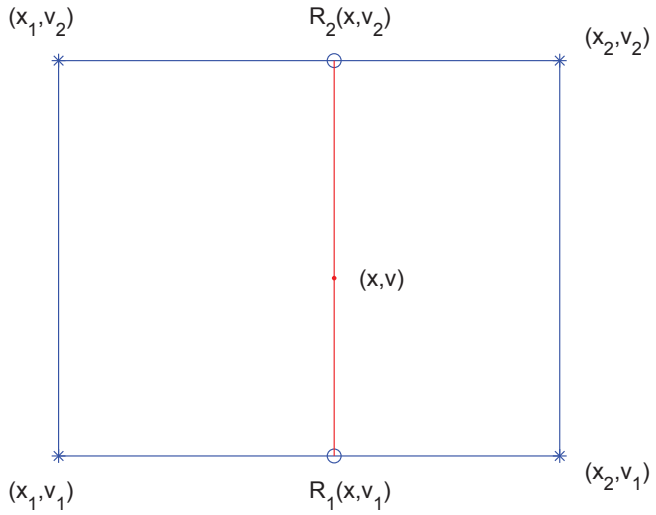


Figure 6: Bilinear interpolation

2. The double integral of $g(x,v)$ with respect to x and v is

$$\int_{v_1}^{b}\int_{x_1}^{x_2} g(x,v)\,dxdv = \frac{(b-v_1)(2v_2-b-v_1)(x_2-x_1)(f_{11}+f_{21})}{4(v_2-v_1)} + \frac{(b-v_1)^2(x_2-x_1)(f_{12}+f_{22})}{4(v_2-v_1)}.$$

3. Setting $b=v_2$ in the previous equation, the integral of $g(x,v)$ over $[x_1,x_2]\times[v_1,v_2]$ is

$$\int_{v_1}^{v_2}\int_{x_1}^{x_2} f(x,v)\,dxdv = \frac{1}{4}(f_{11}+f_{12}+f_{21}+f_{22})(x_2-x_1)(v_2-v_1).$$

E Appendix E: Vega

$$
\begin{aligned}
\frac{\partial \Pi}{\partial v_t} &= \frac{\partial}{\partial v_t} E^{\mathbb{Q}}\left[e^{\int_t^n r(w)\,dw} C^*\left(\mathbf{X}, n\right) \middle| \mathcal{F}_t\right] \\
&= e^{\int_t^n r(w)\,dw} \frac{\partial}{\partial v_t} \int \cdots \int_{\Omega_{\mathbf{xv}}} C^*\left(\mathbf{x}, n\right) p\left(x_i, v_i | x_t, v_t\right) \prod_{j=i+1}^{n} p\left(x_i, v_i | x_{i-1}, v_{i-1}\right)\, d\mathbf{x} d\mathbf{v} \\
&= e^{\int_t^n r(w)\,dw} \frac{\partial}{\partial v_t} \int_{\Omega_{\mathbf{v}}} \left(\int_{\Omega_{\mathbf{x}}} C^*\left(\mathbf{x}, n\right) p\left(x_i, v_i | x_t, v_t\right) \prod_{j=i+1}^{n} p\left(x_j, v_j | x_{j-1}, v_{j-1}\right)\, d\mathbf{x} \right) d\mathbf{v} \qquad \text{(E.1)} \\
&= e^{\int_t^n r(w)\,dw} \frac{\partial}{\partial v_t} \int_{\Omega_{\mathbf{v}}} \left(\sum \int_{S_{\mathbf{x}}^j} f^j\left(\mathbf{x}\right) p\left(x_i, v_i | x_t, v_t\right) \prod_{j=i+1}^{n} p\left(x_j, v_j | x_{j-1}, v_{j-1}\right)\, d\mathbf{x} \right) d\mathbf{v} \qquad \text{(E.2)} \\
&= e^{\int_t^n r(w)\,dw} \int_{\Omega_{\mathbf{v}}} \frac{\partial}{\partial v_t} \left(\sum \int_{S_{\mathbf{x}}^j} f^j\left(\mathbf{x}\right) p\left(x_i, v_i | x_t, v_t\right) \prod_{j=i+1}^{n} p\left(x_j, v_j | x_{j-1}, v_{j-1}\right)\, d\mathbf{x} \right) d\mathbf{v} \qquad \text{(E.3)} \\
&= e^{\int_t^n r(w)\,dw} \int_{\Omega_{\mathbf{v}}} \left(\sum \int_{S_{\mathbf{x}}^j} f^j\left(\mathbf{x}\right) \prod_{j=i+1}^{n} p\left(x_j, v_j | x_{j-1}, v_{j-1}\right) \frac{\partial}{\partial v_t} p\left(x_i, v_i | x_t, v_t\right)\, d\mathbf{x} \right) d\mathbf{v} \qquad \text{(E.4)} \\
&= e^{\int_t^n r(w)\,dw} \int_{\Omega_{\mathbf{v}}} \left(\int_{\Omega_{\mathbf{x}}} C^*\left(\mathbf{x}, n\right) \prod_{j=i+1}^{n} p\left(x_j, v_j | x_{j-1}, v_{j-1}\right) \frac{\partial}{\partial v_t} p\left(x_i, v_i | x_t, v_t\right)\, d\mathbf{x} \right) d\mathbf{x} d\mathbf{v} \qquad \text{(E.5)} \\
&= e^{\int_t^n r(w)\,dw} \int \cdots \int_{\Omega_{\mathbf{xv}}} C^*\left(\mathbf{x}, n\right) \prod_{j=i+1}^{n} p\left(x_j, v_j | x_{j-1}, v_{j-1}\right) \frac{\partial}{\partial v_t} p\left(x_i, v_i | x_t, v_t\right)\, d\mathbf{x} d\mathbf{v} \qquad \text{(E.6)} \\
&= e^{\int_t^n r(w)\,dw} E^{\mathbb{Q}}\left[C^*\left(\mathbf{X}, n\right) \frac{\frac{\partial}{\partial v_t} p\left(x_i, v_i | x_t, v_t\right)}{p\left(x_i, v_i | x_t, v_t\right)} \middle| \mathcal{F}_t \right]. \qquad \text{(E.7)}
\end{aligned}
$$

(E.1) holds because of Fubini's theorem: Since we evaluated the integral or the expectation so that it can be proved that the absolute value of the integrand is integrable. Then we can apply Fubini's theorem to change the order of the integrals. In (E.2), $\bigcup S_{\mathbf{x}}^j = \Omega_{\mathbf{x}}$ and this is a finite partition for n is finite. We make this change to ensure that on each $S_{\mathbf{x}}^j$, the integrand can get rid of the indicator functions and can be written as $f^j\left(\mathbf{x}\right) \prod_{i=1}^{n} p\left(x_i, v_i | x_{i-1}, v_{i-1}\right)$, where $f^j\left(\mathbf{x}\right)$ is a smooth function of $\mathbf{x}$. Hence each inte-

grand on $S^j_{\mathbf{x}}$ should be a smooth function of $\mathbf{x}$, $\mathbf{v}$ and v_0. Further, the integral w.r.t. $\mathbf{x}$ is also a smooth function of $\mathbf{v}$ and v_0 then we can apply Lebesgue's dominated convergence theorem to change the derivative and the integrals which leads to (E.3) and (E.4). Then (E.5) holds as bringing $\frac{\partial}{\partial v_0}$ inside the inner integral does not make any changes to $f^j(\mathbf{x})$ (the partition of $\Omega_{\mathbf{x}}$ only depends on $f^j(\mathbf{x})$). Applying Fubini's theorem again leads to (E.6).