Commodity Derivatives: Modeling and Pricing

Zaizhi Wang

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

MINES ParisTech

École doctorale n° 396 : Économie, Organisation, Société

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École nationale supérieure des mines de Paris

Spécialité " Finance Quantitative "

présentée et soutenue publiquement par

Zaizhi WANG

le 14/12/2011

Produits Dérivés des Matières Premières: Modélisation et Evaluation

Commodity Derivatives: Modeling and Pricing

Directeur de thèse : **Alain GALLI**
Co-encadrement de la thèse : **Eric BENHAMOU**

Jury

Mme. Delphine LAUTIER, Professeur, Université Paris Dauphine	Rapporteur
M. Bernard LAPEYRE, Professeur, Ecole des Ponts ParisTech	Rapporteur
M. René AID, Directeur Exécutif, EDF - Finance for Energy Market Research Centre	Examinateur
M. Eric BENHAMOU, CEO, Pricing Partners	Examinateur
M. Alain GALLI, directeur de recherche, ENSMP	Examinateur
M. Michel SCHMITT, ingénieur général des Mines, ENSMP	Examinateur

T H È S E

MINES ParisTech
Centre CERNA
60, Boulevard Saint-Michel 75272 PARIS cedex 06

Abstract

Commodity prices have been rising at an unprecedented pace over the last years making commodity derivatives more and more popular in many sectors like energy, metals and agricultural products. The quick development of commodity market as well as commodity derivative market results in a continuously uprising demand of accuracy and consistency in commodity derivative modeling and pricing.

The specification of commodity modeling is often reduced to an appropriate representation of convenience yield, intrinsic seasonality and mean reversion of commodity price. As a matter of fact, convenience yield can be extracted from forward strip curve and then be added as a drift term into pricing models such as Black Scholes model, local volatility model and stochastic volatility model. Besides those common models, some specific commodity models specially emphasize on the importance of convenience yield, seasonality or mean reversion feature. By giving the stochasticity to convenience yield, Gibson Schwartz model interprets the term structure of convenience yield directly in its model parameters, which makes the model extremely popular amongst researchers and market practitioners in commodity pricing. Gabillon model, in the other hand, focuses on the feature of seasonality and mean reversion, adding a stochastic long term price to correlate spot price.

In this thesis, we prove that there is mathematical equivalence relation between Gibson Schwartz model and Gabillon model. Moreover, inspired by the idea of Gyöngy, we show that Gibson Schwartz model and Gabillon model can reduce to one-factor model with explicitly calculated marginal distribution under certain conditions, which contributes to find the analytic formulas for forward and vanilla options. Some of these formulas are new to our knowledge and other formulas confirm with the earlier results of other researchers.

Indeed convenience yield, seasonality and mean reversion play a very important role, but for accurate pricing, hedging and risk management, it is also critical to have a good modeling of the dynamics of volatility in commodity markets as this market has very fluctuating volatility dynamics. While the formers (seasonality, mean reversion and convenience yield) have been highly emphasized in the literature on commodity derivatives pricing, the latter (the dynamics of the volatility) has often been forgotten. The family of stochastic volatility model is introduced to strengthen the dynamics of the volatility, capturing the dynamic smile of volatility surface thanks to a stochastic process on volatility itself. It is a very important characteristic for pricing derivatives of long maturity. Stochastic volatility model also corrects the problem of opposite underlying-volatility correlation against market data in many other models by introducing correlation parameter explicitly. The most popular stochastic volatility models include Heston model, Piterbarg model, SABR model, etc.

As pointed out by Piterbarg, the need of time-dependent parameters in stochastic volatility models is real and serious. It is because in one hand

stochastic volatility models with constant parameters are generally incapable of fitting market prices across option expiries, and in the other hand exotics do not only depend on the distribution of the underlying at the expiry, but on its dynamics through all time. This contradiction implies the necessity of time-dependent parameters. In this thesis, we extend Piterbarg's idea to the whole family of stochastic volatility model, making all the stochastic volatility models having time-dependent parameters and show various formulas for vanilla option price by employing various techniques such as characteristic function, Fourier transform, small error perturbation, parameter averaging, etc.

Résumé

Les prix des matières premières ont augmenté à un rythme sans précédent au cours des dernières années rendant les dérivés sur matières premières de plus en plus populaires dans de nombreux secteurs comme l'énergie, les métaux et les produits agricoles. Le développement rapide du marché des produits dérivés sur matières premières a aussi induit une recherche vers toujours plus de précision et cohérence dans la modélisation et l'évaluation de produits dérivés des matières premières.

Les points les plus importants dans la modélisation des matières premières sont la bonne représentation du rendement d'opportunité appelé communément «convenience yield », la prise en compte de la saisonnalité et la capture du phénomène de retour à la moyenne pour les prix des matières premières. Il est à noter que le rendement d'opportunité (convenience yield) peut être induit du prix des la courbe des forwards et être simplement ajouté au terme d'évolution(terme de drift) dans les modèles canoniques, comme le modèle de Black Scholes, le modèle à volatilité locale et les modèles à volatilité stochastique. An delà de ces modèles, d'autres modèles ont été conçus pour modéliser spécifiquement l'évolution du convenience yield, la saisonnalité ou le phénomène de retour à la moyenne des prix. Il s'agit par exemple du modèle de Gibson Schwartz qui suppose que le terme de convenience yield est aléatoire. Cette approche prend donc en compte l'évolution non déterministe du convenience yield et l'interprète comme un paramètre critique du modèle. Ceci explique sa grande popularité et son adoption important par les praticiens du marché. Un autre modèle fréquemment utilisé est le modèle de Gabillon. Celui se concentre sur la saisonnalité des prix et l'effet de retour a la moyenne, en modélisant un prix à long terme stochastique corrélé aux prix du spot. Dans cette thèse, nous prouvons que ces deux approches ne sont en fait qu'une et qu'il y a une relation d'équivalence entre le modèle de Gibson Schwartz et le modèle de Gabillon. Reposant sur le principe de diffusion équivalente introduite par Gyöngy, nous montrons que le modèle de Gibson Schwartz et le modèle de Gabillon peuvent se réduire à un modèle à un facteur dont la distribution marginale peut être explicitement calculée sous certaines conditions. Ceci nous permet en particulier de trouver des formules analytiques pour l'ensemble des options vanilles. Certaines de ces formules sont nouvelles à notre connaissance et d'autres confirment des résultats antérieurs.

Dans un second temps, nous nous intéressons à la bonne modélisation de la dynamique de la volatilité des marchés des matières premières. En effet, les marchés de matières premières sont caractérisés par des volatilités très fluctuantes et importante. Alors que les effets sur la saisonnalité, la modèlisation du convenience yield et l'effet de retour à la moyenne des prix ont été fortement soulignés dans la littérature, la bonne modélisation de la dynamique de la volatilité a souvent été oubliée. La famille de modèle à volatilité stochastique est introduite pour renforcer la dynamique de la volatilité, capturant le phénomène de smile de la surface de volatilité grâce à un processus stochastique pour la volatilité. C'est une caractéristique très im-

portante pour les dérivés à maturité longue où l'effet volatilité stochastique conduit à des résultats très différents de ceux obtenus avec des modèles plus conventionnels. Les modèles à volatilité stochastique permettent aussi de prendre en compte le phénomène de corrélation négative entre le sous-jacent et la volatilité en introduisant de manière explicite ce paramètre de corrélation. Les modèles à volatilité stochastique les plus populaires comprennent le modèle d'Heston, le modèle de Piterbarg, le modèle de SABR, etc.

Dans cette thèse, nous étendons les idées de Piterbarg à la famille des modèles à volatilité stochastique en rendant le concept plus général. Nous montrons en particulier comment introduire des paramètres dépendant du temps dans les modèles à volatilité stochastique et explicitons différentes formules de calcul explicite du prix d'options vanilles, permettant ainsi une calibration des paramètres du modèles extrêmement efficace.

Contents

Acknowledgements

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, I am heartily thankful to my supervisor, Alain Galli, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject.

Doctor Eric Benhamou, CEO of Pricing Partners, who is also the enterprise tutor of the CIFFRE thesis project, I would like to thank for his support in the project with his outstanding insight and useful advices.

Margaret Armstrong, professor of Department of Quantitative Finance at Mines ParisTech, for her excellent teaching in financial market class and continuous encouragement during the project.

Matthieu Glachant, Director of Cerna, Professor at Mines ParisTech, for his interesting seminar and valuable advice in my internal presentation.

Vladimir Piterbarg, Head of Quantitative Analytics at Barclays Capital, for his quick reply in email to help me understand his paper.

Changyin Huang and Sylvain Le Corff, trainees in Pricing Partners, for their help on implementing and testing the algorithms and formulas during the project.

Mohammed Miri and Pierre Gauthier, quantitative analysts in Pricing Partners, for sharing their insights.

Sésaria Ferreira, administrative head of CERNA, Mines ParisTech, for her untiring effort in encouraging the teaching staff to pursue professional growth. Likewise the staff of the Dean's Office for their relaying every communication.

Last but not the least, I offer my regards and blessings to my family, my friends and all of those who supported me in any respect during the completion of the project.

Chapter 1

Introduction

Résumé du chapitre

Les matières premières sont un sous-jacent important en finance, cette importance s'est accrue avec l'émergence du processus de globalisation dans le monde. Les marchés des matières premières telles que le pétrole, les métaux précieux et l'agriculture ont une influence très large sur l'économie mondiale.

Le premier marché à terme organisé, moderne a commencé en 1710 à la Bourse de Dojima Rice, à Osaka, au Japon. L'échange a été utilisé pendant près de 300 ans jusqu'à la Seconde Guerre mondiale. Il a ensuite été dissous entièrement en 1939, étant absorbé par le riz Agence gouvernement.

Au 19ème siècle, un contrat à terme est apparu aux États-Unis. Des produits agricoles comme le maïs et le bétail, ont été échangés par contrat à terme de Chicago et du Midwest. Le Chicago Board of Trade (CBOT) a été fondé en 1848. Le premier contrat était un contrat à terme sur le maïs, écrit le 13 Mars, 1851. Le Chicago Mercantile Exchange (CME) a été créé en 1874, sous le nom de Chicago Produit Exchange, puis réorganisé en 1919 ou il a pris son nom actuel.

Ce marché initialement des matières premières agricoles s'étend progressivement, couvrant d'autres produits comme le charbon, le pétrole, les métaux précieux, etc. Ces produits sont commercialisés dans différentes bourses partout dans le monde. En terme de volume, le pétrole brut et ses dérivés dépasse un tiers du volume traité.

Commodity is an important kind of financial asset among others, especially under the process of globalization all over the world. The markets of commodity such as fossil fuels, precious metal and agricultural have a very wide and radical influence on the world's economy.

The first modern organized futures exchange began in 1710 at the Dojima Rice Exchange in Osaka, Japan, according to West [2000]. The exchange was used for nearly 300 years till the World War II. It was then dissolved entirely in 1939, being absorbed into the Government Rice Agency.

In the 19th century, forward contract appeared in the United States. Agricultural commodity, such as corn and cattle, was traded by forward contract in Chicago and Midwest. Chicago Board of Trade (CBOT) was founded in 1848. The first contract was a forward contract on corn, written on March 13, 1851. Chicago Mercantile Exchange (CME) was established in 1874. The original name was "Chicago Produce Exchange" and then reorganized in 1919 with the name Chicago Mercantile Exchange.

The commodity market gradually extends beyond agricultural, covering other commodities such as coal, petroleum, precious metal, etc. Nowadays, the trading commodity in exchange are shown in table 1.1. These commodities are traded in various exchanges all over the world. In term of volume, crude oil and its derivative exceeds one third of total volume in all kinds of commodity. We list hereby some of the major commodity exchanges in the world in table 1.2. The major indices in commodity market are shown in table 1.3.

Table 1.1: The commodity categories

Category	Description
Agricultural	soybean, corn, wheat, cocoa, coffee, sugar, cotton, citrus, orange juice, cattle, hogs, pork bellies, etc.
Metal	gold, silver, platinum, palladium, copper, tin, zinc, nickel, aluminum, etc.
Oil and oil production	crude oil, Liquefied Petroleum Gas (LPS), gasoline, naphtha, kerosene, diesel, fuel oil, etc.
Gas and gas production	gas and Liquefied Natural Gas (LNG).
Electricity	
Coal and CO_2 emission	

The commodity trading volume has grown to a very large scale. The trading volume of Intercontinental Exchange (ICE), a leading operator of regulated global futures exchanges, clearing houses and over-the-counter (OTC) markets, reported strong futures volume growth in November 2010. Average daily volume (ADV) across ICE's futures exchanges was 1,366,669 contracts, an increase of 26% from November 2009. Year-to-date through November 30, 2010, ICE's

Table 1.2: Some commodity exchanges

Abuja Securities and Commodities Exchange
Africa Mercantile Exchange
Bhatinda Om & Oil Exchange Bathinda
Brazilian Mercantile and Futures Exchange
Chicago Board of Trade
Chicago Mercantile Exchange
Commodity Exchange Bratislava, JSC
Dalian Commodity Exchange
Dubai Mercantile Exchange
Dubai Gold & Commodities Exchange
Euronext.liffe
Hong Kong Mercantile Exchange
Indian Commodity Exchange
Intercontinental Exchange
Iranian Oil Bourse
Kansas City Board of Trade
London Metal Exchange
Minneapolis Grain Exchange
Multi Commodity Exchange
National Commodity and Derivatives Exchange
National Multi-Commodity Exchange of India Ltd
National Food Exchange
New York Mercantile Exchange
New York Board of Trade
Rosario Board of Trade
Steelbay
Winnipeg Commodity Exchange
National Spot Exchange

futures ADV was 1.325 million contracts, up 27% compared to the same period of 2009. Total futures volume in November 2010 was 28.7 million contracts. ICE futures volume exceeded 300 million contracts through November 30; the previous annual record, established in 2009, was 262 million contracts. ICE also reported OTC energy average daily commissions (ADC) of $1.33 million US dollars in the fourth quarter of 2010 and a record $1.37 million for the full year.

The derivative market of commodity contributes the majority part in commodity trading. Unlike stock market or currency market, most commodities involve physical transaction and therefore shipping cost or so-called freight cost, which make spot trading very rare in commodity market. Instead, forward trading is much more common. Consequently, the underlying commodity for the com-

Table 1.3: Some commodity indices

Continuous Commodity Index (CCI)
Astmax Commodity Index(AMCI)
Commin Commodity Index
Dow Jones-AIG Commodity Index
Goldman Sachs Commodity Index
Thomson Reuters/Jefferies CRB Index
Rogers International Commodity Index
Standard & Poor's Commodity Index
NCDEX Commodity Index
Deutsche Bank Liquid Commodity Index (DBLCI)
UBS Bloomberg Constant Maturity Commodity Index (CMCI)
Merrill Lynch Commodity index eXtra (MLCX) , etc.

modity option is not the commodity itself, but rather a standardized forward contract, or commonly known as futures contract, for that commodity. For example, a November soybean option will actually be an option for a November delivery soybean futures contract. In this sense, the options are on futures and not on the physical commodity. This fact results in a somehow slightly different form of formula for option price from other asset classes even under classic Black-Scholes model.

Nevertheless, the carry cost as well as potential benefit of holding a commodity, also known as convenience yield, changes the fundamental relation between spot price and future price. Besides, commodity market has other distinguish characters, among which seasonality and mean reversion patterns play a very important role. Both patterns describe the periodic behavior of commodity price in both short term and long term. Certain specific designed commodity models can explicitly reflect these characters in their model parameters, which presents a better financial interpretation. In these models, convenience yield or long term price themselves are assigned to follow a distinguish stochastic process and eventually make the drift term of spot price to be stochastic.

Quite differently, stochastic volatility models introduce the second stochastic process on the volatility parameter. The models in this family allows for a different and as well complex marginal distribution, which in one way has a better fit to the forward volatility surface, but in the other way adds difficulties on even the simple vanilla option price. Thus, it is crucial to have the stable and precise option formula and calibration to enjoy the advantage of the models.

In the first two chapters of this thesis, two most popular specific commodity models, namely Gibson-Schwartz model and Gabillon model, will be studied. In chapter two we will prove the mathematical equivalence between the two

models. And in the chapter three we develop a so-called model factor reduction technique and apply on both models to achieve the formulas for vanilla options based on spot and future.

In chapter four, we will focus on the stochastic volatility model. We extend the popular stochastic volatility model such as Heston model by allowing the model parameter to be time-dependent. This extension gives the more degree of freedom to stochastic volatility model in term structure, but in the other hand introduces difficulty for formula of option price and process of calibration. We will study different methods including parameter averaging, perturbation and vol of vol expansion to get the option price formula in the extended models. Some of calibration algorithm are also listed in the last part of this chapter.

Chapter 2

Modeling commodity with convenience yield

Résumé du chapitre

Contrairement à d'autres actifs classiques, les matières premières montrent des caractéristiques spécifiques : importance du convenience yield, de la saisonnalité et du retour à la moyenne. Le concept clé est ici convenience yield, qui est clairement expliqué par exemple dans Lautier [2009]. Celui-ci dicte la relation entre les prix futurs et les prix au comptant. Il ya beaucoup de modèles sur des produits de base spécifiques qui mettent l'accent sur ces concepts, citons les principaux : Gibson and Schwartz [1990], Brennan [1991], Gabillon [1991], Schwartz [1997], Hilliard and Reis [1998], Schwartz and Smith [2000], Casassus and Collin-Dufresne [2005]. A présent les modèles de Gibson et Schwartz et de Gabillon sont les plus populaires chez les chercheurs et les praticiens du marché, non seulement parce que leur interprétation en terme de rendement de commodité ou de prix à long terme, mais aussi de part leur simplicité et leur clarté.

Dans ce chapitre, nous montrerons qu'il ya équivalence entre les modèles de Gibson Schwartz et Gabillon. Cette conclusion n'est pas surprenante à si l'on considère que dans le modèle de Gibson et Schwartz le convenience yield apparaît comme un terme de dérive stochastique qui n'est pas loin d'un comportement à long terme comme dans le modèle de Gabillon.

2.1 Introduction

Unlike other asset classed, commodity has a strong presence of convenience yield, seasonality and mean reversion. The key concept here is convenience yield, which is clearly stated in Lautier [2009]. It dictates the relation between future price and spot price. There are many commodity-specialized models emphasize on these concepts, including Gibson and Schwartz [1990], Brennan [1991], Gabillon [1991], Schwartz [1997], Hilliard and Reis [1998], Schwartz and Smith [2000], Casassus and Collin-Dufresne [2005], etc. Among others, Gibson Schwartz model and Gabillon model are the most popular in researchers and market practitioners, not only because their appropriate interpretation on convenience yield or long term price, but also because their simplicity and clarity.

In this chapter, we will point out that there is a mathematical equivalence relation between Gibson Schwartz model and Gabillon model. This conclusion is not surprising provided the consideration that the former model suggests convenience yield as its drift term while the latter uses long term price as its drift term, which is an equivalent counterpart of convenience yield. The close relation between convenience yield and long term price results in the equivalence between two models.

2.2 Review of research and background

2.2.1 Convenience yield

Almost all the commodities, apart from the exception of electricity, are normally considered as storable. That is to say, people can hold their commodities in the stock for economic purpose. The following discussion will rule out electricity and focus on the storable commodity. The concept of convenience yield is closely linked to storage, which can be referred to Lautier [2005] and Lautier [2009].

Storability is essentially a very important feature of commodities. Various scholars and market participants, such as Kaldor [1939], Working [1949], Brennan [1958] and Telser [1958], have worked on the theory of storage. There are two major economic effects based on storage theory:

1. The time value of holding a commodity. It is because the inventory enables holders to store the commodity when its price is low and sell it on the market when the price goes high. It also avoids the disruption of manufacturing.

2. The holding cost. Holding cost is also called carry cost or cost of storage. It includes the land and facility to store the commodity, as well as the operation management cost.

Convenience yield is then defined as the total effect of these two aspects, which generally reflects the economic value of holding a commodity.

Of course, this description only helps understand the economic meaning of convenience yield. More elaborated and mathematical definition, as shown in Lautier [2009] and Geman [2005], will involve market structure, precisely the relation between spot price and future price.

Definition 1 (Convenience yield)**.** *Note $F_{t,T}$ to be the future price of a commodity at time t for maturity T. Note S_t to be the spot price at time t. Note r_t to be the short rate, which is the (annualized) interest rate at which an entity can borrow money for an infinitesimally short period of time from time t. Then we define convenience yield y_t by satisfying the arbitrage-free condition:*

$$F_{t,T} = S_t e^{\int_t^T (r_s - y_s)\mathrm{d}s} \tag{2.1}$$

or equivalently,

$$y_s = r_s - \frac{\mathrm{d}}{\mathrm{d}s}\left(\log \frac{F_{t,s}}{S_t}\right) \tag{2.2}$$

with $t < s < T$.

The definition of convenience yield in equation (2.2) is in continuous form, which presents the average of convenience yield from time t to T. In real world application, since the future values are normally available at some discretion time points, we can then defined convenience yield y_t accordingly in piecewise constant form by the following definition:

Definition 2 (Convenience yield in discrete form)**.** *Let $t = t_0 < t_1 < t_2 < ... < t_n = T$. Note F_{t,t_i} to be the future price at time t for maturity t_i. Note S_t to be the spot price. Note $D(t_i, t_j)$ to be the discount factor between time t_i and t_j, which stands for the present value at time t_i of a future cash flow of 1 unit currency at time t_j.*

Convenience yield y_s is then given by the following equation:

$$y_s = \frac{1}{t_{i+1} - t_i} \log\left(\frac{F_{t,t_i}}{F_{t,t_{i+1}}} \frac{1}{D(t_i, t_{i+1})}\right), t_i \leqslant s < t_{i+1} \tag{2.3}$$

It is very easy to verify that this definition fulfills the arbitrage-free condition in equation (2.1) by using $D(t_i, t_{i+1}) = e^{-\int_{t_i}^{t_{i+1}} r_s \mathrm{d}s}$. This definition of convenience yield in discrete form is practically used as the formula to calculate convenience yield from market data.

The notion of discount factor is very useful in following paragraphs. We here separately write it as a definition.

Remark For a fixed discount rate, r, continuously compounded discount rate from time t to T, then we have the relation $D(t, T) = e^{-r(T-t)}$. Normally discount factor is provided by interest rate market data and very easy to access. So we will treat it as a known variable hereafter.

Convenience yield v.s. stock dividend

We compare arbitrage-free condition (2.1) to its counterpart in stock market. We continue use the same symbol: $F_{t,T}$, S_t and r_t for future, spot and interest rate respectively. Note d_t to be the dividend of the stock. Then we have

$$F_{t,T} = S_t e^{\int_t^T (r_s - d_s)\mathrm{d}s} \tag{2.4}$$

Notice that it is exactly the same form as equation (2.1). This fact encourages us to migrate the stock market models to commodity. All we need to do is to simply replace dividend term by convenience yield.

Term structure: backwardation and contango

We copy here the arbitrage-free condition, equation (2.1), to facilitate reading. The relation between future and spot is,

$$F_{t,T} = S_t e^{\int_t^T (r_s - y_s)\mathrm{d}s} \tag{2.5}$$

The term $\int_t^T (r_s - y_s)\mathrm{d}s$ can be negative or positive. We call these two different term structures *backwardation* and *contango* respectively.

Backwardation means $\int_t^T (r_s - y_s)\mathrm{d}s$ is negative. In this case, future price $F_{t,T}$ is smaller than spot S_t. It happens when interest rate in low and convenience yield is high. For example, when the Gulf war broke up, holding oil became particularly profitable at that time. And convenience yield of oil became so high that caused backwardation in the future market.

Contrarily, when $\int_t^T (r_s - y_s)\mathrm{d}s$ is positive, we have contango. In this case, future price $F_{t,T}$ is greater than spot S_t. The structure of future curve becomes an increasing function against maturity time.

2.2.2 Gibson Schwartz model

Gibson Schwartz model, first introduced by Gibson and Schwartz [1990], is based on a two-factor diffusion. The spot price S_t is assumed to follow a lognormal diffusion but with a stochastic drift. Gibson and Schwartz [1990] assume that convenience yield, δ_t, is a stochastic process, explicitly following a mean reverting normal diffusion, which leads to the following diffusion for the commodity spot price on risk-neutral measure:

$$\frac{\mathrm{d}S}{S} = (r - \delta)\mathrm{d}t + \sigma_1 \mathrm{d}z_1 \tag{2.6}$$

$$\mathrm{d}\delta = (k(\alpha - \delta) - \lambda\sigma_2)\mathrm{d}t + \sigma_2 \mathrm{d}z_2 \tag{2.7}$$

where r is the risk free rate, z_1 and z_2 are two correlated Brownian motions with correlation ρ.

2.2.3 Gabillon model

Gabillon [1991] created this model to account for long term trend. He argued that it was beneficial to use a two-factor diffusion to capture the difference between short and long term effects. He introduced a long term trend factor that would influence the convenience yield. This led him to assume that the spot prices, S, follows a log-normal diffusion but with a stochastic drift term. Instead of assuming a stochastic convenience yield like in Gibson and Schwartz [1990] model, Gabillon assumed that the drift term is influenced by a long term factor that is itself stochastic. This long term factor denoted by L follows also a log-normal diffusion.

In the original paper the model parameters are assumed to be constant. To allow an easy calibration across time, we will extend the hypothesis of Gabillon and assume that these parameters are time dependent. This setting is very interesting as we capture the long term trend stochasticity through the second factor and capture time varying value through time dependent model parameters. We will take the traditional assumption in mathematical finance and assume that the uncertainty is represented by a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a two dimensional Brownian motion. The two components of the Brownian motion will be denoted by z_1, z_2 and will be assumed to be correlated with a constant correlation.

$$\mathrm{d}z_1 \mathrm{d}z_2 = \rho \mathrm{d}t$$

This leads us to assume that the spot price S follow under the historical probability,

$$\frac{\mathrm{d}S}{S} = k \ln \frac{L}{S} \mathrm{d}t + \sigma_S \mathrm{d}z_1 \tag{2.8}$$

$$\frac{\mathrm{d}L}{L} = \mu_L \mathrm{d}t + \sigma_L \mathrm{d}z_2 \tag{2.9}$$

where in the above equations, we have removed any obvious reference to time to simplify the notation, like for instance S_t, L_t, z_1^t, z_2^t.

2.2.4 Previous research

As two of the most popular models in commodity world, Gibson Schwartz model and Gabillon model are actually mathematically equivalent to each other. The previous researches includes Gabillon [1991] and Schwartz and Smith [2000]. The former studied and compared the economic interpretation of the parameters of both models. The latter built a Gabillon style model, namely short-term/long-term model, and proved that the model was equivalent to Gibson Schwartz model by change of variable. And our result in this thesis goes in the other

direction: the equivalence relation between Gabillon model and short-term/long-term model. Our result together with the result of Schwartz and Smith [2000] complete the proof of the equivalence relation between Gibson Schwartz model and Gabillon model.

2.3 Equivalence between Gibson Schwartz model and Gabillon model

We establish the equivalence relation through a third model. The idea is to respectively prove Gibson Schwartz model and Gabillon model are both equivalent to a third model. Consequently, they are equivalent to each other. Moreover, we interpret the equivalence relation in term of the economic significance of the parameters in both models.

The significance of this equivalent relation is to establish a close connection between two most popular models in commodity derivative pricing. Although the two models are mathematically equivalent, the market practitioners have their preference on choosing between the two. It is because they have good knowledge about the economic meaning of the parameters of the model, which guarantees them to better understand and use the model.

In fact, the equivalence between the third model and Gibson Schwartz model were already proved in Schwartz and Smith [2000]. The model is called short-term/long-term model in the paper. The definition is as follows.

Definition 3 (Short-term/long-term model)**.** *Let S_t denote the spot price of a commodity at time t. We define,*

$$\log(S_t) = \chi_t + \xi_t, \tag{2.10}$$

where χ_t and ξ_t present short term price and long term price respectively. They follow the stochastic process:

$$d\chi_t = -\kappa\chi_t dt + \sigma_\chi dz_\chi \tag{2.11}$$
$$d\xi_t = \mu_\xi dt + \sigma_\xi dz_\xi \tag{2.12}$$

Here z_χ and z_ξ are two Brownian motion with correlation ρ. Parameters κ, μ_ξ, σ_χ and σ_ξ are constant.

The first step is to prove the equivalence relation between short-term/long-term mode and Gibson Schwartz model. It is done in Schwartz and Smith [2000]. The key here is the change of variable .

$$\chi_t = \frac{1}{\kappa}(\delta_t - \alpha) \tag{2.13}$$

$$\xi_t = \log S - \frac{1}{\kappa}(\delta_t - \alpha) \tag{2.14}$$

Then simply inject the stochastic diffusion of Gibson Schwartz model, equation (2.6) and (2.7), into the equations above. We have the derivation as follows.

$$\frac{\mathrm{d}S}{S} = (\mu - \delta)\mathrm{d}t + \sigma_1 \mathrm{d}z_1 \tag{2.15}$$

$$\mathrm{d}\delta = (k(\alpha - \delta) - \lambda\sigma_2)\,\mathrm{d}t + \sigma_2 \mathrm{d}z_2 \tag{2.16}$$

Note $X_t = \log S_t$, we have

$$\mathrm{d}X_t = \left(\mu - \delta_t - \frac{1}{2}\sigma_S^2\right)\mathrm{d}t + \sigma_S \mathrm{d}z_1 \tag{2.17}$$

Using 2.13, we have

$$\chi_t = \frac{1}{\kappa}(\delta_t - \alpha) \tag{2.18}$$

$$\Rightarrow \mathrm{d}\delta_t = \kappa(\alpha - \delta_t)\mathrm{d}t + \kappa\sigma_\chi \mathrm{d}z_\chi \tag{2.19}$$

$$\xi_t = \log S - \frac{1}{\kappa}(\delta_t - \alpha) \tag{2.20}$$

$$\Rightarrow \mathrm{d}X_t = (\mu_\xi + \alpha - \delta_t)\mathrm{d}t + \sigma_\xi \mathrm{d}z_\xi + \sigma_\chi \mathrm{d}z_\chi \tag{2.21}$$

Assume $z_\chi = \rho z_\xi + \sqrt{1-\rho^2} z_\xi^\perp$, where $z_\xi^\perp$ is independent to z_ξ.

$$\Rightarrow \mathrm{d}\sigma_\xi \mathrm{d}z_\xi + \sigma_\chi \mathrm{d}z_\chi = (\sigma_\xi + \rho\sigma_\chi)\mathrm{d}z_\xi + \sqrt{1-\rho^2}\sigma_\chi \mathrm{d}z_\xi^\perp \tag{2.22}$$

Then we note

$$\|\sigma\|_t^2 = \int_0^t \left((\sigma_\xi + \rho\sigma_\chi)^2 + (1-\rho^2)\sigma_\chi^2\right)\mathrm{d}s \tag{2.23}$$

So

$$\|\sigma\|_t = \sigma_1 \tag{2.24}$$

$$z_1(t) = \int_0^t \left(\frac{\sigma_\xi + \rho\sigma_\chi}{\|\sigma\|_t}\mathrm{d}z_\xi + \frac{\sqrt{1-\rho^2}\sigma_\chi}{\|\sigma\|_t}\mathrm{d}z_\xi^\perp\right) \tag{2.25}$$

According to Levy Theorem, $z_1(t)$ is Brownian motion. So the two models are equivalent.

$$\begin{aligned} \mathrm{d}\chi_t &= \frac{1}{\kappa}\mathrm{d}\delta_t && (2.26)\\ &= (\alpha - \delta_t)\mathrm{d}t + \frac{\sigma_2}{\kappa}\mathrm{d}z_2 && (2.27)\\ &= -\kappa\chi_t\mathrm{d}t + \frac{\sigma_2}{\kappa}\mathrm{d}z_2 && (2.28) \end{aligned}$$

and

$$\begin{aligned} \mathrm{d}\xi_t &= \mathrm{d}\log S - \frac{1}{\kappa}\mathrm{d}\delta_t && (2.29)\\ &= (\mu - \delta_t - \frac{1}{2}\sigma_1^2)\mathrm{d}t + \sigma_1\mathrm{d}z_1 - (\alpha - \delta_t)\mathrm{d}t - \frac{\sigma_2}{\kappa}\mathrm{d}z_2 && (2.30)\\ &= (\mu - \alpha - \frac{1}{2}\sigma_1^2)\mathrm{d}t + \sigma_1\mathrm{d}z_1 - \frac{\sigma_2}{\kappa}\mathrm{d}z_2 && (2.31) \end{aligned}$$

Here we use the risk neutralized version of both short term/long term model and Gibson Schwartz model to exclude the risk premium parameter. Under non-arbitrage principle, the drift term of both models should equal to $\int_0^t r_s\mathrm{d}s$. So we just need to take care of stochastic term for the spot diffusion. Comparing these forms with short term/long term model, we can see that the two models are equivalent if we relate the parameters of the two models.

In the other hand, the equivalence relation between long term/short term model and Gabillon model is also easy to prove. The change of variables here is

$$\begin{aligned} \xi_t &= \log L && (2.32)\\ \chi_t &= \log S - \log L && (2.33) \end{aligned}$$

Inspired by the above result, we can similarly prove the equivalence relation between long term/short term model and Gabillon model. We inject the diffusion of Gabillon model, equation (2.8) and (2.9), in the equations above. We can re-find the diffusion of long term/short term model.

$$\begin{aligned} \mathrm{d}\chi_t &= \mathrm{d}\log S - \mathrm{d}\log L && (2.34)\\ &= -\kappa\chi_t\mathrm{d}t - \sigma_S\mathrm{d}W_1 + \sigma_L\mathrm{d}W_2 && (2.35) \end{aligned}$$

and

$$d\xi_t = d\log L \tag{2.36}$$
$$= \left(\mu_L - \frac{1}{2}\sigma_L^2\right)dt + \sigma_L dW_2 \tag{2.37}$$

Therefore, we prove the equivalence relation between long term/short term model and Gabillon model. This result is new to our knowledge.

Now we put the two parts together and we can get the equivalence relation between Gibson Schwartz model and Gabillon model. It is given by the change of variables as follows:

$$\log S = X_t \tag{2.38}$$
$$\log L = X_t - \frac{1}{\kappa}(\delta_t - \alpha) \tag{2.39}$$

For the equivalence of parameters, the calculation is a bit lengthy but very straightforward. We just list the result in table 2.1. Given the high similarity between Gabillon model and short-term/long-term model, we observe the the result in table 2.1 is in the same form as the result for the equivalence relation between Gibson Schwartz model and short-term/long-term model, which is shown in Schwartz and Smith [2000].

Table 2.1: The Relationships Between Parameters in Gabillon Model and Gibson Schwartz model

Gabillon model parameter		
Symbols	Description	Equivalence in Gibson Schwartz model
κ	Short-term mean-reversion rate	κ
σ_S	Short-term volatility	σ_2/κ
μ_L	Equilibrium drift rate	$(\mu - \alpha - \frac{1}{2}\sigma_1^2)$
σ_L	Equilibrium volatility	$(\sigma_1^2 + \sigma_2^2/\kappa^2 - 2\rho\sigma_1\sigma_2/\kappa)^{1/2}$
ρ	Correlation in increments	$(\rho\sigma_1 - \sigma_2/\kappa)(\sigma_1^2 + \sigma_2^2/\kappa^2 - 2\rho\sigma_1\sigma_2/\kappa)^{1/2}$

2.4 Summary

In this chapter, we revisited the concept of convenience yield and the models specialized on commodity, namely Gibson Schwartz model and Gabillon model, which emphasizes on stochastic convenience yield and stochastic long term price respectively. As the main result of the chapter, we proved that the two models are eventually equivalent in mathematics.

The explanation of the equivalence of the two models involves the relation of the concept of long term price and convenience yield. As we know, convenience yield is the total benefit and carry cost on holding a commodity, which consequently decides the long term price of the commodity. Gibson Schwartz model assumes that convenience yield follows a Brownian motion with mean reversion, while Gabillon model suggests that the logarithm of long term price is also a Brownian motion with mean reversion. Apparently, these assumptions lead to the equivalence of the two models.

Chapter 3

Model factor reduction technique

Résumé du chapitre

Le résultat principal de ce chapitre est l'introduction et l'application de la technique de réduction de modèle factoriel. C'est une technique qui permet de trouver –sous certaines conditions-un équivalent à un facteur qui a la même distribution marginale de spot et donc du prix d'une option Européenne, que le modèle initial à plusieurs facteurs. Nous appliquons ensuite cette la technique aux modèles de Gibson Schwartz et de Gabillon. En outre, nous allons donner explicitement la distribution marginale des spots et futures de ces deux modèles, et en déduire une formule explicite du prix de l'option vanille. La technique s'applique généralement à tout modèle multi-facteurs de prix sous l'hypothèse que la dérive est un processus de diffusion gaussien. C'est précisément le cas des modèles de commodités que sont les deux modèles mentionnés ci-dessus. La technique peut également s'appliquer à des modèles pour les autres classes d'actifs telles que les modèles stochastiques avec dividendes pour les marchés des actions.

L'organisation de ce chapitre commence en présentant notre technique de réduction de modèle factoriel sur un modèle de tarification à deux facteurs généraux. La première étape consiste à obtenir l'équivalent à un facteur puis la distribution marginale des prix au comptant, ce qui permet de dériver la formule de prix de l'option de vanille. Ensuite, nous appliquons le résultat général aux deux modèles ci-dessus modèle.

3.1 Introduction

The main result in this chapter is the introduction and application of model factor reduction technique. It is a technique to find an equivalent one-factor model that has the same marginal distribution of spot and therefore vanilla option price with the original two-factor model under certain condition. Interestingly, we apply the technique to Gibson Schwartz model and Gabillon model. We show that they are equivalent to one-factor model when pricing futures and vanilla option. In addition, we will explicitly express the marginal distribution of spot and future price of both models, and derive the formula of vanilla option on spot and future thusly.

The equivalent one-factor model admits a weak solution, which has the same one-dimensional marginal probability distribution and therefore same price for vanilla options. Moreover, this distribution can be explicitly expressed under certain conditions. The result can consequently induct closed formulas for future and vanilla option price. The technique generally applies to any multi-factor pricing model under assumption of normally distributed drift, which are precisely, in commodity modeling, the stochastic drift such as convenience yield in Gibson Schwartz model or long term price in Gabillon model. Obviously, the technique can also apply to models for other asset classes such as stochastic dividend models in stock market.

The organization of this chapter begins with introducing our model factor reduction technique on a general two-factor pricing model. The first step is to get the equivalent one-factor model and then the marginal distribution of spot price, which can derive the formula for vanilla option price. Next we apply the general result to Gabillon model and Gibson Schwartz model, verifying that both models fit the assumption condition. Once the equivalent reduced one-factor model and the marginal distribution are obtained, it is straightforward to get the formulas for vanilla option price by applying the result of general two-factor model.

3.2 Review of research

The idea of model factor reduction technique is inspired by the paper of Gyöngy [1986]. In his paper, Gyöngy considered a general form of a stochastic process $\mathrm{d}x(t) = \delta(t,\omega)\ \mathrm{d}W(t) + \beta(t,\omega)\mathrm{d}t$. He proposed an equivalent process, replacing stochastic parameters by non-random parameters. The new process admits a weak solution having the same one-dimensional probability distribution.

Proposition 4 (Gyöngy [1986]). *Let $\xi(t)$ be a stochastic process starting from 0 with*

Ito differential

$$\mathrm{d}\xi(t) = \delta(t,\omega)\mathrm{d}W(t) + \beta(t,\omega)\mathrm{d}t, \tag{3.1}$$

where $(W(t), \mathscr{F}_t)$ *is a Wiener process,* δ *and* β *are bounded* $\mathscr{F}_t$*-non-anticipative processes such that* $\delta\delta^T$ *is uniformly positive definite. Then it is proved that there exists a stochastic differential equation*

$$\mathrm{d}x(t) = \sigma(t,x(t))\mathrm{d}W(t) + b(t,x(t))\mathrm{d}t \tag{3.2}$$

with non-random coefficients which admits a weak solution $x(t)$ *having the same one-dimensional probability distribution as* $\xi(t)$ *for every* t*. The coefficients* σ *and* b *have a simple interpretation:*

$$\sigma(t,x(t)) = \left(\mathbb{E}[\delta\delta^T(t)|\xi(t) = x]\right)^{\frac{1}{2}} \tag{3.3}$$

$$b(t,x) = \mathbb{E}[\beta(t)|\xi(t) = x] \tag{3.4}$$

From this result, we can see the process $x(t)$ is with non-random coefficients, which means the model is reduced to one-factor. However, the condition that δ and β are bounded limits the application of this result, since two-factor models usually cannot satisfy this constraint. For example, in Gibson Schwartz model and Gabillon model, the drift term β is normally distributed, hence not bounded. Follow this line, we make one step further. We will impose normal distribution on $\beta(t,\omega)$, which holds for a large class of commodity pricing models. Although normal distribution isn't bounded, we still can calculate the non-random coefficients for reduced equivalent one-factor model and express this one-dimensional probability distribution explicitly. Then, we show an equivalent one-factor model that presents the same marginal distribution at any time t, as shown in the definition 6. Once we know the distribution of the spot price, it is then possible to compute a closed formula for the future price and vanilla option price, shown and proved in proposition 10, 11 and 12 in following sections for the general case. These results of general case are then applied to Gibson Schwartz model and Gabillon model in the following sections.

3.3 Model factor reduction technique in general form

Assume that S_t is a stochastic process with the following dynamic:

$$\frac{\mathrm{d}S_t}{S_t} = (r + \alpha_{t,\omega})\mathrm{d}t + \sigma(t)\mathrm{d}W_t \tag{3.5}$$

where W_t is a (potentially multi-dimensional) Wiener process under filtration $\mathscr{F}_t$ and risk-neutral measure Q. To be clear, all the calculation in this section is always on risk-neutral measure Q. r is risk free rate. Drift term $\alpha_{t,\omega}$ is stochastic.

And we assume that it follows another diffusion process, which is the case in pratical for models such as Gibson Schwartz model and Gabillon model. ω represents the full stochastic information. In particular, the path-wise behavior of X_t may depend, via the random coefficients $\alpha_{t,\omega}$ in a complicated way on the past filtration $\mathscr{F}_t$.

Assume that $X_t = \log S_t$. So X_t is a stochastic process with the following dynamic:

$$\mathrm{d}X_t = \left(r - \frac{1}{2}\sigma(t)^2 + \alpha_{t,\omega}\right)\mathrm{d}t + \sigma(t)\mathrm{d}W_t \tag{3.6}$$

Equation (3.6) depicts a general form for pricing models under the assumption of a deterministic volatility $\sigma(t)$. Since the parameters $\alpha_{t,\omega}$ is a random variable, the process is multi-factor. Some examples can be found in section 3.4.

Lemma 5. *In process* $\mathrm{d}X_t = (r - \frac{1}{2}\sigma(t)^2 + \alpha_{t,\omega})\mathrm{d}t + \sigma(t)\mathrm{d}W_t$*, as* $\alpha_{t,\omega}$ *is a normal diffusion process for all* $t > 0$*, then* $X_t = X_0 + rt - \frac{1}{2}\int_0^t \sigma(s)^2\mathrm{d}s + \int_0^t \alpha_{s,\omega}\mathrm{d}s + \int_0^t \sigma(s)\mathrm{d}W_s$ *is a normal variable for any* $t > 0$*.*

Proof. Since $\alpha_{t,\omega}$ is normal, so $\int_0^t \alpha_{s,\omega}\mathrm{d}s$ is normal. The Ito integral $\int_0^t \sigma_s \mathrm{d}W_s$ is normal as the integrand σ_s is deterministic. X_t, being the sum of normals, is normal for all t. □

Under the assumption that $\alpha_{t,\omega}$ is a normal variable, we note $\sigma_\alpha(t)$ as the standard deviation (square root of variance) of $\alpha_{t,\omega}$ for future convenience.

Definition 6. *For afterward convenience, we define here* $\mu(t,T)$ *and* $\xi(t,T)$*. Denote* $\mu(t,T)$ *the expectation of* $X_T - X_t$ *at time* t*. Denote* $\xi(t,T)$ *the variance of* $X_T - X_t$ *at time* t*. They are simply defined by:*

$$\mu(t,T) = \mathbb{E}[X_T - X_t|\mathscr{F}_t] = \mathbb{E}\left[\int_t^T \alpha_{s,\omega}\mathrm{d}s \middle| \mathscr{F}_t\right] + r(T-t) - \frac{1}{2}\int_t^T \sigma(s)^2\mathrm{d}s \tag{3.7}$$

$$\xi(t,T) = \mathrm{var}(X_T - X_t|\mathscr{F}_t) = \mathrm{var}\left(\int_t^T \alpha_{s,\omega}\mathrm{d}s + \int_t^T \sigma(s)\mathrm{d}W_s \middle| \mathscr{F}_t\right) \tag{3.8}$$

Remark We can write $\mathbb{E}[X_T|\mathscr{F}_t] = X_t + \mu(t,T)$ and $\mathrm{var}(X_T|\mathscr{F}_t) = \xi(t,T)$.

Now we can construct a new model with non-random time-dependent parameter. This one-factor model should have the same distribution.

$$Y_T = X_t + \mu(t,T) + \int_t^T \sqrt{\xi'(t,s)}\mathrm{d}W_s \tag{3.9}$$

where $\mu(t,T)$ and $\xi(t,T)$ are given by equation (3.7) and (3.8). And $\xi'(t,s) = \frac{\partial\xi(t,s)}{\partial s}$. It is easily to verify that $\mathbb{E}[Y_T|\mathscr{F}_t] = X_t + \mu(t,T)$ and $\mathrm{var}(Y_T|\mathscr{F}_t) = \xi(t,T)$. From lemma 7, we know that $\xi'(t,s)$ is positive, so the process Y_T is well defined.

Lemma 7. *$\xi(t,T)$ is given by equation (3.8). Assume α_t has form $\mathrm{d}\alpha_t = \mu_\alpha(t)\mathrm{d}t + \sigma_\alpha(t)\mathrm{d}W_t^\alpha$. $\mu_\alpha(t)$ and $\sigma_\alpha(t)$ are deterministic. Then we have $\forall t$, $\frac{\partial\xi(t,T)}{\partial T} \geqslant 0$.*

Proof. Note $\mu_\alpha(t,T) = \int_t^T \mu_\alpha(s)\mathrm{d}s$.

$$\int_t^T \alpha_s \mathrm{d}s = \mu_\alpha(t,T) + \int_t^T \mathrm{d}u \int_t^T \mathbf{1}_{u\geqslant s}\sigma_\alpha(s)\mathrm{d}s\mathrm{d}W_s^\alpha \tag{3.10}$$

$$= \mu_\alpha(t,T) + \int_t^T (T-s)\sigma_\alpha(s)\mathrm{d}W_s^\alpha \tag{3.11}$$

Therefore,

$$\begin{aligned} X_T =& X_t + r(T-t) - \frac{1}{2}\int_t^T \sigma(s)^2\mathrm{d}s + \mu_\alpha(t,T) \\ &+ \int_t^T (T-s)\sigma_\alpha(s)\mathrm{d}W_s^\alpha + \int_t^T \sigma(s)\mathrm{d}W_s \end{aligned}$$

$$\mathbb{E}(X_T|\mathscr{F}_t) = X_t + r(T-t) - \frac{1}{2}\int_t^T \sigma(s)^2\mathrm{d}s + \mu_\alpha(t,T) \tag{3.12}$$

$$\mathrm{var}(X_T|\mathscr{F}_t) = \mathbb{E}\left[\left(\int_t^T (T-s)\sigma_\alpha(s)\mathrm{d}W_s^\alpha + \sigma(s)\mathrm{d}W_s\right)^2\right] \tag{3.13}$$

Note ρ the correlation between W_t^α and W_t.
For any $t < T$, we have

$$\begin{aligned} \xi(t,T) &= \mathrm{var}(X_T|\mathscr{F}_t) \\ &= \int_t^T (T-s)^2\sigma_\alpha^2(s)\mathrm{d}s + \int_t^T \sigma^2(s)\mathrm{d}s + 2\rho\int_t^T (T-s)\sigma_\alpha(s)\sigma(s)\mathrm{d}s \\ &\geqslant \int_t^T [(T-s)^2\sigma_\alpha^2(s) + \sigma^2(s) - 2\sigma_\alpha(s)\sigma(s)]\mathrm{d}s \\ &= \int_t^T [(T-s)\sigma_\alpha(s) - \sigma(s)]^2\mathrm{d}s \\ &\geqslant 0 \end{aligned}$$

So we have,

$$\begin{aligned} \frac{\partial\xi(t,T)}{\partial T} &= \lim_{\delta\to 0}\xi(t,T+\delta) - \xi(t,T) \\ &= \lim_{\delta\to 0}\xi(T,T+\delta) \\ &\geqslant 0 \end{aligned}$$

This proves lemma 7. □

The process $Y_T = X_t + \mu(t,T) + \int_t^T \sqrt{\xi'(t,s)}\mathrm{d}W_s$ is a one-factor model with non-random time-dependent parameters. At any time $s > t$, Y_s is a normally distributed variable with $\mathbb{E}[Y_t] = X_0 + \mu(0,t)$, $\mathrm{var}(Y_t) = \xi(0,t)$. So it has the same one-dimensional marginal distribution as our original process $\mathrm{d}X_t = \alpha_{t,\omega}\mathrm{d}t + \sigma_t \mathrm{d}W_t$ at any time t.

From the calculation above, we know

$$\mu(t,T) = \mu_\alpha(t,T) + r(T-t) - \frac{1}{2}\int_t^T \sigma(s)^2 \mathrm{d}s. \tag{3.14}$$

Therefore from $Y_T = X_t + \mu(t,T) + \int_t^T \sqrt{\xi'(t,s)}\mathrm{d}W_s$, we can have $\forall s > t$,

$$\mathrm{d}Y_s = \left(r - \frac{1}{2}\sigma(s)^2 + \mu_\alpha(s)\right)\mathrm{d}s + \sqrt{\xi'(t,s)}\mathrm{d}W_s \tag{3.15}$$

The advantage of the one-factor model, $Y_T = X_t + \mu(t,T) + \int_t^T \sqrt{\xi'(t,s)}\mathrm{d}W_s$, is that there are existing closed formula for future and vanilla option price. We show them in proposition 10 and 11. Theoretically, with the normal distribution in definition 6, all kinds of non-path-dependent options can be priced as in the equivalent one-factor model.

In order to use the Black Scholes formula, we list the result of Black Scholes here for convenience. Black Scholes model is initially introduced by Black and Scholes [1973]. The model suggests the marginal distribution of underlying follows a log-normal distribution. It is now still popular in the market thanks to its simplicity and stability. Black and Scholes [1973] also presented Black Scholes formula for the price of a vanilla option.

Definition 8 (Black Scholes model with dividend)**.** *Assume S_t to be the spot price of an underlying (such as commodity) at time t. The diffusion of Black Scholes model follows:*

$$\frac{\mathrm{d}S_t}{S_t} = (r - q(t))\,\mathrm{d}t + \sigma(t)\mathrm{d}W_t \tag{3.16}$$

where $\mu_S(t)$, $q(t)$ and $\sigma_S(t)$ are deterministic. $q(t)$ is dividend. Note $X_t = \log S_t$. Applying Ito lemma, we have

$$\mathrm{d}X_t = \left(r - \frac{1}{2}\sigma(t)^2 - q(t)\right)\mathrm{d}t + \sigma(t)\mathrm{d}W_t \tag{3.17}$$

Proposition 9 (Black Scholes formula - price of a vanilla option)**.** *Under Black Scholes model in Definition 8, we can express the price of a vanilla option as follows:*

$$V(t) = S_t N(d_1) e^{-\int_t^T q(s)\mathrm{d}s} - KN(d_2) e^{-r(T-t)} \tag{3.18}$$

where r is interest rate. S_t is the spot price observed at time t. $N(x) = \frac{1}{2\pi}\int_{-\infty}^{x} e^{-t^2/2}\mathrm{d}t$ is the cumulative distribution function of standard normal distribution. And $d_1 = \frac{\log(S_t/K)+r(T-t)-\int_t^T q(s)\mathrm{d}s}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}$, $d_2 = d_1 - \sigma\sqrt{T-t}$.

When $\sigma_S(t)$ is time dependent, d_1 and d_2 are changed to $d_1 = \frac{\ln(S_t/K)+r(T-t)-\int_t^T q(s)\mathrm{d}s}{\sqrt{v}} + \frac{1}{2}\sqrt{v}$, $d_2 = d_1 - \sqrt{v}$, with $v = \int_t^T \sigma(s)^2\mathrm{d}s$

Since the distribution S_t is log-normal and the average and variance of $\log S_t$ are known, it needs only straightforward calculation to get Black Scholes formula by integrating S_t from K to ∞. The detailed proof can be found in Black and Scholes [1973] and we won't list here. Example calculation for Black Scholes model with dividend can also be found in page 237 - 238 of Shreve [2004].

Proposition 10 (Formula for future)**.** *If X_t follows equation $\mathrm{d}X_t = \alpha_{t,\omega}\mathrm{d}t + \sigma_t \mathrm{d}W_t$ with $\alpha_{t,\omega}$ being a normal, assuming that S_t is a stochastic process defined by $S_t = e^{X_t}$, then the price of the future $F(t,T)$, observed at time t, expiring at time T, and given by $\mathbb{E}[S_T|\mathscr{F}_t]$ has the following closed formula:*

$$F(t,T) = \mathbb{E}[S_T|\mathscr{F}_t] = S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)} \tag{3.19}$$

Here $\mu(t,T) = \mathbb{E}\left[\int_t^T \alpha_{s,\omega}\mathrm{d}s \middle| \mathscr{F}_t\right]$ and $\xi(t,T) = \mathrm{var}\left(\int_t^T \alpha_{s,\omega}\mathrm{d}s + \int_t^T \sigma_s \mathrm{d}W_s \middle| \mathscr{F}_t\right)$ are given by equation 3.7 and 3.8.

Proof. Since S_t is a log-normal variable, the proposition is obviously proved by employing the property of a log-normal variable. □

Proposition 11 (Formula of vanilla option)**.** *If X follows equation 3.9 under risk neutral measure, with $\alpha_{t,\omega}$ being a normal, then the price of a vanilla call on the spot $S_t = e^{X_t}$ with strike K is given by:*

$$V = S_t N(d_1) e^{-p(t,T)} - KN(d_2) e^{-r(T-t)} \tag{3.20}$$

where $N(\cdot)$ is standard normal cumulative distribution function;

$$\begin{aligned} d_1 &= \frac{\log(S_t/K) + r(T-t) - p(t,T)}{\sqrt{\xi(t,T)}} + \frac{1}{2}\sqrt{\xi(t,T)} \\ d_2 &= d_1 - \sqrt{\xi(t,T)} \\ p(t,T) &= r(T-t) - \frac{1}{2}\xi(t,T) - \mu(t,T) \end{aligned}$$

with $\mu(t,T)$ and $\xi(t,T)$ in equation 3.8 and $S_t = e^{X_t}$.

Proof. Simply use Black Scholes formula with on the stochastic process 3.15

$$\mathrm{d}Y_s = \left(r - \frac{1}{2}\sigma(s)^2 + \mu_\alpha(s)\right)\mathrm{d}s + \sqrt{\xi'(t,s)}\mathrm{d}W_s$$

We have

$$\begin{aligned} \nu &= \int_t^T \left(\sqrt{\xi'(t,s)}\right)^2 \mathrm{d}s \\ &= \xi(t,T) \end{aligned}$$

and $\forall s > t$, the dividend term in Black-Scholes model is

$$\begin{aligned} q(s) &= \left(r - \frac{1}{2}\left(\sqrt{\xi'(t,s)}\right)^2\right) - \left(r - \frac{1}{2}\sigma(s)^2 + \mu_\alpha(s)\right) \\ \Rightarrow p(t,T) &\overset{\text{def}}{=} \int_t^T q(s)\mathrm{d}s \\ &= \frac{1}{2}\left(\int_t^T \sigma(s)^2\mathrm{d}s - \xi(t,T)\right) - \mu_\alpha(t,T) \\ &= r(T-t) - \frac{1}{2}\xi(t,T) - \mu(t,T) \end{aligned}$$

Then simply use the result of proposition 9 to get the result of this proposition. □

Proposition 12 (Vanilla option on future)**.** *If X follows equation 3.6 with $\alpha(t,\omega)$ being a normal and the spot process $S = e^X$ is a log normal process. Assume a vanilla call option has maturity at time t. The underlying is future F_t^T and strike is K. Then the price V of the option at time t is given by:*

$$V = e^{-r(T-t)}\left(S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)}N(d_1) - KN(d_2)\right)$$

where $N(\cdot)$ is standard normal cumulative distribution function;

$$\begin{aligned} d_1 &= \frac{\log\left(S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)}/K\right)}{\sqrt{\xi(t,T)}} + \frac{1}{2}\sqrt{\xi(t,T)} \\ d_2 &= d_1 - \sqrt{\xi(t,T)} \end{aligned}$$

with $\mu(t,T)$ and $\xi(t,T)$ in equation 3.8 and $S_t = e^{X_t}$.

Proof. We have to compute the following expectation:

$$\begin{aligned} e^{-r(T-t)}\mathbb{E}\left[\left(F_t^T - K\right)^+ \middle| \mathscr{F}_t\right] &= e^{-r(T-t)}\mathbb{E}\left[\left(\mathbb{E}\left[S_T|\mathscr{F}_t\right] - K\right)^+ \middle| \mathscr{F}_t\right] \\ &= e^{-r(T-t)}\mathbb{E}\left[\left(S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)} - K\right)^+ \middle| \mathscr{F}_t\right] \end{aligned}$$

where

$$\mu(t,T) = \mathbb{E}\left[\int_t^T \alpha_{s,\omega} \mathrm{d}s \middle| \mathscr{F}_t\right]$$
$$\xi(t,T) = \mathrm{var}\left(\int_t^T \alpha_{s,\omega} \mathrm{d}s + \int_t^T \sigma_s \mathrm{d}W_s \middle| \mathscr{F}_t\right)$$

leading to

$$e^{-r(t-t)}\mathbb{E}\left[\left(F_t^T - K\right)^+ \middle| \mathscr{F}_t\right] = e^{-r(T-t)}\left(S_0 e^{\mu(t,T)+\frac{1}{2}\xi(t,T)} N(d_1) - KN(d_2)\right)$$

where

$$d_1 = \frac{\log\left(S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)}/K\right)}{\xi(t,t)} + \frac{1}{2}\sqrt{\xi(t,T)}$$
$$d_2 = d_1 - \sqrt{\xi(t,T)}$$

which is the result of the proposition. □

Remark The propositions assume that:

1. the drift term $\alpha_{t,\omega}$ in equation 3.6 is normally distributed for any t,
2. the volatility σ_t is deterministic.

These condition holds for at least two common models for commodity derivatives, which are Gibson Schwartz model and Gabillon model.

3.4 Reduced one-factor Gabillon model

Commonly, the process 3.6, X_t, is the log price of the spot. Now, we show two applications of the above propositions to the modeling of commodity derivatives, namely the models of Gibson Schwartz and Gabillon. These models are commonly used in commodity market. And both of them are two factors models with a normal stochastic drift.

We take the traditional assumption in mathematical finance and assume that the uncertainty is represented by a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a two dimensional Brownian motion. The two components of the Brownian motion will be denoted by z_1, z_2 and will be assumed to be correlated with a constant correlation.

$$\mathrm{d}z_1 \mathrm{d}z_2 = \rho \mathrm{d}t$$

This leads us to assume that the spot price S follow under the historical diffusion

$$\frac{\mathrm{d}S}{S} = k\ln\frac{L}{S}\mathrm{d}t + \sigma_S \mathrm{d}z_1 \tag{3.21}$$

$$\frac{\mathrm{d}L}{L} = \mu_L \mathrm{d}t + \sigma_L \mathrm{d}z_2 \tag{3.22}$$

where in the above equations, we have removed any obvious reference to time to simplify the notation, like for instance $S_t, L_t, k_t, z_1^t, z_2^t, \ldots$.

Denoting by λ_S the risk premium price for the spot risk, we can easily derive the diffusion under the risk neutral probability as follows (see for instance Hull and White [1990] or Gabillon [1991]):

$$\frac{\mathrm{d}S}{S} = (k\ln\frac{L}{S}\mathrm{d}t - \lambda_S\sigma_S) + \sigma_S \mathrm{d}\tilde{z}_1 \tag{3.23}$$

$$\frac{\mathrm{d}L}{L} = \sigma_L \mathrm{d}\tilde{z}_2 \tag{3.24}$$

where $\mathrm{d}\tilde{z}_1$ and $\mathrm{d}\tilde{z}$ are two Brownian motion with a correlation $\mathrm{d}\tilde{z}_1\mathrm{d}\tilde{z}_2 = \rho\mathrm{d}t$.

Now let us change the variables and take $X = \ln S$, $Y = \ln L$. Using Ito's lemma in (3.21) and (3.22), we have:

$$\mathrm{d}X = \left(k(Y-X) - \frac{1}{2}{\sigma_S}^2 - \lambda_S\sigma_S\right)\mathrm{d}t + \sigma_S \mathrm{d}z_1 \tag{3.25}$$

$$\mathrm{d}Y = -\frac{1}{2}{\sigma_L}^2\mathrm{d}t + \sigma_L \mathrm{d}z_2 \tag{3.26}$$

It can be rewritten in the following form:

$$\mathrm{d}X = a(t,\omega)\mathrm{d}t + \sigma_S \mathrm{d}z_1 \tag{3.27}$$

with $a(t,\omega) = k\left(Y_0 - \frac{1}{2}{\sigma_L}^2 t + \int_0^t \sigma_L \mathrm{d}z_2 - X\right) - \frac{1}{2}{\sigma_S}^2 - \lambda_S\sigma_S$

To apply the Distribution Match Method, we just need to prove that $a(t,\omega)$ is a normal. Then, propositions 10 and 11 provides us closed formula for the future and European options. To be clear, we work on risk neutral measure.

Lemma 13. *If X follows equation 3.27, then $a(t,\omega)$ is normally distributed. In addition, the term $\mu(0,t)$ and $\xi(0,t)$ defined in equation 3.7 and 3.8 can be computed explicitly as follows:*

$$\begin{aligned}\mu(0,t) =& X_0\left(e^{-kt}-1\right) + \left(1-e^{-kt}\right)Y_0 + \frac{1}{k}\left(-\lambda_S\sigma_S + \frac{1}{2}\sigma_L^2 - \frac{1}{2}\sigma_L^2 kt - \frac{1}{2}\sigma_S^2\right.\\ &\left. + e^{-kt}\lambda_S\sigma_S - \frac{1}{2}e^{-kt}\sigma_L^2 + \frac{1}{2}e^{-kt}\sigma_S^2\right)\end{aligned} \tag{3.28}$$

$$\begin{aligned}\xi(0,t) =& \frac{1}{k}\left(1-2e^{-kt}+e^{-2kt}\right)\sigma_S\sigma_L\rho + \frac{1}{2k}\left(\sigma_S^2 + 4e^{-kt}\sigma_L^2\right.\\ &\left. -3\sigma_L^2 - e^{-2kt}\sigma_S^2 - e^{-2kt}\sigma_L^2 + 2\sigma_L^2 kt\right)\end{aligned} \tag{3.29}$$

Proof. Re-write equation 3.27 here,

$$\mathrm{d}X = a(t,\omega)\mathrm{d}t + \sigma_S \mathrm{d}z_1$$

with $a(t,\omega) = k\left(Y_0 - \frac{1}{2}{\sigma_L}^2 t + \int_0^t \sigma_L \mathrm{d}z_2 - X\right) - \frac{1}{2}{\sigma_S}^2 - \lambda_S\sigma_S$

Denote $a_1 = k\left(Y_0 - \frac{1}{2}{\sigma_L}^2 t + \int_0^t \sigma_L \mathrm{d}z_2\right) - \frac{1}{2}{\sigma_S}^2 - \lambda_S\sigma_S$, it is obvious a_1 is normal. So we can write $a(t,\omega) = a_1 - kX$. Change the variable, $H = e^{kt}X$,

$$\begin{aligned}\mathrm{d}H &= e^{kt}\mathrm{d}X - ke^{kt}X\mathrm{d}t \\ &= e^{kt}a_1\mathrm{d}t + e^{kt}\sigma_S\mathrm{d}z_1 \end{aligned} \tag{3.30}$$

So $H = H_0 + \int_0^t e^{ks}a_1\mathrm{d}t + \int_0^t e^{ks}\sigma_S\mathrm{d}z_1$ is normal since a_1 is normal. Therefore, $X = e^{-kt}H$ is normal. $a(t,\omega) = a_1 - kX$ is normal. This proves that the drift is normal.

Let us now calculate the terms $\mu(0,t)$ and $\xi(0,t)$. In equation 3.30, we can derive the expectation and variance of H.

$$\begin{aligned}\mathbb{E}(H) &= H_0 + \mathbb{E}\left[\int_0^t e^{ks}a_1\mathrm{d}s\right] \\ &= H_0 + \int_0^t e^{ks}a_2\mathrm{d}s \\ \mathrm{var}(H) &= \int_0^t e^{2ks}{\sigma_S}^2\mathrm{d}s + \int_0^t \left(e^{kt} - e^{ks}\right)^2{\sigma_L}^2\mathrm{d}s + 2\rho\int_0^t e^{ks}\left(e^{kt} - e^{ks}\right)\sigma_S\sigma_L\mathrm{d}s\end{aligned}$$

with $a_2 = kY_0 - \frac{1}{2}k{\sigma_L}^2 t - \frac{1}{2}{\sigma_S}^2 - \lambda_S\sigma_S$

Knowing $X = e^{-kt}H$, we have

$$\begin{aligned}\mathbb{E}(X) &= X_0e^{-kt} + \int_0^t e^{k(s-t)}a_2\mathrm{d}s \\ \mathrm{var}(X) &= \int_0^t e^{2k(s-t)}{\sigma_S}^2\mathrm{d}s + \int_0^t \left(1 - e^{k(s-t)}\right)^2{\sigma_L}^2\mathrm{d}s + 2\rho\int_0^t e^{k(s-t)}\left(1 - e^{k(s-t)}\right)\sigma_S\sigma_L\mathrm{d}s\end{aligned}$$

finally we get

$$\begin{aligned}\mu(0,t) =& \mathbb{E}(X) - X_0 \\ =& X_0\left(e^{-kt} - 1\right) + \int_0^t e^{k(s-t)}a_2\mathrm{d}s \\ =& X_0\left(e^{-kt} - 1\right) + \int_0^t e^{k(s-t)}\left(kY_0 - \frac{1}{2}k{\sigma_L}^2 s - \frac{1}{2}{\sigma_S}^2 - \lambda_S\sigma_S\right)\mathrm{d}s \\ \xi(0,t) =& \mathrm{var}(X) \\ =& \int_0^t e^{2k(s-t)}{\sigma_S}^2\mathrm{d}s + \int_0^t \left(1 - e^{k(s-t)}\right)^2{\sigma_L}^2\mathrm{d}s \\ &+ 2\rho\int_0^t e^{k(s-t)}\left(1 - e^{k(s-t)}\right)\sigma_S\sigma_L\mathrm{d}s\end{aligned}$$

The calculation of the integrals is straightforward, and we finally get the result of lemma 13. □

The distribution match method summarized by propositions 10 enables us to retrieve the original result of Gabillon for the futures price. The result of proposition 11 and 12 introduces a new result for the price of European options in Gabillon model.

Proposition 14. *The future price at time t is*

$$F(0,t) = A(t){S_0}^{B(t)}{L_0}^{1-B(t)}$$

$$\begin{aligned} with,\ A(t) &= \exp\left\{\lambda_S\sigma_S\left(\frac{e^{-kt}-1}{k}\right)+\frac{v}{4k}\left(2e^{-kt}-e^{-2kt}-1\right)\right\} \\ B(t) &= e^{-kt} \\ v &= {\sigma_S}^2+{\sigma_L}^2-2\rho\sigma_S\sigma_L \end{aligned}$$

Proof. Application of proposition 10 with equation 3.28 and 3.29. The calculation is straightforward. □

Note that the above result gives another proof of the result of the Gabillon model which is originally done using PDE arguments.

Proposition 15. *The price (at time t) V of a European call option on spot with strike K and option expiry T is given by:*

$$V = S_t N(d_1)e^{-p(t,T)} - KN(d_2)e^{-r(T-t)}$$

where $N(\cdot)$ is standard normal cumulative distribution function;

$$\begin{aligned} d_1 &= \frac{\log(S_t/K)+r(T-t)-p(t,T)}{\sqrt{\xi(t,T)}}+\frac{1}{2}\sqrt{\xi(t,T)} \\ d_2 &= d_1-\sqrt{\xi(t,T)} \\ p(t,T) &= r(T-t)-\frac{1}{2}\xi(t,T)-\mu(t,T) \end{aligned}$$

with $\mu(t,T)$ and $\xi(t,T)$ in equation 3.28 and 3.29. $S_t = e^{X_t}$.

Proof. Straightfuture application of proposition 11. □

Lemma 16. *Under the hypothesis of lemma 13. The variance of process X from t to T is*

$$\begin{aligned} \xi(t,T) = &\frac{2\sigma_L}{k}\left(e^{-kT}-e^{-kt}\right)(\sigma_L-\sigma_S\rho)-\frac{1}{2k}\left(e^{-2kT}-e^{-2kt}\right)\left(\sigma_S^2+\sigma_L^2-2\rho\sigma_S\sigma_L\right) \\ &+\sigma_L^2(T-t) \end{aligned} \tag{3.31}$$

Proof. Substitute lemma 13 into definition 3.8. □

Proposition 17. *Assume a vanilla call option has maturity at time T. The underlying is future $F(t,T)$ and strike is K. Then the price V of the option at time t is given by:*

$$V = e^{-r(T-t)}\left(S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)}N(d_1) - KN(d_2)\right)$$

where $N(\cdot)$ is standard normal cumulative distribution function;

$$d_1 = \frac{\log\left(S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)}/K\right)}{\sqrt{\xi(t,T)}} + \frac{1}{2}\sqrt{\xi(t,T)}$$
$$d_2 = d_1 - \sqrt{\xi(t,T)}$$

with $\mu(t,T)$ and $\xi(t,T)$ in equation 3.28 and 3.31. $S_t = e^{X_t}$.

Proof. Straightfuture application of proposition 12. □

The results in proposition 15 and 17 are new to our knowledge.

3.5 Reduced one-factor Gibson Schwartz Model

Like the Gabillon model, Gibson and Schwartz [1990] model is based on a two-factor diffusion. The spot price S is assumed to follow a lognormal diffusion like in the Black Scholes model but with a stochastic drift. The difference between Gibson Schwartz model and Gabillon model lies in the second factor assumption. Gibson Schwartz Gibson and Schwartz [1990] assume that it is directly the convenience yield, δ, that is stochastic, following a mean reverting log normal diffusion. This leads to the following diffusion for the commodity spot price under measure Q:

$$\frac{\mathrm{d}S}{S} = (r-\delta)\mathrm{d}t + \sigma_1 \mathrm{d}z_1 \tag{3.32}$$
$$\mathrm{d}\delta = (k(\alpha-\delta) - \lambda\sigma_2)\,\mathrm{d}t + \sigma_2 \mathrm{d}z_2 \tag{3.33}$$

where r is the risk free rate, z_1 and z_2 are two correlated Brownian motions. To apply the Distribution Match Method, we will do some change of variables. Let $X = \ln S$. Using Ito's lemma in equation 3.32 and 3.33, we have:

$$\mathrm{d}X = (r-\delta-\frac{1}{2}{\sigma_1}^2)\mathrm{d}t + \sigma_1 \mathrm{d}z_1 \tag{3.34}$$
$$\mathrm{d}\delta = (k(\alpha-\delta) - \lambda\sigma_2)\mathrm{d}t + \sigma_2 \mathrm{d}z_2 \tag{3.35}$$

These equations can be rewritten in the following form:

$$dX = a(t,\omega)dt + \sigma_1 dz_1 \tag{3.36}$$

with $a(t,\omega) = r - (k(\alpha-\delta) - \lambda\sigma_2)t - \int_0^t \sigma_2 dz_2 - \frac{1}{2}{\sigma_1}^2$.

Lemma 18. *If X follows equation 3.36, then $(a(t,\omega), W_t)$ is a normal. In addition, the term $\mu(0,t)$ and $\xi(0,t)$ can be computed explicitly as follows:*

$$\begin{aligned}\mu(0,t) = rt - \frac{1}{k}\delta_0\left(1-e^{-kt}\right) - \frac{k\alpha-\lambda\sigma_2}{k}t \\ + \frac{k\alpha-\lambda\sigma_2}{k}\left(1-e^{-kt}\right) - \frac{1}{2}{\sigma_1}^2 t\end{aligned} \tag{3.37}$$

$$\begin{aligned}\xi(0,t) = {\sigma_1}^2 t + \frac{{\sigma_2}^2}{2k^3}\left(2kt - e^{-2kt} + 4e^{kt} - 3\right) \\ + \frac{2\rho\sigma_1\sigma_2}{k^2}\left(-kt + 1 - e^{-kt}\right)\end{aligned} \tag{3.38}$$

Proof. We re-write the process of convenience yield δ here:

$$d\delta = (k(\alpha-\delta) - \lambda\sigma_2)dt + \sigma_2 dz_2 \tag{3.39}$$

Let $H = e^{kt}\delta$, then

$$\begin{aligned}dH &= e^{kt}d\delta + ke^{kt}\delta dt \\ &= e^{kt}(k\alpha - \lambda\sigma_2)\,dt + e^{kt}\sigma_2 dz_2\end{aligned}$$

So, $H = H_0 + \int_0^t e^{ks}(k\alpha - \lambda\sigma_2)ds + \int_0^t e^{ks}\sigma_2 dz_2$. Consequently, we have:

$$\begin{aligned}\delta &= e^{-kt}H \\ &= e^{-kt}\delta_0 + \int_0^t e^{k(s-t)}(k\alpha - \lambda\sigma_2)ds + \int_0^t e^{k(s-t)}\sigma_2 dz_2\end{aligned} \tag{3.40}$$

$$\delta(t) = e^{-kt}\delta_0 + \left(1-e^{-kt}\right)\frac{k\alpha-\lambda\sigma_2}{k} + \sigma_2 e^{-kt}\int_0^t e^{ks}dz_2 \tag{3.41}$$

Denote $Y(t) = \int_0^t \delta ds$.
In the other hand, we can derive from equation 3.39,

$$\delta(t) - \delta(0) = (k\alpha - \lambda\sigma_2)t - kY(t) + \sigma_2\int_0^t dz_2 \tag{3.42}$$

Substitute equation 3.41 into equation 3.42, we can get the expression for $Y(t)$.

$$Y(t) = \frac{k\alpha - \lambda\sigma_2}{k}t + \frac{1}{k}\sigma_2 \int_0^t \mathrm{d}z_2 - \frac{1}{k}\left(\left(e^{-kt}-1\right)\delta_0 + \left(1-e^{-kt}\right)\frac{k\alpha-\lambda\sigma_2}{k} + \sigma_2 e^{-kt}\int_0^t e^{ks}\mathrm{d}z_2\right)$$

Therefore, we can rewrite X as following:

$$\begin{aligned} X &= \int_0^t \left(r - \delta - \frac{1}{2}{\sigma_1}^2\right)\mathrm{d}s + \int_0^t \sigma_1 \mathrm{d}z_1 \\ &= rt - Y(t) - \frac{1}{2}{\sigma_1}^2 t + \int_0^t \sigma_1 \mathrm{d}z_1 \end{aligned}$$

It is obvious that the drift is normally distributed since $Y(t)$ is a normal. Now we calculate $\mu(0,t)$ and $\xi(0,t)$.

$$\begin{aligned} \mu(0,t) &= \mathbb{E}(X) - X_0 \\ &= rt - \frac{1}{k}\delta_0\left(1-e^{-kt}\right) - \frac{k\alpha-\lambda\sigma_2}{k}t + \frac{k\alpha-\lambda\sigma_2}{k}\left(1-e^{-kt}\right) - \frac{1}{2}{\sigma_1}^2 t \end{aligned} \tag{3.43}$$

$$\begin{aligned} \xi(0,t) &= \mathrm{var}(X) \\ &= {\sigma_1}^2 t + \frac{{\sigma_2}^2}{2k^3}\left(2kt - e^{-2kt} + 4e^{kt} - 3\right) + \frac{2\rho\sigma_1\sigma_2}{k^2}\left(-kt + 1 - e^{-kt}\right) \end{aligned} \tag{3.44}$$

This proves lemma 18. □

Now we can use proposition 10, 11 and 12 to derive following result.

Proposition 19. *The future price at time t is*

$$F(0,t) = S_0 e^{\mu(0,t)+\frac{1}{2}\xi(0,t)}$$

with $\mu(0,T)$ and $\xi(0,T)$ in equation 3.37 and 3.38. $S_0 = e^{X_0}$.

Proof. Straightfuture application of proposition 10. □

Proposition 20. *The price (at time T) of a vanilla call option with strike K and maturity T is*

$$V = S_0 N(d_1) e^{-p(0,T)} - KN(d_2)e^{-rT} \tag{3.45}$$

where $N(\cdot)$ is standard normal cumulative distribution function;

$$d_1 = \frac{\log(S_0/K) + rT - p(0,T)}{\sqrt{\xi(0,T)}} + \frac{1}{2}\sqrt{\xi(0,T)}$$
$$d_2 = d_1 - \sqrt{\xi(0,T)}$$
$$p(0,T) = rT - \frac{1}{2}\xi(0,T) - \mu(0,T)$$

with $\mu(0,T)$ and $\xi(0,T)$ in equation 3.37 and 3.38 and $S_0 = e^{X_0}$.

Proof. Straightfuture application of proposition 11. □

Lemma 21. *Under the same hypothesis of lemma 18. The variance of process X from t to T is*

$$\begin{aligned}\xi(t,T) =& \frac{1}{k}\left(1 - 2e^{-k(T-t)} + e^{-2k(T-t)}\right)\sigma_S\sigma_L\rho + \frac{1}{2k}\left(\sigma_S^2 + 4e^{-k(T-t)}\sigma_L^2\right. \\ & \left. -3\sigma_L^2 - e^{-2k(T-t)}\sigma_S^2 - e^{-2k(T-t)}\sigma_L^2 + 2\sigma_L^2 k(T-t)\right)\end{aligned} \tag{3.46}$$

Proof. Same calculation with lemma 18. □

Proposition 22. *Assume a vanilla call option has maturity at time T. The underlying is future $F(t,T)$ and strike is K. Then the price V of the option at time t is given by:*

$$V = e^{-r(T-t)}\left(S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)}N(d_1) - KN(d_2)\right)$$

where $N(\cdot)$ is standard normal cumulative distribution function;

$$d_1 = \frac{\log(S_t e^{\mu(t,T)+\frac{1}{2}\xi(t,T)}/K)}{\sqrt{\xi(t,T)}} + \frac{1}{2}\sqrt{\xi(t,T)}$$
$$d_2 = d_1 - \sqrt{\xi(t,T)}$$

with $\mu(t,T)$ and $\xi(t,T)$ in equation 3.37 and 3.46. $S_t = e^{X_t}$.

Proof. Straightfuture application of proposition 12. □

Remark The results in proposition 19 and 20 are equivalent to those of Jamshidian and Fein [1990] and Bjerksund [1991]. The result in proposition 22 is new to our knowledge.

3.6 Summary

In this chapter, we introduce model factor reduction method. We show that a diffusion process with deterministic log-normal volatility and a normal stochastic

drift can be approximated by an equivalent one dimensional process with the same final marginal distribution. This enables us to compute explicitly futures, vanilla call options on the spot and futures allowing for fast calibration. In section 3.4, we show how to apply the model factor reduction technique on Gabillon models and Gibson Schwartz model. This enables us to find a formula for the Gabillon model for vanilla option allowing fast calibration. For Gibson Schwartz model, the model factor reduction method provides another proof for the closed form solution for futures and vanilla options. In addition, as model factor reduction technique provides explicitly the marginal distribution, we can easily extend our result to other common vanilla options like digitals. Model factor reduction technique in addition, implies that for future and vanilla option, the Gabillon model and Gibson Schwartz model are indeed equivalent to a one-factor model. This contrasts with the common belief that these models were explicit two-factor models.

An important consequence of model factor reduction technique is that it shows, in the sense of marginal distribution, the similarity between the original two-factor model and one-factor model. This contrasts slightly with the general opinion that commodity derivatives should be priced with two-factor models. Our result shows that in the particular case of vanilla options, Gibson Schwartz model and Gabillon model are indeed equivalent to one-factor models. The equivalence holds only for this specific case and not for path dependent options.

One of the limitation of the model factor reduction method is the loss of information on the dynamic of the original process. Therefore, it can not work for path-dependent products such as Asian option and snowball option. A possible future work could be on the joint distribution at different time $0 < t_1 < t_2 < \cdots < t_n$ that would provide a way to calculate closed forms solutions for path dependent options.

Chapter 4

Stochastic volatility model

Résumé du chapitre

Dans le domaine de la finance quantitative pour évaluer des titres dérivés, comme les options, les produits sont souvent chers et couvert à l'aide du modèle Black-Scholes, dans lequel il ya une relation univoque entre le prix de l'option et la volatilité σ_B. En théorie, la valeur de volatilité σ_B dans le modèle de Black est constante sur la durée du produit dérivé, et ne dépend pas des changements du niveau des prix du sous-jacent. Cependant, en pratique, ce n'est pas le cas. Donc modèle Black-Scholes ne peut pas expliquer les caractéristiques de la surface de volatilité implicite comme le « smile » de volatilité et le « skew », qui indiquent que la volatilité implicite tend à varier par rapport aux niveaux de prix, du strike et de la distance de l'expiration.

Le développement de modèles de volatilité locale par Dupire et Derman-Kani a été une avancée majeure. Ces modèles à volatilité locale sont auto-cohérents, sans arbitrage, et peuvent être calibrés pour correspondre exactement aux volatilités observées. Ces modèles sont très populaires actuellement. Cependant, le comportement dynamique des « smiles » et « skew » prédite par des modèles à volatilité locale est différent du comportement observé dans le marché. Sur le marché, les prix des actifs et des smiles évoluent dans dans la même direction, mais le modèle de volatilité locale prédit le mouvement inverse. Cette contradiction crée des difficultés pour la couverture en delta et vega dérivées du modèle de volatilité locale. Ceci est prouvé dans l'article « Gestion du risque Smile » par Hagan et al. [2002].

Les modèles cités auparavant ne permettent pas de respecter une dynamique réaliste pour la volatilité. Les modèles à volatilité stochastiques ainsi que leur calibration sont bien examinées dans Javaheri [2005].

Les modèles à volatilité stochastique découlent du traitement de la volatilité

du titre sous-jacent comme un processus aléatoire, gouverné par des variables d'état tels que le niveau des prix du sous-jacent, la tendance de la volatilité à revenir vers une certaine valeur à long terme et la variance du processus de volatilité elle-même, la corrélation entre le processus aléatoire pour le sous-jacent et la volatilité du processus aléatoires, et d'autres. Ces variables peuvent décrire précisément les « smiles » et « skew » du marché. En supposant que la volatilité du prix sous-jacent est un processus stochastique plutôt que d'une constante qui peut être adapté à l'exigence du marché en changeant les paramètres et les modèles de formulaires, il devient possible de modéliser les dérivés plus de précision. Les modèles tels que ceux de Heston, Piterbarg et SABR sont les modèles les plus populaires actuellement pour leurs performances dans la manipulation des smiles et skew. Dans ce chapitre on les étend à des fonctions qui dépendent du temps, afin de leur donner encore plus de souplesse. Ensuite on les calibre et on discute leurs performances.

4.1 Introduction

In the field of quantitative finance to evaluate derivative securities, such as options, products are often priced and hedged using Black-Scholes model, in which there is a one-to-one relation between the price of option and the volatility σ_B. In theory, the volatility value σ_B in Black's model is constant over the life of the derivative, and unaffected by the changes in the price level of the underlying. However in practice, options with different strikes K require different valatilities to match their market prices. So Black-Scholes model cannot explain long-observed features of the implied volatility surface such as volatility smile and skew, which indicate that implied volatility does tend to vary with respect to strike price and expiration.

The development of local volatility models by Dupire and Derman-Kani was a major advance in managing smiles and skews. Local volatility models are self-consistent, arbitrage-free, and can be calibrated to match exactly observed market smiles and skews. These models are very popular currently in most financial institution. However, the dynamic behavior of smiles and skews predicted by local volatility models is different from the behavior observed in the marketplace. In the market, asset prices and market smiles move in the same direction, but the local volatility model predicts the reverse movement. This contradiction creates difficulties for delta and vega hedges derived from the local volatility model. This is proved in the article "Managing Smile Risk" by Hagan et al. [2002]).

Commodity specialized model is able to better capture the features such as convenience yield, seasonality and mean reversion. Various drift terms are suggested in these models. These models for commodity derivatives pricing relies on a specific dynamics of some state variables from whom one can derives the corresponding futures prices. The states variables are often assumed to follow a mean reverting process to capture the mean reversion nature of commodity prices. Typical examples include Gibson and Schwartz [1990], Brennan [1991], Gabillon [1991], Schwartz [1997], Hilliard and Reis [1998], Schwartz and Smith [2000], and Casassus and Collin-Dufresne [2005]. But these models also lack a realistic dynamics for the volatility. These models as well as their calibration are well reviewed in Javaheri [2005].

Stochastic volatility models derive from the models' treatment of the underlying security's volatility as a random process, governed by state variables such as the price level of the underlying, the tendency of volatility to revert to some long-run mean value, and the variance of the volatility process itself, the correlation between the random process for the underlying and the volatility random process, and others. These variables can describe precisely the market smiles and skews. By assuming that the volatility of the underlying price is a stochastic process rather than a constant which can be adapted to the market requirement by changing the parameters and model forms, it becomes possible

to model derivatives more accurately. The models such as Heston model, Piterbarg model and SABR model are the most popular stochastic volatility models studied currently for their performance in handling the smiles and skews.

4.2 Review of research

Indeed, various authors had already noticed the importance of stochastic volatility for commodity modeling. This includes the work of Eydeland and Geman [1998], Richter and Sørensen [2002], Nielsen and Schwartz [2004], and Trolle and Schwartz [2008] who explicitly allow for stochastic volatility for the corresponding state variables. Their models are often an adapted version of some existing stochastic volatility models already developped for another asset than commodity. For instance, Eydeland and Geman [1998] adapt an Heston model to commodity. Trolle and Schwartz [2008] adapt their corresponding interest rates HJM model.

This suggest first to study the various stochastic models already developed in equity option pricing and see how to adapt them to commodity issue. Then we need to benchmark these models in the case of commodity to assess the fit or the consistency of these models for this market. These two questions are the motivations of this work.

Stochastic volatility modeling has been studied for quite some time. Among the precursors, Hull and White [1990] and Wiggins [1987] studied lognormal distributed stochastic volatility. The Hull-White model was developed using a trinomial lattice, although closed-form solutions for European-style options and bond prices are possible. They assumed the correlation between spot price and volatility is zero. But from the market data, evidence showed that the two processes are strongly correlated.

Wiggins studied the same model with Hull and White, but with the non-zero correlation. The advantage of the model is that the volatility is always positive by assuming geometric Brownian motion.

Scott [1987], and then Stein and Stein [1991] studied an Ornstein Uhlenbeck process (OU process) for volatility. OU process adds mean reversion into volatility. It is more natural to assume the volatility mean reversion rather than lognormal. The disadvantage of the model is that, the model cannot guarantee positiveness. That is to say, if volatility follows an OU process, then there is chance that volatility goes to negative.

Heston [1993] studied a special case of stochastic volatility model. He supposed the square of volatility follows an OU process. This assumption resolved the problem that volatility goes to negative. Another advantage of Heston model is that, it allows correlation between spot and volatility. The model has also a semi-analytic formula, which makes it practical to use.

Richter and Sørensen [2002] and Trolle and Schwartz [2008] studied the unspanned stochastic volatility model with abundant data analysis. It is an application of perturbation theory. They expand the formula of price around vol of vol near zero.

In dynamic stochastic volatility models, parameters depend on the time t. The process for the forward and the volatility become much more complicate than in constant parameter case. In reality, we only need to price products for given dates, so we intend to consider the parameters as piecewise constant ones. With this assumption, it is possible to find out closed pricing formulas which are used later for the calibration. There are several methods as we noted before: firstly, compute the effective mean value of the parameters so that we can use the closed formula of the constant stochastic volatility model by replacing the time dependent parameters with their effective mean values. This method consists to find out the effective medium value of each parameter who depends on time. This method is employed by Piterbarg [2005]. Secondly as another method, by resolving the PDEs, we can compute the option price directly by an integration of a function which is continuous on time t. Theoretically, we can use both these two methods for each model. But as each model has its different characteristics, one method should be easier than another. For SABR model in the case of constant β, the first method is easier that the second. But we can use the second method to price an option with SABR model in the case of time dependent β. For Heston model and Piterbarg model with correlation ρ, we use the second method, and for Piterbarg model without correlation, we can use both of them. In the following, we will present one by one the methods used for each model.

Stochastic volatility models with time dependent parameters are derived from stochastic volatility models with constant parameters to be able to price products with long maturities. Different periods have variant market smiles and skews, which need different parameter values to match exactly the curves. In reality, the parameters in the models, especially correlation ρ between the process of underlying and the process of volatility and parameters related to the smile and skew, depend on time t. As we only need the model values for certain dates, we can consider the parameters as piecewise functions of time t. This is the approach taken for all the models.

Although jump can relax the need of time-dependent parameter, it cannot fulfill the need completely, especially for the options with long maturity. It is because the expectation of jumps tends to flatten in long time. We don't consider jump in this article, interested readers can refer to Andersen and Andreasen [2000] and Benhamou et al. [2008].

However in practical, SV models are usually difficult to use due to difficulties of calibration. The difficulty is obvious: 1. too many parameters to calibrate: an SV model usually has more than 5 parameters; 2. some of the parameters

cannot be observed from the market. Therefore we cannot perform the same procedure of calibration as we do for BS model or LV models. To calibrate an SV model, the common way is to minimize the vector distance between theoretical price (or implied BS vol) and the market price(or implied BS vol). One of the key points here is how to calculate the theoretical price fast and precisely. The numerical algorithms such as partial differential equations (PDEs) or Monte Carlo is too time-consuming to use in calibration, even for the simplest European-style options. For this reason, we investigate closed form formula for European options.

A closed form formula is typically available for SV models with constant parameters. A series of references are available on this topic, such as Andersen and Brotherton-Ratcliffe [1998], Hagan et al. [2002], Zhou [2003] and Andersen and Andreasen [2002]. Lewis summarizes the Fourier formed solution for static models in his book Lewis [2000]. There are closed formulas for all the common SV models including square root model, GARCH model and 3/2 model.

Here we have some clarification on the term "closed form formula". The definition of the term "closed form formula" can vary from person to person. A solution is said to be a closed form solution if it is in terms of functions and mathematical operations from a given generally accepted set. However, the choice of what to call closed form is rather arbitrary. For example, an infinite sum is be called a closed form in certain context. In this paper, we call the Fourier transform integration a closed form function. This kind of infinite integration is also called semi-analytical solution in some references as well.

For an SV model with time-dependent parameters, saying dynamic SV model, the closed form formula is more difficult to derive. Several different methods have been reported, such as asymptotic expansion and character function. References for the former one includes Labordere [2005] and Osajima [2007]. For the latter, readers can refer to Mikhailov and Nogel [2003a]. Some references suggest piecewise constant parameters since market data in practical is only available on certain expiries. It is a natural assumption because there is no information about how the parameters behave between two adjacent expiries. This assumption can considerably simplify the closed form formula without losing any generality.

4.3 Heston model

Heston [1993] supposes that spot price S_t follows a log-normal process and its variance V_t also follows a log-normal process.

$$\frac{\mathrm{d}S_t}{S_t} = \mu \mathrm{d}t + \sqrt{V_t}\mathrm{d}W_1 \tag{4.1}$$

$$dV_t = k(\theta - V_t)\mathrm{d}t + \xi\sqrt{V_t}\mathrm{d}W_2 \tag{4.2}$$

where W_1 and W_2 are two standard Brownian motion with correlation ρ; μ is drift term; θ is the mean reversion of variance process; k is speed of mean reversion; ξ is vol of vol parameter. Another model parameter is V_0, the initial value of the variance process.

Using Fourier transform and characteristic function, we can get the formula for vanilla option of Heston model as in Heston [1993] and Mikhailov and Nogel [2003b]. The idea is to write the solution in the same form as Black Scholes formula and assume the characteristic function has the form as $e^{C+DV_0+i\phi f}$. By solving C and D , we can have the final result.

4.3.1 Formula for vanilla option

Case 1: Heston formula This is the method developed by Steven L. Heston in "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options" Heston [1993].

Firstly we guess that the option value of a call has the form

$$C(S,V,t) = SP_1 - KP_2$$

Note $x = \log F(t,T)$, from the PDEs that C satisfies (that we can get by applying the Ito's formula), we get the PDE for P_1 and P_2:

$$\frac{1}{2}V\frac{\partial^2 P_j}{\partial x^2} + \rho\xi V\frac{\partial^2 P_j}{\partial x \partial V} + \frac{1}{2}\xi^2 V\frac{\partial^2 P_j}{\partial V^2} + u_j V\frac{\partial P_j}{\partial x} + (a_j - b_j V)\frac{\partial P_j}{\partial V} + \frac{\partial P_j}{\partial t} = 0$$

for $j = 1, 2$, where $u_1 = 1/2$, $u_2 = -1/2$, $a = k\theta$, $b_1 = k - \rho\xi$, $b_2 = k$ with the initial condition $P_j(x,V,T;\log K) = 1_{x\geq \log K}$ when $t = T$.

P_j can be considered as the conditional probability of x_j that the option expires in the money:

$$P_j = \text{Proba}\left[x_j(T) \geq \log K \middle| x_j(0) = \log f, v(0) = V(0)\right]$$

where x_j follows:

$$\begin{aligned} \mathrm{d}x_j(t) &= u_j v(t)\mathrm{d}t + x_j(t)\sqrt{v(t)}\mathrm{d}W_1 \\ \mathrm{d}v(t) &= (a - b_j)\mathrm{d}t + \xi\sqrt{v(t)}\mathrm{d}W_2 \\ <\mathrm{d}W_1, \mathrm{d}W_2> &= \rho\mathrm{d}t \end{aligned}$$

We can't get the solutions for P_j immediately. So we introduce the characteristic functions $f_j(x,V,t;\phi)$ of their probability density function, we have by using Lévy's theorem to get the distribution function from characteristic function:

$$P_j(x,V,t,T;\log(K)) = 1 - \text{Proba}\left[x(T) \le \log(K) \middle| x(t)=x, v(t)=v\right] \tag{4.3}$$

$$= 1/2 + \frac{1}{\pi}\int_0^\infty \text{Re}\left[\frac{e^{-i\phi ln(K)} f_j(x,V,t,T;\phi)}{i\phi}\right] d\phi \tag{4.4}$$

where $f_j(x,V,t;\phi)$ satisfies:

$$\begin{cases} \frac{1}{2}V\frac{\partial^2 f_j}{\partial x^2} + \rho\xi V\frac{\partial^2 f_j}{\partial x \partial V} + \frac{1}{2}\xi^2 V\frac{\partial^2 f_j}{\partial V^2} + u_j V\frac{\partial f_j}{\partial x} + (a_j - b_j V)\frac{\partial f_j}{\partial V} + \frac{\partial f_j}{\partial t} = 0 & \text{if } 0 \le t < T \\ f_j(x,V,T;\log K) = e^{i\phi x} & \text{if } t = T \end{cases}$$

The solutions for $f_j(x,V,t,T;\phi)$ can be found by introducing:

$$f_j(x,V,t,T;\phi) = e^{C(T-t;\phi)+D(T-t;\phi)V+i\phi x}$$

where $C_j(\tau;\phi)$ and $D_j(\tau;\phi)$ with $\tau = T-t$ satisfy the PDEs following:

$$\begin{cases} \frac{dC_j(\tau;\phi)}{d\tau} - aD(\tau,\phi) = 0 \\ \frac{dD_j(\tau;\phi)}{d\tau} - \frac{\xi^2 D_j^2(\tau;\phi)}{2} + (b_j - \rho\xi\phi i)D_j(\tau,\phi) - u_j\phi i + \frac{1}{2}\phi^2 = 0 \end{cases}$$

And $C_j(0;\phi) = 0$, $D_j(0;\phi) = 0$.

The solutions for $C_j(\tau;\phi)$ and $D_j(\tau;\phi)$ are:

$$C_j(\tau;\phi) = \frac{a}{\xi^2}\left(\left(b_j - \rho\xi\phi i + d_j\right)\tau - 2\log\left[\frac{1-g_j e^{d_j\tau}}{1-g_j}\right]\right)$$

$$D_j(\tau;\phi) = \frac{b_j - \rho\xi\phi i + d_j}{\xi^2}\cdot\frac{1-e^{d_j\tau}}{1-g_j e^{d_j\tau}}$$

$$g_j = \frac{b_j - \rho\xi\phi i + d_j}{b_j - \rho\xi\phi i - d_j}$$

$$d_j = \sqrt{(\rho\xi\phi i - b_j)^2 - \xi^2(2u_j\phi i - \phi^2)}$$

Case 2: Lewis formula Another method using Fourier Transform is the method developed by Lewis [2000]. Instead of adapting to the Black Scholes form as Heston, we do directly Fourier Transform to the PDE of the option value C following:

$$\begin{cases} -C_t = \frac{1}{2}VF^2 C_{FF} + (\omega - kV)C_V + \frac{1}{2}\xi^2 V C_{VV} + \rho\xi V F C_{FV} & \text{if } 0 \le t < T \\ C = (S-K)^+ & \text{if } t = T \end{cases}$$

Here C_{FF}, C_V, C_{VV} and C_{FV} are the partial derivation of option price C with respect to forward $F(t,T)$ and variance V.

It can be written by noting $x = \log F(t,T)$ and $\tau = T - t$:

$$\begin{cases} C_\tau = \frac{1}{2}VC_{xx} - \frac{1}{2}VC_x + (\omega - kV)C_V + \frac{1}{2}\xi^2 VC_{VV} + \rho\xi VC_{xV} & \text{if } \tau > 0 \\ C = (S(T) - K)^+ & \text{if } \tau = 0 \end{cases}$$

Then by computing the Fourier Transform $\hat{C} = \int_{-\infty}^{\infty} e^{i\phi x} C(x,V,t)\mathrm{d}x$:

$$\begin{aligned} \hat{C}_\tau &= -\frac{1}{2}V\phi^2\hat{C} - i\frac{1}{2}V\phi\hat{C} + (\omega - kV)\hat{C}_V + \frac{1}{2}\xi^2 V\hat{C}_{VV} + i\rho\xi V\phi\hat{C}_V \\ &= \left(-\frac{1}{2}V\phi^2 + i\frac{1}{2}V\phi\right)\hat{C} + \left(\omega - kV + i\rho\xi V^{\frac{1}{2}}\phi\right)\hat{C}_V + \frac{1}{2}\xi^2 V\hat{C}_{VV} \end{aligned} \tag{4.5}$$

for $\tau > 0$. And the initial condition for $\tau = 0$ is:

Financial Claim	Payoff Function	Payoff Transform	ϕ-plane Restricitons
Call option	$\max[S(T) - K, 0]$	$-\frac{K^{i\phi+1}}{\phi^2 - i\phi}$	$\mathrm{Im}(\phi) > 1$
Put option	$\max[K - S(T), 0]$	$-\frac{K^{i\phi+1}}{\phi^2 - i\phi}$	$\mathrm{Im}(\phi) < 0$

Here we introduce the fundamental transform $H(\phi,\tau,V)$ who satisfies the same equation as $\hat{C}$ but with an initial condition: $H(\phi,0,V) = 1$. Then $\hat{C}$ can be written as : $\hat{C} = H(0,V)$*Payoff Transform, and:

$$C(f,V,\tau) = f - \frac{K}{2\pi}\int_{i\phi_i - \infty}^{i\phi_i + \infty} e^{-i\phi X}\frac{H(\phi,\tau,V)}{\phi^2 - i\phi}\mathrm{d}\phi \tag{4.6}$$

To get the solution for $H(\phi,\tau,V)$, we guess the form $H(\phi,\tau,V) = \exp(f_1 + f_2 V)$, and we have the PDEs for f_1 and f_2:

$$\begin{aligned} \frac{df_1}{d\tau} &= \omega f_2 \\ \frac{df_2}{d\tau} &= -\frac{1}{2}\phi^2 + i\frac{1}{2}\phi + (-k + i\rho\xi\phi)f_2 + \frac{1}{2}\xi^2 f_2^2 \end{aligned}$$

with initial conditions $f_1(0) = 0$ and $f_2(0) = 0$. Here $\omega = k\theta$.

The solutions for f_1 and f_2 are:

$$\begin{aligned} f_1 &= \tilde{\omega}\left[tg - \log\left(\frac{1-h\exp(dt)}{1-h}\right)\right] \\ f_2 &= g\left(\frac{1-\exp(dt)}{1-h\exp(dt)}\right) \\ d &= [\hat{\theta}^2 + 4\tilde{c}]^{1/2} \\ g &= \frac{1}{2}(\hat{\theta}+d) \\ h &= \frac{\hat{\theta}+d}{\hat{\theta}-d} \end{aligned}$$

where $\hat{\theta}(k) = \frac{2}{\xi^2}\left[(1-\gamma+i\phi)\rho\xi + \sqrt{k^2-\gamma(1-\gamma)\xi^2}\right]$, $t = \frac{1}{2}\xi^2\tau$, $\tilde{\omega} = \frac{2}{\xi^2}\omega$, and $\tilde{c} = \frac{\phi^2-i\phi}{\xi^2}$.

The advantage of Lewis formula is that it has a better numerical convergence. The reason is obvious: it has only one integral instead of two in Mikhailov formula. The following figure shows a test between two formulas. We calculate the same option and integral interval is set to $(0,M)$. $M = 2,3,\cdots,12$. For Lewis formula, an interval $(0,6)$ is enough to converge, while Mikhailov formula needs an interval $(0,12)$ to converge.

Figure 4.1: Option price with different formula under Heston model

4.3.2 Control variate in Heston model

This section focuses on the instability of the semi-analytical solution of Heston model. We notice the numerical difficulty on the infinitive integral. When we cut the integral to finite, it will cause a systematic error. We therefore propose a more stable solution to overcome this difficulty. Let us consider a similar model

with deterministic volatility. Instead of integrating the original function, we integrate the difference between the original function and the function in case of deterministic. As the new integrand is smoother, we get a better convergence. Following this idea, we will show more detail in both Mikhailov formula and Lewis formula.

Our improvement is supported in the section of empirical result. It solves the failure in calibration caused by instability, especially with in-the-money options.

Here we consider a modified Heston model with deterministic volatility. In Heston model we take vol of vol $\xi = 0$ and get the following model:

$$\frac{\mathrm{d}S(t)^{\mathrm{DET}}}{S(t)^{\mathrm{DET}}} = \mu \mathrm{d}t + \sqrt{V(t)^{\mathrm{DET}}}\mathrm{d}W_1 \tag{4.7}$$

$$\mathrm{d}V(t)^{\mathrm{DET}} = k\Big(\theta - V(t)^{\mathrm{DET}}\Big)\mathrm{d}t \tag{4.8}$$

The objective is to improve the closed formula solution of the option price with Heston Stochastic Model. As the volatility follows a process stochastic, the convergency of the price formula is not very strong. We try to find out a variate that increases the convergency of the pricing formula. The variate that we choose is the price of an option with Heston model but in which the volatility of volatility is 0. This method is used both for Heston formula and Lewis formula.

Case1: Heston Formula The Heston model without volatility of volatility has the process for the square of volatility:

$$\mathrm{d}V(t) = k[\theta - V(t)]\mathrm{d}t \tag{4.9}$$

which is a deterministic process. Noting $\tilde{V}$ the option price of Heston model without volatility of volatility ξ, $\tilde{V}$ can be written in two ways:

$$\tilde{V} = S\tilde{P}_1 - K\tilde{P}_2$$

where $\tilde{P}_j$ satisfies the same PDEs as P_j by setting ξ to 0. And:

$$\tilde{V} = S\mathcal{N}(d_1) - K\mathcal{N}(d_2) \tag{4.10}$$

where N is the Gaussian distribution, and by noting σ_T the total volatility (definition in equation 4.11) of the Heston model without volatility of volatility, we have

$$d_1 = \frac{\log S/K}{\sigma_T} + \frac{1}{2}\sigma_T$$

$$d_2 = \frac{\log S/K}{\sigma_T} - \frac{1}{2}\sigma_T$$

To get the total volatility σ_T, as we know $\mathrm{d}V(t) = k[\theta - V(t)]\mathrm{d}t$ and $V(t)$ is the square of volatility, then:

$$\begin{aligned}(\sigma_T)^2 &= \int_0^T V(t)\mathrm{d}t \\ &= \theta T + \frac{V_0-\theta}{k}(1-e^{-kT})\end{aligned}$$

And so

$$\sigma_T = \sqrt{\theta T + \frac{V_0-\theta}{k}(1-e^{-kT})} \tag{4.11}$$

By resolving the PDEs of C_j and D_j with $\xi = 0$, let D_j^{DET} and C_j^{DET} the solutions in the case $\xi = 0$, we have:

$$\begin{aligned}D_j^{\mathrm{DET}} &= \frac{u_j\phi i - 1/2\phi^2}{b_j}(1-e^{-b_j\tau}) \\ C_j^{\mathrm{DET}} &= a\left(\frac{u_j\phi i - 1/2\phi^2}{b_j^2}(e^{-b_j\tau}-1) + \frac{(u_j\phi i - 1/2\phi^2)\tau}{b_j}\right)\end{aligned}$$

And we can have the characteristic functions f_j^{DET} of the density function of $\tilde{P}_j$:

$$f_j^{\mathrm{DET}}(x,V,\tau;\phi) = e^{C_j^{\mathrm{DET}}(\tau;\phi)+D_j^{\mathrm{DET}}(\tau;\phi)V+i\phi x} \tag{4.12}$$

We have two different way to calculate the option price with vol of vol = zero. One is on closed formula in equation 4.10. The other is on integral form with characteristic functions in equation 4.12. Now the option price can be expressed by adding and minus the two form on the original formula.

$$C(f,V(0),t) = f\tilde{\tilde{P}}_1 - K\tilde{\tilde{P}}_2 + [S\mathcal{N}(d_1) - K\mathcal{N}(d_2)] - \frac{S-K}{2} \tag{4.13}$$

where

$$\tilde{\tilde{P}}_j = 1/2 + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi ln[K]}(f_j(x,V,T;\phi) - f_j^{DET})}{i\phi}\right]\mathrm{d}\phi$$

The control variate method improve much the calibration process. For Heston static model, without control variate, we usually can't success the calibration process. We know the vega : $vega = \frac{\mathrm{d}C}{\mathrm{d}V}$ is very small when we are in the money as we can see in a graph (figure 4.2) for vega with *strike* = 110.

The Heston formula without control variate calculates the option price in the money with a difference which results a greater error for the implied volatility. Because we calibrate to the market implied volatility, we can't reach the given calibration level with this formula since the error is accumulated in the calibration loop to make the whole process fail. To illustrate the effect of control variate,

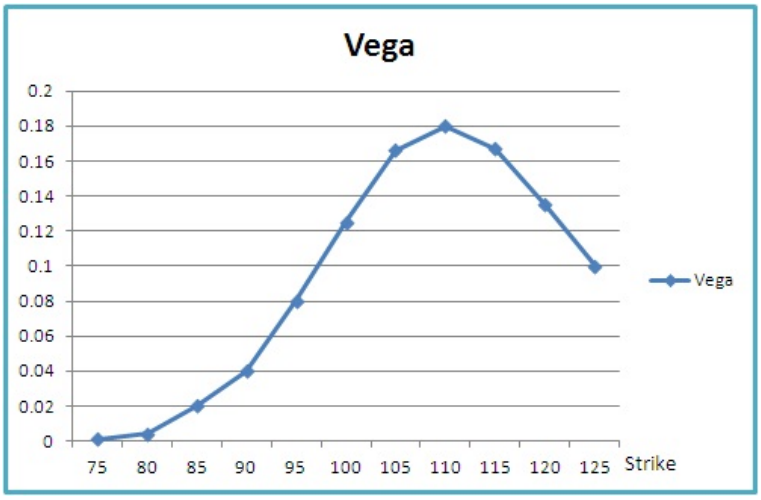


Figure 4.2: vega and spot

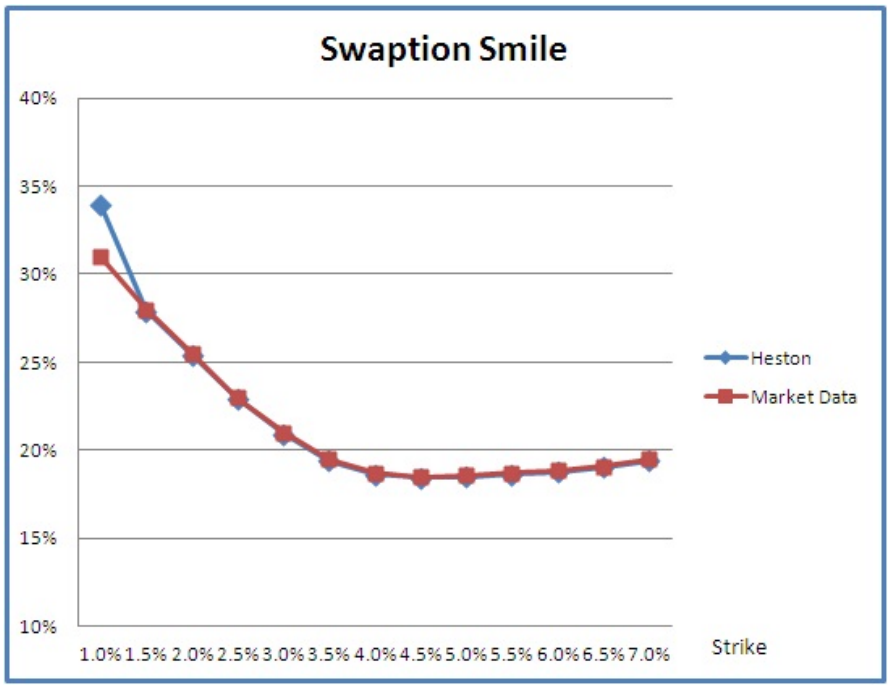


Figure 4.3: Effect of control variate to implied volatility - without control variate

by taking the same calibrated parameters, here we have the implied volatility calculated with the Heston formula without control variate at left and the Heston formula with control variate in figure 4.3 and 4.4.

We see from the figure above that the Heston formula gives a very big error for the case when the *spot* is very small. The error is bigger if we have even smaller spots. This difference is enough for that we can't success the calibration.

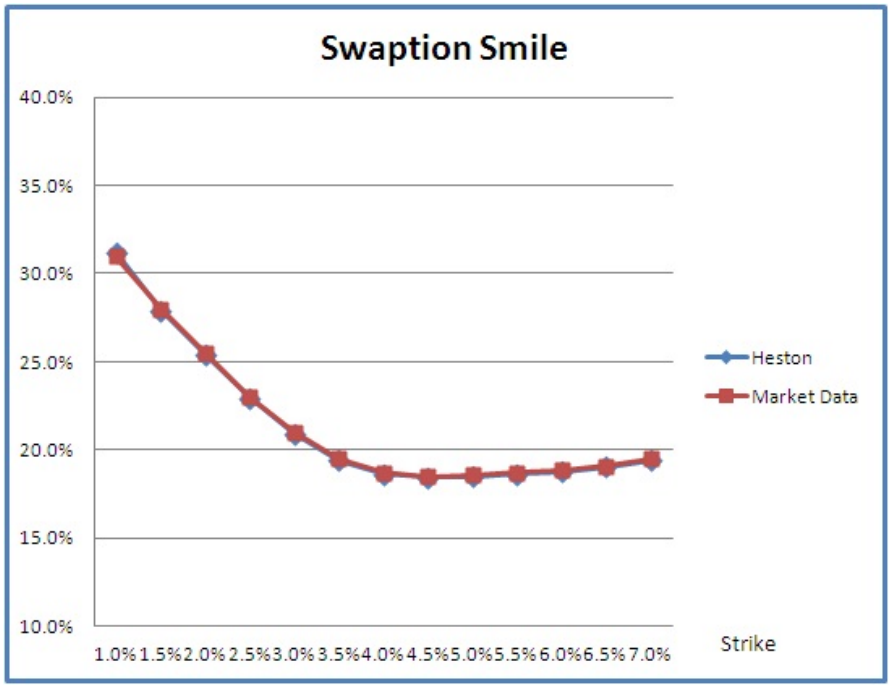


Figure 4.4: Effect of control variate to implied volatility - with control variate

The error shows that it is more difficult for in-the-money option since vega is much smaller in this area.

Case2: Lewis Formula We have the same process for the volatility as in the case of Heston formula. Note $\tilde{C}$ the option price of Heston model with $\xi = 0$, then we get:

$$\begin{cases} \tilde{C}_\tau = \frac{1}{2}V\tilde{C}_{xx} - \frac{1}{2}V\tilde{C}_x + (\omega - kV)\tilde{C}_V & \text{if } \tau > 0 \\ \tilde{C} = (f(T) - K)^+ & \text{if } \tau = 0 \end{cases}$$

where $\omega = k\theta$, by doing Fourier Transform of $\tilde{C}$: $\hat{\tilde{C}} = \int_{-\infty}^{\infty} e^{i\phi x}\tilde{C}(x,V,t)\mathrm{d}x$, we get:

$$\hat{\tilde{C}}_\tau = -\frac{1}{2}Vk^2\hat{\tilde{C}} + i\frac{1}{2}Vk\hat{\tilde{C}} + (\omega - kV)\hat{\tilde{C}}_V$$

for $\tau > 0$ and the initial condition for $\tau = 0$ is the same as in the case with $\xi \neq 0$.

Suppose $\hat{\tilde{C}} = \check{H}(\phi,\tau,V)$*Payoff Transform, $\check{H}(\phi,\tau,V) = \exp(\tilde{f}_1 + \tilde{f}_2 V)$ and $\check{H}(\phi,0,V) = 1$ as initial value, we get easily:

$$\begin{cases} \dfrac{d\tilde{f_1}}{d\tau} = \omega \tilde{f_2} \\ \dfrac{d\tilde{f_2}}{d\tau} = -\dfrac{1}{2}\phi^2 + i\dfrac{1}{2}\phi - k\tilde{f_2} \end{cases}$$

with $\omega = k\theta$ And their solutions:

$$\tilde{f_1} = \frac{\phi i - \phi^2}{2k}(1 - e^{-k\tau})$$
$$\tilde{f_2} = \omega\left(\frac{\phi i - \phi^2}{2k^2}(e^{-k\tau} - 1) + \frac{(\phi i - \phi^2)\tau}{2k}\right)$$

The same as for Heston Formula, we can get the option price by computing the total volatility and use Black's formula:

$$\tilde{C} = f\mathcal{N}(d_1) - K\mathcal{N}(d_2)$$

Finally, we have the option price formula:

$$C(f, V, \tau) = -\frac{K}{2\pi}\int_{i\phi_i - \infty}^{i\phi_i + \infty} e^{-i\phi X}\frac{\hat{H}}{\phi^2 - i\phi}d\phi + [f\mathcal{N}(d_1) - K\mathcal{N}(d_2)]$$

where:

$$\hat{H} = H(\phi, \tau, V) - \check{H}(\phi, \tau, V)$$

and $H(\phi, \tau, V)$ is the fundamental function that we calculated for Lewis formula(See the section 3.2.2).

The figure 4.5 shows the numerical result of using control variate. We can spot the zig-zag effect of the price for the formula without control variate. The control variate makes the price curve smoother and more quickly convergence.

4.3.3 Discussion of numerical stability

As function $\log(\cdot)$ is not continuous in C++ for complex numbers, we have to pay attention during the implementation by using simple a method to track the angle in order to keep it continuous. For doing this, we can use the method explained in Kahl and Jackel [2005].

The formulas we have here can be written in another way by changing d to $-d$. This is because d is the square root of a complex number and the principal square root value is returned for d. The principal square root value makes f pass across the negative real axis when increasing ϕ and hence leads to a discontinuous function causing numerical trouble. In contrast, when we change d to $-d$, we can go around this difficulty. The proof is given in Albrecher et al. [2007].

The numerical problem of solution (4.4) and (4.6) is that it is a semi-analytic solution. In practical use, we must replace the infinitive integral to an integral

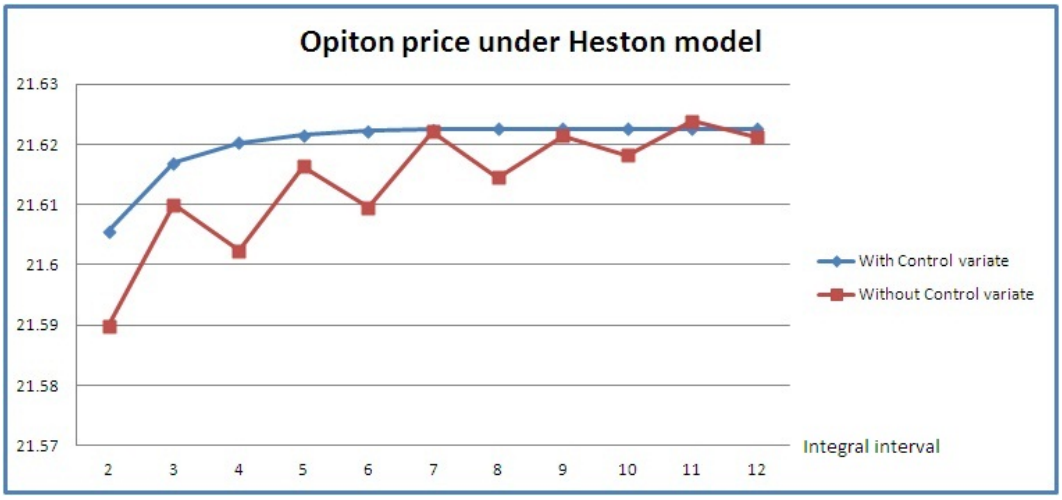


Figure 4.5: Option price with different formula under Heston model

from 0 by a large number M. It causes a systematic error. The structure of Vega also amply the pricing error, especially for the case of in the money options. The empirical result is also given in Albrecher et al. [2007].

4.3.4 Dynamic Heston Model

For Heston stochastic volatility model with piecewise parameters, we use the same method as in the case of Heston static stochastic volatility model by Mikhailov and Nogel [2003b]. The model is presented as:

$$\begin{aligned} \mathrm{d}F(t) &= F(t)\sqrt{V(t)}\mathrm{d}W_1 \\ \mathrm{d}V(t) &= k(t)(\theta(t) - V(t))\mathrm{d}t + \xi(t)\sqrt{V(t)}\mathrm{d}W_2 \\ <\mathrm{d}W_1, \mathrm{d}W_2> &= \rho(t)\mathrm{d}t \end{aligned} \tag{4.14}$$

Suppose that we have N periods with time intervals $[T_0, T_1]$, $[T_1, T_2]$, $[T_2, T_3]$, ..., $[T_{N-1}, T_N]$ where $T_N = T$, and $T_0 = t$, and in each subinterval of time the parameters are constant. We note the parameters: $k_1, k_1, k_1, ..., k_1$, $\theta_1, \theta_2, ..., \theta_N$, ρ_1 , $\rho_2, ..., \rho_N$, and $\xi_1, \xi_2, ..., \xi_N$. As we use the inverse time $\tau = T - t$, the subintervals for τ become $[0, T_N - T_{N-1}]$, $[T_N - T_{N-1}, T_N - T_{N-2}]$, ..., $[T_N - T_1, T_N - t]$.

Return to the Heston static model, the option price $V = fP_1 - KP_2$, to calculate P_1 and P_2, we introduce their characteristic functions f_1 and f_2, and f_j ($j = 1$ or 2) is presented as $f_j = e^{C_j(T-t;\phi)+D_j(T-t;\phi)V+i\phi x}$. Our final task is to resolve the equations of $C_j(\tau, \phi)$ and $D_j(\tau, \phi)$:

$$\begin{cases} \dfrac{dC_j(\tau;\phi)}{d\tau} - a(T-\tau)D(\tau,\phi) = 0 \\ \dfrac{dD_j(\tau;\phi)}{d\tau} - \dfrac{\xi^2 D_j^2(\tau;\phi)}{2} + \Big(b_j(T-\tau) - \rho(T-\tau)\xi(T-\tau)\phi i\Big)D_j(\tau,\phi) - u_j\phi i + \dfrac{1}{2}\phi^2 = 0 \end{cases}$$

where $a(T-\tau) = k(T-\tau)\theta(T-\tau)$, $b_1(T-\tau) = k(T-\tau) - \rho(T-\tau)\xi(T-\tau)$ and $b_2(T-\tau) = k(T-\tau)$.

As the parameters are constant in each subinterval of τ, note i the *ith* subinterval of τ, $1 \le i \le N$, we rewrite the equations for C_j and D_j in the *ith* subinterval $[T_N - T_{N+1-i}, T_N - T_{N-i}]$ of τ where we note C_j^i and D_j^i and $\tilde{\tau} = \tau - (T_N - T_{N+1-i})$:

$$\begin{cases} \dfrac{dC_j^i(\tilde{\tau};\phi)}{d\tilde{\tau}} - a_{N+1-i}D_j^i(\tilde{\tau},\phi) = 0 \\ \dfrac{dD_j^i(\tilde{\tau};\phi)}{d\tilde{\tau}} - \dfrac{\xi^2 (D_j^i(\tilde{\tau};\phi))^2}{2} + (b_j^{N+1-i} - \rho_{N+1-i}\xi_{N+1-i}\phi i)D_j^i(\tilde{\tau},\phi) - u_j\phi i + \dfrac{1}{2}\phi^2 = 0 \end{cases}$$

where $b_1^{N+1-i} = k_{N+1-i} - \rho_{N+1-i}\xi_{N+1-i}$, $b_2^{N+1-i} = k_{N+1-i}$ and $0 \le \tilde{\tau} \le T_{N+1-i} - T_{N-i}$.

The initial conditions are $C_j^i(0)$, $D_j^i(0)$ are to be determined later.

The solution for the previous equations are:

$$\begin{aligned} C_j^i(\tilde{\tau};\phi) &= \frac{a_{N+1-i}}{\xi_{N+1-i}^2}\left\{\Big(b_j^{N+1-i} - \rho_{N+1-i}\xi_{N+1-i}\phi i + d_j\Big)\tilde{\tau} - 2\log\left[\frac{1-g_j e^{d_j\tilde{\tau}}}{1-g_j}\right]\right\} + C_j^i(0) \\ D_j^i(\tilde{\tau};\phi) &= \frac{b_j^{N+1-i} - \rho_{N+1-i}\xi_{N+1-i}\phi i + d_j - (b_j^{N+1-i} - \rho_{N+1-i}\xi_{N+1-i}\phi i - d_j)g_j e^{d_j\tilde{\tau}}}{(1-g_j e^{d_j\tilde{\tau}})\xi_{N+1-i}^2} \\ g_j &= \frac{b_j^{N+1-i} - \rho_{N+1-i}\xi_{N+1-i}\phi i + d_j - D_j^i(0)\xi_{N+1-i}^2}{b_j^{N+1-i} - \rho_{N+1-i}\xi_{N+1-i}\phi i - d_j - D_j^i(0)\xi_{N+1-i}^2} \\ d_j &= \sqrt{(\rho_{N+1-i}\xi_{N+1-i}\phi i - b_j^{N+1-i})^2 - \xi_{N+1-i}^2(2u_j\phi i - \phi^2)} \end{aligned}$$

For the first interval $i = 1$, the same as in the constant case, the initial conditions are:

$$\begin{aligned} C_j^1(0) &= 0, \\ D_j^1(0) &= 0. \end{aligned}$$

C_j and D_j are continuous on τ, so for the second subinterval of τ, we take the final value of the first subinterval as initial value. We do the same procedure for

each subinterval of τ, every time we take the final value of previous interval as initial value, this means for $i \geqslant 2$:

$$C_j^i(0) = C_j^{i-1}(T_{N+2-i} - T_{N+1-i}; \phi),$$
$$D_j^i(0) = D_j^{i-1}(T_{N+2-i} - T_{N+1-i}; \phi).$$

The final values of the last subinterval of τ: $C_j^N(T_1 - t; \phi)$ and $D_j^N(T_1 - t; \phi)$ equal to $C_j(T - t; \phi)$ and $D_j(T - t; \phi)$ that we searched for.

Control variate on Heston Dynamic Model

In the Heston static model, we used control variate to improve the closed formula solution, and we used some methods to resolve the discontinuity problems during the implementation.

The discontinuity problems exist also in Heston dynamic model, and we use the same methods as we cited in section 3.2.4 to handle them.

The control variate that we use here is an option which follows the Heston dynamic model without the volatility of volatility, which means $\xi(t) = 0$. So the volatility part is a deterministic part instead of a volatility process:

$$\mathrm{d}V(t) = k(t)(\theta(t) - V(t))\mathrm{d}t$$

In this part, we use the same notation as for the Heston static model. As parameters are piecewise constant, we compute C_j^{DET} and D_j^{DET} in each subinterval of τ. The only difference is that the initial values for each subinterval ($D_j^{\mathrm{DET}}(0)$, $C_j^{\mathrm{DET}}(0)$) are not 0 except for the first subinterval. So we have the solutions for each subinterval of τ as following:

$$D_j^{\mathrm{DET}} = \frac{u_j\phi i - 1/2\phi^2}{b_j}\left(1 - e^{-b_j\tau}\right) + D_j^{\mathrm{DET}}(0)e^{-b_j\tau}$$
$$C_j^{\mathrm{DET}} = a\left[\left(\frac{u_j\phi i - 1/2\phi^2}{b_j} - D_j^{\mathrm{DET}}(0)\right)\frac{1}{b_j}\left(e^{-b_j\tau} - 1\right) + \frac{(u_j\phi i - 1/2\phi^2)\tau}{b_j}\right] + C_j^{\mathrm{DET}}(0)$$

And the final values of the last subinterval are the values of $C_j^{\mathrm{DET}}(T - t; \phi)$ and $D_j^{\mathrm{DET}}(T - t; \phi)$ that we need to compute f_j^{DET}.

The other way to calculate the option value by using the total volatility σ_T is quite the same as in the Heston static model. We can write:

$$(\sigma_T)^2 = \int_0^T V(t)\mathrm{d}t$$
$$= \sum_{i=1}^{N} \theta_i(T_i - T_{i-1}) + (V_i^0 - \theta_i)\left(-\frac{1}{k_i}\right)\left(e^{-k_i T_i} - e^{-k_i T_{i-1}}\right)$$

where

$$V_i^0 = V(T_{i-1}) = \theta_{i-1} + (V_{i-1}^0 - \theta_{i-1})e^{-k_{i-1}(T_{i-1}-T_{i-2})}$$

and $V_1^0 = V_0$.

The same way as in section 3.2.3, the option price can be expressed by:

$$C(f, V(0), t) = f\tilde{\tilde{P}}_1 - K\tilde{\tilde{P}}_2 + [f\mathcal{N}(d_1) - K\mathcal{N}(d_2)] - \frac{f-K}{2} \tag{4.15}$$

where

$$\tilde{\tilde{P}}_j = 1/2 + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log K}\left(f_j(x, V, T; \phi) - f_j^{\mathrm{DET}}\right)}{i\phi}\right]\mathrm{d}\phi$$

4.4 Piterbarg model

In this section we extend the closed form formula for square root stochastic volatility model with time-dependent parameters. An earlier result by Piterbarg has two constraints on parameters: 1. the correlation between the stochastic processes of underlying and volatility must be zero; 2. mean reversion parameter is constant. By employing our new calculation, we are able to remove the two constraints. The formula can facilitate the calibration of the model.

In the second part, we show an application. We use the result that we have obtained and extend the closed form formula of another stochastic volatility model, SABR model, by removing the constraint on the parameter “skew".

One of the recent results on square root model is the method of “effective" parameters, introduced by Piterbarg [2005] Piterbarg [2003]. The author suggests replacing the time-dependent parameters by their average so that the model can be converted into a static one. The kernel calculation shows how to get the averages from the piecewise parameters in order to approximate the original model as closely as possible. The method is lightweight comparing to asymptotic expansion method and still have good performance.

However, there are two constraints on parameters in the calculation: 1. the correlation between the stochastic processes of underlying and volatility must

be zero; 2. mean reversion parameter is constant. To remove these constraints, we employ Malliavin calculus and perform integration by parts on random variables.

The main result of this paper includes two different closed form formula for square root model. In the first formula, we follow the same line with Piterbarg, and find all the effective parameters. The other way is to transform the model to Heston model, of which the closed form formula is already known. Interestingly, when changing the parameters, we have a constraint: to keep the continuity of the underlying process, the parameter "skew" must be constant, and to average the skew we once more find the need of zero correlation. In a word, it turns out to be the same question with the previous formula. Therefore, the solution for the first formula can apply to the second.

This article is organized in the following way. In chapter 2 we present square root SV model. In chapter 3 we see the transform to Heston model. In chapter 4 we review Piterbarg's result. Chapter 5 is the main result: we show the formula without the previous constraint. In the last chapter, we show one application on SABR model.

4.4.1 Diffusion and formula for vanilla option

In this section, we present the square root stochastic volatility model. Suppose the price of an underlying $S(t)$ follows a shifted log-normal diffusion.

$$\mathrm{d}S(t) = \sigma(\beta S(t) + (1-\beta)S(0))\sqrt{z(t)}\mathrm{d}U(t) \tag{4.16}$$

$$\mathrm{d}z(t) = \theta(z(0) - z(t))\mathrm{d}t + \gamma\sqrt{z(t)}\mathrm{d}V(t) \tag{4.17}$$

where $\mathrm{d}\langle U, V\rangle = \rho\mathrm{d}t$.

In this model, $z(t)$ is the variance of $S(t)$; it follows a CIR process. $z(0)$ is the initial value of variance $z(t)$. σ is the deterministic part of the volatility. β is called skew parameter; usually, $0 \leqslant \beta \leqslant 1$; when $\beta = 1$, the process is log-normal; when $\beta = 0$, the process is normal. θ is the mean reversion of the volatility process. γ is the volatility of volatility. $U(t)$ and $V(t)$ are two Brownian motions and their correlation is ρ.

We can extend the model (4.16), (4.17) with time-dependent parameters. To distinguish the extended model from the original one, we call model (4.16), (4.17) static model and call extended model dynamic model. The dynamic model is the primary target in this article.

$$\mathrm{d}S(t) = \sigma(t)(\beta(t)S(t) + (1-\beta(t))S(0))\sqrt{z(t)}\mathrm{d}U(t) \tag{4.18}$$

$$\mathrm{d}z(t) = \theta(t)(z(0) - z(t))\mathrm{d}t + \gamma(t)\sqrt{z(t)}\mathrm{d}V(t) \tag{4.19}$$

where $\mathrm{d}\langle U, V\rangle = \rho(t)\mathrm{d}t$.

The parameters of the model are continuous. But for practical use, we prefer to supposing the parameters to be piecewise constant. It is because in calibration we fit the price of options on certain maturities. There is no information about how the parameters evolve between the maturities. So the assumption of piecewise constant parameters won't damage the generality of the model.

More precisely and mathematically, we can present the assumption of piecewise constant parameters in the following way. Suppose that we have n different maturities $t_1, t_2, ..., t_n$. Therefore we divide the parameters in n periods: $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$. When $t_{i-1} < t \leqslant t_i$, the parameters are constant on the interval, we note $\sigma(t) = \sigma_i$, $\gamma(t) = \gamma_i$, $\rho(t) = \rho_i$, $\theta(t) = \theta_i$,$\beta(t) = \beta_i$.

Now we show the closed formulas of European options on our model. First we give the formula for the static model, then for the dynamic model. In both cases, we suppose that the option has strike K and maturity T.

Closed formula for static model

We list the closed formula for static model in this subsection. It will be referenced in following sections.

We rewrite the static model (4.16), (4.17) here.

$$\begin{aligned} \mathrm{d}S(t) &= \sigma(\beta S(t) + (1-\beta)S(0))\sqrt{z(t)}\mathrm{d}U(t) \\ \mathrm{d}z(t) &= \theta(z(0) - z(t))\mathrm{d}t + \gamma\sqrt{z(t)}\mathrm{d}V(t) \end{aligned}$$

where $\mathrm{d}\langle U, V\rangle = \rho \mathrm{d}t$.

This model can be regarded as a shifted Heston model. The closed formula of Heston model can refer the result of Lewis [2000] and Mikhailov and Nogel [2003a] in last section. So all we need to do is to change the variables to convert the model to Heston model.

Let $\tilde{S}(t) = S(t) + \frac{1-\beta}{\beta}S(0)$. $\tilde{z}(t) = \sigma^2\beta^2 z(t)$. $\tilde{\gamma} = \sigma\beta\gamma$. Then we get a Heston model.

$$\mathrm{d}\tilde{S}(t) = \tilde{S}(t)\sqrt{\tilde{z}(t)}\mathrm{d}U(t) \tag{4.20}$$

$$\mathrm{d}\tilde{z}(t) = \theta(\tilde{z}(0) - \tilde{z}(t))\mathrm{d}t + \tilde{\gamma}\sqrt{\tilde{z}(t)}\mathrm{d}V(t) \tag{4.21}$$

where $\mathrm{d}\langle U, V\rangle = \rho \mathrm{d}t$, with $\tilde{S}(0) = \frac{1}{\beta}S(0)$.

Here we list the closed formula for Heston model. It bases on the Fourier transform of characteristic function.

Proposition 23. *Suppose we have a Heston model.*

$$\begin{aligned} \mathrm{d}S(t) &= S(t)\sqrt{V(t)}\mathrm{d}W_1 \\ \mathrm{d}V(t) &= k(\theta - V(t))\mathrm{d}t + \xi\sqrt{V(t)}\mathrm{d}W_2 \\ \mathrm{d}\langle W_1, W_2\rangle_t &= \rho\mathrm{d}t \end{aligned} \tag{4.22}$$

Then the price of a European call option with maturity T and strike K is,

$$Call = S(0)P_1(0) - KP_2(0)$$

where

$$\begin{aligned} P_j(t) &= 1/2 + \frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-i\phi\log(K)}f_j(x,V,t,T;\phi)}{i\phi}\right]\mathrm{d}\phi \\ f_j(x,V,t,T;\phi) &= e^{C_j(T-t;\phi)+D_j(T-t;\phi)V+i\phi x} \\ C_j(\tau;\phi) &= \frac{a}{\xi^2}\left\{(b_j - \rho\xi\phi i + d_j)\tau - 2\log\left[\frac{1-g_j e^{d_j\tau}}{1-g_j}\right]\right\} \\ D_j(\tau;\phi) &= \frac{b_j - \rho\xi\phi i + d_j}{\xi^2}\cdot\frac{1-e^{d_j\tau}}{1-g_j e^{d_j\tau}} \\ g_j &= \frac{b_j - \rho\xi\phi i + d_j}{b_j - \rho\xi\phi i - d_j} \\ d_j &= -\sqrt{(\rho\xi\phi i - b_j)^2 - \xi^2(2u_j\phi i - \phi^2)} \end{aligned}$$

with $j = 1,2$, $u_1 = 1/2$, $u_2 = -1/2$, $a = k\theta$, $b_1 = k - \rho\xi$, $b_2 = k$.

4.4.2 Piterbarg Dynamic model

Dynamic model with zero correlation

In this subsection we show the result of Piterbarg [2005]. His idea is to find a static model to mimic the original dynamic model, in the objective of having the same price for a European options. This technique is called parameter averaging.

Here two constraints are added on the dynamic model: mean reversion θ is constant and the correlation ρ is zero. The latter assumption is strong because that correlation controls the form of volatility smile. How to release this constraint is the main purpose of this article and will be presented in the following sections.

$$\mathrm{d}S(t) = \sigma(t)(\beta(t)S(t) + (1-\beta(t))S(0))\sqrt{z(t)}\mathrm{d}U(t) \tag{4.23}$$

$$\mathrm{d}z(t) = \theta(z(0) - z(t))\mathrm{d}t + \gamma(t)\sqrt{z(t)}\mathrm{d}V(t) \tag{4.24}$$

where $\mathrm{d}\langle U, V\rangle = 0$.

Piterbarg gave three approximate formula to convert $\gamma(t)$, $\beta(t)$ and $\sigma(t)$ to three constants, namely η, b and λ. After applying these three formulas, we have a static model. The closed formula of a static model have already solved in section (4.4.1).

$$\mathrm{d}S(t) = \lambda(bS(t) + (1-b)S(0))\sqrt{z(t)}\mathrm{d}U(t) \tag{4.25}$$

$$\mathrm{d}z(t) = \theta(z(0) - z(t))\mathrm{d}t + \eta\sqrt{z(t)}\mathrm{d}V(t) \tag{4.26}$$

where $\mathrm{d}\langle U, V\rangle = 0$.

Parameter averaging formula To approximate the same price for European options with maturity T, the parameters in model (4.25) (4.26) should satisfy the following equations. Basically, we use η, b and λ in the static model (4.25) (4.26) to replace $\gamma(t)$, $\beta(t)$ and $\sigma(t)$ in model (4.23) (4.24).

Proposition 24. *(Piterbarg [2005]) For parameter η,*

$$\eta^2 = \frac{\int_0^T \gamma^2(t)p(t)\mathrm{d}t}{\int_0^T p(t)\mathrm{d}t}$$

with $p(t) = \int_r^T \mathrm{d}s \int_s^T \mathrm{d}t\sigma^2(s)e^{-\theta(t-s)}e^{-2\theta(s-r)}$

Proof. $\sigma(t)$ is the determinist part of the variance; $z(t)$is the volatility part of the variance on the other hand. It can change the curvature of the implied volatility smile. The curvature is$\int_0^T \sigma^2(t)z(t)\mathrm{d}t - \sigma^2(\bar{t})z(\bar{t})$; as the term$\sigma^2(\bar{t})z(\bar{t})$ is constant, therefore, $\int_0^T \sigma^2(t)z(t)\mathrm{d}t$ should be the same in both two models. This gives the following equation:

$$\mathbb{E}\left(\int_0^T \sigma^2(t)\bar{z}(t)dt\right)^2 = \mathbb{E}\left(\int_0^T \sigma^2(t)z(t)dt\right)^2 \tag{4.27}$$

Calculation can be detailed in the following way. Fubini's theorem is employed.

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_0^T \sigma(t)^2 z(t)\mathrm{d}t\right)^2\right] \\
=&\mathbb{E}\left[\left(\int_0^T \sigma(t)^2 z(t)\mathrm{d}t\right)\left(\int_0^T \sigma(s)^2 z(s)\mathrm{d}s\right)\right] \\
=&\mathbb{E}\left[\int_0^T\int_0^T \sigma(t)^2\sigma(s)^2 z(t)z(s)\mathrm{d}s\mathrm{d}t\right] \\
=&2\int_0^T\int_0^t \sigma(t)^2\sigma(s)^2\mathbb{E}\left[z(t)z(s)\right]\mathrm{d}s\mathrm{d}t \qquad (4.28)\\
=&2\int_0^T\int_0^t \sigma(t)^2\sigma(s)^2\left(z_0\int_0^s e^{\theta(2r-s-t)}\gamma(r)^2\mathrm{d}r + z_0^2\right)\mathrm{d}s\mathrm{d}t \\
=&2z_0\int_0^T\int_r^T\int_s^T \sigma(t)^2\sigma(s)^2 e^{\theta(2r-s-t)}\gamma(r)^2\mathrm{d}t\mathrm{d}s\mathrm{d}r + 2z_0^2\int_0^T\int_0^t \sigma(t)^2\sigma(s)^2\mathrm{d}s\mathrm{d}t \\
=&2z_0\int_0^T \gamma(r)^2\rho(r)\mathrm{d}r + z_0^2\int_0^T\int_0^T \sigma(t)^2\sigma(s)^2\mathrm{d}s\mathrm{d}t
\end{aligned}
$$

Where, $\rho(r) = \int_r^T \mathrm{d}s\int_s^T \mathrm{d}t\sigma(t)^2\sigma(s)^2 e^{-\theta(t-s)}e^{-2\theta(s-r)}$

We substitute this result to (4.27), and get $2z_0\int_0^T \gamma(r)^2\rho(r)\mathrm{d}r = 2z_0\int_0^T \eta^2\rho(r)\mathrm{d}r$

This gives that, $\eta^2 = \frac{\int_0^T \gamma(r)^2\rho(r)\mathrm{d}r}{\int_0^T \rho(r)\mathrm{d}r}$

In the above equation we use the covariance of the $z(s)$, which can be derived in the following way.

To calculate the covariance of $z(s)$, this follows:

$$\mathrm{d}z(s) = \theta(z_0 - z(s))\mathrm{d}s + \gamma(s)\sqrt{z(s)}\mathrm{d}V(s) \quad , \quad z(0) = z_0$$

Let $x(s) = e^{\theta s}(z(s) - z_0)$.

So we can derive that $x(0) = 0$,$\mathbb{E}[x(s)] = 0$ and $z(s) = x(s)e^{-\theta s} + z_0$.

Therefore,

$$
\begin{aligned}
\mathrm{d}x(s) &= \theta e^{\theta s}(z(s) - z_0)\mathrm{d}s + e^{\theta s}\mathrm{d}z(s) \\
&= e^{\theta s}\gamma(s)\sqrt{x(s)e^{-\theta s} + z_0}\mathrm{d}V(s)
\end{aligned}
$$

So the covariance of $\mathbb{E}(x(s)x(t))$, $s \leqslant t$ is

$$\begin{aligned}\mathbb{E}(x(s)x(t)) &= \mathbb{E}\left[\int_0^s e^{\theta r}\gamma(r)\sqrt{x(r)e^{-\theta r}+z_0}\mathrm{d}V(r)\int_0^t e^{\theta r}\gamma(r)\sqrt{x(r)e^{-\theta r}+z_0}\mathrm{d}V(r)\right]\\ &= \mathbb{E}\left[\int_0^s e^{2\theta r}\gamma(r)^2\left(x(r)e^{-\theta r}+z_0\right)\mathrm{d}r\right]\\ &= z_0\int_0^s e^{2\theta r}\gamma(r)^2\mathrm{d}r\end{aligned}$$

Now we substitute $x(s) = e^{\theta s}(z(s)-z_0)$into the left-hand side of the formula above.

$$\mathbb{E}\left[\left(e^{\theta s}(z(s)-z_0)\right)\left(e^{\theta t}(z(t)-z_0)\right)\right] = z_0\int_0^s e^{2\theta r}\gamma(r)^2\mathrm{d}r$$

$$\mathbb{E}\left[(z(s)-z_0)(z(t)-z_0)\right] = z_0e^{-\theta(s+t)}\int_0^s e^{2\theta r}\gamma(r)^2\mathrm{d}r$$

We note that,$\mathbb{E}[z(s)] = z_0$

So

$$\mathbb{E}[z(s)z(t)] = z_0^2 + z_0e^{-\theta(s+t)}\int_0^s e^{2\theta r}\gamma(r)^2\mathrm{d}r \tag{4.29}$$

which proves the result.

□

Proposition 25. *For parameter b,*

$$\begin{aligned}b &= \int_0^T \beta(t)w(t)\mathrm{d}t\\ \text{with } w(t) &= \frac{v^2(t)\sigma^2(t)}{\int_0^T v^2(t)\sigma^2(t)\mathrm{d}t}\\ v^2(t) &= z_0^2\int_0^t \sigma^2(s)\mathrm{d}s + z_0\eta^2e^{-\theta t}\int_0^t \sigma^2(s)\frac{e^{\theta s}-e^{-\theta s}}{2\theta}\mathrm{d}s\end{aligned}$$

Proposition 26. *For parameter λ, we need to solving an equation numerically.*

$$\mathbb{E}\left[g\left(\int_0^T \sigma^2(t)z(t)\mathrm{d}t\right)\right] = \mathbb{E}\left[g\left(\lambda^2\int_0^T z(t)\mathrm{d}t\right)\right]$$

with $g(x) = \frac{S_0}{b}\left(2N\left(\frac{1}{2}b\sqrt{x}\right)-1\right)$

$N(\cdot)$ is standard normal cumulative distribution function.

Proof. The proof of proposition 25 and 26 can be found in appendix A.1 and A.2. □

Remark An interesting fact is that the zero correlation condition is only used in the proof of proposition 25. In another word, proposition 24 and 26 can work with a non-zero correlation model, such as model (4.18) (4.19). So we can target to improve this calculation by relaxing the need of the zero correlation assumption.

Remark 2 The formulas in Proposition 24, 25 and 26 can apply to both continuous parameters and piecewise constant parameters.

Dynamic model with constant skew parameter In this part, we follow the idea in section 4.4.1, and regard the square root SV model as a shifted Heston model. Since we know the closed formula for dynamic Heston model, all we have to do is to find a change of variables to connect these two models.

Suppose the skew parameter is constant; that is to say $\beta(t) \equiv b$ in model (4.18) (4.19). We re-write square root SV model here.

$$\mathrm{d}S(t) = \sigma(t)(bS(t) + (1-b)S(0))\sqrt{z(t)}\mathrm{d}U(t) \tag{4.30}$$

$$\mathrm{d}z(t) = \theta(t)(z(0) - z(t))\mathrm{d}t + \gamma(t)\sqrt{z(t)}\mathrm{d}V(t) \tag{4.31}$$

where $\mathrm{d}\langle U, V\rangle = \rho(t)\mathrm{d}t$.

Then we take the change of variables in section 4.4.1, saying $\tilde{S}(t) = S(t) + \frac{1-b}{b}S(0)$. The model can change to a Heston model.

$$\mathrm{d}\tilde{S}(t) = \sigma(t)\tilde{S}(t)b\sqrt{z(t)}\mathrm{d}U(t) \tag{4.32}$$

$$\mathrm{d}z(t) = \theta(t)(z(0) - z(t))\mathrm{d}t + \gamma(t)\sqrt{z(t)}\mathrm{d}V(t) \tag{4.33}$$

where $\mathrm{d}\langle U, V\rangle = \rho(t)\mathrm{d}t$. $\tilde{S}(0) = \frac{S(0)}{b}$.

The hypothesis of constant $\beta(t)$ is necessary because it keeps the continuity of $\tilde{S}(t)$. Otherwise we cannot use the result of Heston model on $\tilde{S}(t)$. The following proposition shows a closed formula for Heston model.

Proposition 27. *The price for a European call option with maturity T, strike K under model (4.32) (4.33) is given by the following equation.*

$$V = \tilde{S}(0)P_1(0) - \tilde{K}P_2(0)$$

with $\tilde{K} = K + \frac{1-b}{b}S(0)$

Where, for $j = 1,2$, *note* $u_1 = 1/2$, $u_2 = -1/2$, $a = \theta(t)z(0)$, $b_1 = \theta(t) - \rho(t)\gamma(t)\sigma(t)b$, $b_2 = \theta(t)$

$$P_j(t) = 1/2 + \frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-i\phi ln(\tilde{K})} f_j(x,V,t,T;\phi)}{i\phi}\right]\mathrm{d}\phi$$

$$f_j(x,V,t,T;\phi) = e^{C_j(T-t;\phi)+D_j(T-t;\phi)V+i\phi x}$$

With

$$C_j(\tau;\phi) = \frac{a}{\gamma^2(T-\tau)}\left\{(b_j - \rho(T-\tau)\gamma(T-\tau)\sigma(T-\tau)b\phi i + d_j)\tau - 2\log\left(\frac{1-g_j e^{d_j\tau}}{1-g_j}\right)\right\}$$

$$D_j(\tau;\phi) = \frac{b_j - \rho(T-\tau)\gamma(T-\tau)\sigma(T-\tau)b\phi i + d_j}{\gamma^2(T-\tau)}\cdot\frac{1-e^{d_j\tau}}{1-g_j e^{d_j\tau}}$$

$$g_j = \frac{b_j - \rho(T-\tau)\gamma(T-\tau)\sigma(T-\tau)b\phi i + d_j}{b_j - \rho\gamma\phi\sigma(T-\tau)\beta i - d_j}$$

$$d_j = -\sqrt{\left(\rho\gamma\sigma(T-\tau)b\phi i - b_j\right)^2 - \gamma^2\left(2u_j\sigma^2(T-\tau)b^2\phi i - \sigma^2(T-\tau)b^2\phi^2\right)}$$

Note that $i = \sqrt{-1}$ *is the unit of pure imaginary number.*

Proof. The proof is very similar to the proof in Mikhailov and Nogel [2003a] on dynamic Heston model. □

Remark The proposition can apply to both continuous parameters and piecewise constant parameters.

Dynamic model without constraint

We rewrite the model here.

$$\mathrm{d}S(t) = \sigma(t)(\beta(t)S(t) + (1-\beta(t))S(0))\sqrt{z(t)}\mathrm{d}U(t) \tag{4.34}$$

$$\mathrm{d}z(t) = \theta(t)(z(0) - z(t))\mathrm{d}t + \gamma(t)\sqrt{z(t)}\mathrm{d}V(t) \tag{4.35}$$

where $\mathrm{d}\langle U,V\rangle = \rho(t)\mathrm{d}t$.

First we analyze why we introduce the constraints in previous sections. In section 4.4.2, we use the condition of zero correlation in the proof of proposition 24. More precisely, in equation (A.6) of the proof of proposition 25 we use the condition of zero correlation to calculate $v^2(t)$. The definition of $v^2(t)$ is in (A.4).

For the closed formula in section 4.4.2, we have a constraint of constant skew $\beta(t)$. If we first use parameter averaging technique on parameter $\beta(t)$, we can remove this constraint. However it introduces the constraint of zero correlation. We find the same problem as in the previous case.

In conclusion, both problems need a new way to average the parameter $\beta(t)$. More precisely, we need to calculate $v^2(t)$ in case of non-zero correlation. The method we propose here is to use the formula of integration by parts.

Proposition 28. *Suppose all the parameters of the model are piecewise constant. Suppose we have n periods:* $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$. *The parameters are piecewise time-dependent. When* $t_{i-1} < t < t_i$, *we note* $\sigma(t) = \sigma_i$, $\gamma(t) = \gamma_i$, $\rho(t) = \rho_i$, $\theta(t) = \theta_i$,$\beta(t) = \beta_i$, $B_i = 4\gamma_i\sigma_i\rho_i - \theta_i$.

Then the value of $v^2(t)$ *in equation (A.4) with* $t_{i-1} < t \leqslant t_i$ *is given by:*

$$\begin{aligned} v^2(t) =& v^2(t_{i-1})e^{(t+1)B_i} + \left(e^{(t-t_{i-1})B_i} - 1\right)\sigma_i^2 z_0 \frac{1}{B_i}\left(z_0 + \frac{1}{2\theta_i}\gamma_i^2\right) \\ &+ z_0^2\sigma_i^2\theta_i\frac{1}{B_i}\left(e^{(t-t_{i-1})B_i}B_i t_{i-1} + e^{(t-t_{i-1})B_i} - B_i t - 1\right) \\ &- \frac{1}{2\theta_i(B_i+2\theta_i)}\sigma_i^2 z_0\gamma_i^2\left(e^{(t-t_{i-1})B_i - 2\theta_i t_{i-1}} - e^{-2\theta_i t}\right) \end{aligned}$$

with $v^2(t_0) = 0$

Proof. See the proof in appendix A.3. □

Remark Obviously, the proposition can only apply to the case of piecewise constant parameters.

Piterbarg Dynamic Model with $\rho = 0$ and constant k

In this case, we use effective medium values of the parameters and we change the problem into the constant case.

The effective values for each parameter are computed in Piterbarg [2005] and Piterbarg [2003]. Note $\bar{\xi}$ the effective volatility of volatility, $\bar{\beta}$ the effective skew, and $\bar{\sigma}$ the effective volatility, the final constant model that replaces the dynamic model is:

$$\begin{aligned} \mathrm{d}S(t) &= \bar{\sigma}\left[\bar{\beta}S(t) + (1-\bar{\beta})F(0)\right]\sqrt{V(t)}\mathrm{d}W_1 \\ \mathrm{d}V(t) &= k(V(0) - V(t))\mathrm{d}t + \bar{\xi}\sqrt{V(t)}\mathrm{d}W_2 \\ <\mathrm{d}W_1, \mathrm{d}W_2> &= 0 \end{aligned}$$

The effective volatility of volatility $\bar{\xi}$:

$$(\bar{\xi})^2 = \frac{\int_0^T \xi^2(s)g(s)\mathrm{d}s}{\int_0^T g(s)\mathrm{d}s}$$

where $g(.)$ satisfies:

$$g(t) = \int_t^T \mathrm{d}s \int_s^T \mathrm{d}r\sigma^2(r)\sigma^2(s)e^{-\theta(r-s)}e^{-2\theta(s-t)}$$

The effective skew $\bar{\beta}$:

$$\bar{\beta} = \int_0^T \beta(s)w(s)\mathrm{d}s$$

where the weights $w(.)$ is given by:

$$\begin{aligned} w(t) &= \frac{\nu^2(t)\sigma^2(t)}{\int_0^T \nu^2(t)\sigma^2(t)\mathrm{d}t} \\ \nu^2(t) &= V_0^2 \int_0^t \sigma^2(s)\mathrm{d}s + V_0(\bar{\xi})^2 e^{-\theta t} \int_0^t \sigma^2(s)\frac{e^{\theta s} - e^{-\theta s}}{2\theta}\mathrm{d}s \end{aligned}$$

The effective volatility $\bar{\sigma}$ is the solution of:

$$\psi_0\left(-\frac{g''(\zeta)}{g'(\zeta)}\bar{\sigma}^2\right) = \psi\left(-\frac{g''(\zeta)}{g'(\zeta)}\right)$$

where

$$\zeta = V_0 \int_0^T \sigma^2(t)\mathrm{d}t$$

$$\psi(\mu) = \mathbb{E}\left[\exp\left(-\mu V(T)\right)\right]$$

$$\psi_0(\mu) = \mathbb{E}\left[\exp\left(-\mu \int_0^T V(t)\mathrm{d}t\right)\right]$$

Piterbarg Dynamic Model with constant β

In this case, by doing the following transformations:

$$L(t) = F(t) + \frac{1-\beta}{\beta}F(0)$$

$$Z(t) = \sigma^2(t)\beta^2 V(t)$$

we have:

$$\begin{cases} \mathrm{d}L(t) = L(t)\sqrt{Z(t)}\mathrm{d}W_1 \\ \mathrm{d}Z(t) = k(t)(\theta(t) - Z(t))\mathrm{d}t + \tilde{\xi}(t)\sqrt{Z(t)}\mathrm{d}W_2 \\ < \mathrm{d}W_1, \mathrm{d}W_2 > = \rho(t)\mathrm{d}t \\ L(0) = \dfrac{F(0)}{\beta} \\ Z(0) = \sigma^2(0)\beta^2 V(0) \end{cases}$$

where:

$$\theta(t) = \sigma^2(t)\beta^2 V(0)$$
$$\tilde{\xi}(t) = \xi(t)\sigma(t)\beta$$

Now as the parameters are piecewise constant, we use the method that we introduced for Heston dynamic model to get the option price.

Note: In the case when β is piecewise constant too, the tranformation: $L(t) = F(t) + \frac{1-\beta}{\beta}F(0)$ introduces jumps for the forward at the end of each constant parameter period. So $L(t)$ is not continuous. We can't use exactly the same method as we used in the Heston model.

4.4.3 Summary

Finally, let us summarize the result here. We showed two paths to find a closed formulas without constraints on parameters.

First path, we use proposition 24, 25 and 26 with replacing the calculation of $v^2(t)$ by proposition 28. As we convert the dynamic model to a static model, we can use proposition 23 to get a closed formula. In this case, the correlation is piecewise constant. This constraint may be relaxed by averaging a time-dependent correlation to a constant correlation. That is left for future research.

Second path, we use the calculation of $\beta(t)$ in proposition 23 to convert $\beta(t)$ to b, where the calculation of $v^2(t)$ is replaced by proposition 28. Then we use proposition 27 to get a closed formula. In this case, correlation and skew are both time-dependent.

As we can see, proposition 28 enables us to avoid the constraint of zero correlation in both paths.

4.5 SABR model

4.5.1 Diffusion

SABR model is a stochastic volatility model which name stands for "Stochastic-α-β-ρ model". The model and its dynamic variation are first introduced by Hagan

et al. [2002]. It is mainly used for forward interest rate and other asset classes including commodity. The diffusion of SABR model is defined as follows.

Definition 29 (SABR model)**.** *Denote $F_t = F(t,T)$ the forward of underlying (i.e. commodity) at time T, observing at time t, and denote α_t the volatility of underlying at date t, then SABR model is given by:*

$$\mathrm{d}F_t = \alpha_t F_t^{\beta} \mathrm{d}W_1 \tag{4.36}$$
$$\mathrm{d}\alpha_t = \gamma \alpha_t \mathrm{d}W_2 \tag{4.37}$$

Where W_1 and W_2 are two standard Brownian motion with correlation ρ. Skew parameter β and vol of vol parameter γ are constant.

The parameters in this model are

1. $\alpha(0)$ - the initial value of volatility at time $t = 0$,
2. β - skew parameter. When $\beta = 0$, the model becomes to normal; when $\beta = 1$, the model is log-normal model,
3. γ - vol of vol parameter. When $\gamma = 0$, the model reduces to a CEV model.
4. ρ - the correlation between the two Brownian processes W_1 and W_2.

Remark By using forward F_t instead of spot price at time t, the model already includes the information of convenience yield and discounting term in forward. Therefore, there is no drift term in the stochastic diffusion. Moreover, we have $F_T = F(T,T) = S_T$ the forward at time T observed at time T, which is obviously the spot price S_T at time T.

If we let parameters depend on time, namely β to $\beta(t)$, γ to $\gamma(t)$ and ρ to $\rho(t)$, then we extend SABR model to dynamic SABR model.

Definition 30 (Dynamic SABR model)**.** *Denote $F_t = F(t,T)$ the forward of underlying (i.e. commodity) at time T, observing at time t, and denote α_t the volatility of underlying at date t, then dynamic SABR model is given by:*

$$\mathrm{d}F_t = \alpha_t F_t^{\beta(t)} \mathrm{d}W_1 \tag{4.38}$$
$$\mathrm{d}\alpha_t = \gamma(t) \alpha_t \mathrm{d}W_2 \tag{4.39}$$

Where W_1 and W_2 are two standard Brownian motion with correlation $\rho(t)$. Skew parameter $\beta(t)$ and vol of vol parameter $\gamma(t)$ are time-dependent.

4.5.2 Static SABR model

Proposition 31 (Vanilla option price under SABR model, Hagan et al. [2002])**.** *Under SABR model, the price (at time $t = 0$) V of a vanilla call option with strike K and maturity T is defined as $V = D(0,T)\mathbb{E}[(S_T - K)^+|\mathscr{F}_0]$. Here $D(0,T)$ is the discount factor, which equals to the value of 1 unit bond at time 0 with maturity at time T. Forward F_t follows the same definition in SABR model. $\mathscr{F}_0$ is the filtration of probability space at time 0.*

Then we can have the formula for V in form of Black Scholes formula:

$$V = D(0,T)\left(F_0 \mathcal{N}(d_1) - K\mathcal{N}(d_2)\right) \tag{4.40}$$

where $F_0 = F(0,T)$ is the forward observed at time 0 with maturity T. $\mathcal{N}(x) = \frac{1}{2\pi}\int_{-\infty}^{x} e^{-t^2/2}\mathrm{d}t$ is the cumulative distribution function of standard normal distribution. $d_{1,2} = \log(F_0/K) \pm \frac{1}{2}\frac{{\sigma_{BS}}^2 T}{\sigma_{BS}\sqrt{T}}$. Implied Black Scholes volatility σ_{BS} is given by

$$\begin{aligned}\sigma_{BS} = &\frac{\alpha}{(F_0K)^{(1-\beta)/2}} \cdot \frac{1}{1 + \frac{(1-\beta)^2}{24}\ln^2 F_0/K + \frac{(1-\beta)^4}{1920}\ln^4 F_0/K + \ldots} \cdot \frac{\zeta}{x(\zeta)} \\ &\cdot \left\{1 + \left[\frac{(1-\beta)^2\alpha^2}{24(F_0K)^{1-\beta}} + \frac{\rho\alpha\gamma\beta}{4(F_0K)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\gamma^2\right]T + \ldots\right\}\end{aligned} \tag{4.41}$$

where

$$\zeta = \frac{\gamma}{\alpha}\sqrt{F_0K}\log(F_0/K) \tag{4.42}$$

$$x(\zeta) = \log\left(\frac{\sqrt{1 - 2\rho\zeta + \zeta^2} - \rho + \zeta}{1 - \rho}\right) \tag{4.43}$$

Remark The pricing of SABR model with constant parameters is well presented by Hagan et al. [2002]. In practise, Benhamou and Croissant [2007] use Taylor series to get a more precise implied volatility when we are at the money. Closed formula for static SABR model is given in Hagan et al. [2002].

4.5.3 Dynamic SABR model

For dynamic model, Hagan et al. [2002] showed a closed formula with constant skew $\beta(t) \equiv \beta$. In Wang [2009], the parameter averaging technique is employed; it shows another closed formula with zero correlation $\rho(t) \equiv 0$.

We follow the same line with Wang [2009], applying the parameter averaging technique. We loosen the constraint of zero correlation $\rho(t) \equiv 0$ to the condition of $\rho(t) \equiv \rho$ being constant. We target to convert the dynamic model in equation 4.38 and 4.39 to a static model in the following form.

$$\begin{aligned} \mathrm{d}\alpha(t) &= \nu\alpha(t)\mathrm{d}V(t) \\ \mathrm{d}S(t) &= \alpha(t)S(t)^b\mathrm{d}U(t) \\ \mathrm{d}\langle U, V\rangle_t &= \rho\mathrm{d}t \end{aligned}$$

For parameters $\gamma(t)$ of the dynamic model, the calculation doesn't need the zero correlation condition in Wang [2009]; therefore we take directly the result from the paper and we don't list the result in this thesis. Here we focus on the calculation of $\beta(t)$.

Proposition 32. *Suppose all the parameters of the model are piecewise constant. Suppose we have n periods:* $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$. *The parameters are piecewise time-dependent. When* $t_{i-1} < t < t_i$, *we note* $\gamma(t) = \gamma_i$, $\rho(t) = \rho_i$, $\beta(t) = \beta_i$, $B_i = 4\gamma_i\sigma_i\rho_i$. *For parameter b,*

$$\begin{aligned} b &= \int_0^T \beta(t)w(t)\mathrm{d}t \\ \textit{with } w(t) &= \frac{v^2(t)\sigma^2(t)}{\int_0^T v^2(t)\sigma^2(t)\mathrm{d}t} \\ v^2(t) &= v^2(t_{i-1})e^{(t+1)B_i} + \frac{z_0}{4\rho_i}\left(\frac{1}{4\rho_i} + \sigma_i\gamma_i t + \frac{\sigma_i z_0}{\gamma_i}\right)\left(e^{B_i(t-t_{i-1})} - 1\right) \end{aligned}$$

Proof. We use the same calculation with the proof of proposition 28. The only difference is that the volatility process (equation 4.37) of SABR has no drift. Therefore, we take the limit of the result in proposition 28 when drift term $\theta(t)$ tends to zero. Then we get the result of proposition 32. □

Remark With proposition 32, we convert the dynamic model to a static SABR model. Then we can use the closed formula for a static SABR model.

In proposition 32, correlation is actually piecewise constant. It means that the constraint of correlation being constant can be removed if we can find a way to averaging a piecewise constant correlation. That is left for future research.

Vol of vol expansion for SABR with constant β

For a SABR model with constant β, we can use the effective medium parameter method. Firstly, the model is:

$$\begin{cases} \mathrm{d}F(t) = \alpha(t)F(t)^\beta\mathrm{d}W_1 \\ \mathrm{d}\alpha(t) = \gamma(t)\alpha(t)\mathrm{d}W_2 \\ < \mathrm{d}W_1, \mathrm{d}W_2 >= \rho(t)\mathrm{d}t \end{cases}$$

as in the constant case, we will rewrite $\alpha \longrightarrow \varepsilon\alpha$ and $\gamma \longrightarrow \varepsilon\gamma$, with $\varepsilon \ll 1$ which is the distinguished limit. And we get the answers in terms of the original variables

by replacing $\varepsilon\alpha \longrightarrow \alpha$ and $\varepsilon\gamma \longrightarrow \gamma$ after obtaining the results.

Let $V(t,f,\alpha)$ be the value of a European call option at date t, with the economy in state $F(t) = f$, $\alpha(t) = \alpha$. Let also T be the option's expiry, and K be its strike.

As in the constant parameter case, by defining the probability density $p(t,f,\alpha;T,F,A)$:

$$p(t,f,\alpha;T,F,A)\mathrm{d}F\mathrm{d}A = \mathrm{prob}\{F < F(T) < F + dF, A < \alpha(T) < A + dA \quad | \quad F(t) = f, \alpha(t) = \alpha\}$$

And

$$P(t,f,\alpha;T,K) = \int_{-\infty}^{\infty} A^2 p(t,f,\alpha;T,F,A)\mathrm{d}A$$

we have:

$$V(t,f,\alpha) = [f-K]^+ + \frac{1}{2}\varepsilon^2 K^{2\beta}\int_t^T P(t,f,\alpha;s',K)\mathrm{d}s'$$

where $P(t,f,\alpha;T,K)$ satisfies the forward problem:

$$P_t + \frac{1}{2}\varepsilon^2\left\{\alpha^2 C^2(f)P_{ff} + 2\eta(t)\alpha^2 C(f)P_{f\alpha} + v^2(t)\alpha^2 P_{\alpha\alpha}\right\} = 0 \qquad \text{for } t < T$$

$$P = \alpha^2\delta(f-K) \qquad \text{for } t = T$$

where

$$\eta(t) = \rho(t)\gamma(t)$$
$$v(t) = \gamma(t)$$

We solve this problem by using an effective media strategy. Our objective is to determine constant values $\bar{\eta}$ and $\bar{v}$ yield the same option price as the time dependent coefficients $\eta(t)$ and $v(t)$. Then the problem is reduced to the non-dynamic SABR model solved before.

Noting $C(f) = f^\beta$, and $C(0) = K^\beta$. By defining the variate z:

$$z = \frac{1}{\varepsilon\alpha}\int_K^f \frac{\mathrm{d}f'}{C(f')}$$

we have the probability density P is Gaussian in z.

We carry out this strategy to resolve the near-Hamiltonian system by applying the same series of time-independent transformations $\mathcal{L}$ that was used to solve the non-dynamic SABR model before. The transformations here are in terms of the constants $\bar{\eta}$ and $\bar{v}$ to be determined later.

We name $\mathcal{L}$ all the transformations used:

1. $z = \frac{1}{\varepsilon\alpha}\int_K^f \frac{\mathrm{d}f'}{C(f')}$;
2. $P = \frac{\alpha}{\varepsilon}K^{-\beta}\hat{P}$;
3. $\hat{P} = \sqrt{\frac{f^\beta}{K^\beta}}H$;
4. $H = e^{\varepsilon^2\rho\gamma b_1 z^2/4}\hat{H}$;
5. $x = \frac{1}{\varepsilon\gamma}\int_0^{\varepsilon\gamma z}\frac{\mathrm{d}\xi}{\sqrt{1-2\rho\xi+\xi^2}} = \frac{1}{\varepsilon\gamma}\log\left(\frac{\sqrt{1-2\varepsilon\rho\gamma z+\varepsilon^2\gamma^2z^2}-\rho_1+\varepsilon\gamma z}{1-\rho}\right)$;
6. $\hat{H} = \left(1-2\varepsilon\rho\gamma z+\varepsilon^2\gamma^2z^2\right)^{1/4}Q$;

Then, the value $V(t,f,\alpha)$ can be written as:

$$V = [f-K]^+ + \frac{1}{2}\varepsilon\alpha\sqrt{K^\beta f^\beta}I^{1/2}(\varepsilon\bar{v}z)e^{1/4\varepsilon^2\alpha b_1\delta z^2}\int_0^{T-t}Q(\tau,x)\mathrm{d}\tau \qquad (4.44)$$

where $I(\zeta) = \sqrt{1-2\rho\zeta+\zeta^2}$, $f_{av} = \sqrt{fK}$, $b_1 = \frac{\beta}{f_{av}}$, $b_2 = \frac{\beta(\beta-1)}{f_{av}^2}$.

Q satisfies:

$$Q_t + \frac{1}{2}Q_{xx} = \varepsilon[(\eta-\bar{\eta})xQ_{xx}] + \varepsilon^2\left[\frac{1}{2}\left(v^2-\bar{v}^2-3\bar{\eta}(\eta-\bar{\eta})\right)x^2Q_{xx}\right.$$
$$\left.+\frac{1}{2}\alpha b_1(\eta-\delta)(xQ_x-Q)-\frac{3}{4}\alpha b_1\delta Q-\bar{v}^2\left(\frac{1}{4}I''I-\frac{1}{8}I'I\right)Q-\alpha^2\left(\frac{1}{4}b_2-\frac{3}{8}b_1^2\right)Q\right] \quad \text{for } t<T$$
$$Q = \delta(x) \quad \text{for } t=T$$

Using perturbation expansion, we expand $Q(t,x,T)$ as:

$$Q(t,x,T) = Q^{(0)}(t,x,T) + \varepsilon Q^{(1)}(t,x,T) + \varepsilon^2Q^{(2)}(t,x,T) + \ldots$$

Then we get the equations that satisfy Q^0, Q^1 and Q^2:

$$\begin{cases} Q_t^{(0)} + \frac{1}{2}Q_{xx}^{(0)} = 0 & \text{for } t<T \\ Q^{(0)} = \delta(x) & \text{for } t=T \end{cases}$$

$$\begin{cases} Q_t^{(1)} + \frac{1}{2}Q_{xx}^{(1)} = (\eta-\bar{\eta})xQ_{xx}^{(0)} & \text{for } t<T \\ Q^{(1)} = 0 & \text{for } t=T \end{cases}$$

$$\begin{cases} Q_t^{(2)} + \frac{1}{2}Q_{xx}^{(2)} = \frac{1}{2}\left(v^2-\bar{v}^2-3\bar{\eta}(\eta-\bar{\eta})\right)x^2Q_{xx}^{(0)} + \frac{1}{2}\alpha b_1(\eta-\delta)\left(xQ_x^{(0)}-Q^{(0)}\right)-\frac{3}{4}\alpha b_1\delta Q^{(0)} \\ \qquad -\bar{v}^2\left(\frac{1}{4}I''I-\frac{1}{8}I'I\right)Q^{(0)}-\alpha^2\left(\frac{1}{4}b_2-\frac{3}{8}b_1^2\right)Q^{(0)}+(\eta-\bar{\eta})xQ_{xx}^{(1)} & \text{for } t<T \\ Q^{(2)} = 0 & \text{for } t=T \end{cases}$$

By resolving these standard diffusion equations, we set some parts to 0 and find values for $\bar{\eta}$, $\bar{v}$, δ, $\bar{\tau}$:

$$\begin{aligned}
\bar{\eta} =& \frac{\int_t^T (T-s)\eta(s)\mathrm{d}s}{\frac{1}{2}(T-t)^2} \\
\delta =& \bar{\eta} \\
\bar{v}^2 =& \frac{1}{\frac{1}{3}(T-t)^3}\left\{\int_t^T (T-s)^2 v^2(s)\mathrm{d}s - 3\bar{\eta}\int_t^T (T-s)^2[\eta(s)-\bar{\eta}]\mathrm{d}s\right. \\
& \left. -6\int_t^T \int_t^{s_1} (s_1-s_2)[\eta(s_1)-\bar{\eta}][\eta(s_2)-\bar{\eta}]\mathrm{d}s_2\mathrm{d}s_1\right\} \\
\bar{\tau} =& T-t+\int_0^{T-t} s[v^2(s)-\bar{v}^2]\mathrm{d}s
\end{aligned}$$

Then the option price is identical to the static model's option price by making the identifications:

$$\rho \to \frac{\bar{\eta}}{\bar{v}}, \qquad \gamma \to \bar{v}, \qquad \tau \to \bar{\tau}$$

SABR: Dynamic β

For a SABR model with a non-constant β, it is not easy to find out an effective medium value for β. But as we have seen during the search for the effective medium values, we have the formula to calculate the option price in the form of an integration. The integrand is a function continuous on time. Inspired from this, we can compute the option price with β depending on time. The model is given as:

$$\begin{cases} \mathrm{d}F(t) = \alpha(t)F(t)^{\beta(t)}\mathrm{d}W_1 \\ \mathrm{d}\alpha(t) = \gamma(t)\alpha(t)\mathrm{d}W_2 \\ < \mathrm{d}W_1, \mathrm{d}W_2 >= \rho(t)\mathrm{d}t \end{cases}$$

We consider firstly the most simple case where we have just two periods:

Hypotheses: Suppose that all the parameters are piecewise, and we just have two periods: from 0 to T_1 and from T_1 to T_2.

In the case of SABR model, we have parameters as

$$\alpha(t) = \alpha_0 \quad \text{if } t = 0$$

$$\beta(t) = \begin{cases} \beta_1 & \text{if } 0 \leqslant t \leqslant T_1 \\ \beta_2 & \text{if } T_1 \leqslant t \leqslant T_2 \end{cases}$$

$$\rho(t) = \begin{cases} \rho_1 & \text{if } 0 \leqslant t \leqslant T_1 \\ \rho_2 & \text{if } T_1 \leqslant t \leqslant T_2 \end{cases}$$

$$\gamma(t) = \begin{cases} \gamma_1 & \text{if } 0 \leqslant t \leqslant T_1 \\ \gamma_2 & \text{if } T_1 \leqslant t \leqslant T_2 \end{cases}$$

We define the probability density p and P as before, we can write the value as:

$$V(t,f,\alpha) = [f-K]^+ + \frac{1}{2}\varepsilon^2 \int_0^{T_2} K^{2\beta(T_2-\tau)} P(\tau,f,\alpha)\mathrm{d}\tau$$

where $P(\tau,f,\alpha)$ satisfies the forward problem:

$$P_\tau = \frac{1}{2}\varepsilon^2\left\{\alpha^2 f^{2\beta(T_2-\tau)}P_{ff} + 2\rho(T_2-\tau)\alpha^2 f^{\beta(T_2-\tau)}P_{f\alpha} + \gamma^2(T_2-\tau)\alpha^2 P_{\alpha\alpha}\right\} \qquad \text{for } \tau > 0$$

$$P = \alpha_0^2\delta(f-K) \qquad \text{for } \tau = 0$$

As the parameters are piecewise constant, we rewrite the equation above as following:

$$V(t,f,\alpha) = [f-K]^+ + \frac{1}{2}\varepsilon^2 \int_0^{T_2-T_1} K^{2\beta_2} P(\tau,f,\alpha)\mathrm{d}\tau + \frac{1}{2}\varepsilon^2 \int_{T_2-T_1}^{T_2} K^{2\beta_1} P(\tau,f,\alpha)\mathrm{d}\tau$$

and

$$P_\tau = \frac{1}{2}\varepsilon^2\left\{\alpha^2 f^{2\beta_2}P_{ff} + 2\rho_2\alpha^2 f^{\beta_2}P_{f\alpha} + \gamma_2^2\alpha^2 P_{\alpha\alpha}\right\} \qquad \text{for } 0 < \tau < T_2 - T_1$$

$$P_\tau = \frac{1}{2}\varepsilon^2\left\{\alpha^2 f^{2\beta_1}P_{ff} + 2\rho_1\alpha^2 f^{\beta_1}P_{f\alpha} + \gamma_1^2\alpha^2 P_{\alpha\alpha}\right\} \qquad \text{for } T_2 - T_1 < \tau < T_2$$

$$P = \alpha_0^2\delta(f-K) \qquad \text{for } \tau = 0$$

P is a function continuous on time t. The option price computing problem can be reduced to:

$$\begin{aligned} V(t,f,\alpha) &= V^{(2)} + V^{(1)} \\ V^{(2)} &= [f-K]^+ + \frac{1}{2}\varepsilon^2 \int_0^{T_2-T_1} K^{2\beta_2} P^{(2)}(\tau,f,\alpha)\mathrm{d}\tau \\ V^{(1)} &= \frac{1}{2}\varepsilon^2 \int_{T_2-T_1}^{T_2} K^{2\beta_1} P^{(1)}(\tau,f,\alpha)\mathrm{d}\tau \end{aligned} \tag{4.45}$$

where

$$P_\tau^{(2)} = \frac{1}{2}\varepsilon^2\left\{\alpha^2 f^{2\beta_2}P_{ff}^{(2)} + 2\rho_2\alpha^2 f^{\beta_2}P_{f\alpha}^{(2)} + \gamma_2^2\alpha^2 P_{\alpha\alpha}^{(2)}\right\} \qquad \text{for } 0 < \tau < T_2 - T_1$$

$$P^{(2)} = \alpha_0^2\delta(f-K) \qquad \text{for } \tau = 0 \tag{4.46}$$

and

$$P_\tau^{(1)} = \frac{1}{2}\varepsilon^2\left\{\alpha^2 f^{2\beta_1}P_{ff}^{(1)} + 2\rho_1\alpha^2 f^{\beta_1}P_{f\alpha}^{(1)} + \gamma_1^2\alpha^2 P_{\alpha\alpha}^{(1)}\right\} \qquad \text{for } T_2 - T_1 < \tau < T_2$$

$$P^{(1)} = P^{(2)}(T_2 - T_1) \qquad \text{for } \tau = T_2 - T_1 \tag{4.47}$$

Remark1: $V^{(2)}$ is the value of a portfolio in state $\hat{F}(0) = f$, $\hat{\alpha}(0) = \alpha_0$ at date 0, with parameters of the second period and maturity equal to $T_2 - T_1$.

$P^{(2)}(T_2 - T_1)$ can be got easily regarding to the constant case. By doing some simplifications and transformations (see detail in **Appendix A**), we can write $V^{(1)}$ in the form of:

$$\begin{aligned} V^{(1)} &= M * \frac{1}{2}\varepsilon\alpha_0 \sqrt{K^{\beta_1} f^{\beta_1}} I_{(1)}^{1/2}(\varepsilon\gamma_1 z) e^{1/4\varepsilon^2\rho_1\gamma_1\alpha_0 b_1^{(1)} z^2} \int_{T_2-T_1}^{T_2} \tilde{Q}(\tilde{\tau}, x)\mathrm{d}\tau \\ &= M * N^1 \left(\int_0^{T_2-\tau_0} \tilde{Q}(\tilde{\tau}, x) d\tilde{\tau} - \int_0^{T_2-T_1-\tau_0} \tilde{Q}(\tilde{\tau}, x)\mathrm{d}\tilde{\tau} \right) \\ &= M * (V^{11} - V^{12}) \end{aligned} \tag{4.48}$$

where

$$V^{11} = [f-K]^+ + N^1 \int_0^{T_2-\tau_0} \tilde{Q}(\tilde{\tau}, x)\mathrm{d}\tilde{\tau} \tag{4.49}$$

$$V^{12} = [f-K]^+ + N^1 \int_0^{T_2-T_1-\tau_0} \tilde{Q}(\tilde{\tau}, x)\mathrm{d}\tilde{\tau} \tag{4.50}$$

and M is a constant with $M = \frac{\sqrt{2}}{\sqrt{2K^{2(\beta_1-\beta_2)} + \varepsilon^2 J(T_2-T_1)}} K^{\beta_1-\beta_2} e^{\varepsilon^2(k^{(2)}-k^{(1)})(T_2-T_1)} e^{\varepsilon^2 k^{(1)}\tau_0}$, $N^1 = \frac{1}{2}\varepsilon\alpha_0 \sqrt{K^{\beta_1} f^{\beta_1}} I_{(1)}^{1/2}(\varepsilon\gamma_1 z) e^{1/4\varepsilon^2\rho_1\gamma_1\alpha_0 b_1^{(1)} z^2}$.

Remark2: V^{11} is the value of a portfolio in state $\hat{F}(0) = f$, $\hat{\alpha}(0) = \alpha_0$ at date 0, and with parameters of the first period and maturity equal to $T_2 - \tau_0$.

Remark3: V^{12} is the value of a portfolio in state $\hat{F}(0) = f$, $\hat{\alpha}(0) = \alpha_0$ at date 0, and with parameters of the first period and maturity equal to $T_2 - T_1 - \tau_0$.

So the value of the option is obtained by:

$$\begin{aligned} V(t, f, \alpha) &= V^2 + V^1 \\ &= V^2 + M * (V^{11} - V^{12}) \end{aligned} \tag{4.51}$$

4.5.4 Summary

In this section we discuss the Fourier formed closed formula for square root SV model. We reviewed the existing result on square root SV model with constraint of zero correlation. We figured out another path by changing of variables, and found a closed formula with constraint of constant skew. The constraints in both cases can be removed by employing the integration by parts on random

variables. We show this result in proposition 28. The proposition helps two mentioned closed formulas relax the constraint.

An application on another SV model, SABR model, is also shown. It removes the constraint of constant skew in the original paper of Hagan.

4.6 Vol of vol expansion

In order to calibrate stochastic volatility models, it is convenient to have an accurate and fast-computing analytical formula for call options. However, deriving such a formula is not always an easy task. For instance, as regards Heston model, the most popular technique involves numerical integration, which is necessarily time-consuming. This article presents the volatility of volatility series expansion, a powerful technique for deriving fast-computing analytical formulae for European options. The main idea is to apply perturbation method to the parameter vol of vol, calculating the first and second order of the difference between a stochastic volatility model and a Black Scholes model. In general case, we can reduce the integration of the exact formula to some simpler integration. For certain models, for example Heston model, the new integrations can be express into analytic form. The result is significant: the time of calculation can be only 1% of exact formula with optimization in integral; and the difference in implied vol is often smaller than 0.05% in most cases except too short maturity. The method can be applied successfully to several stochastic volatility models including Heston model, GARCH model, SABR model, etc, to enhance the calibration and pricing routines.

4.6.1 Framework

Consider the following two-factor stochastic volatility model

$$\mathrm{d}S = (r-d)S\mathrm{d}t + \sigma S\mathrm{d}B_t \tag{4.52}$$

$$\mathrm{d}V = b(V)\mathrm{d}t + \epsilon\nu(V)\mathrm{d}W \tag{4.53}$$

$$\mathrm{d}B\mathrm{d}W = \rho(V)\mathrm{d}t \tag{4.54}$$

where r is the short rate, d the dividend yield, ϵ a constant, and $b(V)$ and $\nu(V)$ are independent of ϵ. We assume parameters r and d to be constant for the sake of simplicity. The series expansion consists in writing the option price formula as a series in ϵ.

Fourier transform integration tells us that the call option price is given by

$$C(S,V,T) = Se^{-dT} - \frac{Ke^{-rT}}{2\pi}\int_{i/2-\infty}^{i/2+\infty}\exp(-iuX)\frac{\Phi(u,V,T)}{u(u-i)}\mathrm{d}u, \tag{4.55}$$

where $X = \log(S/K) + (r-d)T$.

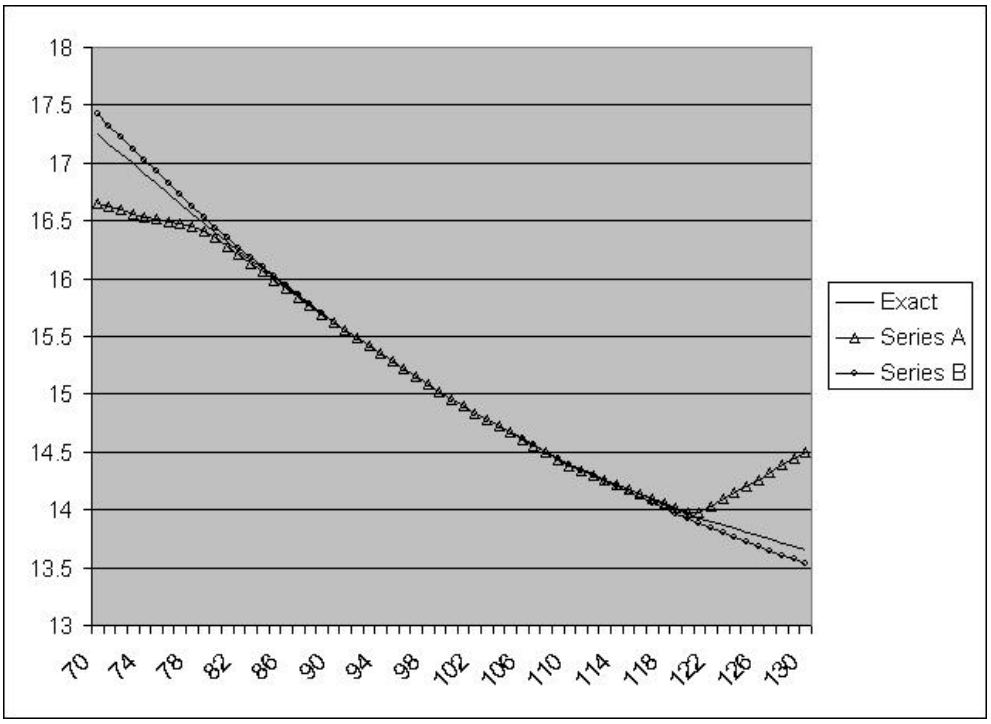


Figure 4.6: Comparison of series A (expasion on price), series B (expansion on implied volatility) and exact volatility

4.6.2 The expansion technique

We are looking for solutions which can be written as a power series of ϵ. Thus, we can obtain the power series of the call price using either 4.55 directly

$$C(u,V,T) = C^{(0)}(u,V,T) + \epsilon C^{(1)}(u,V,T) + \epsilon^2 C^{(2)}(u,V,T)$$

or expanding first the implied volatility

$$V_{\text{imp}}(u,V,T) = V_{\text{imp}}^{(0)}(u,V,T) + \epsilon V_{\text{imp}}^{(1)}(u,V,T) + \epsilon^2 V_{\text{imp}}^{(2)}(u,V,T)$$

and then plugging it into the option price with Black-Scholes formula.

As a matter of fact, these two methods differ significantly. The former - denoted by series A in the remainder of this article - gives the call price first and implied volatility while the latter - denoted by series B - gives the implied volatility and call price is obtained afterwards. Nevertheless, in most cases, numerical difference is slight between the two series. However, regarding far out-of-the money options, the two series give different results as shown in figure 4.6. Empirical evidence shows that series B is very often better than series A. Even though this is not a general rule, series B should be usually preferred over series A.

4.6.3 Deriving the series

After we have presented the expansion technique, we will now explicit the two series for the following model. In particular, this model encompasses the Heston model for the special case $\phi = 0.5$.

$$\mathrm{d}S = (r-d)S\mathrm{d}t + \sigma S\mathrm{d}B_t \tag{4.56}$$
$$\mathrm{d}V = (\omega - \theta V)\mathrm{d}t + \epsilon V^{\phi}\mathrm{d}W \tag{4.57}$$
$$\mathrm{d}B\mathrm{d}W = \rho(V)\mathrm{d}t \tag{4.58}$$

The expansion is first on the fundamental transform of the closed formula, which is presented by $\Phi(u, V, T)$ in equation 4.55. The idea is that we can expand this function into simpler form, so that the integration in the equation 4.55 can be reduced to analytic form. We don't jump into the detailed derivation. Interested readers can refer to the book of Lewis [2000]. Here we just list the result.

For series A, the expansion on price, is

$$\begin{aligned} C(S,V,\tau) &= c(S,v,\tau) + \epsilon\tau^{-1}J^{(1)}\tilde{R}^{(1,1)}c_V(S,v,\tau) \\ &+ \epsilon^2\left\{\tau^{-1}J^2 + \tau^{-2}J^{(3)}\tilde{R}^{(2,0)} + \tau^{-1}J^{(4)}\tilde{R}^{(1,2)} + \frac{\tau^{-2}}{2}(J^1)^2\tilde{R}^{(2,2)}\right\}c_V(S,v,\tau) \\ &+ O(\epsilon^3) \end{aligned}$$

For series B, the expansion on price, is

$$\begin{aligned} V_{\text{imp}} &= v(S,v,\tau) + \epsilon\tau^{-1}J^{(1)}\tilde{R}^{(1,1)} \\ &+ \epsilon^2\left\{\tau^{-1}J^2 + \tau^{-2}J^{(3)}\tilde{R}^{(2,0)} + \tau^{-1}J^{(4)}\tilde{R}^{(1,2)} + \frac{\tau^{-2}}{2}(J^{(1)})^2\left[\tilde{R}^{(2,2)} - (\tilde{R}^{(1,1)})^2\tilde{R}^{(2,0)}\right]\right\} \\ &+ O(\epsilon^3) \end{aligned}$$

In the above formulae, term $c(S,v,\tau)$ presents corresponding Black Scholes price. When vol of vol $\epsilon = 0$, the stochastic model reduces to a Black Scholes model. v here is the equivalent variance for Black Scholes, which is basically the the integration of the variance from 0 to τ.

The function $\tilde{R}^{(p,q)}$ and $J^{(s)}$ are the derivative ratios and integration respec-

tively. Here list their expressions for practical use purpose.

$$\tilde{R}^{(2,0)} = \tau\left[\frac{1}{2}\frac{X^2}{Y^2} - \frac{1}{2Z} - \frac{1}{8}\right],$$
$$\tilde{R}^{(1,1)} = \left[-\frac{X}{Z} + \frac{1}{2}\right],$$
$$\tilde{R}^{(1,2)} = \left[\frac{X^2}{Z^2} - \frac{X}{Z} - \frac{1}{4Z}(4-Z)\right],$$
$$\tilde{R}^{(2,2)} = \tau\left[\frac{1}{2\frac{X^4}{Z^4}} - \frac{1}{2}\frac{X^3}{Z^3} - 3\frac{X^2}{Z^3} + \frac{1}{8}\frac{X}{Z^2}(12+Z) + \frac{1}{32}\frac{1}{Z^2}(48-Z^2)\right]$$

with $Z = V\tau$

And

$$J^{(1)}(V,\tau) = \frac{\rho}{\theta}\int_0^\tau \left(1-e^{-\theta(\tau-s)}\right)\left[\frac{\omega}{\theta} + e^{-\theta s}(V-\frac{\omega}{\theta})\right]^{\phi+\frac{1}{2}} \mathrm{d}s,$$
$$J^{(2)}(V,\tau) = 0$$
$$J^{(3)}(V,\tau) = \frac{1}{2\theta^2}\int_0^\tau \left(1-e^{-\theta(\tau-s)}\right)^2\left[\frac{\omega}{\theta} + e^{-\theta s}(V-\frac{\omega}{\theta})\right]^{2\phi} \mathrm{d}s$$
$$J^{(4)}(V,\tau) = (\phi+\frac{1}{2})\int_0^\tau \left[\frac{\omega}{\theta} + e^{-\theta(\tau-s)}(V-\frac{\omega}{\theta})\right]^{\phi+\frac{1}{2}} J^{(6)}(V,\tau)\mathrm{d}s$$
$$\text{with } J^{(6)}(V,\tau)\mathrm{d}s = \int_0^\tau \left(e^{-\theta(\tau-s)} - e^{-\theta s}\right)\left[\frac{\omega}{\theta} + e^{-\theta(\tau-u)}(V-\frac{\omega}{\theta})\right]^{\phi-\frac{1}{2}} \mathrm{d}u$$

4.6.4 Example of vol of vol expansion: Heston model

In this section we will present an interesting example to show how the expansion of vol of vol works with Heston model. By the asymptotic expansion, we can finally obtain an approximate analytic formula for the European call option. This work comes from the result of Benhamou et al. [2009].

Suppose we take a Heston model,

$$\mathrm{d}X_t = \sqrt{v_t}\mathrm{d}W_t - \frac{v_t}{2}\mathrm{d}t, \qquad X_0 = x_0, \tag{4.59}$$
$$\mathrm{d}v_t = \kappa(\theta_t - v_t)\mathrm{d}t + \xi_t\sqrt{v_t}\mathrm{d}B_t, \qquad v_0, \tag{4.60}$$
$$\mathrm{d}\langle W,B\rangle_t = \rho_t\mathrm{d}t,$$

Here we have adjusted to risk neutral probability; as consequence it introduces the drift term in spot process by change of probability.

To expand the model, we add ϵ in the model.

$$\mathrm{d}X_t^\epsilon = \sqrt{v_t^\epsilon}\mathrm{d}W_t - \frac{v_t^\epsilon}{2}\mathrm{d}t, \qquad X_0^\epsilon = x_0,$$
$$\mathrm{d}v_t^\epsilon = \kappa(\theta_t - v_t^\epsilon)\mathrm{d}t + \epsilon\xi_t\sqrt{v_t^\epsilon}\mathrm{d}B_t, \qquad v_0^\epsilon = v_0, \tag{4.61}$$

Now we will expand the European call option price formula respect to ϵ. Note that, when $\epsilon = 0$, we have a Black Scholes model; while $\epsilon = 1$, we have a Heston model. We have already the closed formula of Black Scholes for $\epsilon = 0$. We expand at $\epsilon = 0$, and let $\epsilon = 1$ to obtain the approximate formula. This can write in mathematics language as follows.

$$P_{\text{Heston}} = P_{\text{BS}} + \mathbb{E}\left[\frac{\partial P_{\text{BS}}}{\partial \epsilon}\right] + \frac{1}{2}\mathbb{E}\left[\frac{\partial^2 P_{\text{BS}}}{\partial \epsilon^2}\right] + \mathcal{E}.$$

Another approximation we will take here is to simulate the partial derivatives in the equation above by the linear combination of the Greek letters of Black Scholes. The idea is that using the chain rule in derivative can make $\frac{\partial P_{\text{BS}}}{\partial \epsilon} = \frac{\partial P_{\text{BS}}}{\partial S}\frac{\partial S}{\partial \epsilon} + \frac{\partial P_{\text{BS}}}{\partial \sigma}\frac{\partial \sigma}{\partial \epsilon}$. Same idea for the second derivative.

$$P_{\text{Heston}} = P_{\text{BS}}(x_0, var_T) + \sum_{i=1}^{2} a_{i,T}\frac{\partial^{i+1} P_{\text{BS}}}{\partial x^i y}(x_0, var_T) + \sum_{i=0}^{1} b_{2i,T}\frac{\partial^{2i+2} P_{\text{BS}}}{\partial x^{2i} y^2}(x_0, var_T) + \mathcal{E}, \tag{4.62}$$

We omit the lengthy derivation but just listing the result here. Readers can refer to Benhamou et al. [2009] for proofs and intermediate derivation. The parameters in the formula are:

$$\begin{aligned}
var_T &= m_0 v_0 + m_1 \theta, & a_{1,T} &= \rho\xi(p_0 v_0 + p_1 \theta),\\
a_{2,T} &= (\rho\xi)^2(q_0 v_0 + q_1 \theta), & b_{0,T} &= \xi^2(r_0 v_0 + r_1 \theta).\\
var_T &= \int_0^T v_{0,t}\mathrm{d}t,
\end{aligned}$$

$$
\begin{aligned}
m_0 &= \frac{e^{-\kappa T}\left(-1+e^{\kappa T}\right)}{\kappa},\\
m_1 &= T - \frac{e^{-\kappa T}\left(-1+e^{\kappa T}\right)}{\kappa},\\
p_0 &= \frac{e^{-\kappa T}\left(-\kappa T+e^{\kappa T}-1\right)}{\kappa^2},\\
p_1 &= \frac{e^{-\kappa T}\left(\kappa T+e^{\kappa T}(\kappa T-2)+2\right)}{\kappa^2},\\
q_0 &= \frac{e^{-\kappa T}\left(-\kappa T(\kappa T+2)+2e^{\kappa T}-2\right)}{2\kappa^3},\\
q_1 &= \frac{e^{-\kappa T}\left(2e^{\kappa T}(\kappa T-3)+\kappa T(\kappa T+4)+6\right)}{2\kappa^3},\\
r_0 &= \frac{e^{-2\kappa T}\left(-4e^{\kappa T}\kappa T+2e^{2\kappa T}-2\right)}{4\kappa^3},\\
r_1 &= \frac{e^{-2\kappa T}\left(4e^{\kappa T}(\kappa T+1)+e^{2\kappa T}(2\kappa T-5)+1\right)}{4\kappa^3}.
\end{aligned}
$$

The advantage is that there is no integration in the approximate formula. So that the calculation speed is much faster than that of the exact formula. We will see this point in section numerical results.

The error in the approximation is estimated as $\mathcal{E} = O\left(\left[\xi_{\text{sup}}\sqrt{T}\right]^3\sqrt{T}\right)$.

Numerical results

Benhamou et al. [2009] tested the approximate formula with strikes from 70% to 130% for short maturity, and 10% to 730% for long maturity. Implied Black-Scholes volatilities of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = 0\%$. Except short maturity with very small strike, where we observe the largest difference (18.01 bp), the difference is less than 5 bp (1 bp = 0.01%) in almost all other cases. For the calculation speed, the approximate formula is about 100 times quicker than the exact formula (with the optimization in integral).

4.7 Calibration of Stochastic Volatility Model

4.7.1 Introduction

A calibration procedure is a procedure in which we use market data (usually prices) as input to determine the parameters of a financial model. We search

for those parameters which enable the model to fit the best the input market data. Calibration can be time consuming, this issue depending on the size of the input data, on the number of parameters to calibrate and also on the calibration algorithm itself.

Because a model parameter can be either time dependent or constant, we have two categories of calibration methods: (1) the calibration procedure of time dependent parameters is called "bootstrap", while (2) the calibration procedure of the constant parameters is called "optimization".

Usually, the model we use for pricing contains one or several parameters. These parameters give the model some flexibility to better fit the characteristic of the market. Before using a model to price, we need to get the parameters from the market data. This procedure is often called calibration.

There are at least two methods to find the parameters. The quick way is to estimate manually the parameters. This method is good when the number of parameters are small. For example, in Black-Scholes model Black and Scholes [1973], we need volatility. User of this model can estimate the parameter and then use the model with his estimation. But this method is inaccurate, and useless to the models with many parameters. Stochastic volatility models usually have 4 or more parameters. So we need to calibrate the models.

Description

Generally, the calibration process is a process using certain algorithm to find the minimum of the difference of the target function by changing the parameter (or parameter vector if there are more than one parameter).

Target function

Usually we use the option prices listed in the market as the target function. For each given parameter vector, we need the option price calculated by the model.

For a given financial instrument, there are different maturities. So we need to minimize the

$$\sum_{\text{all maturities}} (p_{\text{model}} - p_{\text{market}})^2$$

where p_{model} and p_{market} are respectively the option price calculated by the model and the option price listed in the market.

Algorithm

The algorithm are usually the minimum finding algorithm.

Levenberg Marquadt

L-BFGS-B

Nelder and Mead (Simplex)

The calibration of stochastic volatility model is a

We need analytic formula to evaluate the price of options, in order to calibrate the model.

The Calibration Scheme in General

To calibrate the static stochastic volatility models, we search for ways to compute the value of option with the corresponding model. Several methods are introduced for the calibration of stochastic volatility models according to their forms. For SABR model, by introducing the probability density, we compute the value of a call option by resolving the PDEs. For the square root model, two different ways using Fourier Transform are used, one of which will be applied in the dynamic case.

The calibration of stochastic volatility models require analytic solution of the option price or implied volatility. We calibrate the models to implied volatility that we get from the market with an optimization algorithm "Levenberg Marquardt". Let T be the time to maturity with fwd the corresponding forward. Let $K1$, $K2$, ... K_M be a set of strikes, and V_1, V_2, ... V_M be the corresponding market implied volatiliy. The aim of the calibration is to minimize the least square error:

$$R = \sum_{i=1}^{M} (V_i - V_i^{\text{im}})^2$$

where V_i^{im} is the implied volatility corresponding to the model(calculated with the Black's formula after having obtained the option price). So the speed of the calculation of the option price is very important for the speed of the calibration.

By using methods to increase the convergency of the closed formulas and to resolve the discontinuities during the implementation, we don't dispose much difficulty in the case of models with constant parameters. Based on these solutions, we try to find out solutions for time dependent models. Several approaches are used in the research. These methods are adapted to their process of spot (or forward) and volatility. The effective medium parameter method is to find out a mean value for each parameter. The effective medium parameter depends on time and the values of all time dependent parameters in each period. By doing this, we change a time dependent model into a constant parameter model where we can use the analytical formulas that we developed in the constant parameter case for the calibration. Another method is to compute the option price by resolving the PDEs and to use the final value of one constant parameter period as the initial value of the following constant parameter period. According to the form of each model, we will choose the easiest method.

Note: In the following, we work with "forward" instead of "spot" of the underlying. So we are in the neutral risk environment. This will simplify our calculus. The option price that we calculate is the expected value of the payoff, which equals to the real option price divided by the discount factor. The discount factor is generally $DF = e^{-rT}$ where r is the constant interest rate and T is the expiry.

Bootstrap

The so called "Bootstrap Technique" may be used for parameters which are time dependent, that is, deterministic functions depending on time. Below we explain in which manner the bootstrap technique is performed.

Consider a financial model $\mathcal{M}$, having an unknown function parameter f depending on time. Suppose we want to calibrate this function f. This means that we want to exactly say who $f(t)$ is for any t belonging to $[0,T]$. As a base hypothesis, suppose that you are also knowing the shape of f; typically, your function will be piecewise constant, or piecewise linear, or quadratic.

Recall that we have to recover $f(t)$ over the time interval $[0,T]$, and suppose that we get on the market n prices $p_1, \ldots, p_n$ of financial instruments $\pi_1, \ldots, \pi_n$:

$$\begin{aligned} p_1 &= \text{Market price}(\pi_1; 0, t_1), \\ p_2 &= \text{Market price}(\pi_2; 0, t_2), \\ &\cdots \\ p_i &= \text{Market price}(\pi_i; 0, t_i), \\ &\cdots \\ p_n &= \text{Market price}(\pi_n; 0, t_n), \end{aligned}$$

where each price p_i is computed at time $t = 0$ for the maturity t_i, and we have:

$$0 < t_1 < t_2 < \cdots < t_n = T.$$

Because we know the interpolation type of f, to compute $f(t)$ over the time interval $[0,T]$, it suffices to compute f at the points

$$t_1, t_2, \ldots, t_N,$$

as f is completely defined by the vector of couples:

$$\begin{aligned} &\{t_1, f(t_1)\} \\ &\{t_2, f(t_2)\} \\ &\cdots \\ &\{t_i, f(t_i)\} \\ &\cdots \\ &\{t_n, f(t_n)\}. \end{aligned}$$

Indeed, if $f_1 = f(t_1)$, $f_2 = f(t_2), \ldots, f_n = f(t_n)$ are known, to compute $f(t)$, for t in $[0,T]$, we have only to find the index k, such that : t belongs to $[t_k, t_{k+1}]$; once we have k, we will compute $f(t)$ interpolating f between t_k and t_{k+1}.

We explain now how we can compute $f_1, f_2, \ldots, f_n$. This will be done iteratively, solving numerically n equations; first we compute f_1; then, using that f_1 is known, we compute f_2; then, using that f_1 and f_2 are known, we compute f_3; and so on; at the end, using that $f_1, f_2, \ldots, f_{n-1}$ are known, we compute f_n.

Step 1 To compute f over the first interval $[0,t_1]$, we have to use the price p_1; as this information does not suffice to recover $f(t)$, for any $t \in [0,t_1]$, we suppose in addition that f is constant over this interval: $f(t) \equiv f_1$, for any $t \in [0,t_1]$.

Then, on $[0,t_1]$, f is defined by the 2-dimensional vector

$$[\{0,f_1\},\{t_1,f_1\}],$$

which will be known, once f_1 will be known. Denote by $f(t_1, f_1)$ the function f, constructed with the vector above. We will compute f_1 solving the following equation in y:

$$\mathcal{M}(\pi_1; f(t_1,y)) = p_1,$$

where $\mathcal{M}(\pi_1; f(t_1,y))$ is the price of the financial instrument π_1, computed with the model $\mathcal{M}$, which uses the function $f(t_1,y)$. The equation above is solved using Newton Raphson algorithm.

Step 2 We will use now that f is already constructed over $[0,t_1]$, and we will compute it over $(t_1,t_2]$. Our goal function f is defined by the 3-dimensional vector

$$[\{0,f_1\},\{t_1,f_1\},\{t_2,f_2 = f(t_2)\}],$$

where f_1 is already know, and we have only to compute f_2.

This will be done solving the following equation in y:

$$\mathcal{M}(\pi_2; f(t_2,y)) = p_2,$$

where $\mathcal{M}(\pi_2; f(t_2,y))$ is the price of the financial instrument π_2, computed with the model $\mathcal{M}$, which uses the function $f(t_2,y)$. The above equation is well defined because we know to compute $\mathcal{M}(\pi_2; f(t_2,y))$ using: (1), that we have already computed the function on the interval $[0,t_1]$, and, (2), that f has a known interpolation type.

Step 3 to n Are analogous to Step 2.

In this way we iteratively compute $f(t_1), f(t_2), \cdot, f(t_n)$.

Optimization

Consider a financial model $\mathcal{M}$, depending on m unknown parameters

$$\alpha_1, \ldots, \alpha_m,$$

where each α_i is a constant real number: $\alpha_i \in \mathbb{R}$. Suppose we want to calibrate these numbers and that we get on the market n prices $p_1, \ldots, p_n$ of financial instruments $\pi_1, \ldots, \pi_n$:

$$\begin{aligned} p_1 &= \text{Market price}(\pi_1), \\ p_2 &= \text{Market price}(\pi_2), \\ &\ldots \\ p_i &= \text{Market price}(\pi_i), \\ &\ldots \\ p_n &= \text{Market price}(\pi_n). \end{aligned}$$

We consider the following Least Square Problem: find the numbers $x_1, \ldots, x_m$ that minimize the following function F:

$$F(x_1, \ldots, x_m) = \sum_{i=1}^{n} [\mathcal{M}(\pi_i; x_1, \ldots, x_m) - p_i]^2,$$

where $\mathcal{M}(\pi_i; x_1, \ldots, x_m)$ is the price of the financial instrument π_i, computed with the model $\mathcal{M}$, which uses the parameters $x_1, \ldots, x_m$. We suppose that $m \leq n$.

To resolve this Least Square Problem, we can employ the following numerical methods

(1) Levenberg-Marquardt algorithm.

(2) Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm.

Find details on these numerical methods in following subsection.

4.7.2 Calibration order

When we calibrate a model its parameters are calibrated one by one, but there are two possible calibration procedures:

(1) Successive calibration.

(2) Nested calibration.

These calibration procedures use essentially an optimization algorithm to compute $(x_1^*, ..., x_m^*)$, which is the optimal value for the parameter $x = (x_1, ..., x_m)$:

$$(x_1^*, ..., x_m^*) = Calibrate\{x = (x_1, ..., x_m) | (y_1, ..., y_n), K\}$$

using the known values $(y_1, ..., y_n)$ and performing at most K steps.

Successive calibration

This method calibrates the parameters independently: each parameter is calibrated independently with respect to the other parameters.

It is the end user who chooses the calibration order in the parameter vector. Suppose that you have to calibrate tree parameters called m, v, w. If you want to calibrate first v, then m, and then w you have to construct the vector: (v, m, w). Writing (v, w, m) means that v will be calibrated first, then w and then m.

To explain the successive calibration procedure, let us suppose the model to calibrate has only two parameters: (a_1, a_2), having the initial value (a_1^0, a_2^0). The successive calibration procedure has only 2 steps and is as follows:

(1) Compute a_1^*, that is the optimum value of the first parameter:

$$Calibrate\left(a_1 | a_2^0, K\right).$$

Calibrate a_1 using the value a_2^0 for the second parameter. a_1^0 is used as the starting point of the algorithm.

(2) Compute a_2^*, that is the optimum value of the second parameter:

$$Calibrate\left(a_2 | a_1^*, K\right)$$

calibrate a_2 using the value a_1^* for the first parameter. Note that to compute a_2^* we used a_1^*.

The maximum total number of steps of the 2-dimensional successive calibration procedure is $K + K$. When calibrating N parameters, the maximum total number of steps is $N \times K$.

Example

Let consider that we want to calibrate a Hull-White 1 factor model to the following portfolio of swaptions. We denote by $f(start, end, \sigma, a)$ the price of an ATM swaption with Hull-White model. σ can be a time-dependent parameter.

Initial volatility (σ^0) and initial mean reversion are both set to (a^0) are both set to 1%. The prices of portfolio 1 are given by:

First, we start by resizing the volatility parameter to the size of the portfolio. The volatility is now a curve model parameter with size 10. Abscissas are equal to 14-May-08,...,15-May-17 and ordinates are equal to $\sigma_1^0 = 1\%, \sigma_2^0 = 1\%, ..., \sigma_{10}^0 = 1\%$.

Then we bootstrap the volatility. We look for the value σ_1^* such that

$$f(\text{14-May-08}, \text{14-May-18}, \sigma^1, a^0) = 1.16186\% \tag{4.63}$$

where σ^1 is the volatility parameters with ordinates

$$\sigma_1^*, \sigma_2^0 = 1\%, ..., \sigma_{10}^0 = 1\%. \tag{4.64}$$

Once we have found σ_1^*, we look for the value σ_2^*. It is the value of σ_2 such that

$$f(\text{14-May-09}, \text{14-May-18}, \sigma, a^0) = 2.06796\% \tag{4.65}$$

where σ is the volatility parameters with ordinates

$$\sigma_1^*, \sigma_2^*, \sigma_3^0 = 1\%, ..., \sigma_{10}^0 = 1\%. \tag{4.66}$$

At the end of this process, we obtain the following volatility parameter σ^* :

Once we have finished the bootstrap of volatility, we won't modify it anymore. We will optimize the parameter a^*. It is the mean reversion value that minimizes $\sum_{i=1}^{10}[f(start_i, end_i, \sigma^*, a) - f^M(start_i, end_i)]^2$ where f^M is the market price of a swaption, $start_i$ are the start dates of the swaption portfolio and end_i are the end dates.

The linked calibration algorithm is now finished and output parameters are σ^* and a^*.

Nested calibration

This method nests calibration loops; which means we go through each loop of parameter 1 by finishing a full calibration on parameter 2 with the present value of parameter 1. We also call this kind of calibration order as "linked".

To explain the nested calibration procedure, let us suppose the model to calibrate has only two parameters: (a_1, a_2), having the initial value (a_1^0, a_2^0). The nested calibration procedure has at most $K \times K$ steps and is as follows:

$$Calibrate\left(a_2 \,|\, Calibrate\left(a_1 | a_2, K\right), K\right),$$

that is, compute a_2^*, doing at most K steps, and using at each step

$$(a_1^i)^* = Calibrate\left(a_1^i | a_2^i, K\right)$$

the optimal value of a_1^i; a_1^i corresponds, at the ith step, to a_2^i, that is the current value of the second parameter.

The maximum total number of steps of the 2-dimensional nested calibration procedure is $K \times K$. When calibrating N parameters, the maximum total number of steps is K^N.

We take the same example as in the linked calibration.Let consider that we want to calibrate a Hull-White 1 factor model to the following portfolio of swaptions. We denote by $f(start, end, \sigma, a)$ the price of an ATM swaption with Hull-White model. σ can be a time-dependent parameter.

Initial volatility (σ^0) and initial mean reversion are both set to (a^0) are both set to 1%. The prices of portfolio 1 are given by :

First, we start by resizing the volatility parameter to the size of the portfolio. The volatility is now a curve model parameter with size 10. Abscissas are equal to 14-May-08,...,15-May-17 and ordinates are equal to $\sigma_1^0 = 1\%, \sigma_2^0 = 1\%, ..., \sigma_{10}^0 = 1\%$.

Then we bootstrap the volatility. We look for the value σ_1^{*1} such that

$$f(\text{14-May-08}, \text{14-May-18}, \sigma^1, a^0) = 1.16186\% \tag{4.67}$$

where σ^1 is the volatility parameters with ordinates

$$\sigma_1^*, \sigma_2^0 = 1\%, ..., \sigma_{10}^0 = 1\%. \tag{4.68}$$

Once we have found σ_1^*, we look for the value σ_2^*. It is the value of σ_2 such that

$$f(\text{14-May-09}, \text{14-May-18}, \sigma, a^0) = 2.06796\% \tag{4.69}$$

where σ is the volatility parameters with ordinates

$$\sigma_1^*, \sigma_2^*, \sigma_3^0 = 1\%, ..., \sigma_{10}^0 = 1\%. \tag{4.70}$$

At the end of this process, we obtain the following volatility parameter σ^*1 :

Like in the successive calibration scheme, we will now optimize the parameter a. It is the mean reversion value that minimizes $\sum_{i=1}^{10}[f(start_i, end_i, \sigma^*, a) - f^M(start_i, end_i)]^2$ where f^M is the market price of a swaption, $start_i$ are the start dates of the swaption portfolio and end_i are the end dates. The optimization algorithm returns that $a^{*1} = 0.87\%$.

Contrary to the successive calibration, the nested calibration is not finished yet. We will now redo the bootstrap process with a mean reversion equal to a^{*1}. Thus, we get a bootstrapped volatility σ^{*2}. We iterate this process until we get a good calibration precision. This nested calibration generally finds better results than the successive calibration but it can be much slower.

4.7.3 Numerical method

Newton Raphson

Newton Raphson is an efficient algorithm for finding approximations to the zeros (or roots) of a real-valued function. As such, it is an example of a root-finding algorithm. It can also be used to find a minimum or maximum of such a function, by finding a zero in the function's first derivative.

The idea of the method is as follows: one starts with an initial guess which is reasonably close to the true root, then the function is approximated by its

tangent line, and one computes the x-intercept of this tangent line (which is easily done with elementary algebra). This x-intercept will typically be a better approximation to the function's root than the original guess, and the method can be iterated.

Suppose $f : [a,b] \to R$ is a differentiable function defined on the interval $[a,b]$ with values in the real numbers R. The formula for converging on the root can be easily derived. Suppose we have some current approximation x_n. Then we can derive the formula for a better approximation x_{n+1} by referring to the Fig 4.7. We know from the definition of the derivative at a given point that it is the slope of a tangent at that point.

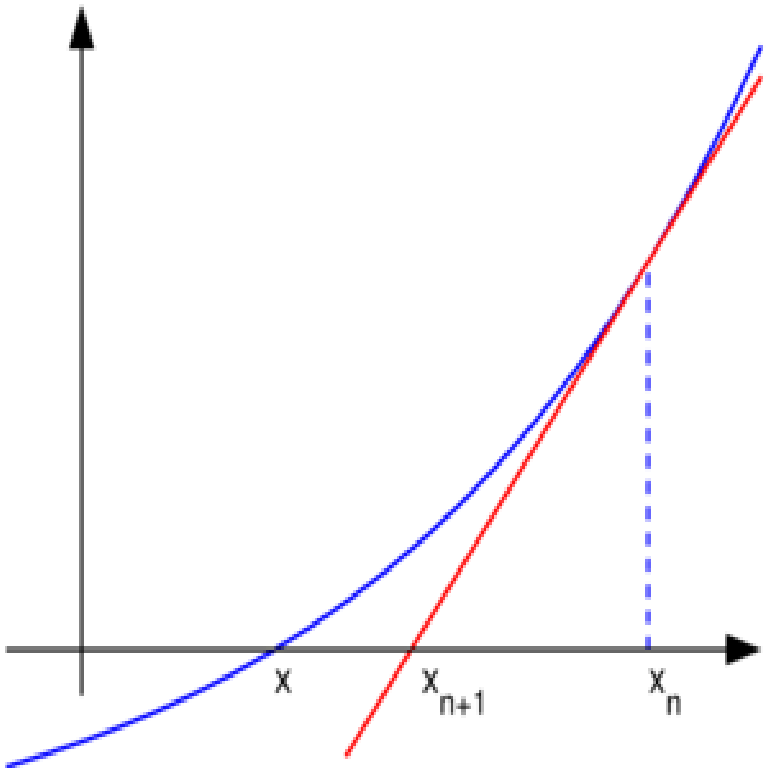


Figure 4.7: An illustration of one iteration of Newton's method (the function f is shown in blue and the tangent line is in red). We see that x_{n+1} is a better approximation than x_n for the root x of the function f.

That is

$$f'(x_n) = \frac{\Delta y}{\Delta x} = \frac{f(x_n) - 0}{x_n - x_{n+1}} = \frac{0 - f(x_n)}{x_{n+1} - x_n}$$

were f' denotes the derivative of the function f. Then by simple algebra we can derive

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

We start the process off with some arbitrary initial value x_0. (The closer to the zero, the better. But, in the absence of any intuition about where the zero might lie, a "guess and check" method might narrow the possibilities to a reasonably small interval by appealing to the intermediate value theorem.) The method will usually converge, provided this initial guess is close enough to the unknown zero, and that $f'(x_0) \neq 0$. Furthermore, for a zero of multiplicity 1, the convergence is at least quadratic in a neighborhood of the zero, which intuitively means that the number of correct digits roughly at least doubles in every step.

Levenberg Marquardt

Levenberg-Marquardt (LM) algorithm provides a numerical solution to the problem of minimizing a function, generally nonlinear, over a space of parameters of the function.These minimization problems arise especially in least squares curve fitting and nonlinear programming.

Minimization algorithm We reformulate the Least Square Problem: find the vector $\mathbf{p} = (p_1, \ldots, p_m)$ that minimize the following function F:

$$F(\mathbf{p}) = \sum_{i=1}^{n} [f_i(\mathbf{p})]^2, \tag{4.71}$$

where $f : \mathbb{R}^m \to \mathbb{R}^n$, $\mathbf{p} \mapsto f(\mathbf{p}) = (f_1(\mathbf{p}), \ldots, f_n(\mathbf{p})) \in \mathbb{R}^n$, and f is of class C^1.

Like any other numeric minimization algorithms, the LM algorithm is an iterative procedure. To start a minimization, the user has to provide an initial guess for the vector $\mathbf{p}$. In many cases, an uninformed standard guess like $\mathbf{p}_0 = (1, 1, ..., 1)$ will work fine; in other cases, the algorithm converges only if the initial guess is somewhat close to the final solution.

At each iteration step, the vector $\mathbf{p}$ is replaced by a new estimate $\mathbf{p} + \mathbf{q}$. To determine $\mathbf{q}$, the value $f(\mathbf{p} + \mathbf{q})$ is approximated by its linearization:

$$f(\mathbf{p} + \mathbf{q}) \approx f(\mathbf{p}) + \mathbf{J}\mathbf{q}.$$

At a minimum of the sum of squares F in eq. (4.71), the gradient of F is zero. Differentiating F the square of the right hand side of the equation (4.71) and setting it to zero leads to:

$$(\mathbf{J}^T\mathbf{J})\mathbf{q} = \mathbf{J}^T[\mathbf{y} - f(\mathbf{p})],$$

from which $\mathbf{q}$ can be obtained by inverting $\mathbf{J}^T\mathbf{J}$. The key hallmark of the LM alogorithm is to replace this equation by a “damped version”:

$$(\mathbf{J}^T\mathbf{J} + \lambda\mathbf{I})\mathbf{q} = \mathbf{J}^T[\mathbf{y} - f(\mathbf{p})],$$

where $\mathbf{I}$ is the identity matrix, and

$$\mathbf{q} = (\mathbf{J}^T\mathbf{J} + \lambda\mathbf{I})^{-1}\mathbf{J}^T[\mathbf{y} - f(\mathbf{p})].$$

The non-negative damping factor λ is adjusted at each iteration. If the reduction of F is rapid, then a smaller value can be used bringing the algorithm close to the Gauss-Newton alogorithm. On the other hand, if an iteration gives insufficient reduction in the residual, then λ can be increased, giving a step close to the gradient descent direction. A similar damping factor appears in Tikhonov regularization, which is used to solve linear ill-posed problems.

The solution of our minimization problem is found when one of the following conditions is satisfied:

(1) The calculated step length **q** falls below a predefined limit.

(2) The reduction of the squares sum from the latest parameter vector **p**, falls below predefined limits.

In these cases, the iteration is stoped and the last parameter vector **p** is considered to be the solution.

Advantages The advantage of this algorithm is that it introduces a damping parameter μ, which influences both the direction and the size of the step. The step h^* is defined by following formula:

$$(\mathbf{J}^T\mathbf{J} + \mu\mathbf{I})h^* = -\mathbf{J}^T f,$$

where f is the objective function and $\mathbf{J}$ is the Jaccobian of f.

Large damping parameter: When the damping parameter μ is large

$$h^* \approx -\frac{1}{\mu}\mathbf{J}^T f,$$

the algorithm takes the steepest descent direction. This is advantageous if the current search is far from the solution.

Small damping parameter: When the damping parameter μ is small, the algorithm becomes the Gauss-Newton algorithm, which is very advantageous when the current search is close to the solution. The damping parameter is updated at each loop of the algorithm.

In short, LM algorithm is the combination of the Steepest Descent method and the Gauss-Newton algorithm. As a consequence it converges more quickly than these two algorithms.

Quasi-Newton (BFGS)

In mathematics, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is a method to solve an unconstrained nonlinear optimization problem.

The BFGS method is derived from the Newton's method in optimization, a class of hill-climbing optimization techniques that seeks the stationary point of

a function, where the gradient is 0. Newton's method assumes that the function can be locally approximated as a quadratic in the region around the optimum, and use the first and second derivatives to find the stationary point.

In Quasi-Newton methods the Hessian matrix of second derivatives of the function to be minimized does not need to be computed at any stage. The Hessian is updated by analyzing successive gradient vectors instead. Quasi-Newton methods are a generalization of the secant method to find the root of the first derivative for multidimensional problems. In multi-dimensions the secant equation is under-determined, and quasi-Newton methods differ in how they constrain the solution.

Rationale The search direction $\mathbf{p}_k$ at stage k is given by the solution of the analogue of the Newton equation

$$B_k \mathbf{p}_k = -\nabla f(\mathbf{x}_k). \tag{4.72}$$

A line search in the direction $\mathbf{p}_k$ is then used to find the next point $\mathbf{x}_{k+1}$.

Instead of requiring the full Hessian matrix at the point $\mathbf{x}_{k+1}$ to be computed as B_{k+1}, the approximate Hessian at stage k is updated by the addition of two matrices.

$$B_{k+1} = B_k + U_k + V_k \tag{4.73}$$

Both U_k and V_k are rank-one matrices but have different bases. The rank one assumption here means that we may write

$$C = \mathbf{a}\mathbf{b}^T \tag{4.74}$$

So equivalently, U_k and V_k construct a rank-two update matrix which is robust against the scale problem often suffered in the gradient descent searching.

(as in Broyden's method, the multidimensional analogue of the secant method). The quasi-Newton condition imposed on this update is

$$B_{k+1}(\mathbf{x}_{k+1} - \mathbf{x}_k) = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k). \tag{4.75}$$

Algorithm From an initial guess $\mathbf{x}_0$ and an approximate Hessian matrix B_0 the following steps are repeated until x converges to the solution.

1. Obtain $\mathbf{s}_k$ by solving: $B_k \mathbf{s}_k = -\nabla f(\mathbf{x}_k)$.
2. Perform a line search to find the optimal k in the direction found in the first step, then update $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{s}_k$.
3. $\mathbf{y}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)$.
4. $B_{k+1} = B_k + \frac{\mathbf{y}_k \mathbf{y}_k^\top}{\mathbf{y}_k^\top \mathbf{s}_k} - \frac{B_k \mathbf{s}_k (B_k \mathbf{s}_k)^\top}{\mathbf{s}_k^\top B_k \mathbf{s}_k}$.

$f(\mathbf{x})$ denotes the objective function to be minimized. Convergence can be checked by observing the norm of the gradient, $\left|\nabla f(\mathbf{x}_k)\right|$. Practically, B_0 can be initialized with $B_0 = I$, so that the first step will be equivalent to a gradient descent, but further steps are more and more refined by B_k, the approximation to the Hessian.

The first step of the algorithm is carried out using an approximate inverse of the matrix B_k, which is usually obtained efficiently by applying the Sherman-Morrison formula to the fourth line of the algorithm, giving

$$B_{k+1}^{-1} = B_k^{-1} + (\mathbf{s}_k\mathbf{s}_k^\top)(\mathbf{s}_k^\top\mathbf{y}_k + \mathbf{y}_k^\top B_k^{-1}\mathbf{y}_k)/(\mathbf{s}_k^\top\mathbf{y}_k)^2 - (B_k^{-1}\mathbf{y}_k\mathbf{s}_k^\top + \mathbf{s}_k\mathbf{y}_k^\top B_k^{-1})/(\mathbf{s}_k^\top\mathbf{y}_k). \quad (4.76)$$

Credible intervals or confidence intervals for the solution can be obtained from the inverse of the final Hessian matrix.

Simplex algorithm

The algorithm begins with $n+1$ randomly sampled vectors, $\mathbf{x}_i$ $i \in [0,n]$, when we want to calibrate n parameters in our model. Each component $x_{i,j}$ $j \in [0,n-1]$ represents a parameter of our model. The aim is to find the best parameters to compute implied volatilities with our model. We minimize the following function

$$f(\mathbf{x}) = \sum_{(T,K)\in\Gamma} (\sigma(\mathbf{x},T,K) - \sigma_{Market}(T,K))^2$$

where Γ is the set of all the strikes and maturities in our market data and $\sigma(\mathbf{x},T,K)$ is the implied volatility computed when we consider a model whose parameters are the components of $\mathbf{x}$. When it's possible, $\sigma(\mathbf{x},T,K)$ is directly computed with a closed formula, otherwise we can get it with a closed formula for the price.

The Nelder-Mead algorithm (also known as the simplex algorithm) is usually used for multidimensional unconstrained optimization. It remains a famous solution for at least two reasons : it's one of the methods which do not require derivatives and it's considered as reliable when it comes to deal with noisy functions. If one has to find an optimal solution to a problem in n dimensions, this algorithm creates and modifies a polytope of $n+1$ vertices. Then, one has just to compute objective function values on these vertices to know the next modification of the simplex : reflexion, expansion, contraction, shrinkage. The main idea of these distorsions is to progress toward areas where the objective function takes the best values.

First, we need to find $n+1$ vertices to create a simplex. We use a first set of parameters given by the user and the others are choosen randomly. Algorithm's end occures when one of these condition is satisfied :

- A vertex giving the good precision, x_0, is found :
  $|f(x_0)| < \epsilon_1$
- New vertices don't modify objective function values anymore :
  $\sqrt{\sum_{i=0}^{n} \frac{(f(x_i)-\mu)^2}{n+1}} < \epsilon_2$ where $\mu = \frac{\sum_{i=0}^{n} f(x_i)}{n+1}$
- Time given to the algorithm has passed.
- The execution has rised above iteration limit.

We usually choose :
$\alpha = 1$ (reflexion parameter)
$\beta = 2$ (expansion parameter)
$\gamma = \frac{1}{2}$ (contraction parameter)
$\delta = \frac{1}{2}$ (shrink parameter)
This algorithm consists in the following pseudo code algorithm 1.

Modified simplex algorithm

Previously, we generated $n+1$ vertices to create a simplex. However, it's not enough to find the minimum we're looking for and we compute more objective function values to increase our hope to reach a global solution. That's why we begin with the computation of $4n$ values so as to keep n interesting vertices. This is the first modification of the initial Nelder-Mead algorithm. Then, we also have to look for the minimum in every direction : the initial vertices mustn't be linearly dependent, this would prevent us from transforming the simplex in a part of the space (and from computing objective function values in this part). This step is also called degeneration test. After every initialization a test checks if the simplex is degenerated and resets it if need. The modified simplex algorithm consists in algorithm 2.

Advantages / Drawbacks The simplex algorithm is really famous for many reasons. First of all, it does not need derivatives and we can use it for noisy functions. Moreover, it does not compute many objective function values (indeed, one by iteration in general) so it's not a time-consuming algorithm.

The biggest difficulty with this algorithm appears when we have to minimize multi-modal functions. It's difficult to prevent the algorithm from finding a 'bad' local minimum and giving it as a solution of the problem.

Differential evolution algorithm

The algorithm begins with p randomly sampled vectors to initialize the first generation : $\mathbf{x}_{i,0}$ $i \in [0, p-1]$. Each component $x_{i,0,j}$ $j \in [0, n-1]$ represents a

if *User hasn't provided initial value of* $x_1, x_2, \ldots, x_{n+1}$ **then**
  Randomly generate $n+1$ vertices in n dimension: $x_1, x_2, \ldots, x_{n+1} \in \mathbb{R}^N$;
end
$\alpha = 1$ `// reflexion parameter`
$\beta = 2$ `// expansion parameter`
$\gamma = \frac{1}{2}$ `// contraction parameter`
$\delta = \frac{1}{2}$ `// shrink parameter`
while *condition function is true* **do**
  Sort $x_1, x_2, \ldots, x_{n+1}$ so that $f(x_1) < f(x_2) < \cdots < f(x_{n+1})$;
  $x_{center} \leftarrow \frac{1}{n} \sum_{k=1}^{n} x_k$;
  $x_r \leftarrow x_{center} + \alpha(x_{center} - x_{n+1})$;
  if $f(x_r) < f(x_1)$ **then**
    $x_e \leftarrow x_{center} + \beta(x_r - x_{center})$;
    if $f(x_e) < f(x_r)$ **then**
      $x_{n+1} \leftarrow x_e$;
    else
      $x_{n+1} \leftarrow x_r$;
    end
  else if $f(x_1) \leq f(x_r) < f(x_n)$ **then**
    $x_{n+1} \leftarrow x_r$;
  else if $f(x_n) \leq f(x_r) < f(x_{n+1})$ **then**
    $x_{cout} \leftarrow x_{center} + \gamma(x_r - x_{center})$;
    if $f(x_{cout}) < f(x_r)$ **then**
      $x_{n+1} \leftarrow x_{cout}$;
    else
      $x_{n+1} \leftarrow x_r$;
      Call **shrink function**;
    end
  else if $f(x_{n+1}) \leq f(x_r)$ **then**
    $x_{cin} \leftarrow x_{center} + \gamma(x_{n+1} - x_{center})$;
    if $f(x_{cin}) < f(x_{n+1})$ **then**
      $x_{n+1} \leftarrow x_{cin}$;
    else
      Call **shrink function**;
    end
  end
end

Algorithm 1: Simplex

begin condition function
  if *time limit reached or iteration limit reached* **then**
    return false;
  end
  if $|f(x_1)| < \epsilon_1$ **then**
    return false;
  end
  $\mu \leftarrow \frac{1}{n}\sum_{i=1}^{n+1} f(x_i)$;
  if $\sqrt{\sum_{i=1}^{n+1} \frac{(f(x_i)-\mu)^2}{n+1}} < \epsilon_2$ **then**
    return false;
  end
  return true;
end

Function for simplex

begin shrink function
  for $i = 2$ *to* $n+1$ **do**
    $x_i \leftarrow x_1 + \delta(x_i - x_1)$;
    compute $f(x_i)$;
  end
end

Function for simplex

parameter of our model (we want to calibrate n parameters). One iteration starts with a generation $\mathbf{x}_{i,g}$ $i \in [0, p-1]$ and creates a new set of vectors $\mathbf{x}_{i,g+1}$ $i \in [0, p-1]$. The aim is to find the best parameters to compute implied volatilities with our model. We minimize the following function

$$f(\mathbf{x}) = \sum_{(T,K)\in\Gamma} (\sigma(\mathbf{x}, T, K) - \sigma_{Market}(T, K))^2$$

where Γ is the set of all the strikes and maturities in our market data and $\sigma(\mathbf{x}, T, K)$ is the implied volatility computed when we consider a model whose parameters are the components of $\mathbf{x}$. When it's possible, $\sigma(\mathbf{x}, T, K)$ is directly computed with a closed formula, otherwise we can get it with a closed formula for the price.

The differential evolution algorithm is a population-based optimizer. An initial population is sampled in the domain and objective function values are computed to start the optimization. Then it produces a new generation of vectors with perturbations of the existing points. To generate a new population the scaled difference of two randomly chosen vectors is added to a third vector. The next step consists in a recombination between this new population and the previous

repeat
  if *User hasn't provided initial value of* $x_1, x_2, \ldots, x_{n+1}$ **then**
    Randomly generate $n+1$ vertices in n dimension:
    $x_1, x_2, \ldots, x_{n+1} \in \mathbb{R}^N$;
  end
  $m \leftarrow 4n$;
  Generate $x_{n+2}, x_{n+3}, \ldots, x_m \in \mathbb{R}^N$;
  Sort $x_1, x_2, \ldots, x_m$ so that $f(x_1) < f(x_2) < \cdots < f(x_m)$;
until *The volume of polyhedron* $(x_1, x_2, \ldots, x_{n+1}) > \epsilon$;
$\alpha = 1$; `// reflexion parameter`
$\beta = 2$; `// expansion parameter`
$\gamma = \frac{1}{2}$; `// contraction parameter`
$\delta = \frac{1}{2}$; `// shrink parameter`
while *condition function is true* **do**
  Sort $x_1, x_2, \ldots, x_{n+1}$ so that $f(x_1) < f(x_2) < \cdots < f(x_{n+1})$;
  $x_{center} \leftarrow \frac{1}{n} \sum_{k=1}^{n} x_k$;
  $x_r \leftarrow x_{center} + \alpha(x_{center} - x_{n+1})$;
  if $f(x_r) < f(x_1)$ **then**
    $x_e \leftarrow x_{center} + \beta(x_r - x_{center})$;
    if $f(x_e) < f(x_r)$ **then**
      $x_{n+1} \leftarrow x_e$;
    else
      $x_{n+1} \leftarrow x_r$;
    end
  else if $f(x_1) \leq f(x_r) < f(x_n)$ **then**
    $x_{n+1} \leftarrow x_r$;
  else if $f(x_n) \leq f(x_r) < f(x_{n+1})$ **then**
    $x_{cout} \leftarrow x_{center} + \gamma(x_r - x_{center})$;
    if $f(x_{cout}) < f(x_r)$ **then**
      $x_{n+1} \leftarrow x_{cout}$;
    else
      $x_{n+1} \leftarrow x_r$;
      Call **shrink function**;
    end
  else if $f(x_{n+1}) \leq f(x_r)$ **then**
    $x_{cin} \leftarrow x_{center} + \gamma(x_{n+1} - x_{center})$;
    if $f(x_{cin}) < f(x_{n+1})$ **then**
      $x_{n+1} \leftarrow x_{cin}$;
    else
      Call **shrink function**;
    end
  end
end

Algorithm 2: Modified simplex

one. To achieve the algorithm, we compare objective function values between the two sets of vectors to determine which vectors are part of the new generation.

Algorithm's end occures when one of these condition is satisfied :

- A vector giving the good precision, $x_{0,g}$, is found :

  $|f(x_{0,g})| < \epsilon_1$

- New generations don't modify objective function values anymore :

  $\sqrt{\sum_{i=0}^{n} \frac{(f(x_{i,g})-\mu)^2}{n+1}} < \epsilon_2$ where $\mu = \frac{\sum_{i=0}^{n} f(x_{i,g})}{n+1}$

- Time given to the algorithm has passed.

- The execution has rised above iteration limit.

We use these parameters throughout the evolution of the optimization :

$F = 0.5$ (scale factor used in the mutation step)
$C = 0.5$ (criterion used in the crossover step)
The index g refers to the gth generation.

This algorithm consists in the following algorithm 3.

Advantages / Drawbacks An interesting advantage of this algorithm is the use of difference vectors. The difference distribution adapts the algotithm to the objective function landscape. When we have to cope with a multi-modal function, this distribution is also multi-modal and gives us the opportunity to look for the minimum in each region and to move vectors from a region to another. We have a better chance to find a globalm minimum (in comparison with the simplex algorithm for example).

However, we are not sure to be in a position to afford this opportunity. The main difficulty is the time consumed in each iteration because many objective function values have to be computed for each generation.

MGB algorithm

Call price approximation This algorithm uses an approximation of call prices under the Heston model. We can compute them in the constant case and with a

if *User hasn't provided initial value of* $x_1, x_2, \ldots, x_n$ **then**
  Randomly generate n vertices in n dimension: $x_1, x_2, \ldots, x_n \in \mathbb{R}^N$;
end
$m \leftarrow 5n$;
Generate $x_{n+1}, x_{n+2}, \ldots, x_m \in \mathbb{R}^N$;
$g \leftarrow 0$;
for $i = 1$ *to* m **do**
  $x_{i,g} \leftarrow x_i$
end
$F \leftarrow 0.5$; $C \leftarrow 0.5$;
while *condition function is true* **do**
  `// Mutation`
  for $i = 1$ *to* m **do**
    Randomly choose r_0, r_1, r_2 from 1 to m with $r_0 \neq r_1 \neq r_2$;
    $u_{i,g} \leftarrow x_{r_0,g} + F(x_{r_1,g} - x_{r_2,g})$
  end
  `// Crossover`
  for $i = 1$ *to* m **do**
    Randomly choose j_r from 1 to $n-1$;
    for $j = 0$ *to* n **do**
      randomly choose $u \in [0,1]$;
      if $u < C$ *or* $j == j_r$ **then**
        j-th ordinate of $v_{i,g} \leftarrow$ j-th ordinate of $u_{i,g,j}$
      else
        j-th ordinate of $v_{i,g} \leftarrow$ j-th ordinate of $x_{i,g,j}$
      end
    end
  end
  `// Selection`
  for $i = 1$ *to* m **do**
    if $f(v_{i,g}) < f(x_{i,g})$ **then**
      $x_{i,g+1} \leftarrow v_{i,g}$
    end
    $x_{i,g+1} \leftarrow x_{i,g}$
  end
  $g \leftarrow g + 1$
end

Algorithm 3: Differential evolution

```
begin condition function
    if time limit reached or iteration limit reached then
        return false;
    end
    if |f(x_1)| < ε_1 then
        return false;
    end
    μ ← (1/m) Σ_{i=1}^{m} f(x_i);
    if sqrt( Σ_{i=1}^{m} (f(x_i)−μ)^2 / m ) < ε_2 then
        return false;
    end
    return true;
end
```

Function for differential evolution

time dependent model. From now on we use the following model :

$$\begin{aligned} dS(t) &= r_t S(t)dt + \sqrt{V(t)}S(t)dW_1, \\ dV(t) &= \kappa(\theta_t - V(t))dt + \xi\sqrt{V(t)}dW_2, \\ x_0 &= \ln(S_0), \\ v_0 &= V_0. \end{aligned}$$

ρ_t is the correlation coefficient between W_1 and W_2. These formula are very interesting since they give us the opportunity to write a call price as an easily computable function of ρ and ξ. This will be really useful for a new calibration procedure. In this subsection we will just focus on these analytical formula.

In the constant case, we can derive the following expression:

$$\begin{aligned} var_T &= m_0 v_0 + m_1\theta, & a_{1,T} &= (p_0 v_0 + p_1\theta)\frac{\partial^2 P_{BS}}{\partial xy}(x_0, var_T), \\ a_{2,T} &= (q_0 v_0 + q_1\theta)\frac{\partial^3 P_{BS}}{\partial x^2 y}(x_0, var_T), & b_{0,T} &= (r_0 v_0 + r_1\theta)\frac{\partial^2 P_{BS}}{\partial y^2}(x_0, var_T), \\ b_{2,T} &= \frac{(p_0 v_0 + p_1\theta)^2}{2}\frac{\partial^4 P_{BS}}{\partial x^2 y^2}(x_0, var_T). \end{aligned}$$

where

$$m_0 = \frac{e^{-\kappa T}\left(-1+e^{\kappa T}\right)}{\kappa}, \qquad m_1 = T - \frac{e^{-\kappa T}\left(-1+e^{\kappa T}\right)}{\kappa},$$
$$p_0 = \frac{e^{-\kappa T}\left(-\kappa T + e^{\kappa T} - 1\right)}{\kappa^2}, \qquad p_1 = \frac{e^{-\kappa T}\left(\kappa T + e^{\kappa T}(\kappa T - 2) + 2\right)}{\kappa^2},$$
$$q_0 = \frac{e^{-\kappa T}\left(-\kappa T(\kappa T+2) + 2e^{\kappa T} - 2\right)}{2\kappa^3}, \qquad q_1 = \frac{e^{-\kappa T}\left(2e^{\kappa T}(\kappa T - 3) + \kappa T(\kappa T+4) + 6\right)}{2\kappa^3},$$
$$r_0 = \frac{e^{-2\kappa T}\left(-4e^{\kappa T}\kappa T + 2e^{2\kappa T} - 2\right)}{4\kappa^3}, \qquad r_1 = \frac{e^{-2\kappa T}\left(4e^{\kappa T}(\kappa T+1) + e^{2\kappa T}(2\kappa T - 5) + 1\right)}{4\kappa^3}.$$

and $P_{BS}(x_0, var_T)$ is a call price under a Black-Scholes model with volatility $\sqrt{\frac{var_T}{T}}$. Then, the MGB formula is :

$$P_{MGB}(x_0, T) = P_{BS}(x_0, var_T, T) + a_{1,T}\rho\xi + (a_{2,T} + b_{2,T})\rho^2\xi^2 + b_{0,T}\xi^2$$

If we consider piecewise constant parameters, we write $T_0 = 0 \le T_1 \le \cdots \le T_n = T$ such that θ, ρ, ξ are constant on each interval $]T_i, T_{i+1}[$ and are equal respectively to $\theta_{T_{i+1}}, \rho_{T_{i+1}}, \xi_{T_{i+1}}$. The new coefficients used to derive our analytical formula are

$$\begin{aligned}
a_{1,T_{i+1}} &= a_{1,T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{1,T_i} + \rho_{T_{i+1}}\xi_{T_{i+1}} f^1_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
a_{2,T_{i+1}} &= a_{2,T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\alpha_{T_i} + \rho_{T_{i+1}}\xi_{T_{i+1}}\tilde{\omega}^{0,-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{1,T_i} + (\rho_{T_{i+1}}\xi_{T_{i+1}})^2 f^2_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
b_{0,T_{i+1}} &= b_{0,T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\beta_{T_i} + \tilde{\omega}^{-\kappa,-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{2,T_i} + \xi^2_{T_{i+1}} f^0_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
\alpha_{T_{i+1}} &= \alpha_{T_i} + \rho_{T_{i+1}}\xi_{T_{i+1}}(T_{i+1} - T_i)\tilde{\omega}_{1,T_i} + \rho^2_{T_{i+1}}\xi^2_{T_{i+1}} g^1_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
\beta_{T_{i+1}} &= \beta_{T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{2,T_i} + \xi^2_{T_{i+1}} g^2_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
\tilde{\omega}_{1,T_{i+1}} &= \tilde{\omega}_{1,T_i} + \rho_{T_{i+1}}\xi_{T_{i+1}} h^1_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
\tilde{\omega}_{2,T_{i+1}} &= \tilde{\omega}_{2,T_i} + \xi^2_{T_{i+1}} h^2_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
v_{0,T_{i+1}} &= e^{-\kappa(T_{i+1}-T_i)}(v_{0,T_i} - \theta_{T_{i+1}}) + \theta_{T_{i+1}},
\end{aligned}$$

where

$$f^0_{\kappa,v_0}(\theta,t,T) = \frac{e^{-2\kappa T}\left(e^{2\kappa t}(\theta-2v_0) + e^{2\kappa T}((-2\kappa t + 2\kappa T - 5)\theta + 2v_0) + 4e^{\kappa(t+T)}((-\kappa t + \kappa T + 1)\theta + \kappa(t-T)v_0)\right)}{4\kappa^3},$$
$$f^1_{\kappa,v_0}(\theta,t,T) = \frac{e^{-\kappa T}\left(e^{\kappa T}((-\kappa t + \kappa T - 2)\theta + v_0) - e^{\kappa t}((\kappa t - \kappa T - 2)\theta - \kappa t v_0 + \kappa T v_0 + v_0)\right)}{\kappa^2},$$
$$f^2_{\kappa,v_0}(\theta,t,T) = \frac{e^{-\kappa(t+3T)}\left(2e^{\kappa(t+3T)}((\kappa(T-t)-3)\theta + v_0) + e^{2\kappa(t+T)}((\kappa(\kappa(t-T)-4)(t-T)+6)\theta - (\kappa(\kappa(t-T)-2)(t-T)+2)v_0)\right)}{2\kappa^3},$$
$$g^1_{\kappa,v_0}(\theta,t,T) = \frac{2e^{\kappa T}\theta + e^{\kappa t}\left(\kappa^2(t-T)^2 v_0 - (\kappa(\kappa(t-T)-2)(t-T)+2)\theta\right)}{2\kappa^2},$$
$$g^2_{\kappa,v_0}(\theta,t,T) = \frac{e^{-\kappa T}\left(e^{2\kappa T}\theta - e^{2\kappa t}(\theta - 2v_0) + 2e^{\kappa(t+T)}(\kappa(t-T)(\theta - v_0) - v_0)\right)}{2\kappa^2},$$
$$h^1_{\kappa,v_0}(\theta,t,T) = \frac{e^{\kappa T}\theta + e^{\kappa t}((\kappa t - \kappa T - 1)\theta + \kappa(T-t)v_0)}{\kappa},$$
$$h^2_{\kappa,v_0}(\theta,t,T) = \frac{\left(e^{\kappa t} - e^{\kappa T}\right)\left(e^{\kappa t}(\theta - 2v_0) - e^{\kappa T}\theta\right)}{2\kappa},$$

and $\tilde{\omega}^u_t(T) = \frac{-e^{tu} + e^{Tu}}{u}$, $\tilde{\omega}^{0,u}_t(T) = \frac{e^{Tu}(-tu + Tu - 1) + e^{tu}}{u^2}$, $\tilde{\omega}^{u,u}_t(T) = \frac{\left(e^{tu} - e^{Tu}\right)^2}{2u^2}$.

Thanks to these expressions we can write $P_{MGB}(x_0, T_{i+1})$ as a function of $\rho_{T_{i+1}}$ and $\xi_{T_{i+1}}$ with :

$$\begin{aligned}
a'_{1,T_{i+1}} =& f^1_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1})\frac{\partial^2 P_{BS}}{\partial xy}(x_0, var_{T_{i+1}}) + \tilde{\omega}^{0,-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{1,T_i}\frac{\partial^3 P_{BS}}{\partial x^2 y}(x_0, var_{T_{i+1}}) \\
&+ (a_{1,T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{1,T_i})\frac{\partial^4 P_{BS}}{\partial x^2 y^2}(x_0, var_{T_{i+1}}), \\
a'_{2,T_{i+1}} =& f^2_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1})\frac{\partial^3 P_{BS}}{\partial x^2 y}(x_0, var_{T_{i+1}}), \\
b'_{0,T_{i+1}} =& f^0_{\kappa,v_{0,T_i}}(\theta_{T_{i+1}}, T_i, T_{i+1})\frac{\partial^2 P_{BS}}{\partial y^2}(x_0, var_{T_{i+1}}), \\
b'_{2,T_{i+1}} =& \frac{1}{2}\frac{\partial^4 P_{BS}}{\partial x^2 y^2}(x_0, var_{T_{i+1}}), \\
a'_{0,T_{i+1}} =& (a_{1,T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{1,T_i})\frac{\partial^2 P_{BS}}{\partial xy}(x_0, var_{T_{i+1}}) \\
&+ (a_{2,T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\alpha_{T_i})\frac{\partial^3 P_{BS}}{\partial x^2 y}(x_0, var_{T_{i+1}}) \\
&+ (b_{0,T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\beta_{T_i} + \tilde{\omega}^{-\kappa,-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{2,T_i})\frac{\partial^2 P_{BS}}{\partial y^2}(x_0, var_{T_{i+1}}) \\
&+ \frac{(a_{1,T_i} + \tilde{\omega}^{-\kappa}_{T_i,T_{i+1}}\tilde{\omega}_{1,T_i})^2}{2}\frac{\partial^4 P_{BS}}{\partial x^2 y^2}(x_0, var_{T_{i+1}}).
\end{aligned}$$

$$\begin{aligned}
P_{MGB}(x_0, T_{i+1}) =& a'_{0,T_{i+1}} + P_{BS}(x_0, var_T, T) + a'_{1,T_{i+1}}\rho_{T_{i+1}}\xi_{T_{i+1}} + (a'_{2,T_{i+1}} + b'_{2,T_{i+1}})\rho_{T_{i+1}}{}^2\xi_{T_{i+1}}{}^2 \\
&+ b'_{0,T_{i+1}}\xi_{T_{i+1}}{}^2
\end{aligned}$$

From now on, for each maturity T, this formula gives us the opportunity to write a call price as an easily computable function of ρ_T and ξ_T.

Calibration This algorithm calibrates the Heston model using the new formula to compute prices. We consider a time dependent Heston model with constant v_0 (chosen accordingly to our market data) and κ. With several maturities, Θ represents the values of the longterm vol on $[0, T_1], ..., [T_{N-1}, T_N]$ and Γ and Ξ the values of the correlation and the volatility of volatility on $[0, T_1], ..., [T_{N-2}, T_{N-1}]$. If ρ_N and ξ_N are the corresponding values on $[T_{N-1}, T_N]$, the call price for a strike K is :

$$P_{MGB}(T_N, K) = a(\kappa, \Theta, \Gamma, \Xi) + b(\kappa, \Theta, \Gamma, \Xi)\rho_N\xi_N + c(\kappa, \Theta, \Gamma, \Xi)\xi_N^2 + d(\kappa, \Theta, \Gamma, \Xi)\rho_N^2\xi_N^2$$

It seems really interesting to use this approximation in a calibration procedure since two market prices are enough to solve the system and get the volatility of volatility and the correlation. However, numerical solutions for ξ_N and ρ_N can

be useless for our model (for instance if we have to deal with complex numbers or huge ξ_N). Finding a good value for θ_N in order to get interesting ξ_N and ρ_N is the aim of this calibration.

The algorithm starts with given values for κ, v_0 and θ_1 for the first maturity. Then, we use a simplex algorithm to find a good θ_1. For each value of θ_1, the simplex solves the following system (using Mathemartica):

$$\begin{cases} P_1(T_1) = a(\kappa,\theta_1) + b(\kappa,\theta_1)\rho_1\xi_1 + c(\kappa,\theta_1)\xi_1^2 + d(\kappa,\theta_1)\rho_1^2\xi_1^2 \\ P_2(T_1) = a(\kappa,\theta_1) + b(\kappa,\theta_1)\rho_1\xi_1 + c(\kappa,\theta_1)\xi_1^2 + d(\kappa,\theta_1)\rho_1^2\xi_1^2 \end{cases}$$

where $P_1(T_1)$ and $P_2(T_1)$ are two market prices (at the money and the first strike higher than the spot) and the coefficients a, b, c and d are given in the previous subsection. This system gives four solutions but two of them lead to negative ξ. We decided to remove one of the two remaining solutions since it did not give suitable results with our pricing formula. Then we have to realize our optimization with one of these solution : (ρ_1, ξ_1). The simplex computes ($\overline{\xi}$ is the upper bound chosen for ξ)

$$\begin{aligned} F(\rho_1,\xi_1) = \max(0,|\rho_1|-1) + |\mathrm{Im}(\rho_1)| + |\mathrm{Im}(\xi_1)| + |\min(0,\mathrm{Re}(\xi_1))| \\ + \max(0,\mathrm{Re}(\xi_1)^2 - 2\kappa\theta_1) + \max(0,\mathrm{Re}(\xi_1) - \overline{\xi}) \end{aligned}$$

If $F(\rho_1,\xi_1) = 0$ this is a good solution and the simplex returns the current value of θ_1 and save ρ_1 and ξ_1. Otherwise, it tries to minimize F. To end the calibration on the first maturity, we use these solutions as an initial vector for a Levenberg Marquardt optimization.

When the model is calibrated on several maturities, T_0,...,T_N, we write Γ, Θ and Ξ the calibrated values of ρ, θ and ξ from 0 to T_N and the algorithm begins with a given value θ_{N+1} for T_{N+1}. For each value of θ_{N+1}, the simplex solves the following system :

$$\begin{cases} P_1(T_{N+1}) & = a(\kappa,\Gamma,\Xi,\Theta,\theta_{N+1}) + b(\kappa,\Gamma,\Xi,\Theta,\theta_{N+1})\rho_{N+1}\xi_{N+1} \\ & \quad + c(\kappa,\Gamma,\Xi,\Theta,\theta_{N+1})\xi_{N+1}^2 + d(\kappa,\Gamma,\Xi,\Theta,\theta_{N+1})\rho_{N+1}^2\xi_{N+1}^2 \\ P_2(T_{N+1}) & = a(\kappa,\Gamma,\Xi,\Theta,\theta_{N+1}) + b(\kappa,\Gamma,\Xi,\Theta,\theta_{N+1})\rho_{N+1}\xi_{N+1} \\ & \quad + c(\kappa,\Gamma,\Xi,\Theta,\theta_{N+1})\xi_{N+1}^2 + d(\kappa,\Gamma,\Xi,\Theta,\theta_{N+1})\rho_{N+1}^2\xi_{N+1}^2 \end{cases}$$

where $P_1(T_{N+1})$ and $P_2(T_{N+1})$ are two market prices (at the money and the first strike higher than the spot). Then, the simplex tries to find a θ_{N+1} which gives good values for ρ_{N+1} and ξ_{N+1}. If it fails, this part is replaced by a classical simplex optimizing every parameter. The solution of this subsection is then used as an initial vector for a Levenberg Marquardt optimization.

Table 4.1: Market data of crude oil future

Maturity	Jun09	Jul09	Aug09	Sep09	Oct09	Nov09
Future Value	66.12	67.09	68.02	68.86	69.66	70.46
	Dec09	Jan10	Feb10	Mar10	Apr10	May10
	71.11	71.63	72.11	72.59	73.06	73.52
	Jun10	Jul10	Aug10	Sep10	Oct10	Nov10
	73.97	74.3	74.53	74.79	75.05	75.3
	Dec10	Jan11	Feb11	Mar11	Apr11	May11
	75.56	75.84	76.13	76.42	76.71	77

4.8 Numerical analysis

4.8.1 Market data: crude oil

We take a market data as of the 3rd June 2009. The spot value is 66.12. The future and volatility market data are shown in figure 4.8 , table 4.1 and figure 4.9 respectively.

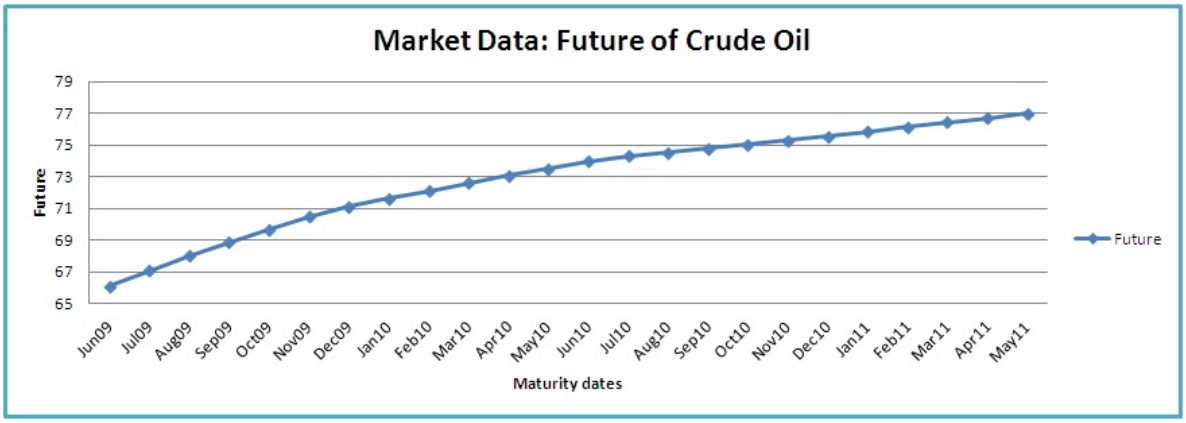


Figure 4.8: Market data of crude oil future

Now we use the spot, futures and volatility surface market data to calibrate a Heston model. All the numerical tests are carried on under financial software Price-it®Excel.

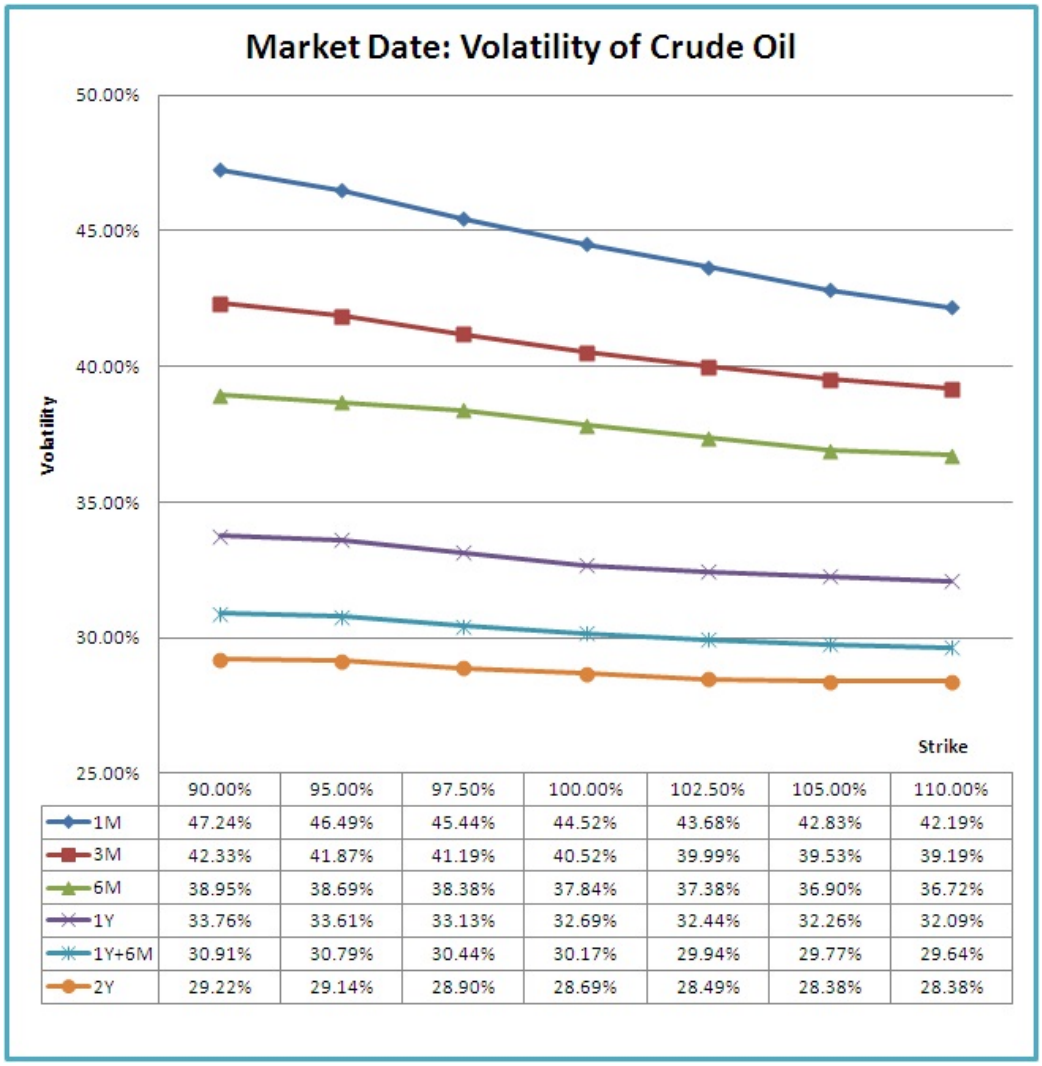
We recall Heston model here for convenience.

$$\frac{\mathrm{d}S_t}{S_t} = r\mathrm{d}t + \sqrt{V_t}\mathrm{d}W_1 \tag{4.77}$$

$$dV_t = k(\theta - V_t)dt + \xi\sqrt{V_t}\mathrm{d}W_2 \tag{4.78}$$

There are five parameters to calibrate:

1. V_0, the initial value of V_t, is initial volatility.



	90.00%	95.00%	97.50%	100.00%	102.50%	105.00%	110.00%
1M	47.24%	46.49%	45.44%	44.52%	43.68%	42.83%	42.19%
3M	42.33%	41.87%	41.19%	40.52%	39.99%	39.53%	39.19%
6M	38.95%	38.69%	38.38%	37.84%	37.38%	36.90%	36.72%
1Y	33.76%	33.61%	33.13%	32.69%	32.44%	32.26%	32.09%
1Y+6M	30.91%	30.79%	30.44%	30.17%	29.94%	29.77%	29.64%
2Y	29.22%	29.14%	28.90%	28.69%	28.49%	28.38%	28.38%

Figure 4.9: Market data of crude oil volatility

2. θ is long term volatility.
3. k is mean reversion.
4. ξ is vol of vol.
5. ρ is the correlation between W_1 and W_2

.

4.8.2 Calibration Heston model with constant parameters

In the first test, we assume all the parameters of Heston model are constant. And we use only one year maturity on the volatility surface. The optimization

Table 4.2: Calibration result

InitialVol	LongTermVol	MeanReversion	VolOfVol	Correlation
0.3981	0.0614	1.0922	0.6451	-0.0853

algorithm is set to be Levenberg-Marquardt. The result of calibration is listed in table 4.2. We use this set of parameter in Heston model and re-calculate the options we use to calibrate. In this way, we can compare the smile from the model and the smile of market data. It is shown in figure 4.10.

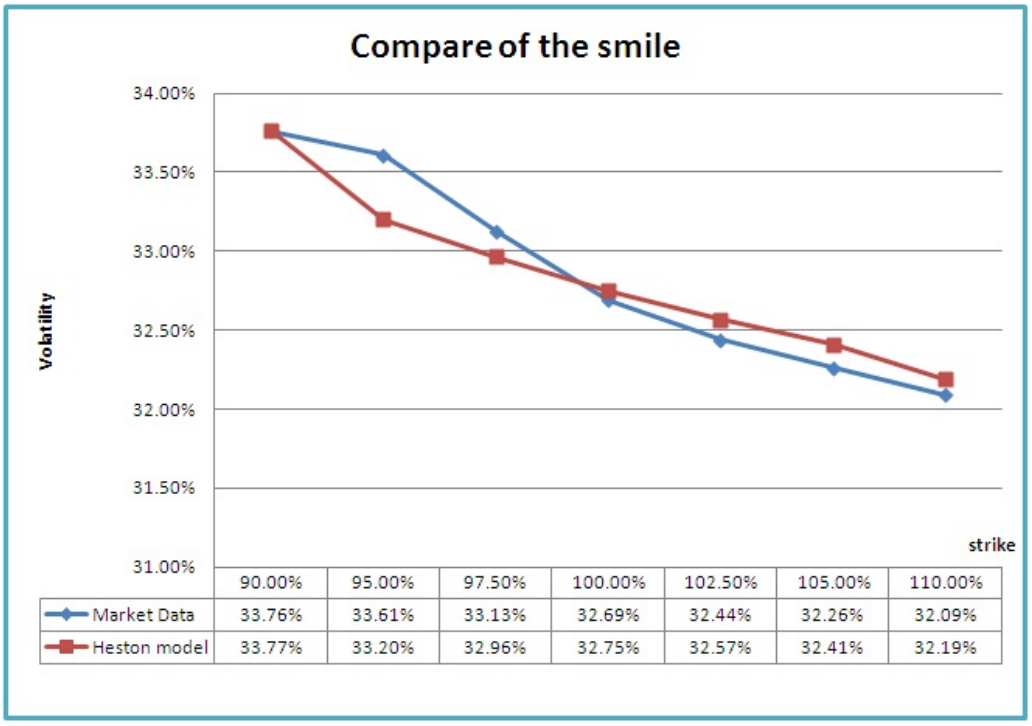


Figure 4.10: Smile of Market data and Heston model

From figure 4.10 we observe the largest difference between two curves at strike 95%. In market data, the smile presents a slight "Z" form instead of normal "U" form, especially at point strike 95%. Heston model cannot match exactly at this point. Therefore it causes a relative larger difference than other points. The difference is shown in figure 4.11.

4.8.3 Calibration Heston model with time-dependent parameters

In this test, we continue to use the same market data in figue 4.8 , table 4.1 and figure 4.9. We let parameters in Heston model be time-dependent, more precisely piecewise constant. Among the five Heston model parameters, we allow three

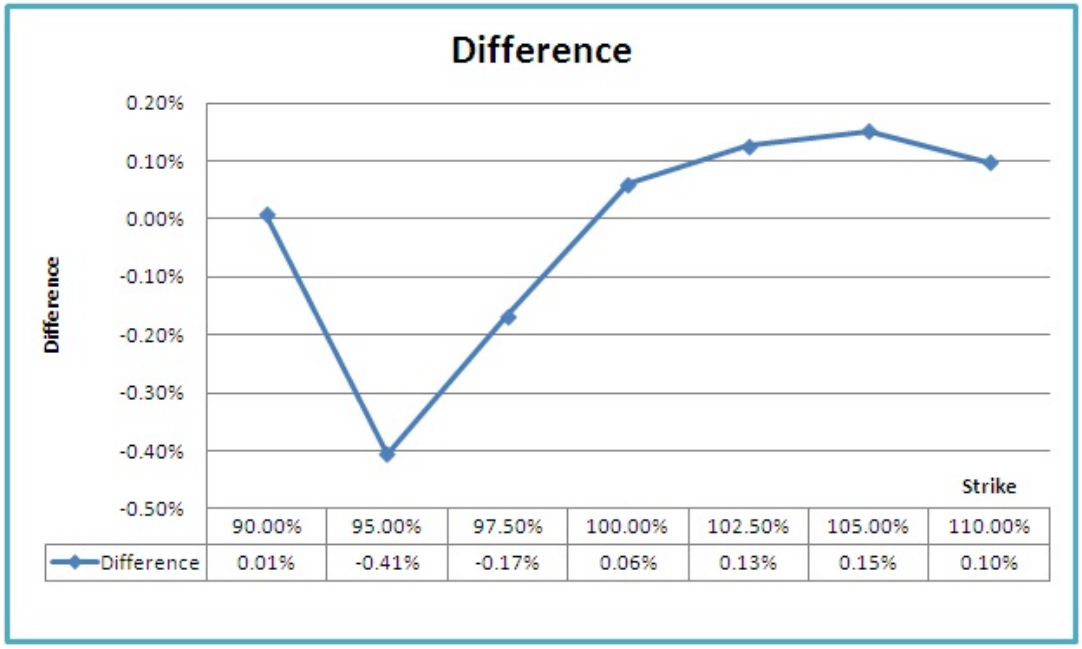


Figure 4.11: Difference on smile

of them to be piecewise constant: long term volatility, vol of vol and correlation. Parameter initial volatility and Mean reversion parameter are set to be constant over different maturities. The reason of setting mean reversion to be constant is to decrease the degree of freedom in calibration, in order to stabilize the result of calibration. We will discuss the stability of calibration in more details in the next subsection.

Here we take maturity 1 year, 18 months and 2 years. So there are 21 input volatility in total, seven strikes for each maturity. The calibration targets to minimize the sum of the square of the difference on all the 21 points. The calibrated parameters are listed in table 4.3 and the difference between market data and Heston model is shown in figure 4.13 and 4.12.

Compare figure 4.11 and 4.12, we can see they present same form. The reason is always that market data presents irregularity at strike 95%. Another observation from table 4.3 is that the correlation parameter changes sign in the second maturity, though there is little difference in the shape of the volatility curves in figure 4.13.

With the observation on correlation parameter, we now set correlation as constant over different maturities. The calibrated parameters are listed in table 4.4 and the difference between market data and Heston model is shown in figure 4.14 and 4.15. We can observe that all the parameters are smoother without big change over different maturities in table 4.4. And the calibration error in figure 4.14 remains as little as that in the case of time-dependent correlation.

Table 4.3: Calibration result for time-dependent Heston

	InitialVol	LongTermVol	MeanReversion	VolOfVol	Correlation
1Y	47.96%	8.68%	408.47%	142.19%	-9.13%
1Y+6M	47.96%	4.72%	408.47%	186.55%	30.82%
2Y	47.96%	7.49%	408.47%	112.43%	-1.90%

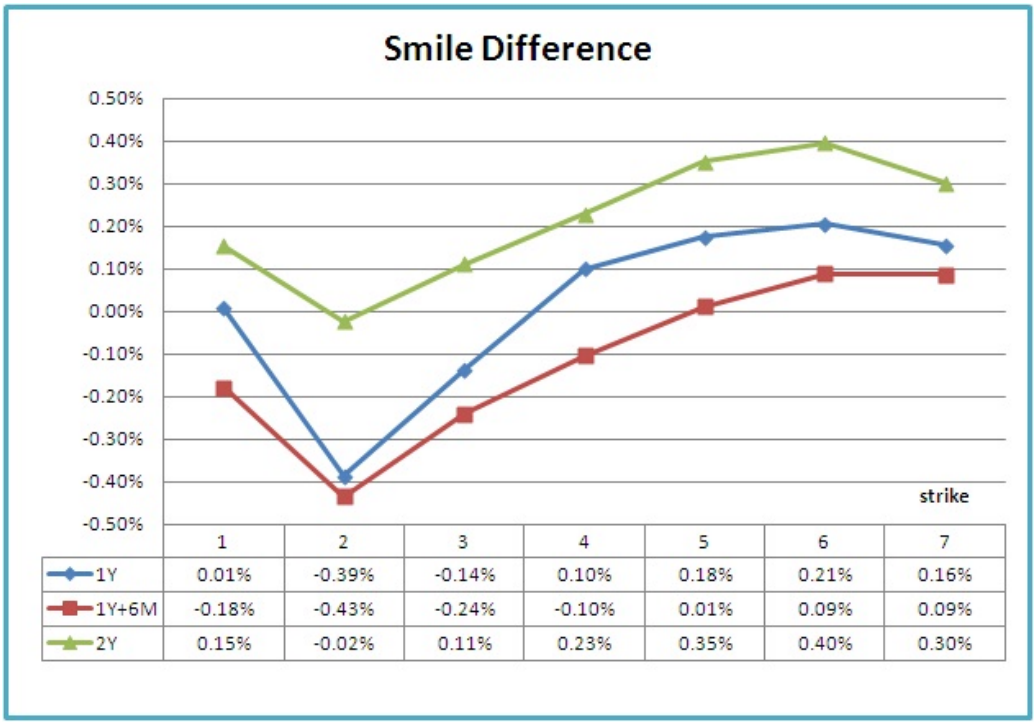
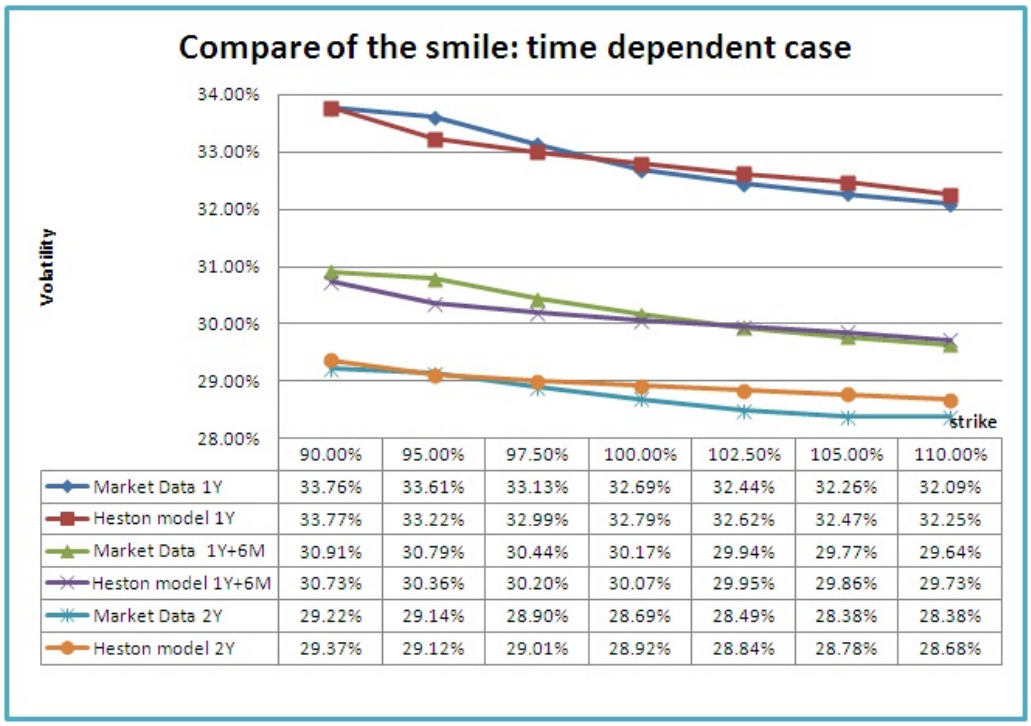


Figure 4.12: Difference on Smile of Market data and time-dependent Heston model

From the comparison, we find that we don't need to allow every parameters to be time-dependent. Two time-dependent parameters can achieve a calibration result as good as the case of three time-dependent parameters.

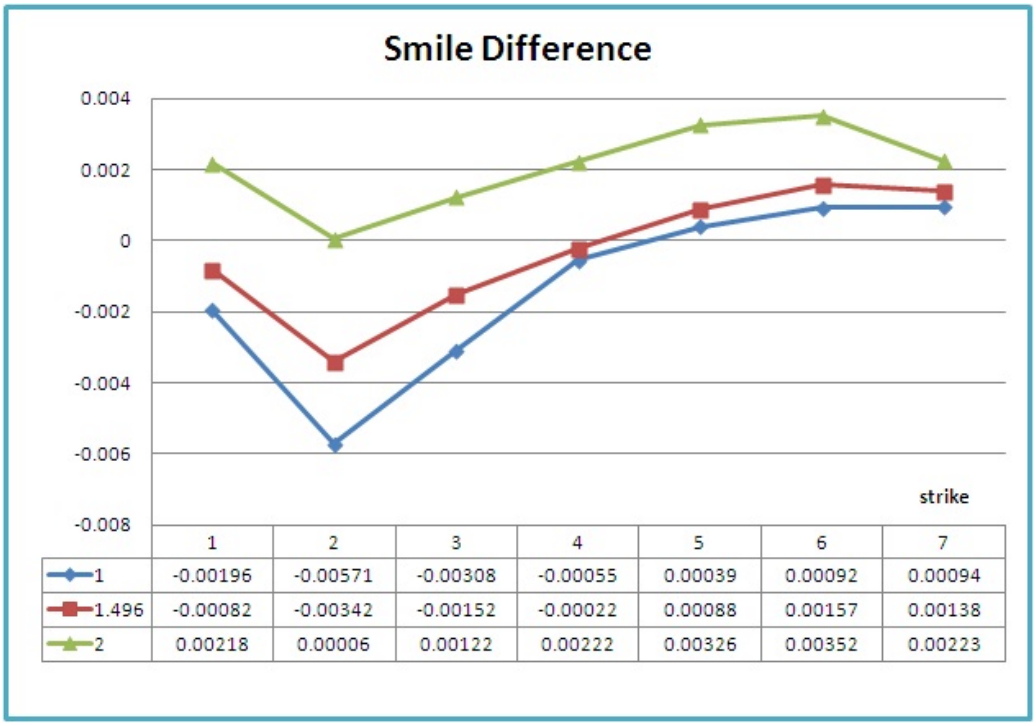


	90.00%	95.00%	97.50%	100.00%	102.50%	105.00%	110.00%
Market Data 1Y	33.76%	33.61%	33.13%	32.69%	32.44%	32.26%	32.09%
Heston model 1Y	33.77%	33.22%	32.99%	32.79%	32.62%	32.47%	32.25%
Market Data 1Y+6M	30.91%	30.79%	30.44%	30.17%	29.94%	29.77%	29.64%
Heston model 1Y+6M	30.73%	30.36%	30.20%	30.07%	29.95%	29.86%	29.73%
Market Data 2Y	29.22%	29.14%	28.90%	28.69%	28.49%	28.38%	28.38%
Heston model 2Y	29.37%	29.12%	29.01%	28.92%	28.84%	28.78%	28.68%

Figure 4.13: Smile of Market data and Heston model for time dependent case

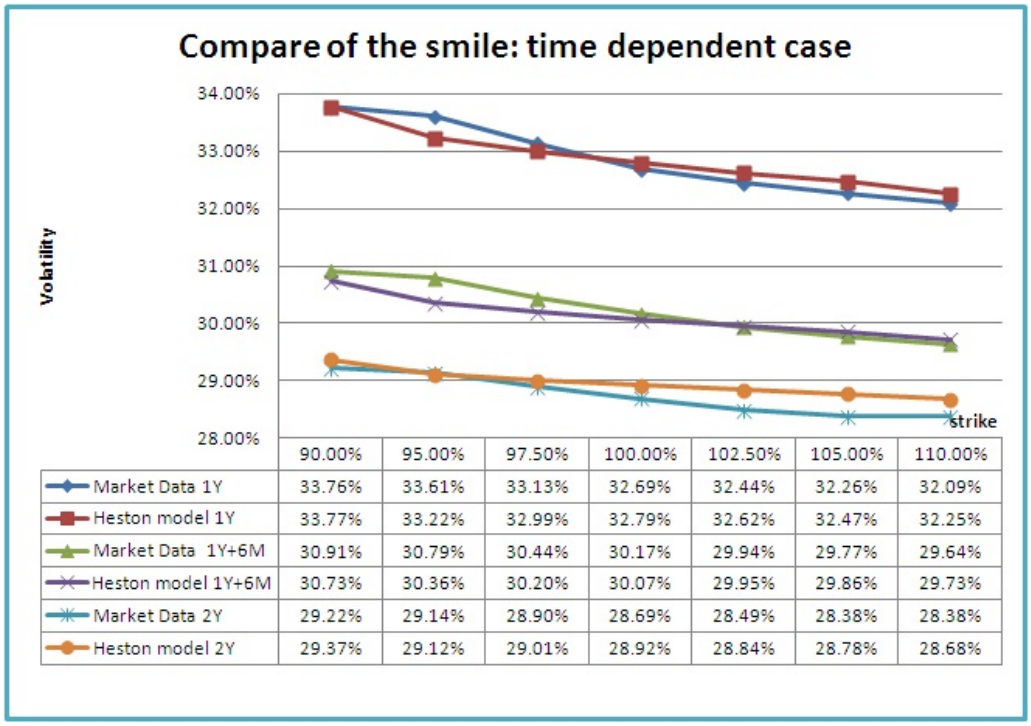
Table 4.4: Calibration result with constant correlation

	InitialVol	LongTermVol	MeanReversion	VolOfVol	Correlation
1Y	59.77%	6.32%	495.80%	171.95%	-5.01%
1Y+6M	59.77%	6.06%	495.80%	171.98%	-5.01%
2Y	59.77%	6.49%	495.80%	169.17%	-5.01%



	1	2	3	4	5	6	7
1	-0.00196	-0.00571	-0.00308	-0.00055	0.00039	0.00092	0.00094
1.496	-0.00082	-0.00342	-0.00152	-0.00022	0.00088	0.00157	0.00138
2	0.00218	0.00006	0.00122	0.00222	0.00326	0.00352	0.00223

Figure 4.14: Difference on Smile of Market data and time-dependent Heston model with constant correlation



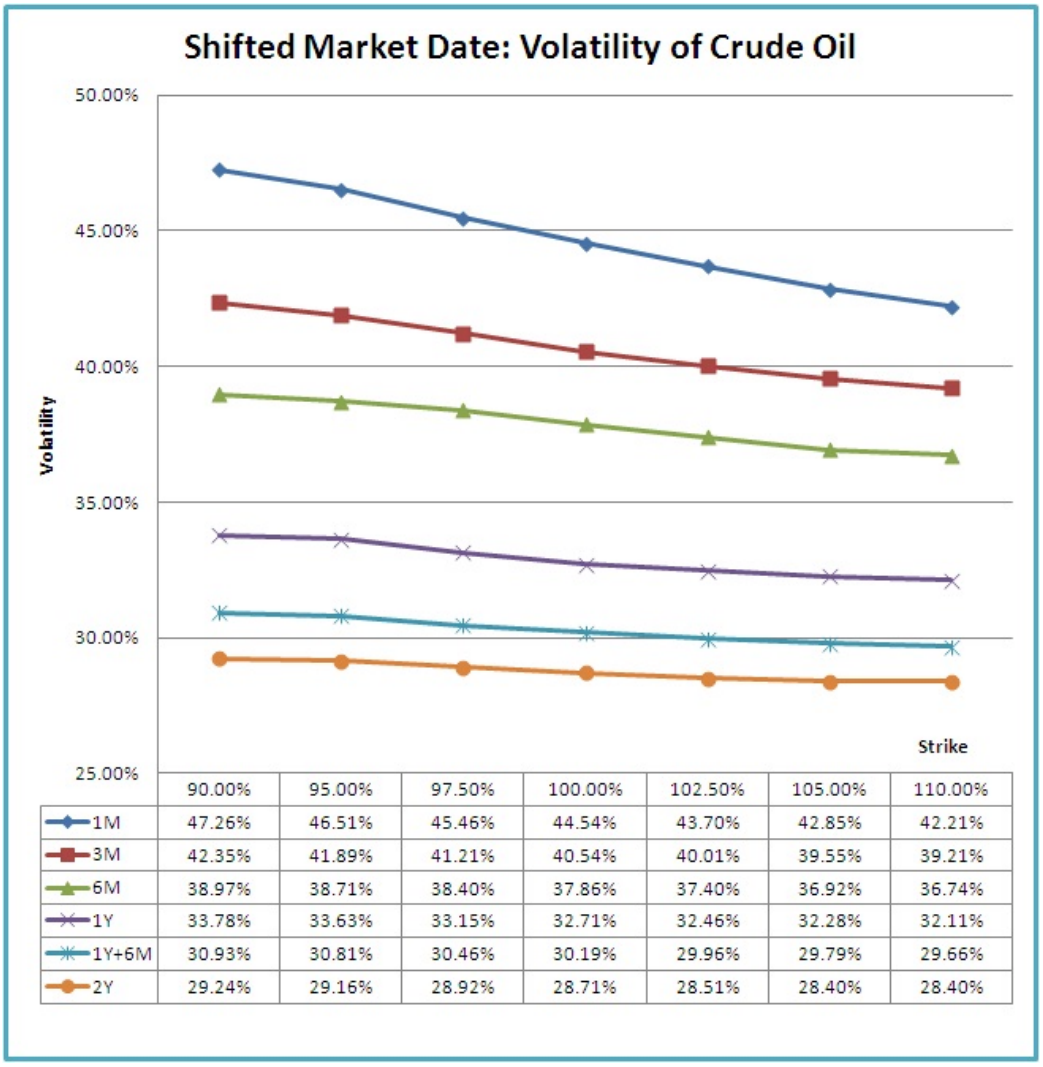
	90.00%	95.00%	97.50%	100.00%	102.50%	105.00%	110.00%
Market Data 1Y	33.76%	33.61%	33.13%	32.69%	32.44%	32.26%	32.09%
Heston model 1Y	33.77%	33.22%	32.99%	32.79%	32.62%	32.47%	32.25%
Market Data 1Y+6M	30.91%	30.79%	30.44%	30.17%	29.94%	29.77%	29.64%
Heston model 1Y+6M	30.73%	30.36%	30.20%	30.07%	29.95%	29.86%	29.73%
Market Data 2Y	29.22%	29.14%	28.90%	28.69%	28.49%	28.38%	28.38%
Heston model 2Y	29.37%	29.12%	29.01%	28.92%	28.84%	28.78%	28.68%

Figure 4.15: Smile of Market data and Heston model for constant correlation case

4.8.4 Sensibility of parameters

In this subsection, we add 0.02% on the volatility surface of market data. Then we use this shifted market data to calibrate Heston model. And we compare against the previous result to see if the parameters largely change or just have small change.

The spot value remains at 66.12. And the futures are still as in table 4.1. But the volatility surface is shifted by 0.02% as shown in figure 4.16.



	90.00%	95.00%	97.50%	100.00%	102.50%	105.00%	110.00%
1M	47.26%	46.51%	45.46%	44.54%	43.70%	42.85%	42.21%
3M	42.35%	41.89%	41.21%	40.54%	40.01%	39.55%	39.21%
6M	38.97%	38.71%	38.40%	37.86%	37.40%	36.92%	36.74%
1Y	33.78%	33.63%	33.15%	32.71%	32.46%	32.28%	32.11%
1Y+6M	30.93%	30.81%	30.46%	30.19%	29.96%	29.79%	29.66%
2Y	29.24%	29.16%	28.92%	28.71%	28.51%	28.40%	28.40%

Figure 4.16: Shifted market data of crude oil volatility

The calibrated parameters are listed in table 4.5 and the difference between market data and Heston model is shown in figure 4.17 and 4.18.

The result in table 4.4 and 4.5 shows that the parameters of Heston model don't change a lot. They are relatively stable. And the calibration result in figure

Table 4.5: Calibration result with constant correlation

	InitialVol	LongTermVol	MeanReversion	VolOfVol	Correlation
1Y	64.48%	5.41%	494.33%	157.08%	-19.32%
1Y+6M	64.48%	6.36%	494.33%	159.48%	-19.32%
2Y	64.48%	6.79%	494.33%	155.21%	-19.32%

4.17 stays at a good level as the case before shifting in figure 4.14.

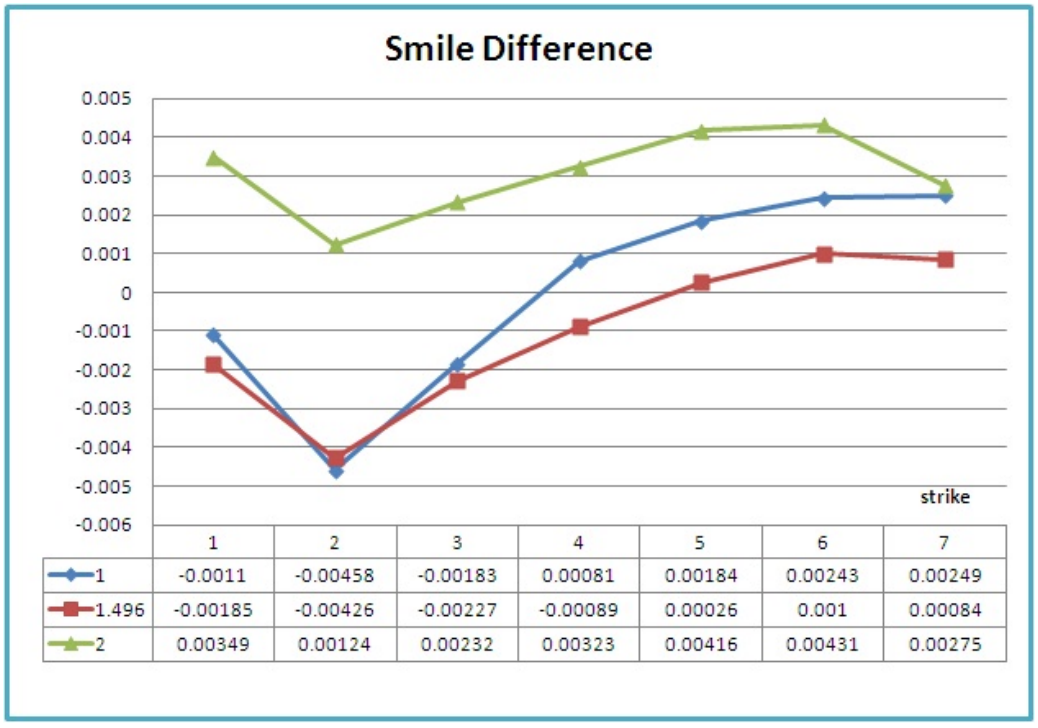
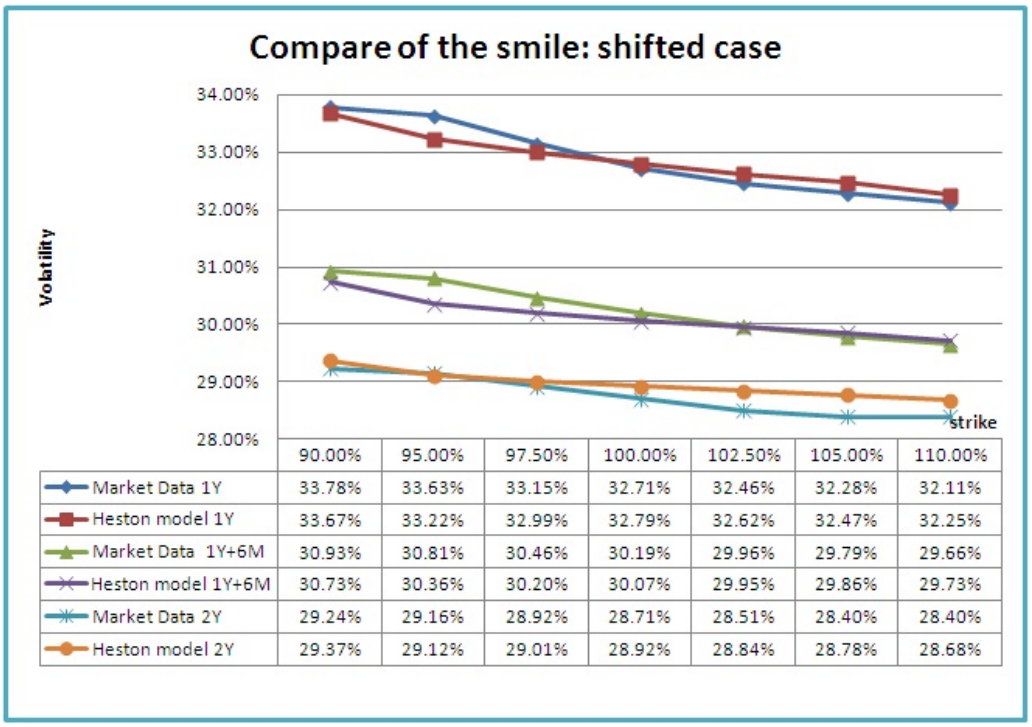


Figure 4.17: Difference on Smile of Market data and time-dependent Heston model with shifted market data



	90.00%	95.00%	97.50%	100.00%	102.50%	105.00%	110.00%
Market Data 1Y	33.78%	33.63%	33.15%	32.71%	32.46%	32.28%	32.11%
Heston model 1Y	33.67%	33.22%	32.99%	32.79%	32.62%	32.47%	32.25%
Market Data 1Y+6M	30.93%	30.81%	30.46%	30.19%	29.96%	29.79%	29.66%
Heston model 1Y+6M	30.73%	30.36%	30.20%	30.07%	29.95%	29.86%	29.73%
Market Data 2Y	29.24%	29.16%	28.92%	28.71%	28.51%	28.40%	28.40%
Heston model 2Y	29.37%	29.12%	29.01%	28.92%	28.84%	28.78%	28.68%

Figure 4.18: Smile of Market data and Heston model with shifted market data

4.9 Summary

In this chapter, we have mainly discussed stochastic volatility model in two parts. In the first part (section 4.1 to 4.5), we replace the constant model parameters to time dependent, which extends the original stochastic volatility model to dynamic stochastic volatility model. It allows more degree of freedom on the implied volatility surface generated by these models. We study the closed formula for vanilla option price for three typical stochastic volatility models, namely Heston model, Piterbarg model and SABR model, by employing control variate, parameter averaging technique and perturbation theory. Instead of the the formulas for the option price, we also have the formula for implied volatility. In section 4.6, we discuss the result of vol of vol expansion. The advantage of the method is that it can get directly the implied volatility which is very convenient for calibration, but the precision is often not good enough for calibration.

In the second part (section 4.7), we discuss the calibration for stochastic volatility models. Substantially, calibration is the process to numerically reverse the closed formula of option price in n-dimension space (n is the number of the parameters to calibrate). As the closed formulas for option price have ready achieved in the first part, we focus on different numerical algorithm, including bootstrap, Newton Raphson, simplex, differential evolution. We list their pseudo-code and comment on their advantage and limitation.

Chapter 5

Conclusion

The modeling of commodity and commodity derivatives consists in three steps:

1. choose the dynamic model of the spot price,

2. calibrate the model,

3. price the products with the calibrated model.

All these three steps are necessarily needed for a complete pricing process and they are related to each other. Step three needs the model parameters and so it depends on the result of step two, the calibration. Step two needs to use the formula of vanilla option price to calibrate the model to the market implied volatility surface, which depends the choice of the dynamic model. Interestingly, the choice of the dynamic model sometimes depends on the final product we want to price. For example, if the product is sensitive to the volatility smile, then the stochastic volatility model family is a better choice than Black-Scholes model.

In this thesis, we focus on the different choices of dynamic model and their formula for vanilla option price, as well as some calibration algorithms. Using the model to price the products by employing numerical method, such as Monte Carlo simulation or PDE grid is beyond the scope of the thesis.

There are basically two different families of dynamic models in our discussion, namely stochastic convenience yield model and stochastic volatility model. We pick up the most popular models, Gibson-Schwartz model and Gabillon model, from the family of stochastic convenience yield model. In chapter 2, we proved the mathematical equivalence relation between Gibson-Schwartz model and Gabillon model. It is not surprising because both models share the similar economy concept of stochastic convenience yield. And in financial mathematics

equation, it shows the stochastic drift term. In chapter 3, we presented that both Gibson-Schwartz model and Gabillon model were equivalent to a one-factor model by introducing model factor reduction technique. Of course, this conclusion is true only when the product is not path dependent because the equivalence is established on the marginal distribution of the stochastic process. Based on the equivalent one-factor model, we give the formula of vanilla option price with the underlying being spot as well as forward.

The family of stochastic volatility model is studied in chapter 4. The main idea is to extend the original version of the models to dynamic models by allowing their parameter depending on time. To make it more practical, the parameters are not set to be continuous function depending on time. We choose piecewise constant parameters. The reason to choose piecewise constant parameters is based on the fact that market data is normally given on a few maturities. Therefore, we simply don't have the information of the parameters between two maturities of market data from the point of view of calibration. These piecewise constant parameters represent for the average value of the parameters between two maturities in term of calibration result. Based on this assumption, we extended Heston model, SABR model and Piterbarg model into dynamic Heston model, dynamic SABR model and dynamic Piterbarg model. To set up the formula of vanilla price for these extended models we employ different techniques including parameter averaging techniques, perturbation theory and vol of vol expansion. These formulas are approximation but with a good numerical precision.

In the last part of chapter 4, we list some calibration algorithms. These algorithms are global optimization in high dimension space. We gave some adjustment to increase the probability of convergence. Although the algorithms are model-free, they can be more helpful for the family of stochastic volatility model, since these models have more parameters.

Bibliography

Hansjorg Albrecher, Philipp Mayer, Wim Schoutens, and Jurgen Tistaert. The little heston trap. *Wilmott magazine*, 1:83–92, 2007. [cited at p. 49, 50]

L. Andersen and J. Andreasen. Jump diffusion process: Volatility smile fitting and numerical methods for pricing. *Review of Operations Research*, 4:231–262, 2000. [cited at p. 39]

L. Andersen and J. Andreasen. Volatile volatilities. *Risk*, pages 163–168, December 2002. [cited at p. 40]

L. B. G. Andersen and R. Brotherton-Ratcliffe. The equity option volatility smile: an implicit finite-difference approach. *Journal Computational Finance*, 1:5–37, 1998. [cited at p. 40]

E. Benhamou and O. Croissant. Local time for the sabr model: Connection with the 'complex' black scholes and application to cms and spread options. *SSRN*, 2007. [cited at p. 66]

Eric Benhamou, Emmanuel Gobet, and Mohammed Miri. Smart expansion and fast calibration for jump diffusion. September 2008. Available at SSRN: http://ssrn.com/abstract=1079627. [cited at p. 39]

Eric Benhamou, Emmanuel Gobet, and Mohammed Miri. Time dependent heston model. *SSRN eLibrary*, 2009. [cited at p. 76, 77, 78]

P. Bjerksund. Contingent claims evaluation when the convenience yield is stochastic: analytical results. *Working paper, Norwegian School of Economics and Business Administration*, 1991. [cited at p. 32]

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973. [cited at p. 22, 23, 79]

Michael Brennan. The price of convenience and the valuation of commodity contingent claims. *Diderik Lund and Bernt Oksendal, eds.: Stochastic Models and Option Values, North-Holland, Amsterdam*, 1991. [cited at p. 7, 8, 37]

M.J. Brennan. The supply of storage. *American Economic Review*, 48:50–72, 1958. [cited at p. 8]

J. Casassus and P. Collin-Dufresne. Stochastic convenience yield implied from commodity futures and interest rates. *Journal of Finance*, (60):2283–2331, 2005. [cited at p. 7, 8, 37]

A. Eydeland and H. Geman. Pricing power derivatives. *Risk*, September 1998. [cited at p. 38]

J. Gabillon. The term structure of oil futures prices. *Oxford Institute for Energy Studies, Working Paper*, (17), 1991. [cited at p. 7, 8, 11, 26, 37]

Helyette Geman. *Commodities and Commodity Derivatives Modelling and Pricing for Agriculturals Metals and Energy*. Wiley Finance, 2005. [cited at p. 9]

R. Gibson and E. S. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, XLV(3):959–976, 1990. [cited at p. 7, 8, 10, 11, 29, 37]

I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an ito differential. *Probability Theory and Related Fields*, 71:501–516, 1986. [cited at p. 18]

P. Hagan, D. Kumar, A. Lesniewski, and D. Woodward. Managing smile risk. *Wilmott Magazine*, pages 84–108, September 2002. [cited at p. 35, 37, 40, 64, 66]

S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2): 327–343, 1993. [cited at p. 38, 40, 41]

Jimmy Hilliard and Jorge Reis. Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions on the spot. *Journal of Financial and Quantitative Analysis*, 33(1):61–86, 1998. [cited at p. 7, 8, 37]

John Hull and Alan White. Pricing interest-rate derivative securities. *The Review of Financial Studies*, 3(4):573–592, 1990. [cited at p. 26, 38]

F. Jamshidian and M. Fein. Closed-form solutions for oil futures and european options in the gibson-schwartz model: A note. 1990. Working paper, Merrill Lynch Capital Markets. [cited at p. 32]

Alireza Javaheri. Inside volatility arbitrage : The secrets of skewness. September 2005. Wiley Finance. [cited at p. 35, 37]

C. Kahl and P. Jackel. Not-so-complex logarithms in the heston model. *Wilmott magazine*, pages 94–103, September 2005. [cited at p. 49]

N. Kaldor. Speculation and economic stability. *The Review of Economic Studies*, 7: 1–27, 1939. [cited at p. 8]

P. H. Labordere. General asymptotic implied volatility for sv models. 2005. Working paper; available at http://arxiv.org/abs/cond-mat/0504317v2. [cited at p. 40]

Delphine Lautier. Term structure models of commodity prices: a review. *The Journal of Alternative Investments*, pages 42–64, Summer 2005. [cited at p. 8]

Delphine Lautier. Convenience yield and commodity markets. *Bankers, Markets & Investors*, (102):59–66, September - October 2009. [cited at p. 7, 8, 9]

Alan L. Lewis. *Option Valuation Under Stochastic Volatility: With Mathematica Code*. Finance Press, February 2000. [cited at p. 40, 42, 55, 75]

S. Mikhailov and U. Nogel. Heston's stochastic volatility model. implementation, calibration and some extensions. *Wilmott magazine*, pages 74–79, July 2003a. [cited at p. 40, 55, 61]

Sergei Mikhailov and Ulrich Nogel. Heston's stochastic volatility model implementation, calibration and some extensions. *Wilmott magazine*, pages 74–79, July 2003b. [cited at p. 41, 50]

M. J. Nielsen and E. S. Schwartz. Theory of storage and the pricing of commodity claims. *Review of Derivatives Research*, 7:5–24, 2004. [cited at p. 38]

Yasufumi Osajima. The asymptotic expansion formula of implied volatility for dynamic sabr model and fx hybrid model. February 2007. Available at http://ssrn.com/abstract=965265. [cited at p. 40]

V. Piterbarg. A stochastic volatility forward libor model with a term structure of volatility smiles. *SSRN*, 2003. [cited at p. 53, 62]

V. Piterbarg. Time to smile. *Risk*, 19(5), May 2005. [cited at p. 39, 53, 56, 57, 62]

M. Richter and C. Sørensen. Stochastic volatility and seasonality in commodity futures and options: The case of soybeans. *Copenhagen Business School*, Working paper, 2002. [cited at p. 38]

E. S. Schwartz and J. E. Smith. Short-term variations and long-term dynamics of commodity prices. *Management Science*, 46:893–911, 2000. [cited at p. 7, 8, 11, 12, 15, 37]

Eduardo S. Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance*, 52(3):923, July 1997. [cited at p. 7, 8, 37]

L. Scott. Option pricing when the variance changes randomly: Theory, estimation and an application. *Journal of Financial and Quantitative Analysis*, 22(4):419–438, 1987. [cited at p. 38]

Steven Shreve. Stochastic calculus for finance ii: Continuous-time models. *Springer Finance*, 2004. [cited at p. 23]

Stein and Stein. Stock price distributions with stochastic volatility: an analytical approach. *Review Finance Study*, 4:727–752, 1991. [cited at p. 38]

L.G. Telser. Futures trading and the storage of cotton and wheat. *Journal of Political Economy*, 66:233–255, 1958. [cited at p. 8]

Anders B. Trolle and Eduardo S. Schwartz. Unspanned stochastic volatility and the pricing of commodity derivatives. *SSRN eLibrary*, January 2008. [cited at p. 38, 39]

Zaizhi Wang. Effective parameters for stochastic volatility models. *Journal of Financial Transformation*, pages 108–115, September 2009. Paris December 2007 Finance International Meeting AFFI-EUROFIDAI Paper; Available at SSRN: http://ssrn.com/abstract=1071768. [cited at p. 66, 67]

Mark D. West. Private ordering at the world's first futures exchange. (Dojima rice exchange in osaka, japan). *Michigan Law Review*, 98(8):2574–2615, August 2000. Symposium: Empirical Research in Commercial Transactions (Aug., 2000); Stable URL: http://www.jstor.org/stable/1290356. [cited at p. 2]

J. B. Wiggins. Option values under stochastic volatilities. *Journal of Financial Economics*, 19:351–372, 1987. [cited at p. 38]

H. Working. Theory of inverse carrying charge in futures markets. *Journal Farm Economics*, 30:1–28, 1949. [cited at p. 8]

F. Zhou. Black smirks. *Risk*, pages 87–91, May 2003. [cited at p. 40]

Appendices

Appendix A

Proof of propositions

A.1 Proof of proposition 25

To porve the proposition, we consider two diffusions more generally,

$$\begin{aligned} \mathrm{d}X(t) &= f(t, X(t))\sqrt{z(t)}\sigma(t)\mathrm{d}W(t), & X(0) &= x_0 \\ \mathrm{d}Y(t) &= \bar{f}(Y(t))\sqrt{z(t)}\sigma(t)\mathrm{d}W(t), & Y(0) &= x_0 \end{aligned}$$

Without loss of generality:

$$f(t, x_0) \equiv \bar{f}(x_0) = 1$$

The objective is to find the best function $\bar{f}(\cdot)$, which minimizes the "distance" between two diffusion $X(t)$ and $Y(t)$.

We note the objective function to minimize as:

$$p(X, Y) = \mathbb{E}\left[(X(T) - Y(T))^2\right]$$

Use Ito's lemma and then expand to the first order,

$$\begin{aligned} p(X, Y) &= \int_0^T \mathbb{E}\left[z(t)(f(t, X(t)) - \bar{f}(Y(t)))^2\right]\sigma^2(t)\mathrm{d}t \\ &= \int_0^T \mathbb{E}\left[z(t)\left(\frac{\partial}{\partial x}f(t, x_0)(X(t) - x_0) - \frac{\partial}{\partial x}\bar{f}(x_0)(Y(t) - x_0)\right)^2\right]\sigma^2(t)\mathrm{d}t \end{aligned}$$

Further approximation:

$$\begin{aligned} X(t) - x_0 &= X_0(t) - x_0 \\ Y(t) - x_0 &= X_0(t) - x_0 \end{aligned}$$

Where $X_0(t)$ is given by,

$$\mathrm{d}X_0(t) = \sqrt{z(t)}\sigma(t)\mathrm{d}W(t),\ X_0(0) = x_0 \tag{A.1}$$

Finally we get,

$$p(X,Y) = \int_0^T \left(\frac{\partial}{\partial x}f(t,x_0) - \frac{\partial}{\partial x}\bar{f}(x_0)\right)^2 \mathbb{E}\left[z(t)(X_0(t)-x_0)^2\right]\sigma^2(t)\mathrm{d}t$$

Minimizing $p(X,Y)$ leads to,

$$\frac{\partial \bar{f}(x_0)}{\partial x} = \int_0^T \frac{\partial f(t,x_0)}{\partial x} w(t)\mathrm{d}t \tag{A.2}$$

where the weights $w(t)$ can calculate by,

$$w(t) = \frac{v^2(t)\sigma^2(t)}{\int_0^T v^2(t)\sigma^2(t)\mathrm{d}t} \tag{A.3}$$

$$v^2(t) = \mathbb{E}\left[z(t)(X_0(t)-x_0)^2\right] \tag{A.4}$$

Now we apply the general result (A.2) to our diffusion,

$$\mathrm{d}S(t) = \sigma(t)(\beta(t)S(t) + (1-\beta(t))S(0))\sqrt{z(t)}\mathrm{d}U(t)$$

And we can get,

$$b = \int_0^T \beta(t)w(t)\mathrm{d}t$$

$$w(t) = \frac{v^2(t)\sigma^2(t)}{\int_0^T v^2(t)\sigma^2(t)\mathrm{d}t}$$

$$v^2(t) = \mathbb{E}\left[z(t)(X_0(t)-x_0)^2\right]$$

Remark In this part of proof, we don't use the zero correlation condition. Therefore, the result (A.2) can also apply to a model with non-zero correlation.

Now we calculate $v^2(t)$. It needs the assumption of zero correlation.

Using the independence between X_0 and z we obtain,

$$v^2(t) = \mathbb{E}\left[z(t)(X_0(t)-x_0)^2\right] \tag{A.5}$$

$$= \mathbb{E}\left[z(t)\mathbb{E}\left[(X_0(t)-x_0)^2\middle|z(\cdot)\right]\right] \tag{A.6}$$

$$= \mathbb{E}\left[z(t)\int_0^T z(s)\sigma^2(s)\mathrm{d}s\right] \tag{A.7}$$

$$= \int_0^T \sigma^2(s)\mathbb{E}[z(t)z(s)]\mathrm{d}s \tag{A.8}$$

Calculate the covariance of $z(t)$ and substitute in equation (A.8), we can get,

$$v^2(t) = z_0^2 \int_0^t \sigma^2(s)\mathrm{d}s + z_0\eta^2 e^{-\theta t} \int_0^t \sigma^2(s)\frac{e^{\theta s} - e^{-\theta s}}{2\theta}\mathrm{d}s$$

This proves proposition 25. □

A.2 Proof of proposition 26

We can first remark that the model given by:

$$\mathrm{d}S(t) = \sigma(t)(bS(t) + (1-b)S(0))\sqrt{z(t)}\mathrm{d}U(t)$$

where the stochastic volatility $z(t)$follows

$$\mathrm{d}\bar{z}(t) = \theta(z_0 - \bar{z}(t))\mathrm{d}t + \eta\sqrt{\bar{z}(t)}\mathrm{d}V(t), z(0) = z_0$$

has closed formula. We can use Fourier transform as in Lewis (2000), with the non-constant sigma directly. In this case, instead of solving a Riccati system by Runge-Kutta method, we have to solve a Riccati system in the fundamental transform to construct the closed formula.

Consider the European style at-the-money call option.

$$\mathbb{E}(S(T) - S_0)^+ = \mathbb{E}\left[\mathbb{E}\left[(S(T) - S_0)^+\right]\middle| z(\cdot)\right] \tag{A.9}$$

Where $(S(T) - S_0)^+ = \begin{cases} S(T) - S_0, \text{ when } S(T) \geq S_0 \\ 0, \text{ when } S(T) < S_0 \end{cases}$

That's to say, we can express$(S(T) - S_0)^+$ as a function of$Z(T)$, saying, in form of $g(Z(T))$. And

$$Z(T) = \int_0^T \sigma^2(t)z(t)\mathrm{d}t \tag{A.10}$$

Now we derive the explicit form of function $g(\cdot)$

The process $S(t)$is a shifted lognormal.

$$\begin{aligned}
\mathrm{d}S(t) &= \sigma(t)(bS(t)+(1-b)S_0)\sqrt{z(t)}\mathrm{d}U(t), S(0)=S_0 \\
\mathrm{d}S(t) &= \sigma(t)b\left(S(t)+\frac{(1-b)}{b}S_0\right)\sqrt{z(t)}\mathrm{d}U(t) \\
\mathrm{d}\left(S(t)+\frac{1-b}{b}S_0\right) &= \sigma(t)b\left(S(t)+\frac{(1-b)}{b}S_0\right)\sqrt{z(t)}\mathrm{d}U(t) \\
S(t)+\frac{1-b}{b}S_0 &= \exp\left\{\int_0^t \sigma(t)b\sqrt{z(t)}\mathrm{d}U(t)+\log\left(S_0+\frac{1-b}{b}S_0\right)\right\} \\
&= \frac{S_0}{b}\exp\int_0^t \sigma(t)b\sqrt{z(t)}\mathrm{d}U(t) \\
S(T)-S_0 &= \frac{S_0}{b}\left(\exp\int_0^t \sigma(t)b\sqrt{z(t)}\mathrm{d}U(t)-1\right) \\
(S(T)-S_0)^+ &= \frac{S_0}{b}\left(\exp\int_0^t \sigma(t)b\sqrt{z(t)}\mathrm{d}U(t)-1\right)^+
\end{aligned}$$

The right-hand side of the equation $(\exp\int_0^t \sigma(t)b\sqrt{z(t)}\mathrm{d}U(t)-1)^+$can be regarded as Black-Scholes model, with the spot price being 1, strike price being 1 and interest rate being 0.

Therefore,

$$(S(T)-S_0)^+ = \frac{S_0}{b}\left(N(d_1)-N(d_2)\right)$$

With

$$\begin{aligned}
d_1 &= \frac{\log(S_0/K)+(r+(vol)^2/2)T}{\sigma\cdot\sqrt{T}} = \frac{\sigma\cdot\sqrt{T}}{2} = \frac{1}{2}b\sqrt{Z(T)} \\
d_2 &= d_1-\sigma\sqrt{T} = -\frac{1}{2}b\sqrt{Z(T)}
\end{aligned} \tag{A.11}$$

$N(y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-x^2/2}\mathrm{d}x$ is cumulative function of standard normal distribution. $Z(T)$is defined in A.10.

$$\begin{aligned}
c(S(T)-S_0)^+ &= \frac{S_0}{b}\left(N\left(\frac{1}{2}b\sqrt{Z(T)}\right)-N\left(-\frac{1}{2}b\sqrt{Z(T)}\right)\right) \\
&= \frac{S_0}{b}\left(2N\left(\frac{1}{2}b\sqrt{Z(T)}\right)-1\right)
\end{aligned}$$

We note $g(x)=\frac{S_0}{b}\left(2N\left(\frac{1}{2}b\sqrt{x}\right)-1\right)$

So we get an equation for λ:

$$\mathbb{E}g\left(\int_0^T \sigma^2(t)z(t)\mathrm{d}t\right) = \mathbb{E}g\left(\lambda^2\int_0^T z(t)\mathrm{d}t\right) \tag{A.12}$$

Now we make an approximation for $g(x)$

We have obtained $g(x) = \frac{S_0}{b}\left(2N\left(\frac{b\sqrt{x}}{2}\right) - 1\right)$

Here, $N(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-x^2/2}\mathrm{d}x$ is cumulative function of standard normal distribution

Note that, $N\left(\frac{b\sqrt{x}}{2}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{b\sqrt{x}}{2}} e^{-t^2/2}\mathrm{d}t$

The derivative respect to x is:

$$\left(N\left(\frac{b\sqrt{x}}{2}\right)\right)' = \frac{1}{\sqrt{2\pi}} e^{-\frac{b^2 x}{8}}$$

So, we can approximate $N\left(\frac{b\sqrt{x}}{2}\right)$ with $p + qe^{-rx}$

That's to say,

$$g(x) \approx a + be^{-cx} \tag{A.13}$$

a, b and $c$are constants, which are to be determined.

Substitute (A.13) into (A.12), we get:

$$a + b\mathbb{E}\exp\left(-c\int_0^T \sigma^2(t)z(t)\mathrm{d}t\right) = a + b\mathbb{E}\exp\left(-c\lambda^2\int_0^T z(t)\mathrm{d}t\right)$$

$$\mathbb{E}\exp\left(-c\int_0^T \sigma^2(t)z(t)\mathrm{d}t\right) = \mathbb{E}\exp\left(-c\lambda^2\int_0^T z(t)\mathrm{d}t\right) \tag{A.14}$$

And the constant $c$can be given by:

$$c = -\frac{g''(\zeta)}{g'(\zeta)},$$

$$\zeta = \mathbb{E}Z(t) = \int_0^T \sigma^2(t)\mathbb{E}z(t)\mathrm{d}t = z_0\int_0^T \sigma^2(t)\mathrm{d}t$$

We note the Laplace transform of $Z(T)$ as:

$$\varphi(\mu) = \mathbb{E}\exp(-\mu Z(T))$$

And the Laplace transform of $z(t)$ as:

$$\varphi_0(\mu) = \mathbb{E}\exp\left(-\mu\int_0^T z(t)\mathrm{d}t\right)$$

Substitute the transforms above into (A.14), we get the equation, from which we can solve the parameterλ.

$$\varphi_0\left(-\frac{g''(\zeta)}{g'(\zeta)}\lambda^2\right) = \varphi\left(-\frac{g''(\zeta)}{g'(\zeta)}\right) \tag{A.15}$$

The Laplace transform can be presented as:

$$\varphi(\mu) = \exp(A(0,T) - z_0 B(0,T))$$

$A(t,T)$ and $B(t,T)$ suffice the Riccati system of ordinary differential equations:

$$A'(t,T) - \theta z_0 B(t,T) = 0 \tag{A.16}$$

$$B'(t) - \theta B(t,T) - \frac{1}{2}\eta^2 B^2(t,T) + \mu\sigma^2(t) = 0 \tag{A.17}$$

$$A(T,T) = B(T,T) = 0$$

When$\sigma(t) \equiv 1$, we arrive from $\varphi(\mu)$ to $\varphi_0(\mu)$. This case can be solved explicitly.

$$\begin{aligned}
\varphi_0(\mu) &= \exp(A_0(0,T) - z_0 B_0(0,T)) \\
B_0(0,T) &= \frac{2\mu(1-e^{-\gamma T})}{(\theta+\gamma)(1-e^{-\gamma T}) + 2\gamma e^{-\gamma T}} \\
A_0(0,T) &= \frac{2\theta z_0}{\eta^2}\log\left(\frac{2\gamma}{\theta+\gamma(1-e^{-\gamma T}) + 2\gamma e^{-\gamma T}}\right) - 2\theta\frac{\mu}{\theta+\gamma}T
\end{aligned}$$

Here,$\gamma = \sqrt{\theta^2 + 2\eta^2\mu}$

To solve (A.15), we use Newton's root searching method. We use recursion to get the solution of$f(x) = 0$.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Because the scope of λ^2 is relatively small, saying [0,2] typically, Newton's method works very fast.

(A.16) and (A.17) are called Riccati system. When $\sigma(t)$ is not constant, Riccati system doesn't have explicit solution. In this case, we employ Runge-Kutta's method to solve it numerically which concludes the result.□

A.3 Proof of proposition 28

According to the definition in equation (A.4)

$$v^2(t) = \mathbb{E}\left[z(t)\left(\int_0^t \sqrt{z(u)}\sigma(u)\mathrm{d}W(u)\right)^2\right]$$

We note $Y(t) = \int_0^t \sqrt{z(u)}\sigma(u)\mathrm{d}W(u)$ and $N(t) = Y(t)^2$.
Apparently, $N(0) = Y(0) = 0$.

$$\mathbb{E}[Y(t)] = 0$$

$$\mathbb{E}[N(t)] = z_0\int_0^t \sigma(s)^2\mathrm{d}s$$

$$\mathrm{d}N(t) = 2Y(t)\mathrm{d}Y(t) + \mathrm{d}\langle Y\rangle_t$$

$$dz(t) = \theta(t)(z_0 - z(s))ds + \gamma(s)\sqrt{z(s)}dV(s)$$

with $z(0) = z_0$

$$\mathbb{E}\left[z(s)^2\right] = z_0^2 + z_0 e^{-2\theta(s)s}\int_0^s e^{2\theta(u)u}\gamma(u)^2 du$$

Here we use the integration by parts ($X_t Y_t = X_0 Y_0 + \int_0^t X_{s-}\,dY_s + \int_0^t Y_{s-}\,dX_s + \langle X, Y\rangle_t$)

$$v^2(t) = \mathbb{E}[z(t)N(t)] \tag{A.18}$$

$$= \mathbb{E}\left[\int_0^t z(s)dN(s) + \int_0^t N(s)dz(s) + \langle z, N\rangle_t\right] \tag{A.19}$$

The first term of equation A.19:

$$\begin{aligned}\mathbb{E}\left[\int_0^t z(s)dN(s)\right] &= \mathbb{E}\left[\int_0^t z(s)\cdot 2Y(s)dY(s) + \int_0^t z(s)d\langle Y\rangle_s\right]\\ &= \mathbb{E}\left[\int_0^t z(s)d\langle Y\rangle_s\right]\\ &= \mathbb{E}\left[\int_0^t z(s)^2\sigma(s)^2 ds\right]\\ &= \int_0^t \mathbb{E}\left[z(s)^2\right]\sigma(s)^2 ds\end{aligned}$$

The second term of equation A.19:

$$\begin{aligned}\mathbb{E}\left[\int_0^t N(s)dz(s)\right] &= \mathbb{E}\left[\int_0^t N(s)\theta(s)\left(z(0) - z(s)\right)ds + \int_0^t N(s)\gamma(s)\sqrt{z(s)}dV(s)\right]\\ &= z_0\int_0^t \mathbb{E}[N(s)]\theta(s)ds - \int_0^t \mathbb{E}[N(s)z(s)]\theta(s)ds\end{aligned}$$

The third term of equation A.19:

$$\begin{aligned}\mathbb{E}[\langle z, N\rangle_t] &= \mathbb{E}\left[\int_0^t \gamma(s)z(s)\cdot 4Y(s)^2\sigma(s)\rho(s)ds\right]\\ &= 4\int_0^t \gamma(s)\sigma(s)\rho(s)\mathbb{E}\left[z(s)N(s)\right]ds\end{aligned}$$

Finally we get,

$$\mathbb{E}[z(t)N(t)] = A(t) + \int_0^t B(s)\mathbb{E}[z(s)N(s)]ds \tag{A.20}$$

Or,

$$v^2(t) = A(t) + \int_0^t B(s)v^2(s)\mathrm{d}s$$

$$\frac{\mathrm{d}v^2(t)}{\mathrm{d}t} = A'(t) + B(t)v^2(t) \tag{A.21}$$

with

$$\begin{aligned} A(t) &= \int_0^t \mathbb{E}\left[z(s)^2\right]\sigma(s)^2\mathrm{d}s + z_0\int_0^t \mathbb{E}[N(s)]\theta(s)\mathrm{d}s \\ B(s) &= 4\gamma(s)\sigma(s)\rho(s) - \theta(s) \\ v^2(0) &= 0 \end{aligned}$$

Note that the ordinary differential equation (ODE) A.21 is easy to solve if all the parameters are constant. The solution in this case is,

$$v^2(t) = e^{\int B(t)\mathrm{d}t}\left(\int A'(t)e^{-\int B(t)dt}\mathrm{d}t + K\right) \tag{A.22}$$

K is the constant of integration,

$$K = -\int A'(t)e^{-\int B(t)\mathrm{d}t}\mathrm{d}t\Big|_{t=0}$$

Now we consider all the parameters of the model is piecewise constant. The ODE A.21 can be solved in each interval. And the initial value for each interval is the last value for the last interval.

Suppose we have n periods: $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$.

When $t_{i-1} < t < t_i$, we note $\sigma(t) = \sigma_i$, $\gamma(t) = \gamma_i$, $\rho(t) = \rho_i$, $\theta(t) = \theta_i$,$\beta(t) = \beta_i$, $B(t) = B_i = 4\gamma_i\sigma_i\rho_i - \theta_i$.

Now solve ODE A.21 on interval $[t_{i-1}, t_i]$, with the initial value $v_{i-1}^2 = v^2(t_{i-1})$. We obtain the solution for $t \in (t_{i-1}, t_i]$,

$$v^2(t) = e^{\int B(t)\mathrm{d}t}\left(\int A'(t)e^{-\int B(t)\mathrm{d}t}\mathrm{d}t + K\right) \tag{A.23}$$

$$K = v_{i-1}^2 e^{-\int B(t)\mathrm{d}t} - \int A'(t)e^{-\int B(t)\mathrm{d}t}\mathrm{d}t\Big|_{t=t_{i-1}} \tag{A.24}$$

For the first period, initial value $v_0^2 = v^2(0) = 0$.

Now we substitute the initial value and develop equation A.23. It yields,

$$\begin{aligned} v^2(t) =& v_{i-1}^2 e^{B_i(t+1)} + \left(e^{B_i(t-t_{i-1})} - 1\right)\sigma_i^2 z_0 \frac{1}{B_i}\left(z_0 + \frac{1}{2\theta_i}\gamma_i^2\right) \\ &+ z_0^2\sigma_i^2\theta_i \frac{1}{B_i}\left(e^{B_i(t-t_{i-1})} B_i t_{i-1} + e^{B_i(t-t_{i-1})} - B_i t - 1\right) \\ &- \frac{1}{2\theta_i(B_i + 2\theta_i)}\sigma_i^2 z_0 \gamma_i^2 \left(e^{B_i(t-t_{i-1}) - 2\theta_i t_{i-1}} - e^{-2\theta_i t}\right) \end{aligned}$$

with $B_i = 4\gamma_i\sigma_i\rho_i - \theta_i$.

It proves proposition 28. □

List of Symbols and Abbreviations

Abbreviation	Description	Definition
ADC	average daily commissions	
ADV	average daily volume	
BS model	Black Scholes model	
BS implied vol	implied Black Scholes volatility	
$D(t,T)$	discount factor	It equals to the value of 1 unit of risk free bond at time t with maturity at time T.
$F_{t,T}$	forward price	The forward price at time t with maturity T.
ICE	Intercontinental Exchange	
LV model	local volatility model	
OTC	over-the-counter	
OU process	Ornstein Uhlenbeck process	
S_t	spot price	Spot price at time t.
SV model	stochastic volatility model	
vol of vol	volatility of volatility	

List of Figures

List of Tables

Index

Produits Dérivés des Matières Premières:

Modélisation et Evaluation

RESUME : Les prix des matières premières ont augmenté à un rythme sans précédent au cours des dernières années rendant les dérivés sur matières premières de plus en plus populaires dans de nombreux secteurs comme l'énergie, les métaux et les produits agricoles. Le développement rapide du marché des produits dérivés sur matières premières a aussi induit une recherche vers toujours plus de précision et cohérence dans la modélisation et l'évaluation de produits dérivés des matières premières.

Reposant sur le principe de diffusion équivalente introduite par Gyöngy, nous montrons que le modèle de Gibson Schwartz et le modèle de Gabillon peuvent se réduire à un modèle à un facteur dont la distribution marginale peut être explicitement calculée sous certaines conditions. Ceci nous permet en particulier de trouver des formules analytiques pour l'ensemble des options vanilles. Certaines de ces formules sont nouvelles à notre connaissance et d'autres confirment des résultats antérieurs.

Dans cette thèse, nous étendons les idées de Piterbarg à la famille des modèles à volatilité stochastique en rendant le concept plus général. Nous montrons en particulier comment introduire des paramètres dépendant du temps dans les modèles à volatilité stochastique et explicitons différentes formules de calcul explicite du prix d'options vanilles, permettant ainsi une calibration des paramètres du modèles extrêmement efficace.

Mots clés : Produits Dérivés des Matières Premières, modèle de Gibson Schwartz, modèle de Gabillon, Théorème de Gyöngy, modèle à volatilité stochastique, parametres dépend du temp.

Commodity Derivatives:

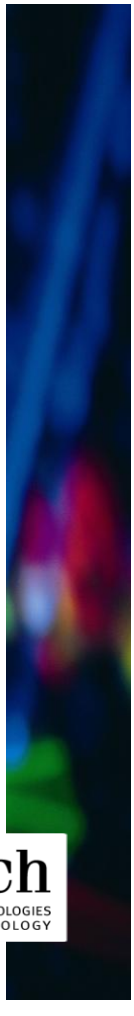
Modeling and Pricing

ABSTRACT : Commodity prices have been rising at an unprecedented pace over the last years making commodity derivatives more and more popular in many sectors like energy, metals and agricultural products. The quick development of commodity market as well as commodity derivative market results in a continuously uprising demand of accuracy and consistency in commodity derivative modeling and pricing.

In this thesis, we prove that there is mathematical equivalence relation between Gibson Schwartz model and Gabillon model. Moreover, inspired by the idea of Gyöngy, we show that Gibson Schwartz model and Gabillon model can reduce to one-factor model with explicitly calculated marginal distribution under certain conditions, which contributes to find the analytic formulas for forward and vanilla options. Some of these formulas are new to our knowledge and other formulas confirm with the earlier results of other researchers.

We extend Piterbarg's idea to the whole family of stochastic volatility model, making all the stochastic volatility models having time-dependent parameters and show various formulas for vanilla option price by employing various techniques such as characteristic function, Fourier transform, small error perturbation, parameter averaging, etc.

Keywords : Commodity derivative, Gibson-Schwartz model, Gabillon model, theory of Gyöngy, Stochastic volatility model, time-dependent parameter.



MINES ParisTech

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY