7-11-2018 11:00 AM

Stochastic modelling of implied correlation index and herd behavior index. Evidence, properties and pricing.

Lin Fang, *The University of Western Ontario*

Supervisor: Escobar, Marcos, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Statistics and Actuarial Sciences

Abstract

In this work, we provide the definition, study properties, and craft new stochastic models for two dependence indices: the implied correlation index and the herd behavior index (HIX). In particular, we model and price financial derivatives on the basic implied correlation index (CIX) as reported by CBOE. Our analysis is the first revealing the presence of heteroscedasticity in the time series of CIX leading to two Correlation Stochastic Volatility (CSV) models. We describe properties of CSV models and use discretization methods for their simulation. A partial estimation methodology is implemented on CBOE S& P 500 CIX historical data treating the stochastic volatility (SV) as a hidden, unobservable process. The impact of the SV parameters is studied for two types of digital CIX options, both motivated by the usage of CIX as an indicator of crisis conditions.

Keywords: Implied correlation index, stochastic volatility models

Contents

List of Figures

List of Tables

Chapter 1

Introduction

An interesting phenomenon in financial markets, confirmed by the Great Recession of 2008, is that numerous stocks crash simultaneously during a crisis. If we consider stock prices as a vector of lognormal random variable, then we see that these random variables are highly positively dependent during a crisis period. A professional portfolio manager will prudently select financial assets to avoid strong positive dependence structures among his portfolio components because they could jeopardize the benefits of diversification, particularly during a crisis. A strong positive dependence structure is also called a comonotonic dependence structure, and was first proposed by [Dhaene and Vyncke, 2002b] and was used to study the distribution function of financial indices. Due to the importance of correlation in portfolio diversification, we give a brief overview of the concept of modern portfolio theory as well as the optimal objective function in a risk minimization problem in Section 1.1. We also summarize some important definitions and properties of comonotonicity in Section 1.2 by referring to [Dhaene and Vyncke, 2002b]. For extensive financial applications of comonotonicity, readers can refer to [Dhaene and Vyncke, 2002a].

The connection between market crisis and assets' comonotonic dependence structures explains that, if one can measure the degree of assets' dependence and model the process of dependence structure, then one would be relatively safe in crisis periods. To be specific, investors can hedge risk during crises by buying some digital options. These digital options use the degree of dependence as the underlying and will give the holder a benefit (i.e 10 million dollars) in a crisis. When the degree of dependence is greater than its danger threshold, investors shall execute the option to receive the benefit. Although investors have to pay a premium for the option contract, they can reduce loss during a crisis by receiving 10 million dollars benefit. In this thesis, we design two kinds of digital options for this purpose. But back to where we were, how can we measure the degree of dependence properly? If we want to measure the degree of dependence between two financial assets, the natural way is to compute their correlation coefficient. As a proxy to the traditional correlation coefficient, the basic implied correlation index (CIX) was proposed by [van Emmerich, 2006] to measure the degree of comovement between two market index components. CIX is defined as the ratio of the sum of the weighted covariance to that of the weighted variance between underlying assets. CBOE calculates and daily publishes the S&P 500 Implied Correlation Index. This index has become a measure of the market's expectation about the future correlation of the S& P 500 index (SPX) components, see [Exchange, 2009]. At different year, CBOE uses symbols KCJ,ICJ and JCJ to name a se-

ries of CIX in rotation. The calculation of KCJ, ICJ and JCJ are same, all of them use SPX option prices and prices of single-stock option on the 50 largest components of the SPX. The number behind symbols indicates the expired year of single-stock option. For example, the calculation of ICJ2010 uses SPX options expiring in December 2009 and single-stock options expiring in January 2010 (see [Exchange, 2009]). In [van Emmerich, 2006]'s theory, the calculation of CIX in discrete time relies on the at-the-money index option prices. [Linders and Schoutens, 2014] introduced a general implied correlation index (ICX). Compared with CIX, the methodology of ICX does not assume that the moneyness of index option equals 1. [Linders and Schoutens, 2014] regarded [van Emmerich, 2006]'s method as a traditional approach to determine correlation and stated that this approach would underestimate the correlation for out-of-the-money index option prices. Instead, [Linders and Schoutens, 2014] proposed a robust measurement model to determine ICX where ICX minimizes the distance between the theoretical index option price and the observed option price.

A comonotonic dependence structure also causes herd behaviour. In finance, herd behaviour is a relatively new concept that is used to describe how individual assets can act collectively. Herd behaviour also indicates the systemic risk in the market. Researchers have used herd behaviour to explain financial bubbles and crashes. For example, the 2008 US housing bubble was caused by herd behaviour and human greed, see [Fenzl and Pelzmann, 2012]. When the bubble burst, individuals were driven by panic to mimic others in an irrational way, which led to a dramatic price drop in the market. By using the concept of comonotonicity, [Dhaene and Vyncke, 2012] proposed the herd behaviour index (HIX) to measure the degree of herd behaviour among financial assets. Their idea behind HIX is that the market's expectation on the degree of herd behaviour in the future should be based on a comparison between the actual dependence structure and the comonotonic dependence structure among future stock prices. Thus, HIX in [Dhaene and Vyncke, 2012] is computed as the ratio of an option-based estimate of the risk-neutral variance of the market index to an option-based estimate of the variance of the comonotonic index.

In this thesis, we focus on modeling CIX and HIX. Since the degree of comovement and herd behaviour may change quickly over time in a random manner, we model CIX and HIX using a continuous-time stochastic process. CIX has the mathematical properties of a correlation thus we have a natural connection to stochastic correlation processes (SCP). The study of SCP has attracted attention in mathematical finance but mostly from the perspective of stochastic covariance models, see [Engle, 2002] and [Gouriéroux, 2006]. The works of [van Emmerich, 2006] and [Ma, 2009] are two of a handful that treats correlation as stand-alone processes in continuous time.

The main challenge in working with the stochastic differential equation for correlation is the bounded $[-1, 1]$ domain for the process itself. This bounded domain affects closed-form analytical solution to key properties like conditional characteristic functions and option prices due to the non-affine nature of the process, but this also impacts the accuracy of the discretization as well as simulation exercises. [van Emmerich, 2006] presented an approach to construct SCP models where the new stochastic process is directly formulated as a convenient function of Brownian motion. In a similar fashion [Teng et al., 2016] developed a more general SPC by modeling correlation as a hyperbolic function of the popular mean-reverting process. Their correlation process also satisfies the properties provided in [van Emmerich, 2006]. Although this

approach benefits from the ease of construction and a better degree of analytical tractability, it lacks intuitive interpretation. Therefore, [van Emmerich, 2006] introduced a second approach where the SPC is described by a stochastic differential equation driven by Brownian motion, a particular case of so-called Jacobi processes. Under this direction, a modified Jacobi process was proposed in [Ma, 2009], and here bounds for the correlation coefficient are more flexible than -1 and 1.

All these models ensure a key stylized fact of stochastic correlations: the mean-reverting property. However, the volatility in these SCP is assumed constant. In this thesis, we provide empirical evidence of heteroscedasticity in the time series of the CBOE S&P 500 Implied Correlation Index (CBOE S&P 500 CIX). Therefore we build two Correlation Stochastic Volatility (CSV) models, which can better describe the evolution of CIX data. One model uses flexible functions of Ito's processes and the other goes along the lines of a Jacobi process. In both models, we assume that the volatility of CIX follows a mean-reverting process, the CIR process for simplicity. Apart from [van Emmerich, 2006]'s and [Ma, 2009]'s approaches, [Da Fonseca et al., 2007] introduced the Wishart Affine Stochastic Correlation (WASC) model, which is a continuous-time process for stock prices with stochastic covariance. Based on [Da Fonseca et al., 2007]'s work, we compute the stochastic correlation process implied by the Wishart in the two asset case. We also display the instantaneous covariance. However, it is not possible to derive a closed-form, non-SDE representation for the correlation process.

The previous modeling focused on the CIX as this takes values in $[-1, 1]$. On the other hand, HIX takes values within the interval $[0, 1]$, where 1 indicates significant herd behavior and no diversification possible. Inspired by [van Emmerich, 2006]'s work, [Dhaene and Vyncke, 2012] modeled HIX as a combination of a mean-reverting process X_t with an algebraic function which can map the domain of definition of X to the unit interval. This method preserves the mean-reverting property while satisfying the fundamental properties of HIX. In this thesis, we present four specific HIX models based on the CIR and the Vasicek process.

In this thesis, we also study the impact of the new CSV models on pricing of two types of derivatives written on the CBOE S& P 500 CIX. Derivatives for correlation processes has been of interest over the past few years, see [Salvi and Swishchuk, 2014] and [Bossu, 2005]. Although traded over-the-counter, derivatives on CIX are not yet popular. We study and propose two types of digital CIX options. Assuming that only high rating insurance companies can issue these options, our options shall help investors hedge risk coming from future crisis in the market. The first digital CIX option uses the value of CIX at maturity while the second uses the maximum value of the CIX index over the life of the option. If by maturity the underlying CIX crosses a dangerously high threshold, indicative of a crisis at play, then investors will execute the first digital option to obtain nonzero benefit. For the second digital option, investors will execute it if and only if the maximum value of CIX exceeds the dangerous threshold over the contract period. Since the second digital CIX option keeps track of the CIX performance, it offers a higher degree of safety but it is more expensive than the first one. In addition, if CIX stays safe all the time, both options will not be executed, and investors will only lose their initial premium which is the total cost of the option.

The main contributions of our thesis are:

- We summarize stochastic models for the implied correlation index and herd behavior index, see Sections 4.1, 4.2 and 4.5.

- We detect Heteroscedastic behavior on a series of the CBOE S& P 500 Implied Correlation Index, see Section 5.2.

- We introduce two choices of Correlation Stochastic Volatility (CSV) models for CIX. The volatility of CIX follows a Cox-Ingersoll-Ross (CIR) process, see Section 4.3.

- We perform the discretization and partial estimation of our CSV models, see Sections 4.3 and 5.2.2.

- We price two digital CIX options and perform a sensitivity analysis. For each option, our analysis shows that both κ (the speed reversion of SV parameter) and ς (the volatility of SV parameter) significantly affect prices for the first CSV model but have minor influence on prices for the second CSV model, see Section 5.2.3.

This thesis is organized as follows, Chapter 1 provides an overview of the thesis with basic knowledge of modern portfolio theory, comonotonicity and the Multivariate Black& Scholes model. Chapter 2 introduces the basic implied correlation index (CIX), two specific types of CIX (JCJ and ICJ) and the general Implied correlation index (ICX). Chapter 3 expands the concept of implied correlation index into the herd behavior index (HIX). In addition, we discuss RHIX in a portfolio of two risky assets at the end of Chapter 3. Chapter 4 focuses on the stochastic modeling of CIX and HIX. For CIX, we first summarize two approaches for the stochastic modeling of these bounded objects, then we introduce two choices of CSV models, together with an algorithm for discretization and some properties of the models. We also adopt the dynamic of correlation in a Wishart Affine Stochastic Correlation (WASC) model as an alternative approach. At the end of Chapter 4, we detail various stochastic models for HIX with different mapping functions. Chapter 5 covers applications of the models to the CIX. We present covariance and correlation swap for two risky asset case and reveal relevant empirical aspects of CIX data. In particular, we perform an ARCH test with the CBOE 500 CIX data in Section 5.2.1 and estimate parameters of CSV models in Section 5.2.2. Finally, we price two digital CIX options and perform a sensitivity analysis to reveal the impact of parameters on option pricing in Section 5.2.3.

1.1 Modern portfolio theory

In 1952, Markowitz published his paper on portfolio selection providing the foundation for modern portfolio theory as a mathematical problem. From the single period case, we can see the importance of correlation parameter in portfolio optimization. Suppose we purchase an asset with S_0 dollars at time $t = 0$ and then sell it for S_1 dollars at time $t = 1$. Then the return on this asset is defined as

$$R_i = \frac{S_{1,i} - S_{0,i}}{S_{0,i}} \tag{1.1}$$

If we consider a portfolio consisting of M risky assets, we have single-period returns $R = [R_1, R_2, ..., R_M]^\top$ with following mean and variance

$$E[R] = \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_M \end{bmatrix} \tag{1.2}$$

$$Cov[R] = \Sigma = \begin{bmatrix} \Sigma_{1,1} & ... & \Sigma_{1,M} \\ \vdots & \ddots & \vdots \\ \Sigma_{M,1} & ... & \Sigma_{M,M} \end{bmatrix} \tag{1.3}$$

With initial budget α_0 (dollars), the amount that we wish to assign to asset i is defined by $w_i\alpha_0$, where w_i is a weighting factor for asset i, $i = 1, 2, ..., M$. The negative weight indicates a short position in the portfolio. To preserve the budget constraint, weights are required to satisfy the condition that $\sum_{i=1}^{M} w_i = 1$. If we use w, a m-vector of weights, to represent the proportion of total wealth allocated in each asset, the portfolio return $R_w = w^\top R = \sum_{i=1}^{m} w_i R_i$ is deemed as a random variable with:

$$\begin{aligned} \text{Mean } \mu_w &= E[R_w] = w^\top \mu \\ \text{Variance } \sigma_w^2 &= Var[R_w] = w^\top \Sigma w \end{aligned} \tag{1.4}$$

By using the mean-variance pair of the portfolio (μ_w, σ_w^2) in (1.4), we can evaluate different portfolios Π^w with preferences for higher expected returns μ_w and lower variance σ_w^2. In fact, we have three kinds of problems- risk minimization, expected return maximization and risk aversion optimization. In order to highlight the importance of correlation and as an example, we focus in the problem of risk minimization. That is, for a given choice of target mean return α_1, we choose the portfolio Π^w that can satisfy the following conditions:

$$\begin{aligned} \text{Minimize :} \quad & \tfrac{1}{2} w^\top \Sigma w \\ \text{Subject to :} \quad & w^\top \mu = \alpha_1 \\ & w^\top 1_m = 1 \end{aligned} \tag{1.5}$$

where 1_m denoted the m-vector of ones. Here we set $\frac{1}{2} w^\top \Sigma w$ to make the first derivative of Lagrangian concise. The solution of (1.5) can be achieved by applying the method of Lagrange multipliers to the convex optimization problem subject to linear constraints. In details, let us define the Lagrangian first:

$$L(w, \lambda_1, \lambda_2) := \frac{1}{2} w^\top \Sigma w + \lambda_1(\alpha_1 - w^\top \mu) + \lambda_2(1 - w^\top 1_m) \tag{1.6}$$

Then, we derive the first-order conditions:

$$\begin{array}{lll} \frac{\partial L}{\partial w} & = 0_m & = \Sigma w - \lambda_1 \mu - \lambda_2 1_m \\ \frac{\partial L}{\partial \lambda_1} & = 0 & = \alpha_1 - w^\top \mu \\ \frac{\partial L}{\partial \lambda_2} & = 0 & = 1 - w^\top 1_m \end{array} \tag{1.7}$$

Note that w can be solved as a function of λ_1, λ_2:

$$w^* = \lambda_1 \Sigma^{-1} \mu + \lambda_2 \Sigma^{-1} 1_m \tag{1.8}$$

By substituting for w, we solve for λ_1, λ_2 :

$$\begin{aligned} \alpha_1 &= w^{*\top}\mu = \lambda_1(\mu^\top\Sigma^{-1}\mu) + \lambda_2(\mu^\top\Sigma^{-1}1_m) \\ 1 &= w^{*\top}1_m = \lambda_1(\mu^\top\Sigma^{-1}1_m) + \lambda_2(1_m^\top\Sigma^{-1}1_m) \end{aligned} \tag{1.9}$$

After using matrix, we have

$$\begin{bmatrix}\lambda_1 \\ \lambda_2\end{bmatrix} = \begin{bmatrix} a & b \\ b & c\end{bmatrix}^{-1}\begin{bmatrix}\alpha_1 \\ 1\end{bmatrix} \tag{1.10}$$

where $a = (\mu^\top\Sigma^{-1}\mu), b = (\mu^\top\Sigma^{-1}1_m)$ and $c = (1_m^\top\Sigma^{-1}1_m)$.

Hence, with weight w^*, our portfolio can achieve the target mean return α_1 with minimum variance σ^{*2}, where

$$\begin{aligned} w^* &= [\Sigma^{-1}\mu, \Sigma^{-1}1_m]\begin{bmatrix}(\mu^\top\Sigma^{-1}\mu) & (\mu^\top\Sigma^{-1}1_m) \\ (\mu^\top\Sigma^{-1}1_m) & (1_m^\top\Sigma^{-1}1_m)\end{bmatrix}^{-1}\begin{bmatrix}\alpha_1 \\ 1\end{bmatrix} \\ \sigma^{*2} &= w^{*\top}\Sigma w^* \\ &= \frac{(1_m^\top\Sigma^{-1}1_m)\alpha_1^2 - 2(\mu^\top\Sigma^{-1}1_m)\alpha_1 + (\mu^\top\Sigma^{-1}\mu)}{(\mu^\top\Sigma^{-1}\mu)(1_m^\top\Sigma^{-1}1_m) - (\mu^\top\Sigma^{-1}1_m)^2} \end{aligned} \tag{1.11}$$

Now let us define π_i as the number of shares held on stock i, then our initial portfolio consisted in m different stocks is given by:

$$\Pi_0^w = \sum_{i=1}^m w_i\alpha_0 = \alpha_0 = \sum_{i=1}^m \pi_i S_{0,i} \tag{1.12}$$

From equation (1.12), we see $\pi_i = \frac{w_i\alpha_0}{S_{0,i}}$ where $w_i\alpha_0$ is the amount assigned to stock i. The expected returns of portfolio is

$$\begin{aligned} w^\top\mu &= \sum_{i=1}^m E[w_i\tfrac{S_{1,i}-S_{0,i}}{S_{0,i}}] \\ &= \sum_{i=1}^m \tfrac{w_i}{S_{0,i}}E[S_{1,i}] - 1 \\ &= \tfrac{1}{\alpha_0}\sum_{i=1}^m \pi_i E[S_{1,i}] - 1 \\ &= \tfrac{E[\Pi_1]}{\alpha_0} - 1 \end{aligned} \tag{1.13}$$

where the value of portfolio at time $t = 1$ is $\Pi_1 = \sum_{i=1}^m \pi_i S_{1,i}$.

In the search for profits, investors want their portfolio to become more valuable at maturity, to make this explicit let $E[\Pi_1] = \alpha_0 + \alpha_2 > \alpha_0 = \Pi_0$ where $\alpha_2 > 0$ is the excess return. From equation (1.13), we see $\alpha_1 = \frac{\alpha_0+\alpha_2}{\alpha_0} - 1$. Hence, the optimal solution w^* in (1.11) depends on α_0 (initial wealth) via α_1, Σ (variance-covariance matrix) and μ (expected mean return). The optimal objective function of risk minimization problem is given by:

$$\sigma_\Sigma^{*2} = f(\mu, \Sigma, \alpha_0, \alpha_2, w^*) = w^{*\top}\Sigma w^* \tag{1.14}$$

Since we are more interested in the role of correlation parameters in portfolio optimization, we denote the optimal objective function (1.14) with respect to Σ . Let Σ, Σ^C represent variance-covariance matrix during normal and crisis periods respectively. w_Σ^* and $w_{\Sigma^C}^*$ are their corresponding optimal solutions. If the allocation of investor's portfolio during crisis is the same as

the one during a normal period, the variance of the existing portfolio during crisis would be $\sigma^2_{\Sigma^C} = f(\mu, \Sigma^C, \alpha_0, \alpha_2, w^*_\Sigma)$. Since w^*_Σ is not an optimal solution of $\sigma^2_{\Sigma^C}$, $\sigma^2_{\Sigma^C}$ must be greater than $\sigma^{*2}_{\Sigma^C} = f(\mu, \Sigma^C, \alpha_0, \alpha_2, w^*_{\Sigma^C})$ which represents the minimum variance during crisis. In other words, there is a smaller initial budget α^* such that

$$f(\mu, \Sigma^C, \alpha^*, \alpha_2, w^*_{\Sigma^C}) = f(\mu, \Sigma^C, \alpha_0, \alpha_2, w^*_\Sigma) \tag{1.15}$$

with given values of $\mu, \alpha_0, \alpha_2, \Sigma^n$ and Σ^C. From equation (1.15), we see the critical importance of variance-covariance matrixes in portfolio evaluation and optimization.

1.2 Comonotonicity ([Dhaene and Vyncke, 2002b])

In this section, we will introduce the definition and properties of comonotonicity by referring to [Dhaene and Vyncke, 2002b].

Consider the sum of dependent random variables $s = \sum_{i=1}^n X_i$. The distribution function of s depends not only on the marginal distribution but also on the joint distribution of $X_1, X_2..., X_n$. In this thesis, we try to find the joint distribution with the given margins such that the sum s is the largest in the convex order sense. The definition and properties of convex order are summarized in [Dhaene and Vyncke, 2002b].

For any two n-vectors $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ in $\mathbb{R}^n$, let $x \leq y$ denote the componentwise order $x_i \leq y_i$ for all $i = 1, 2, ..., n$. The set $A \subseteq \mathbb{R}^n$ is called a comonotonic set if for any $x, y \in A$, we have componentwise order $x \leq y$ (or $y \leq x$). In other words, for any two elements x and y in a comonotonic set, if $x_i < y_i$ for any i, we must have $x \leq y$. Moreover, the subset of a comonotonic set is also a comonotonic set.

The comonotonic random vector is defined by a comonotonic support. Recall that $A \subseteq \mathbb{R}^n$ is called the support of X if $P(X \in A) = 1$. Therefore, an n-dimensional random vector $X = (X_1, ..., X_n)$ is called comonotonic if it has a comonotonic support. We see that comonotonicity displays strong positive dependence structure between elements of the comonotonic support. For example, if $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ are possible outcomes of the comonotonic random vector X, then they must be ordered componentwise.

Theorem 1.2.1 *A random vector X is comonotonic if and only if* (X_i, X_j) *is comonotonic for all* $i \neq j$ *in* $\{1, 2, ..., n\}$.

The theorem states that comonotonicity of a random vector is equivalent with pairwise comonotonicity.

In the case of random variables with continuous marginal distributions, we have a simpler and stronger result.([McNeil et al., 2015])

Corollary 1.2.2 *Let* $X_1, ..., X_n$ *be random variables with continuous distribution functions. They are comonotonic if and only if for every pair* (i, j) *we have* $X_j = T_{ji}(X_i)$ *almost surely for some increasing transformation* T_j.

There are several properties of comonotonic random variables

Properties 1 *A random vector $X = (X_1, ..., X_n)$ is comonotonic if and only if one of the following equivalent conditions holds:*

1). X has a comonotonic support.

2). For all $i \neq j$ in $\{1, 2, ..., n\}$, the couples (X_i, X_j) are comonotonic.

3). For all $x = (x_1, x_2, ..., x_n)$, we have $F_X(x) = P(X \leq x) = \min\{F_{X_1}(x_1), ..., F_{X_n}(x_n)\}$

4). X can be represented as $X \overset{\mathbb{D}}{=} (F^{-1}_{X_1}(U), ..., F^{-1}_{X_N}(U)$ where $U \sim Uniform(0, 1)$ and the notation $\overset{\mathbb{D}}{=}$ is used for equal in distribution.

5). X can be represented as $X \overset{\mathbb{D}}{=} (f_1(Z), f_2(Z), ..., f_n(Z))$ where Z is a random variable and $f_i (i = 1, 2, ..., n)$ are non-decreasing functions $f_i (i = 1, 2, ..., n)$

6). X can be represented as $X \overset{\mathbb{D}}{=} (g_1(Z), g_2(Z), ..., g_n(Z))$ where Z is a random variable and $g_i (i = 1, 2, ..., n)$ are non-increasing functions $f_i (i = 1, 2, ..., n)$

From the second condition in property (1), we see that the comonotonic random vector indicates the pairwise comonotonicity. For the random vector $X = (X_1, ..., X_n)$ with continuous marginal distribution functions, it is comonotonic if and only there exist some increasing transformation T_j such that $X_j = T_{ji}(X_i)$ almost surely for every pair $(i, j), i \neq j$. For any random vector $X = (X_1, ..., X_n)$, the notation $X^C = (X^C_1, ..., X^C_n)$ is used to represent a comonotonic counterpart with the same margins as $(X_1, ..., X_n)$. The outcome of $X^C = (X^C_1, ..., X^C_n)$ is with probability 1 in the following set

$$\{(F^{-1}_{X_1}(p), F^{-1}_{X_2}(p), ..., F^{-1}_{X_n}(p))|0 < p < 1\} \tag{1.16}$$

Since the support of X^C may not be a connected curve due to all horizontal segments of the CDF of X_i ,we can create a new comonotonic connected curve in $\mathbb{R}^n$ by linking the end points of consecutive curves with straight lines. This new set is called the connected support of X^C. It can be parameterized as follows:

$$\{(F^{-1(\alpha)}_{X_1}(p), F^{-1(\alpha)}_{X_2}(p), ..., F^{-1(\alpha)}_{X_n}(p))|0 \leq p \leq 1, 0 \leq \alpha \leq 1\} \tag{1.17}$$

where $F^{-1(\alpha)}_X(p) = \alpha F^{-1}_X(p) + (1 - \alpha)F^{-1+}_X(p)$, and $F^{-1+}_X(p) = sup\{x \in \mathbb{R}|F_X(x) \leq p\}$ is a non-decreasing and right-continuous function. With varying values of α, the parameterization is not unique.

1.2.1 Correlation coefficient and comonotonicity

The Pearson's correlation coefficient for two random variables X and Y is defined by $\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$, where σ^2_X and σ^2_Y is the variance of random variable X and Y respectively. $Cov(X, Y)$ is the covariance of X and Y. $|\rho_{XY} = 1|$ means that X and Y are perfectly linearly dependent. The independence of X and Y imply their zero correlation, not vice versa.

Recall that $\rho_{XY} = 1$ if and only if there exist real numbers $a > 0$ and b such that $Y = aX + b$ holds with probability 1. (see [Dhaene and Vyncke, 2002b])

Hence,

$$\rho_{XY} = 1 \Rightarrow \text{Pairwise comonotonicity of } (X, Y). \tag{1.18}$$

Moreover, correlation is invariant under strictly increasing linear transformations (see [McNeil et al., 2015]). So for $a_1, a_2 > 0$, we have

$$\rho_{a_1X+b_1,a_2Y+b_2} = \rho_{X,Y} \tag{1.19}$$

But correlation is not invariant under nonlinear strictly increasing transformations $T : \mathbb{R} \to \mathbb{R}$. For two real-valued random variables, we have $\rho_{T(X),T(Y))} \neq \rho_{XY}$ in general.

Recall that the random vector X has marginal distribution functions F_{X_i} that belong to the same location-scale family, if there exists a random variable Y, positive real constants a_i and real constants b_i such that $X_i \overset{\mathbb{D}}{=} a_iY + b_i$, for any $i = 1, 2, ..., n$. In other words, for a random vector X with marginal CDFs F_{X_i} belonging to the same location-scale family, there exists a CDF F_Y, positive real constants a_i and real constants b_i such that $F_{X_i}(x) = F_Y((x-b_i)/a_i)$ holds for $i = 1, 2, ..., n$.

Furthermore, we have

$$F^{-1}_{X_i}(p) = a_iF^{-1}_Y(p) + b_i, \ \ p \in (0, 1). \tag{1.20}$$

It is clear that the comonotonic sum has a distribution function that also belongs to the same location-scale family.

$$s^C = X^C_1 + X^C_2 + \cdots + X^C_n \overset{\mathbb{D}}{=} \sum_{i=1}^{n} a_iF^{-1}_Y(U) + \sum_{i=1}^{n} b_i \tag{1.21}$$

where U is a uniform random variable on (0,1).

From equation (1.18), we have the following theorem:

Theorem 1.2.3 *A random vector X with marginal CDFs F_{X_i} belonging to the same location-scale family is comonotonic if and only if $\rho_{X_iX_j} = 1$ for all $i, j \in \{1, 2, ..., n\}$.*

If we consider a random vector X with normal marginals F_{X_i}: $X_i \sim N(\mu_i, \sigma^2_i)$. Since the marginal distribution functions can be represented as $X_i \overset{\mathbb{D}}{=} \sigma_iZ + \mu_i$, where Z is a standard normal random variable, we have

$$F^{-1}_{X_i}(p) = \sigma_i\Phi^{-1}(p) + \mu_i, \ \ p \in (0, 1) \tag{1.22}$$

where Φ is the standard normal CDF. Hence, X is comonotonic if and only if $\rho_{X_iX_j} = 1$ for all $i, j \in 1, 2, ..., n$. We also have that $s^C = X^C_1 + \cdots + X^C_n$ is normally distributed with mean $\sum_{i=1}^n \mu_i$ and variance $(\sum_{i=1}^n \sigma_i)^2$. One should note that if the X_i were independent, we would get the normal distribution with mean $\sum_{i=1}^n \mu_i$ and variance $\sum_{i=1}^n \sigma^2_i \leq (\sum_{i=1}^n \sigma_i)^2$.

In addition, we have following theorem

Theorem 1.2.4 *A random vector X is comonotonic if and only if $\rho_{X_iX_j} = \rho_{X^c_iX^c_j}$ (or $Var(X_i + X_j) = Var(X^C_i + X^C_j)$) for all $i, j \in \{1, 2, ..., n\}$.*

In fact, comonotonicity is an extension of the concept of positive perfect correlation. In two dimensions, it is also possible to consider perfect negative dependence between the components of a random vector. It is called countermonotonicity. Also X_1 and X_2 are countermonotonic if and only if $(X_1, X_2) \overset{\mathbb{D}}{=} (v_1(Z), v_2(Z))$ for some random variable Z with increasing function v_1 and decreasing function v_2, or vice versa.(see [McNeil et al., 2015])

We may consider using given univariate distributions F_1 and F_2 and any correlation value ρ in $[-1, 1]$ to construct a joint distribution F. It is always possible if we only focus on elliptically distributed risk factors, but may be fail in general.

1.2.2 Attainable correlations

[McNeil et al., 2015] presents the so-called attainable correlations which can form a strict subset of the interval $[-1, 1]$. Let (X_1, X_2) be a random vector with finite-variance marginal CDF F_1, F_2 and an unspecified joint CDF; assume also that $Var(X_1) > 0$ and $Var(X_2) > 0$. We have following properties:

Properties 2 *1. The attainable correlations form a closed interval* $[\rho_{\min}, \rho_{\max}]$ *with* $\rho_{\min} < 0 < \rho_{\max}$.

2. The minimum correlation $\rho = \rho_{\min}$ *is attained if and only if* X_1 *and* X_2 *are countermonotonic. The maximum correlation* $\rho = \rho_{\max}$ *is attained if and only if* X_1 *and* X_2 *are comonotonic.*

3. $\rho_{\min} = -1$ *if and only if the distribution of* X_1 *and* $-X_2$ *are from the same location-scale family, and* $\rho_{max} = 1$ *if and only if the distribution of* X_1 *and* X_2 *are from the same location-scale family.*

An example given by [McNeil et al., 2015] calculated the exact value of $\rho_{\max}$ and $\rho_{\min}$ when $lnX_1 \sim N(0, 1)$ and $lnX_2 \sim N(0, \sigma^2)$. For $\sigma \neq 1$, the distribution of random variables X_1 and X_2 do not belong to the same location-scale family so $\rho_{max} < 1$. Likewise, the distribution of X_1 and $-X_2$ do not belong to the same location-scale family so $\rho_{min} > -1$.

Now we calculate the actual boundaries of the attainable interval. Let $Z \sim N(0, 1)$ and observe that if X_1 and X_2 are comonotonic, then $(X_1, X_2) \overset{\mathbb{D}}{=} (e^Z, e^{\sigma Z})$. Clearly, we have $\rho_{max} = \rho_{e^Z, e^{\sigma Z}}$ and $\rho_{min} = \rho_{e^Z, e^{-\sigma Z}}$. After analytical calculation, we obtain

$$\rho_{min} = \frac{e^{-\sigma} - 1}{\sqrt{(e-1)(e^{\sigma^2} - 1)}} \tag{1.23}$$

$$\rho_{max} = \frac{e^{\sigma} - 1}{\sqrt{(e-1)(e^{\sigma^2} - 1)}} \tag{1.24}$$

[McNeil et al., 2015] displays the attainable correlation interval for different values of σ in Figure 1.1. The boundaries of the interval both tend rapidly to zero as σ increases. This implies that we can have situations where comonotonic random vectors have very small correlation values. In the Figure 1.1, once the value of σ exceeds 4, the maximum and minimum attainable correlation values are almost equal to zero. Since comonotonicity is the strongest form of positive dependence, this contradicts the intuition that small correlation implies weak dependence.

1.2.3 Sums of comonotonic random variables

Recall that $(X_1^C, ..., X_n^C)$ is the comonotonic counterpart of random vector $(X_1, ..., X_n)$. Let s^C represent the sum of the components of random vector $(X_1^C, ..., X_n^C)$, we have

$$s^C = X_1^C + X_2^C + \cdots + X_n^C \tag{1.25}$$

[Dhaene and Vyncke, 2002b] states that the distribution function of $s = \sum_{i=1}^n X_i$ can be approximated by the distribution function of the comonotonic sum s^C when s precedes s^C in the convex order sense. We can easily determine the distribution function of s^C from the marginal

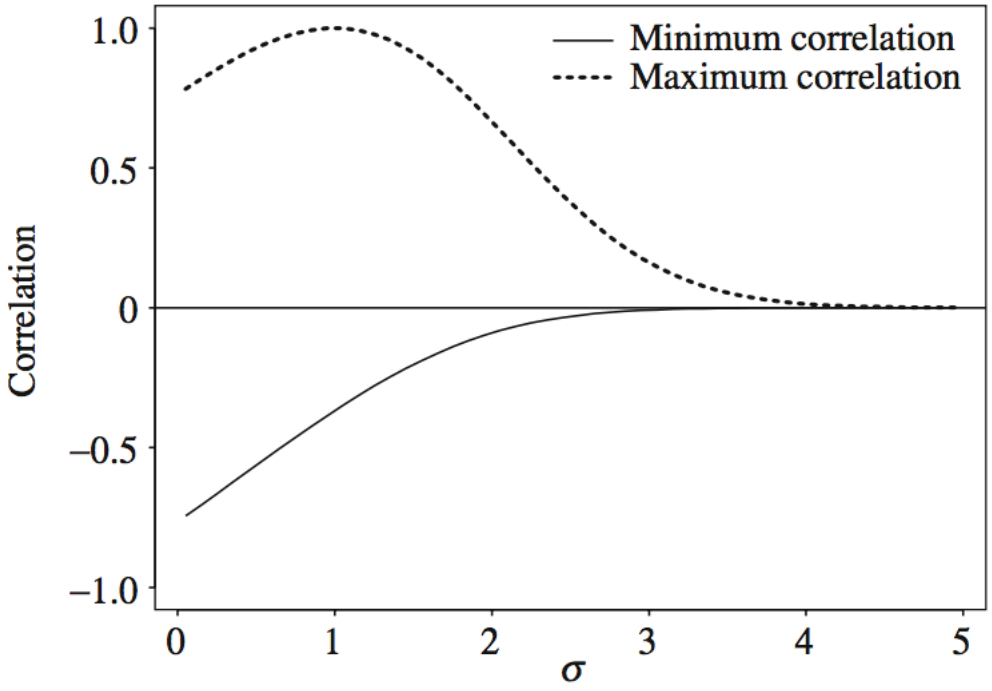

Figure 1.1: Correlation vs σ

distribution of X_i^C. [Dhaene and Vyncke, 2002b] proved that the α-inverse distribution function $F_{s^C}^{-1(\alpha)}$ can be written as the sum of the inverse distribution function of the margins:

$$F_{s^C}^{-1(\alpha)}(p) = \sum_{i=1}^{n} F_{X_i}^{-1(\alpha)}(p),\ 0 < p < 1,\ 0 \leq \alpha \leq 1. \tag{1.26}$$

Thus, we have

$$s^C \overset{\mathbb{D}}{=} \sum_{i=1}^{n} F_{X_i}^{-1(\alpha)}(U),\ U \sim U(0,1) \tag{1.27}$$

and the connected support of s^C is given $\{F_{s^C}^{-1(\alpha)}(p)|0 < p < 1, 0 \leq \alpha \leq 1\} = \{\sum_{i=1}^{n} F_{X_i}^{-1(\alpha)}(p)|0 < p < 1, 0 \leq \alpha \leq 1\}$.

This implies $F_{s^C}^{-1+}(0) = \sum_{i=1}^{n} F_{X_i}^{-1+}(0)$ and $F_{s^C}^{-1}(0) = \sum_{i=1}^{n} F_{X_i}^{-1}(0)$.

Hence, the minimal value of the comonotonic sum equals the sum of the minimal values of each term.

If we consider the random vector X as stock prices at time T with marginal distributions $F_{X_i}, i = 1, 2, ..., n$, X is comonotonic if and only if

$$(X_1, ..., X_n) \overset{\mathbb{D}}{=} (F_{X_1}^{-1}(U), ..., F_{X_n}^{-1}(U)), \tag{1.28}$$

where U is a uniform (0, 1) random variables.

Let S be the weighted sum of X, $S = \sum_{i=1}^{N} w_i X_i$, and S^C be the weighted sum of the comonotonic vector

$$S^C = \sum_{i=1}^{N} w_i X_i^C = \sum_{i=1}^{N} w_i F_{X_i}^{-1}(U)$$

. We see there are another two equivalent statements for the comonotonicity of X in Properties1:

6). $S \overset{\mathbb{D}}{=} S^C$

7). $Var[S] = Var[S^C]$

Chapter 2

Implied Correlation Index

The implied correlation index is meant to be a measure of the average level of correlation on a given portfolio. It is defined as the ratio of the sum of the weighted covariance to that of the weighted variance between stocks. In this chapter, we introduce the basic implied correlation index (CIX) by referring to [Skintzi and Refenes, 2005]. Then, we detail two specific CIXs with S & P 500 historical data. Finally, we learn the general implied correlation index (ICX) and summarize approaches used to find the value of ICX.

2.1 CIX ([Skintzi and Refenes, 2005])

CIX is first introduced by [Skintzi and Refenes, 2005] to represent implied correlation index. Its definition, properties and some special types of CIX will be presented in this section.

Assume that a portfolio consists of N assets, the portfolio variance $\sigma^2_{P,t}$ at time t over a T-day horizon can be calculated as follows:

$$\sigma^2_{P,t} = \sum_{i=1}^{N} w^2_{i,t}\sigma^2_{i,t} + 2\sum_{i=1}^{N-1}\sum_{j>i}^{N} w_{i,t}w_{j,t}\rho_{ij,t}\sigma_{i,t}\sigma_{j,t} \tag{2.1}$$

where $w_{i,t}$ is the weight of asset i on the portfolio at time t, $\sigma_{P,t}$ is the portfolio volatility at time t, $\sigma_{i,t}$ is the volatility of asset i at time t and $\rho_{ij,t}$ is the pairwise correlation between assets i and j at time t. One should note that all volatilities and correlations are measured over the same time horizon (T days).

2.1.1 Definition and properties of CIX

The key assumption in the methodology of CIX is that the correlation between any two assets is constant at time t rather than having $N(N-1)$ different values. We define the correlation coefficient by ρ_t. Then we can rewrite the equation 2.1 as:

$$\sigma^2_{P,t} = \sum_{i=1}^{N} w^2_{i,t}\sigma^2_{i,t} + 2\rho_t\sum_{i=1}^{N-1}\sum_{j>i}^{N} w_{i,t}w_{j,t}\sigma_{i,t}\sigma_{j,t} \tag{2.2}$$

[Skintzi and Refenes, 2005] defines CIX as the coefficient ρ_t assumed to hold for any pair of assets in a portfolio. From the portfolio variance equation (2.2), the value of CIX at time t over a T-day horizon is determined by:

$$CIX_t = \rho_t = \frac{\sigma^2_{P,t} - \sum_{i=1}^{N} w^2_{i,t}\sigma^2_{i,t}}{2\sum_{i=1}^{N-1}\sum_{j>i}^{N} w_{i,t}w_{j,t}\sigma_{i,t}\sigma_{j,t}} \tag{2.3}$$

In order to have more accurate measure of future correlation, CIX is computed using the implied volatility $\sigma_{P,t}$ at time t from the portfolio option and the implied volatilities $\sigma_{i,t}$ from options on each of the portfolio asset i at time t. All implied volatilities should be derived by at-the-money options with the same time-to-maturity equal to T.

In addition, we can rewrite equation 2.3 as follows:

$$CIX_t = \sum_{i=1}^{N}\sum_{j>i} c_{ij,t}\rho_{ij,t} \text{ where } c_{ij,t} = \frac{w_{i,t}w_{j,t}\sigma_{i,t}\sigma_{j,t}}{\sum_{i=1}^{N}\sum_{j>i} w_{i,t}w_{j,t}\sigma_{i,t}\sigma_{j,t}} \tag{2.4}$$

Also $\rho_{ij,t}$ is the pair-wise correlation between assets i and j at time t. $\rho_{ij,t} \in (-1, 1)$. From equation (2.4), we see that CIX can be interpreted as a weighted average of the pair-wise correlations among the portfolio asset returns.

Moreover, equation (2.3) can be expressed as follows:

$$CIX_t = \rho_t = \frac{\sigma^2_{P,t} - \sigma^2_{Zero,t}}{\sigma^2_{Perf,t} - \sigma^2_{Zero,t}} \tag{2.5}$$

where $\sigma^2_{Perf,t}$ is the portfolio variance assuming perfect correlations ($\rho_{ij,t} = 1,\ i,\ j = 1, ..., N$) between each pair of assets in the portfolio, and $\sigma^2_{Zero,t}$ is the portfolio variance assuming zero correlations ($\rho_{ij,t} = 0,\ i,\ j = 1, ..., N, j > i$) between each pair of assets in the portfolio. From equation (2.5), we see that CIX can represent how far lies the portfolio variance between the minimum portfolio variance assuming zero pair-wise correlations and the maximum portfolio variance assuming perfect correlations.

Equation (2.5) also interprets the reason why [Skintzi and Refenes, 2005] use index options to calculate CIX. This value of CIX can provide a measure of the market portfolio diversification in the specific market represented by the underlying index. In details, changes to a correlation index constructed from a stock index option imply changes to the level of portfolio diversification that can be achieved in the market represented by the index.

We conclude several properties of CIX by referring to [Kim and Ahn, 2013]'s work.

Properties 3 *1) Based on the properties of comonotonicity, for any $X(t)$, we can have a wider range of CIX_t:*

$$-\frac{\sum_{i=1}^{N} w^2_{i,t}\sigma^2_{i,t}}{2\sum_{i=1}^{N-1}\sum_{j>i} w_{i,t}w_{j,t}\sigma_{i,t}\sigma_{j,t}} \leq CIX_t \leq \frac{\sigma^2_{S^c,t} - \sum_{i=1}^{N} w^2_{i,t}\sigma^2_{X_i^c,t}}{2\sum_{i=1}^{N-1}\sum_{j>i} w_{i,t}w_{j,t}\sigma_{i,t}\sigma_{j,t}} \tag{2.6}$$

where $\sigma_{S^c,t}$ is the implied volatility of the weighted market index under the comonotonic assumption. $\sigma_{X_i^c,t}$ is the implied volatility of asset i under the comonotonic assumption.

2) The lower bound of equation (2.6) is attained if and only if

$$P\left(\sum_{i=1}^{N} w_i X_i(t) = c\right) = 1 \text{ for some constant } c \in \mathbb{R} \tag{2.7}$$

3) The upper bound of equation (2.6) is attained if $X(t)$ is comonotonic.
4) If elements of $X(t)$ are pairwise uncorrelated, the $CIX_t = 0$.

By applying their methodology to DJIA index option prices, [Skintzi and Refenes, 2005] find evidence of the existence of a long-run dependence in correlation and contemporaneous relationship between the correlation index daily changes and DJIA returns.

2.1.2 Specific CIX - ICJ and JCJ for S&P 500 Index ([Exchange, 2009])

In July 2009, CBOE began disseminating daily values for the CBOE S& P 500 Implied Correlation Index. The calculation of CBOE S& P 500 Implied Correlation Index uses implied volatilities of SPX option and implied volatilities of single-stock options on the 50 largest components of S&P 500. These 50 largest components are measured by market capitalization. According to different maturities of SPX option and individual stock option (LEAPS), CBOE name CIX as ICJ and JCJ alternatively. The number behind ICJ or JCJ represents the expiring year of individual stock LEAPS used in calculation. For example, ICJ 2010 is based on SPX options expiring in December 2009 and individual stock LEAPS expiring in January 2010. JCJ 2011 is based on SPX options expiring in December 2010 and LEAPS expiring in January 2011. There are overlaps between two close ICJ and JCJ. The time horizon of each ICJ (or JCJ) is two years.

[Exchange, 2009] used [Skintzi and Refenes, 2005]'s methodology to compute the specific value of S&P 500 Implied Correlation Index. In this case, the implied variance of portfolio $\sigma^2_{P,t}$ in equation (2.3), is replaced by the implied variance of SPX and the implied volatilities of 50 index components. The weight of i^{th} index component is defined as follows:

$$w_i = \frac{P_i S_i}{\sum_{i=1}^{50} P_i S_i} \tag{2.8}$$

where P_i is the price of the i^{th} index component, S_i is the float-adjusted shares outstanding of the i^{th} index component.

Hence, the value of S&P 500 CIX can be calculated via the following equation:

$$\rho_{Average} = \frac{\sigma^2_{Index} - \sum_{i=1}^{N} w_i^2 \sigma_i^2}{2 \sum_{i=1}^{N-1} \sum_{j>i}^{N} w_i w_j \sigma_i \sigma_j} \tag{2.9}$$

The specific algorithm from [Exchange, 2009] to determine S&P 500 CIX is presented as follows:

Algorithm 1 (Algorithm for determining S&P 500 CIX) *1) Construct the 50-stock SPX tracking basket based on market capitalization (closing price times "float-adjusted shares").*

2) Select the options to be used in the implied correlation calculation; determine the implied volatility for SPX options, σ_{index}, and the implied volatilities for options on the stock comprising the SPX tracking basket, σ_i.

For each stock in SPX tracking basket, the put option with a strike price just below and the call option with a strike price just above the current stock price are selected.

Next, for each option, an implied volatility of put/call pair are then weighted through a linear interpolation to arrive at a single at-the-money implied volatility for each stock.

The SPX options used to calculate ICJ are the put option with a strike price just below and the call option with a strike price just above the forward SPX level to the options' expiration date. The forward index level is determined using at-the-money SPX option prices, where the strike price with the smallest absolute difference between the call and put prices is considered to be the at-the-money strike price. Each option price is deemed to be the average of its bid/ ask quote.

3) Calculate the capitalization weight, w_i, of each component in the 50-stock basket.

4) Calculate the implied correlation, $\rho_{Average}$, by using equation (2.9).

[Zhou, 2013] states that the implied correlation index can be used to monitor the market's overall systematic risk. Systematic risk analysis is important because it cannot be diversified away by any portfolio optimization methods. [Zhou, 2013] explores the forecasting power of the ICJ index for S&P 500 Index returns. They implement different regression in which future S&P 500 Index returns are regressed on the current information set of ICJ index changes. The result shows that the ICJ changes, i.e. current weekly change and changes in the past, are strongly linked to the S&P 500 Index returns in the future, and their model consistently outperforms the random walk model using the Superior Predictive Ability testing procedure.

2.2 General Implied correlation Index (ICX)([Linders and Schoutens, 2014])

[Linders and Schoutens, 2014] introduced a concept of general implied correlation index (ICX) which depends on the value of index option moneyness. CIX introduced in Section 2.1 can seem as a particular case of ICX when the moneyness equals 1. [Linders and Schoutens, 2014] summarized [Skintzi and Refenes, 2005]'s methodology as a traditional approach to determine implied correlation index. They stated that the traditional approach for determining implied correlation will underestimate the real correlation. The error is more pronounced when some stock volatilities are large compared to the other volatility levels. Instead of using the traditional approach, [Linders and Schoutens, 2014] presented a robust measurement model to determine ICX with application to DJIA.

2.2.1 General framework of ICX

In multivariate Black& Scholes model setting, the put option price on stock i ($P_i[K,T]$), call option price on stock i ($C_i[K,T]$) and out-of-money index option price ($Q[K,T]$) with strike price K and maturity T can be derived by Black& Scholes formula. The implied correlation

($\widehat{\rho}[\pi]$) among components of index S is determined by

$$\overline{Q}[K;\widehat{\sigma}[\pi],\ \widehat{\rho}[\pi]] = \widehat{Q}[K] \tag{2.10}$$

where

$\pi = \frac{K}{S(0)}$ is the moneyness of index S, $\widehat{\sigma}[\pi] = (\widehat{\sigma}_1[\pi], \widehat{\sigma}_2[\pi], ..., \widehat{\sigma}_n[\pi])$ are the marginal implied volatilities with moneyness π.

$\widehat{Q}[K]$ is the observed out-of-the-money index option price with strike K and maturity T.

$\overline{Q}[K;\widehat{\sigma}[\pi],\ \widehat{\rho}[\pi]]$ is the approximation of the index option price with marginal implied volatilities $\widehat{\sigma}$.

In high-dimensional cases, it is not convenient to compute the actual index option price $Q[K]$, thus [Linders and Schoutens, 2014] used the approximation of the index option price in formular 2.10. The authors commented that smaller value of $|\overline{Q}[K] - Q[K]|$ implies more accurate estimation of $\widehat{\rho}[\pi]$.

2.2.2 Traditional approach

The traditional approach calculates the implied correlation index by:

$$\widehat{\rho}[\pi] \approx \frac{\widehat{\sigma}^2[\pi] - \sum_{i=1}^{n} \widetilde{w}_i^2 \widehat{\sigma}_i^2[\pi]}{\sum_{i=1}^{n} \sum_{j=1, j\neq i} \widetilde{w}_i \widetilde{w}_j \widehat{\sigma}_i[\pi] \widehat{\sigma}_j[\pi]}. \tag{2.11}$$

where π is given moneyness and the weights are determined by $\widetilde{w}_i = \frac{w_i X_i(0)}{S(0)}$. The CIX defined in equation (2.3) is a particular case of equation (2.11) where the moneyness equals 1.

2.2.3 Robust measurement for Implied correlation Index

Since it is hard to find an explicit formula of implied correlation index to make equation (2.10) holds, [Linders and Schoutens, 2014] presented a robust measurement model where the implied correlation index is the one minimizes the distance between the approximated theoretical index option price and the observed index option price:

$$\widehat{\rho}[\pi] = \arg\min_{0<\rho<1} \frac{|\overline{Q}[K;\widehat{\sigma}[\pi],\ \rho] - \widehat{Q}[K]|}{\widehat{Q}[K]} \tag{2.12}$$

For unavailable strike price K in the market, we choose two traded strike price K_j and K_{j+1} such that $K_j < K < K_{j+1}$ where $j \in \mathbb{Z}$ and $j \in [-l, h)$. Their corresponding index option prices $\widehat{Q}[K_j]$ and $\widehat{Q}[K_{j+1}]$ are reachable in the market so we can use equation (2.12) to compute the value of $\widehat{\rho}[\pi_j]$ and $\widehat{\rho}[\pi_{j+1}]$ with given moneyness $\pi_j = \frac{K_j}{S(0)}$ and $\pi_{j+1} = \frac{K_{j+1}}{S(0)}$. By interpolation, ICX with moneyness π is determined by:

$$\widehat{\rho}[\pi] = \widehat{\rho}[\pi_j]\frac{K_{j+1} - K}{K_{j+1} - K_j} + \widehat{\rho}[\pi_{j+1}]\frac{K - K_j}{K_{j+1} - K_j} \tag{2.13}$$

With low moneyness $\pi = \frac{K}{S(0)}$, ICX may strictly bigger than 1. To make sure ICX bounded in $[-1, 1]$, [Linders and Schoutens, 2014] suggested that the moneyness π should be larger than 0.75.

Chapter 3

Herd behaviour index

3.1 HIX ([Dhaene and Vyncke, 2012])

[Dhaene and Vyncke, 2012] introduced Herd Behaviour Index (HIX) as the ratio of the risk-neutral variance of the real market index to the risk-neutral variance of the comonotonic index. Portfolio managers can use HIX to measure the degree of diversification among portfolio components. HIX takes value in $[0, 1]$, where 1 implies no diversification. For any market index, HIX can be calculated by using a seris of vanilla options which is traded on this index and its components.

In this section, we first introduce the empirical distributions of real market index components as well as approximate variance of the real market index. Then we introduce the concept of perfect herd behaviour and decompose the empirical comonotonic call/put option prices as a linear combination of a series of vanilla call/put option prices. Finally, we present the definition and properties of HIX by using the theory of comonotonicity. Consider a market index S which is a linear combination of the N underlying assets. Let X_i denote the price of asset i. We denote the price of the index at time t ($0 \leq t \leq T$) as follows

$$S(t) = \sum_{i=1}^{N} w_i X_i(t) \tag{3.1}$$

A finite number of traded strike prices for European options written on asset i are denoted by $K_{i,0}, K_{i,1}, ..., K_{i,m_i}$ where $0 = K_{i,0} < K_{i,1} < ... < K_{i,m_i} < F^{-1}_{X_i}(1)$ and K_{i,m_i} is the maximal value of X_i. The actual CDF of F_{X_i} is not completely specified because of a finite number of traded strikes. By referring to [Chen and Vanmaele, 2008], the risk-neutral empirical CDF of X_i can be presented as follows:

$$\widehat{F}_{X_i}(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 + \mathrm{e}^{rT} \frac{C_i[K_{i,j+1}] - C_i[K_{i,j}]}{K_{i,j+1} - K_{i,j}} & \text{if } K_{i,j} \leq x < K_{i,j+1} \\ j = 0, 1, ..., m_i & \\ 1 & \text{if } K_{i,j+1} \leq x \end{cases} \tag{3.2}$$

where $C_i[K_{i,j}]$ is the European call option price on asset i with strike price $K_{i,j}$.

According to put-call parity, we can express $\widehat{F}_{X_i}(x)$ as a function of European put option prices:

$$\widehat{F}_{X_i}(x) = \begin{cases} 0 & \text{if } x < 0 \\ e^{rT} \frac{P_i[K_{i,j+1}] - P_i[K_{i,j}]}{K_{i,j+1} - K_{i,j}} & \text{if } K_{i,j} \le x < K_{i,j+1} \\ j = 0, 1, ..., m_i & \\ 1 & \text{if } K_{i,j+1} \le x \end{cases} \tag{3.3}$$

where $P_i[K_{i,j}]$ is the price of European put option on asset i with strike price $K_{i,j}$.

From equations (3.2) and (3.3), we see that the empirical distribution can be determined by a approxmate convex option curve $C_i[K]$ (or $P_i[K]$) that is the segmented linear convex function connecting observed points $(K_{i,j},\ C_i[K_{i,j}])$ (or $(K_{i,j},\ P_i[K_{i,j}])$), $j = 0, 1, ..., m_i + 1$.

If we know the index option prices with a series of strike prices, we can determine the risk neutral variance of the index. Let us denote the strike prices for index put option as $K_{-i},\ i = 0, 1, ..., l$ with $K_{-l} < K_{-l+1} < ... < K_{-1} < K_0 \le \mathbb{E}[S]$ and K_0 is the first strike price less than $\mathbb{E}[S]$. Similarly, we denote the strike prices for index call option as $K_i,\ i = 0, 1, ..., h$ with $K_h > K_{h-1} > ... > K_1 > \mathbb{E}[S]$. [Dhaene and Vyncke, 2012] estimated the risk-neutral variance of the index as follows:

$$Var[S] \approx 2e^{rT} \sum_{i=-l}^{h} \Delta K_i Q[K_i] - (\mathbb{E}[S] - K_0)^2 \tag{3.4}$$

where ΔK_i is given by:

$$\Delta K_i = \begin{cases} K_{-l+1} - K_{-l} & \text{for the lowest strike } K_{-l} \\ \frac{K_{i+1} + K_{i-1}}{2} & \text{for } i = -l+1, ..., h-1 \\ K_h - K_{h-1} & \text{for the highest strike } K_h \end{cases} \tag{3.5}$$

$Q[K_i]$ is the observed index option price denoted by

$$Q[K_i] = \begin{cases} P[K_i] & \text{if } K_i < K_0 \\ \frac{P[K_i] + C[K_i]}{2} & \text{if } K_i = K_0 \\ C[K_i] & \text{if } K_i > K_0 \end{cases} \tag{3.6}$$

3.1.1 Perfect herd behaviour and comonotonic index option prices

Perfect herd behavior is the extreme situation where stock market is comonotonic. The comonotonic index price is a linear combination of comonotonic vector X^C:

$$S^C(t) = \sum_{i=1}^{N} w_i X_i^C(t) \tag{3.7}$$

Equation (1.28) in Section 1.2.3 clearly shows that comonotonic risks are driven by a single source of randomness and exhibits extreme herd behaviour. Since the pricing distributions $F_{X_i}, i = 1, ..., N$ are unknown, [Linders and Vanmaele, 2012] use the inverse of the empirical distributions $\widehat{F}_{X_i}^{-1}$ to replace $F_{X_i}^{-1}$:

$$\widehat{F}_{X_i}^{-1}(p) = K_{i,j} \quad \text{if } \widehat{F}_{X_i}(K_{i,j-1}) < p \le \widehat{F}_{X_i}(K_{i,j}) \quad j = 0, 1, ..., m_i + 1,\ p \in (0, 1) \tag{3.8}$$

with $K_{i,-1} = -1$

Therefore, the comonotonic sum based on the empirical marginal distributions is given by:

$$\widehat{S}^C = \sum_{i=1}^{N} w_i \widehat{F}_{X_i}^{-1}(U) \tag{3.9}$$

[Linders and Vanmaele, 2012] also presented the following algorithm to determine the distribution of the comonotonic sum.

Algorithm 2 (Algorithm for determining $F_{\widehat{S}^C}(K)$) *1) Using equations (3.2) or (3.3), determine all elements of the following set:*

$$\mathrm{A} = \{\widehat{F}_{X_i}(K_{i,j}) |\ i = 1, ..., N \text{ and } j = 0, 1, ..., m_i\} \setminus \{0\} \tag{3.10}$$

2) With the help of equation (3.8), calculate $\sum_{i=1}^{N} w_i \widehat{F}_{X_i}^{-1}(p)$ for all $p \in \mathrm{A}$.

3) Introducing the following notation:

$$\widehat{F}_{X_i}^{-1+}(0) = \min_j \{K_{i,j} | \widehat{F}_{X_i}(K_{i,j}) > 0\} \tag{3.11}$$

For any $K \in (\sum_{i=1}^{N} w_i \widehat{F}_{X_i}^{-1+}(0),\ \sum_{i=1}^{N} w_i K_{i,m_i+1})$, calculate $F_{\widehat{S}^C}(K)$ from:

$$F_{\widehat{S}^C}(K) = \max\{p \in \mathrm{A} | \sum_{i=1}^{N} w_i \widehat{F}_{X_i}^{-1}(p) \leq K\} \tag{3.12}$$

4) For other values of K, $F_{\widehat{S}^C}(K)$ is given by:

$$F_{\widehat{S}^C}(K) = \begin{cases} 0 & \text{if } K < \sum_{i=1}^{N} w_i \widehat{F}_{X_i}^{-1+}(0), \\ \min_i \widehat{F}_{X_i}(\widehat{F}_{X_i}^{-1+}(0)) & \text{if } K = \sum_{i=1}^{N} w_i \overline{F}_{X_i}^{-1+}(0), \\ 1 & \text{if } K > \sum i = 1^N w_i K_{i,m_i+1} \end{cases} \tag{3.13}$$

In [Linders and Vanmaele, 2012], comonotonic index call and put option prices with strike price K are given by:

$$\begin{aligned} \widehat{C}^C[K] &= \mathrm{e}^{-rT} \mathbb{E}[(\widehat{S}^C - K)_+] \\ \widehat{P}^C[K] &= \mathrm{e}^{-rT} \mathbb{E}[(K - \widehat{S}^C)_+] \end{aligned} \tag{3.14}$$

Starting from equation (3.2), [Chen and Vanmaele, 2008] and [Hobson* et al., 2005] prove that the comonotonic index call option price $\widehat{C}^C[K]$ can be expressed as follows:

$$\widehat{C}^C[K] = \sum_{i \in N_K} w_i C_i[K_{i,j_i}] + \sum_{i \in \overline{N}_K} w_i (\alpha_K C_i[K_{i,j_i}] + (1 - \alpha_K) C_i[K_{i,j_i+1}]) \tag{3.15}$$

where the coefficient α_K, set N_K and its complement $\overline{N}_K$ are defined by

$$\begin{aligned} \alpha_K &= 1 - \frac{K - \sum_{i=1}^{N} w_i K_{i,j_i}}{\sum_{i \in \overline{N}_K} w_i (K_{i,j_i+1} - K_{i,j_i})} \\ N_K &= \{i \in \{1, 2, ..., N\} | \widehat{F}_{X_i}(K_{i,j_i-1}) < F_{\widehat{S}^C}(K) \leq \widehat{F}_{X_i}(K_{i,j_i})\} \\ \overline{N}_K &= \{i \in \{1, 2, ..., N\} | F_{\widehat{S}^C}(K) = \widehat{F}_{X_i}(K_{i,j_i})\} \end{aligned} \tag{3.16}$$

Equation (3.15) holds for any $K \in (\sum_{i=1}^{N} w_i \widehat{F}_{X_i}^{-1+}(0),\ \sum_{i=1}^{N} w_i K_{i,m_i+1})$ where $\widehat{F}_{X_i}^{-1+}(0)$ is defined in equation (3.11). In this form, each $j_i, i = 1, 2, ..., N$, depends on K and is defined as the unique integer in the set $0, 1, ..., m_i + 1$ with $i \in N_K$.

Likewise, the comonotonic index put option price $\widehat{P}^C[K]$ is given by:

$$\widehat{P}^C[K] = \sum_{i \in N_K} w_i P_i[K_{i,j_i}] + \sum_{i \in \overline{N}_K} w_i(\alpha_K P_i[K_{i,j_i}] + (1 - \alpha_K) P_i[K_{i,j_i+1}]) \tag{3.17}$$

where N_K, $\overline{N}_K$ and α_K are defined in equation (3.16). The index j_i should satisfy $i \in N_K$.

Under the condition that $\mathbb{E}[\widehat{S}^C] = \mathbb{E}[S]$, the approximate variance of the comonotonic index price is defined as:

$$Var[\widehat{S}^C] \approx 2e^{rT} \sum_{i=-l}^{h} \Delta K_i \widehat{Q}^C[K_i] - (\mathbb{E}[S] - K_0)^2 \tag{3.18}$$

where $\widehat{Q}^C[K_i]$ is the comonotonic index option price, $\widehat{Q}^C[K_i]$

$$\widehat{Q}^C[K_i] = \begin{cases} \widehat{P}^C[K_i] & \text{if } K_i < K_0 \\ \frac{\widehat{P}^C[K_i] + \widehat{C}^C[K_i]}{2} & \text{if } K_i = K_0 \\ \widehat{C}^C[K_i] & \text{if } K_i > K_0 \end{cases} \tag{3.19}$$

3.1.2 Definition and Properties of HIX

[Dhaene and Vyncke, 2012] defines the HIX as the ratio of the variance of real market index to the variance of comonotonic market index:

$$HIX[T] = \frac{Var[S]}{Var[S^C]} \tag{3.20}$$

where HIX[T] is the notation of T-year implied Herd Behaviour Index. [Dhaene and Vyncke, 2012] replaced $Var[S^C]$ by its approximation $Var[\widehat{S}^c]$, so the HIX[T] is determined by comparing an appropriate linear combination of real index option prices with the same linear combination of corresponding comonotonic index option prices:

$$HIX[T] = \frac{Var[S]}{Var[\widehat{S}^C]} = \frac{2e^{rT} \sum_{i=-l}^{h} \Delta K_i Q[K_i] - (\mathbb{E}[S] - K_0)^2}{2e^{rT} \sum_{i=-l}^{h} \Delta K_i \widehat{Q}^C[K_i] - (\mathbb{E}[S] - K_0)^2} \tag{3.21}$$

where ΔK_i, $Q[K_i]$ and $Q^C[K_i]$ are defined by equations (3.5), (3.6) and (3.19) respectively.

Properties 4 (Properties of HIX) *1) Since variance is nonnegative and* $Var[S] \leq Var[S^C]$,

$$0 \leq HIX[T] \leq 1 \tag{3.22}$$

2) Boundaries can be achieved under the following conditions:

$$HIX[T] = 0 \ \textit{ if and only if } S(t) \textit{ is constant.} \tag{3.23}$$

and

$$HIX[T] = 1 \textit{ if and only if } X(t) \textit{ is comonotonic.} \tag{3.24}$$

Chapter 4

Stochastic Modeling of CIX and HIX

The processes of both CIX and HIX should be characterized in continuous time due to its frequent trading. In this chapter, we first show two main approaches from [van Emmerich, 2006] and [Teng et al., 2016] to model correlation (CIX). Then, we define our two choices of Correlation Stochastic Volatility (CSV) models and show some stylized facts. We also discuss an additional approach which uses the Wishart Affine Stochastic process to model correlation. Finally, we present some stochastic models for HIX.

4.1 First approach to model CIX

[van Emmerich, 2006] introduced the first approach to model CIX. In this approach, the correlation (CIX) is a function of a Brownian motion. Moreover, [van Emmerich, 2006] gave the following properties of their models:

Properties 5 (Properties of CIX model) *1) The model is concentrated on [-1, 1],*
2) The model varies around a mean,
3) The probability mass approaches zero in the boundary values,
4) There is a suitable number of parameters to calibrate the model to market data.

Here we describe the modeling approach in detail. Without loss of generality, let us consider a mean-reverting stochastic process X_t,

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t,\ t \neq 0,\ X_0 = x_0 \tag{4.1}$$

where $W = \{W_t, t \neq 0\}$ is a standard Brownian motion. Next, we consider a function $f(x)$ which is twice continuously differentiable on $\mathbb{R}$. By applying Ito's Lemma to $Y_t = f(X_t)$, we have:

$$dY_t = \tilde{a}(t, X_t)dt + \tilde{b}(t, X_t)dW_t, \tag{4.2}$$

where

$$\tilde{a}(t, X_t) = a(t, X_t)\frac{\partial f}{\partial x}(X_t) + \frac{1}{2}b^2(t, X_t)\frac{\partial^2 f}{\partial x^2}(X_t) \tag{4.3}$$

$$\tilde{b}(t, X_t) = b(t, X_t)\frac{\partial f}{\partial x}(X_t) \tag{4.4}$$

Since the correlation ρ_t takes value in the interval $[-1, 1]$, the choice of f is restricted to functions that can map the domain of X_t to the interval $[-1, 1]$.

Here we consider two kinds of mean-reverting processes for X_t: Vasicek process and CIR process. If X_t is assumed to be a Vasicek process, we have:

$$dX_t = \kappa(\eta - X_t)dt + \varsigma dW_t, \quad X_t \in (-\infty, \infty), \tag{4.5}$$

If X_t is assumed to be a CIR process, we have:

$$dX_t = \kappa(\eta - X_t)dt + \varsigma \sqrt{X_t} dW_t, \ X_t \in (0, \infty), \tag{4.6}$$

Since the domain of Vasicek process is $(-\infty, \infty)$, based on [Teng et al., 2016]'s work, we use the following mapping functions to model the correlation ρ_t:

$$\begin{array}{ll} f_1(x) & := \tanh(x) \\ f_2(x) & := \frac{2}{\pi} \arctan((\pi/2)x) \\ f_3(x) & := 2\Phi(x) - 1, \Phi(x) \text{ is the CDF of the standard normal distribution} \end{array} \tag{4.7}$$

Functions in (4.7) can map an underlying process X_t in the interval $(-\infty, \infty)$ to a stochastic correlation process ρ_t in the interval $(-1, 1)$. It is clear that all these functions are symmetrical and measurable. When $X_t = 0$, the stochastic correlation process $\rho_t = 0$. As X_t gets large, ρ_t gets close to 1 but never reaches it. As X_t gets large and negative, ρ_t gets closet to -1 but never reaches it either. Since perfect correlation rarely occurs in the real market, functions f_1, f_2 and f_3 are still useful to model the correlation ρ_t.

With the CIR process, the domain of mapping function is $(0, \infty)$. We consider some new mapping functions as follows:

$$\begin{array}{ll} f_4(x) & := -2e^{-kx} + 1, \text{ for any } k > 0 \\ f_5(x) & := 2\tanh(x) - 1 \end{array} \tag{4.8}$$

Functions in (4.8) can map X_t in $(0, \infty)$ to ρ_t in $(-1, 1)$. As X_t decreases toward zero, ρ_t with functions $f_4(x)$ or $f_5(x)$ approaches -1 but never reach it. Likewise, as X_t tends to positive infinity, ρ_t approaches 1 but never reach it. Due to the rare occurrences of perfect correlation in reality, functions $f_4(x)$ and $f_5(x)$ are still useful to model ρ_t.

In Table 4.1, we display six model specifications by combing the mapping functions (4.7), (4.8) with a Vasicek and a CIR process. The coefficients $\tilde{a}(t, X_t)$ and $\tilde{b}(t, X_t)$ are obtained by plugging functions (4.7), (4.8) to equation (4.2)

Table 4.1: CIX model specifications

X_t	$f(X_t)$	$\tilde{a}(t, X_t)$	$\tilde{b}(t, X_t)$
Vasicek	$f_1(x)$	$(1 - f_1(X_t)^2)(\kappa(\eta - X_t) - f_1(X_t)\varsigma^2)$	$(1 - f_1(X_t)^2)\varsigma$
Vasicek	$f_2(x)$	$(\frac{\kappa(\eta - X_t)}{1+\tan^2(f_2(X_t)\pi/2)} - \frac{\pi\varsigma^2 \tan(f_2(X_t)\pi/2)}{2(1+\tan^2(f_2(X_t)\pi/2))^2})$	$\frac{\varsigma}{(1+\tan^2(f_2(X_t)\pi/2))}$
Vasicek	$f_3(x)$	$\frac{2\kappa(\eta - X_t)}{\sqrt{2\pi}} e^{-X_t^2/2} - \varsigma^2 \frac{X_t}{\sqrt{2\pi}} e^{-X_t^2/2}$	$\frac{2\varsigma}{\sqrt{2\pi}} e^{-X_t^2/2}$
CIR	$f_4(x)$	$2\kappa(\eta - X_t)ke^{-kX_t} - \varsigma^2 X_t k^2 e^{-kX_t}$	$2\varsigma\sqrt{X_t}ke^{-kX_t}$
CIR	$f_5(x)$	$2sech^2(X_t)(\kappa(\eta - X_t) - \varsigma^2 X_t \tanh(X_t))$	$2\varsigma\sqrt{X_t}sech^2(X_t)$

4.2 Second approach to model CIX

Although the first approach stands out because of the ease of construction and the high degree of analytical tractability, it lacks intuitive interpretation. For this reason, [van Emmerich, 2006] proposed the second approach where the correlation ρ_t is a stochastic process driven by a Brownian motion.

In details, [van Emmerich, 2006] chose a mean-reverting process with a deterministic mean:

$$d\rho_t = \kappa(\theta - \rho_t)dt + \sqrt{\varsigma(1 - \rho_t^2)}dW_t, \quad \rho_0 \in (-1, 1) \tag{4.9}$$

where constant $\kappa \geq 0$ is the speed of reversion, constant $\theta \in (-1, 1)$ is the long term mean level of ρ_t and ς is a positive constant.

Based on [van Emmerich, 2006]'s work, [Teng et al., 2016] improved the second approach, a more general stochastic correlation process, by modelling correlation as a hyperbolic function of any mean-reverting process. Their correlation process also satisfies the Properties 5 in Section 4.1. An interesting finding proposed by [Teng et al., 2016] is that the correlation process (4.9) can be derived as a special case of their models. In fact, [Teng et al., 2016] first defined a special mean-reverting process, then combined it with the function $\rho_t = \tanh(X_t)$. After applying Ito's Lemma and redefining the parameters, the authors derived the correlation process shown in (4.9). Readers can refer to [Teng et al., 2016] for further details.

A modified Jacobi process, which is a general case of the correlation process (4.9), was proposed by [Ma, 2009]. The authors assumed that the correlation ρ_t is a random walk following a square root process:

$$d\rho_t = \beta(\bar{\rho} - \rho_t)dt + \sqrt{\zeta(h - \rho_t)(\rho_t - f)}dW_t, \quad 1 \geq h \geq f \geq -1, \; h > \bar{\rho} > f \tag{4.10}$$

The correlation stochastic process ρ_t is centred around the equilibrium $\bar{\rho}_t$ with the speed of reversion β. The bound for correlation process (4.10) is $h \geq \rho_t \geq f$. When $h = 1, f = -1$, the process (4.10) is exactly the process (4.9).

According to [Wilmott,], the parameters of ρ_t should satisfy the following constraints to make ρ_t never reach its bounds:

$$\begin{array}{ll} \beta(\bar{\rho} - f) & > \zeta(h - f)/2 \\ \beta(h - \bar{\rho}) & > \zeta(h - f)/2 \end{array} \tag{4.11}$$

Boundary classifications In [van Emmerich, 2006], the authors discussed the relation between parameter values and boundary classifications. [van Emmerich, 2006] classified boundaries as attractive or unattractive boundaries. The attractive boundaries are further classified as attainable or unattainable boundaries. The definition of attractive boundaries is presented as follows:

Consider a stochastic process $dY_t = a(Y_t)dt + b(Y_t)dW_t, y_0 \in \mathbb{R}$ with left bound l and right bound r. Let $m(v) = exp(-\int_{v_0}^{v} \frac{2a(\omega)}{b^2(\omega)} d\omega)$. $v_0 \in (l, r)$ and has no relevance to y_0. The left bound l is called attractive if and only if there is a $y^* \in (l, r)$ such that $\lim_{a \to l} \int_a^{y^*} m(v)dv < \infty$. Otherwise it is unattractive.

Similarly, let $M(y) = \int_{y_0}^{y} m(v)dv$. The left bound l is attainable if and only if there is a $y^* \in (l, r)$ such that $\lim_{a \to l} \int_a^{y^*} \int_v^{y^*} \frac{1}{b^2(y)m(y)} dy dM(v) < \infty$. Otherwise it is unattainable. The analysis of the right bond are analogous.

In process (4.9), the boundaries of ρ_t are -1 and 1. By referring to [van Emmerich, 2006], we summarized that

1) The left bound -1 is attractive and attainable if $\frac{\kappa}{\varsigma}(\theta + 1) < 1$ otherwise the left bound is unattractive and unattainable;

2) The right bound 1 is attractive and attainable if $\frac{\kappa}{\varsigma}(1 - \theta) < 1$ otherwise the right bound is unattractive and unattainable;

3) Bounds -1 and 1 are unattainable if $\kappa \geq \frac{\varsigma}{1 \pm \theta}$;

4) The behaviour of bounds is symmetric with respect to θ;

5) Increasing the value of parameter κ concentrates the process around the mean.

Transition density [van Emmerich, 2006] used the Fokker-Planck equation to determine the transition density. Assume that ς in process (4.9) equals to 1. The stochastic differential equation

$$d\rho_t = a(t, \rho_t)dt + b(t, \rho_t)dW_t, \quad \rho_0 \in (-1, 1) \tag{4.12}$$

possesses a transition density $\varphi(t, \rho|\rho_0)$ where φ satisfies the Fokker-Planck equation

$$\frac{\partial}{\partial t}\varphi(t, \rho) + \frac{\partial}{\partial \rho}(a(t, \rho)\varphi(t, \rho)) - \frac{1}{2}\frac{\partial^2}{\partial \rho^2}(b(t, \rho)^2\varphi(t, \rho)) = 0 \tag{4.13}$$

[van Emmerich, 2006] wants to find the transition density of (4.9) when t goes to infinity. The φ is required to fulfill the following structural conditions:

1) $\int_{-1}^{1} \varphi(t, \rho)d\rho = 1$;

2) $\int_{-1}^{1} \varphi(t, \rho)d\rho \to \theta$, as $t \to \infty$.

To find a stationary solution $\varphi(\rho) = \lim_{t \to \infty} \varphi(t, \rho)$, [van Emmerich, 2006] first considered the process (4.9) with $\theta = 0$, then extended it to the general case where $\theta \neq 0$. The author

derived an unique solution

$$\varphi(\rho) = c(\frac{1-\rho}{1+\rho})^{-\kappa\cdot\theta}(1-\rho^2)^{\kappa-1} \tag{4.14}$$

with parameter c which makes $\int_{-1}^{1}\varphi(\rho) = 1$ hold. However, we are not able to compute the constant c analytically but it can be determined numerically.

[van Emmerich, 2006] also discussed some features of the transition densities:

1) φ is concentrated on [-1,1]

2) φ id symmetric with respect to θ in the sense of $\varphi^{\theta}(\rho) = \varphi^{-\theta}(-\rho)$.

3) If $\theta = 0$, the global maximum is attained at $\rho = 0$.

4) The probability mass vanished when getting away from the mean. In particular, it approaches zero in the boundary values.

4.3 New CSV Models and properties of CIX

In this section, we first define two Correlation Stochastic Volatility (CSV) models for $CIX_t = \rho_t$. Then we will describe an Euler discretization methodology and show some stylized facts.

4.3.1 Two CSV models

Recall that [Teng et al., 2016] specified a stochastic correlation process by combining a Vasicek process X_t with a mapping function $\rho_t = f(X_t)$ where $f(x) = \tanh(x)$:

$$d\rho_t = (1-\rho_t^2)\Big((\kappa(\eta - \operatorname{artanh}(\rho_t)) - \rho_t\varsigma^2)dt + \varsigma dW_t\Big) \tag{4.15}$$

where $t \geq 0$, $\rho_0 \in (-1,1)$, $\kappa, \varsigma > 0$, $\eta \in \mathbb{R}$ and W_t is a Brownian motion.

Now we consider the process (4.15) with stochastic volatility.

CSV model 1

$$\begin{cases} dX_t = \vartheta(\eta - X_t)dt + \sqrt{\upsilon_t}dW_t^X \\ d\upsilon_t = \kappa(\theta - \upsilon_t)dt + \varsigma\sqrt{\upsilon_t}dW_t^{\upsilon} \\ d\langle W^X, W^{\upsilon}\rangle_t = \xi_1 dt \end{cases} \tag{4.16}$$

where X_t is a mean-reverting process. In this thesis, $X_t = \operatorname{artanh}(CIX_t)$. η is the long term mean level of X_t, ϑ is the speed of reversion, and $\sqrt{\upsilon_t}$ is the instantaneous volatility of X_t. θ is the long term mean level of υ_t, κ is the speed of reversion of υ_t, and ς is the volatility of υ_t. W_t^X, W_t^{υ} are Brownian motions with correlation ξ_1.

Note that if the parameters of volatility satisfy the Feller condition $2\kappa\theta > \varsigma^2$, the process υ_t is strictly positive. In this thesis, we choose the mapping function $f = tanh(x)$ so we have $\rho_t = f(X_t) = \tanh(X_t)$. Also note that $f \in C^2(\mathbb{R})$ and

$$\rho_t \quad = f(X_t) : (-\infty, \infty) \to (-1, 1) \tag{4.17}$$

Recall that [Ma, 2009] developed a modified Jacobi process to model ρ_t,

$$d\rho_t = \beta(\bar{\rho} - \rho_t)dt + \sqrt{\zeta(h - \rho_t)(\rho_t - f)}dW_t, \quad 1 \geq h \geq f \geq -1, \ h > \bar{\rho} > f \tag{4.18}$$

where $\beta(\bar{\rho} - \rho_t)$ is a drift, $\sqrt{\zeta(h - \rho_t)(\rho_t - f)}$ is a diffusion. h and f are upper and lower bounds of ρ_t. The correlation stochastic process ρ_t is centred around the long term mean level, $\bar{\rho}_t$, with the speed of reversion, β.

In our CSV model 2, we consider the Jacobi process (4.18) with stochastic volatility.

CSV model 2

$$\begin{cases} d\rho_t = \beta(\bar{\rho} - \rho_t)dt + \sqrt{\zeta_t(h - \rho_t)(\rho_t - f)}dW_t^{\rho} \\ d\zeta_t = \kappa(\theta - \zeta_t)dt + \varsigma\sqrt{\zeta_t}dW_t^{\zeta} \\ d\langle W^{\rho}, W^{\zeta}\rangle_t = \xi_2 dt \end{cases} \tag{4.19}$$

where ρ_t is the stochastic correlation process. $\bar{\rho}$ is the long term mean level of ρ_t, and $\sqrt{\zeta_t}$ is the stochastic volatility. h and f are upper and lower bounds of ρ_t. θ is the long term mean level of ζ, κ is the speed of reversion of ζ, and ς is the volatility of volatility. W_t^{ρ}, W_t^{ζ} are Brownian motions with correlation ξ_2.

Feller condition on the volatility process ensures positivity.When $h = 1, f = -1$, the constraint used to prevent ρ_t from hitting its bounds is $\zeta_t < \beta \min\{(\bar{\rho}+1), (1-\bar{\rho})\}$ (see [Ma, 2009] for more on this). Since $\bar{\rho} \approx 0.5$ and $\beta \approx 1.1$ in our analysis, the volatility ζ_t is supposed to be smaller than 0.5. Empirically, the probability of $\max_{0\leq t\leq 2}\{\zeta_t\} \geq 0.5$ is almost zero thus the correlation ρ_t rarely hits its bounds.

4.3.2 Discretization and Stylized facts.

In this section we use Euler-Maruyama method to do a discretization of our advanced CSV models. First we introduce some additional parameters: t_0 is the time origin; T is the horizon of simulation; N is the number of simulation steps; $\Delta t = T/N, \Delta t > 0$ is the partition size and partition points are $t_i = i\Delta t, i = 0, 1, ..., N$. $\Delta W_{t_i} = W_{t_{i+1}} - W_{t_i}, i = 0, 1, ..., N$ are independent and identically distributed normal random variables with mean 0 and variance Δt. Without loss of generality, we choose $t_0 = 0, T = 2, N = 1000$ in this thesis.

Discretization of CSV model 1

Since the correlation of W_t^X and W_t^{υ} is ξ_1, we can rewrite formula (4.16) as follows:

$$\begin{cases} dX_t = \vartheta(\eta - X_t)dt + \sqrt{\upsilon_t}(\xi_1 dW_t^{\upsilon} + \sqrt{1 - \xi_1^2}dW_t^1) \\ d\upsilon_t = \kappa(\theta - \upsilon_t)dt + \varsigma\sqrt{\upsilon_t}dW_t^{\upsilon} \end{cases} \tag{4.20}$$

where W_t^{υ} and W_t^1 are independent Brownian motions.

Algorithm 3 (Algorithm for simulation of correlation (4.16)) *Based on [Lord et al., 2010]'s full truncation method, the discretization for the squared volatility process υ_t in formula (4.20) is:*

$$\upsilon_{t_{i+1}} = \upsilon_{t_i} + \kappa(\theta - \max(0, \upsilon_{t_i}))\Delta t + \varsigma\sqrt{\max(0, \upsilon_{t_i})}\Delta W^{\upsilon}_{t_i}, \quad i = 0, 1, ..., N \tag{4.21}$$

where $\Delta W^{\upsilon}_{t_i}, i = 0, 1, ..., N$ are iid normal random variables with mean 0 and variance Δt.

Now given the value of $\kappa, \theta, \varsigma, \upsilon_0$, we can use Euler-Maruyama method to obtain nonnegative variance vector over time:

$$\boldsymbol{\upsilon} = (\upsilon_0, \upsilon_1, ..., \upsilon_T) \tag{4.22}$$

Further, given the value of ϑ, η, X_0, we can use Euler-Maruyama method to simulate the underlying process X_t over $[0, T]$:

$$X_{t_{i+1}} = X_{t_i} + \vartheta(\eta - X_{t_i})\Delta t + \sqrt{\max(0, \upsilon_{t_i})}(\xi_1\Delta W^{\upsilon}_{t_i} + \sqrt{1 - \xi_1^2}\Delta W^{1}_{t_i}), \quad i = 0, 1, ..., N \tag{4.23}$$

where $\Delta W^{1}_{t_i}, i = 0, 1, ..., N$ are iid normal random variables with mean 0 and variance Δt. $\Delta W^{1}_{t_i}$ should be independent with $\Delta W^{\upsilon}_{t_i}$.

Applying the mapping function $f(x) = tanh(x)$ to an underlying process X_t, the correlation process ρ_t is therefore:

$$\rho_{t_i} = f(X_{t_i}) = tanh(X_{t_i}), \quad i = 0, 1, ..., N \tag{4.24}$$

where $\rho_{t_i} \in (-1, 1)$.

For exemplary purposes, we choose the value of parameters $\kappa, \theta, \varsigma, \upsilon_0$ as in [Heston, 1993], see Table 4.2. Besides, we keep $\upsilon_0 = \theta$ for varying θ and $X_0 = \eta$ for varying η to identify the mean reversion in our diagrams. Figure 4.1 shows the correlation process with three varying volatility parameters: Speeds of reversion (κ), long term mean level (θ) and volatilities of volatility parameter (ς). All other parameters are held constant as shown in Table 4.2. Besides, we hold $\upsilon_0 = \theta$ for different long-run variances (θ) in the middle panel. To keep consistency, we use the same black curve in each panel to depict the process of ρ_t with default parameter values.

Model	Parameter	Value
Advanced model 1	Initial volatility	$\sqrt{v_0} = 0.1$
	Speed of reversion of the squared volatility parameter	$\kappa = 2$
	Long term mean level of the squared volatility parameter	$\theta = 0.01$
	Volatility of the squared volatility parameter	$\varsigma = 0.1$
	Initial value of the underlying process	$X_0 = 0.3$
	Long term mean level of the underlying process	$\eta = 0.3$
	Speed of reversion of the underlying process	$\vartheta = 1.1$
	Correlation of W_t^X and W_t^v	$\xi_1 = -0.7$
Advanced model 2	Initial volatility	$\sqrt{\zeta_0} = 0.1$
	Speed of reversion of the squared volatility parameter	$\kappa = 2$
	Long term mean level of the squared volatility parameter	$\theta = 0.01$
	Volatility of the squared volatility parameter	$\varsigma = 0.1$
	Initial value of correlation process	$\rho_0 = 0.3$
	Long term mean level of the correlation parameter	$\bar{\rho} = 0.3$
	Speed of reversion of the correlation parameter	$\beta = 1.1$
	Bounds of the correlation parameter	$h = 1, f = -1$
	Correlation of W_t^ρ and W_t^ζ	$\xi_2 = -0.7$

Table 4.2: Default parameters for discretization CSV model 1 & 2

Discretization of CSV model 2

Since the correlation of W_t^ρ and W_t^ζ is ξ_2, we rewrite formula (4.19) as follows:

$$\begin{cases} d\rho_t = \beta(\bar{\rho} - \rho_t)dt + \sqrt{\zeta_t(h-\rho_t)(\rho_t - f)}(\xi_2 dW_t^\zeta + \sqrt{1-\xi_2^2}dW_t^2) \\ d\zeta_t = \kappa(\theta - \zeta_t)dt + \varsigma\sqrt{\zeta_t}dW_t^\zeta \end{cases} \tag{4.25}$$

where W_t^ζ and W_t^2 are independent Brownian motions.

Algorithm 4 (Algorithm for simulation of correlation (4.19)) *The discretization for volatility ζ in formula (4.25) is based on [Lord et al., 2010]'s full truncation method:*

$$\zeta_{t_{i+1}} = \zeta_{t_i} + \kappa(\theta - \max(0, \zeta_{t_i}))\Delta t + \varsigma\sqrt{\max(0, \zeta_{t_i})}\Delta W_{t_i}^\zeta, \quad i = 0, 1, ..., N \tag{4.26}$$

where $\Delta W_{t_i}^\zeta, i = 0, 1, ..., N$ are iid normal random variables with mean 0 and variance Δt.

Given the value of $\kappa, \theta, \varsigma, \zeta_0$, we can use Euler-Maruyama method to obtain nonnegative variance vector over time:

$$\boldsymbol{\zeta} = (\zeta_0, \zeta_1, ..., \zeta_T) \tag{4.27}$$

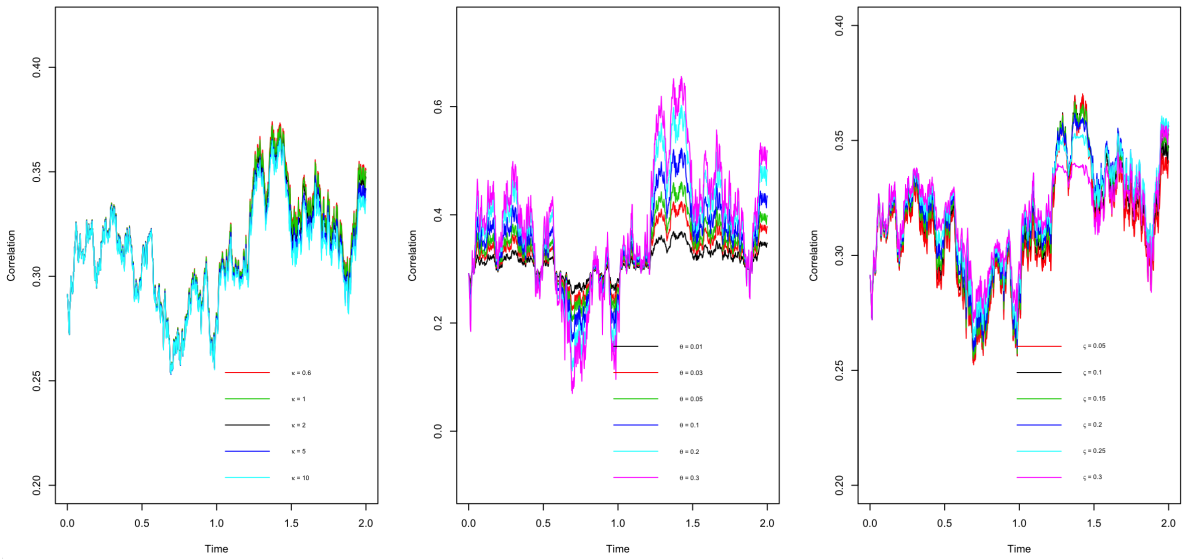

Figure 4.1: Correlation process ρ_t on CSV model 1 for three varying volatility parameters: Speeds of reversion (κ) is on the left, Long term mean level (θ) is in the middle, and Volatilities of volatility parameter (ς) is on the right.

.

Then, we set the following function to make $(h - g(\rho_{t_i}))(g(\rho_{t_i}) - f), i = 0, 1, ..., N$ *positive over* $[0, T]$.

$$g(\rho_{t_i}) = \begin{cases} -1 & \text{if } \rho_{t_i} \leq -1 \\ 1 & \text{if } \rho_{t_i} \geq 1 \\ \rho_{t_i} & \text{otherwise} \end{cases} \tag{4.28}$$

Hence, the correlation process ρ_t *in formula (4.25) can be discretized via Euler-Maruyama method with the given value of* $h, f, \bar{\rho}, \beta, \rho_0$. *For* $i = 0, 1, ..., N$, *we have*

$$\begin{aligned} \rho_{t_{i+1}} &= g(\rho_{t_i}) + \beta(\bar{\rho} - g(\rho_{t_i}))\Delta t \\ &+ \sqrt{\zeta_{t_i}(h - g(\rho_{t_i}))(g(\rho_{t_i}) - f)}(\xi_2 \Delta W^{\zeta}_{t_i} + \sqrt{1 - \xi_2^2}\Delta W^2_{t_i}), \end{aligned} \tag{4.29}$$

where $\Delta W^2_{t_i}, i = 0, 1, ..., N$ *are iid normal random variables with mean 0 and variance* Δt. $\Delta W^2_{t_i}$ *is independent with* $\Delta W^{\zeta}_{t_i}$

We choose the value of parameters $\kappa, \theta, \varsigma, \zeta_0$ by referring to [Heston, 1993], see Table 4.2. $\zeta_0 = \theta$ is held for varying θ, and $\rho_0 = \bar{\rho}$ is held for varying $\bar{\rho}$ to display the mean reversion in our diagrams. Figure 4.2 shows the process of correlation with varying κ, θ and ς respectively. For consistency, we use the same black curve in each panel to present the process of the correlation ρ_t with default parameter values.

Even though patterns in Figures 4.1 and 4.2 look similar, there are slight differences between CIX simulated by these advanced CSV model 1 and CIX simulated by these advanced CSV model 2. To be specific, the CIX estimated by the CSV model 2 is relative greater than its counterpart estimated by the CSV model 1, and the difference is up to 3%. However, these slight differences will be enlarged in pricing, see Figures 5.2, 5.3,5.4 and 5.5.

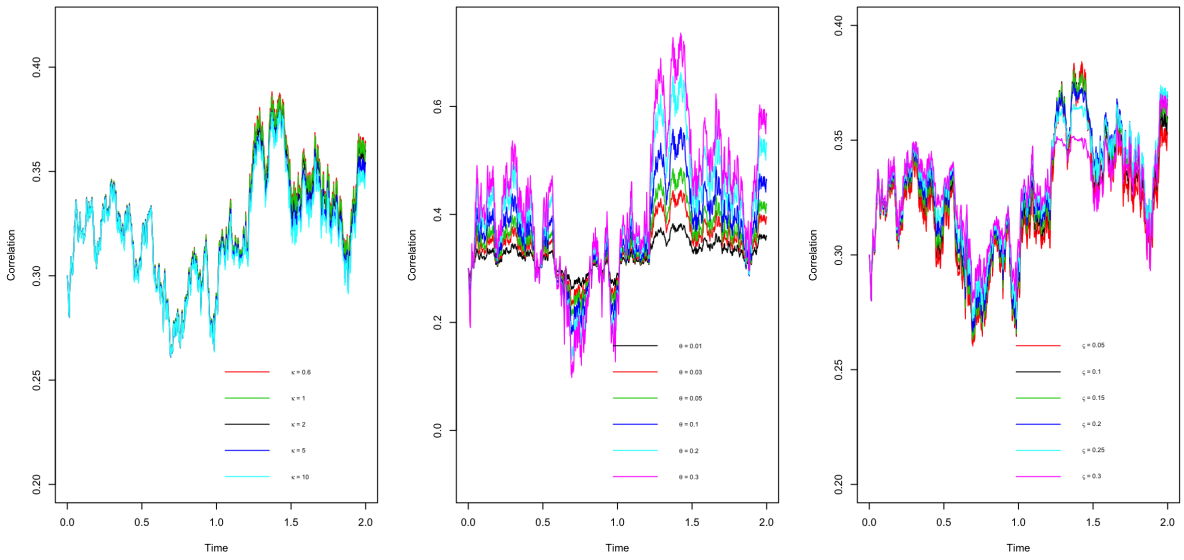

Figure 4.2: Correlation process ρ_t on CSV model 2 with three varying volatilities parameters: Speeds of reversion (κ) is on the left, Long term mean level (θ) is in the middle, and Volatilities of volatility parameter (ς) is on the right. For each panel, all other things being equal, only the parameter of interest can vary. Besides, we hold $\zeta_0 = \theta$ for varying θ in the middle panel.

.

Stylized facts

In this section, we show the volatility of ρ_t, denoted by σ_{ρ_t}, and the leverage for both models using Ito's lemma.

Stylized Facts of CSV model 1 Since we consider $\rho_t = f(X_t) = \tanh(X_t)$ in CSV model 1, a stochastic differential equation for ρ_t can be derived by Ito's lemma

$$\begin{aligned} d\rho_t &= d\tanh(X_t) \\ &= (1-\rho_t^2)\Big((\vartheta(\eta - \text{artanh}(\rho_t)) - \rho_t \upsilon_t)dt + \sqrt{\upsilon_t}dW_t^X\Big) \end{aligned} \tag{4.30}$$

From formula (4.30), we obtain $\sigma_{\rho_t} = (1-\rho_t^2)\sqrt{\upsilon_t}$. By Ito's lemma,

$$\begin{aligned} d\sigma_{\rho_t}^2 &= d(\tfrac{\partial f(X_t)}{\partial X_t}\sqrt{\upsilon_t})^2 \\ &= \upsilon_t dm(X_t) + m(X_t)d\upsilon_t + d\upsilon_t dm(X_t) \end{aligned} \tag{4.31}$$

where $m(X_t) := (\frac{\partial f(X_t)}{\partial X_t})^2$. Moreover, we have the following lemma:

Lemma 4.3.1

$$\begin{aligned} \upsilon_t dm(X_t) &= \Big(\tfrac{\partial m(X_t)}{\partial X_t}\vartheta(\eta - X_t)\upsilon_t + \tfrac{1}{2}\tfrac{\partial^2 m(X_t)}{\partial X_t^2}\upsilon_t^2\Big)dt + \tfrac{\partial m(X_t)}{\partial X_t}\upsilon_t^{3/2}dW_t^X \\ m(X_t)d\upsilon_t &= \kappa(\theta - \upsilon_t)m(X_t)dt + \varsigma\sqrt{\upsilon_t}m(X_t)dW_t^\upsilon \\ d\upsilon_t dm(X_t) &= \tfrac{\partial m(X_t)}{\partial X_t}\varsigma\xi_1\upsilon_t dt \end{aligned} \tag{4.32}$$

Hence,

$$\begin{aligned} d\sigma^2_{\rho_t} &= \Big(\kappa(\theta-\upsilon_t)m(X_t) + (\varsigma\xi_1 + \vartheta(\eta - X_t))\tfrac{\partial m(X_t)}{\partial X_t}\upsilon_t + \tfrac{1}{2}\tfrac{\partial^2 m(X_t)}{\partial X_t^2}\upsilon_t^2\Big)dt + \tfrac{\partial m(X_t)}{\partial X_t}\upsilon_t^{3/2}dW_t^X + \varsigma\sqrt{\upsilon_t}m(X_t)dW_t^\upsilon \\ &= (1-\rho_t^2)^2\Big((\kappa(\theta-\upsilon_t) - 4\rho_t\upsilon_t(\varsigma\xi_1 + \vartheta(\eta - \operatorname{artanh}(\rho_t))) + \upsilon_t^2(10\rho_t^2-2))dt - 4\rho_t\upsilon_t^{3/2}dW_t^X + \varsigma\sqrt{\upsilon_t}dW_t^\upsilon\Big) \end{aligned} \tag{4.33}$$

From equations (4.30) and (4.33), we obtain

$$\begin{aligned} d\langle\rho,\sigma^2_\rho\rangle_t &= (-4\rho_t\upsilon_t + \varsigma\xi_1)(1-\rho_t^2)^3\upsilon_t dt \\ \sqrt{d\langle\rho\rangle_t} &= \sqrt{\upsilon_t(1-\rho_t^2)^2 dt} \\ \sqrt{d\langle\sigma^2_\rho\rangle_t} &= \sqrt{\Big(16\rho_t^2\upsilon_t^2 + \varsigma^2 - 8\xi_1\varsigma\rho_t\upsilon_t\Big)(1-\rho_t^2)^4\upsilon_t dt} \end{aligned} \tag{4.34}$$

where $\langle,\rangle$ is quadratic variation (covariation). Therefore, the leverage of CSV model 1 is given by

$$\begin{aligned} Leverage &= \frac{d\langle\rho,\sigma^2_\rho\rangle_t}{\sqrt{d\langle\rho\rangle_t}\sqrt{d\langle\sigma^2_\rho\rangle_t}} \\ &= \frac{-4\rho_t\upsilon_t + \varsigma\xi_1}{\sqrt{16\rho_t^2\upsilon_t^2 + \varsigma^2 - 8\xi_1\varsigma\rho_t\upsilon_t}} \end{aligned} \tag{4.35}$$

Stylized Facts of CSV model 2 From formula (4.19),

$$d\rho_t = \beta(\bar\rho - \rho_t)dt + \sqrt{\zeta_t(h-\rho_t)(\rho_t - f)}dW_t^\rho \tag{4.36}$$

the volatility of ρ_t is $\sigma_{\rho_t} = \sqrt{\zeta_t(h-\rho_t)(\rho_t-f)}$. For simplicity, we define $\psi(\rho_t) := (h-\rho_t)(\rho_t - f)$ such that $\sigma^2_{\rho_t} = \zeta_t\psi(\rho_t)$. Applying Ito's lemma to $\sigma^2_{\rho_t}$, we obtain

$$\begin{aligned} d\sigma^2_{\rho_t} &= d(\zeta_t\psi(\rho_t)) \\ &= \zeta_t d\psi(\rho_t) + \psi(\rho_t)d\zeta_t + d\psi(\rho_t)d\zeta_t \end{aligned} \tag{4.37}$$

Next, we consider $\zeta_t d\psi(\rho_t)$, $\psi(\rho_t)d\zeta_t$ and $d\psi(\rho_t)d\zeta_t$ respectively. Since we have the following lemma

Lemma 4.3.2

$$\begin{aligned} \psi(\rho_t)d\zeta_t &= \kappa(\theta-\zeta_t)\psi(\rho_t)dt + \varsigma\sqrt{\zeta_t}\psi(\rho_t)dW_t^\zeta \\ \zeta_t d\psi(\rho_t) &= \Big(\beta(\bar\rho-\rho_t)\tfrac{\partial\psi(\rho_t)}{\partial\rho_t}\zeta_t + \tfrac{1}{2}\tfrac{\partial^2\psi(\rho_t)}{\partial\rho_t^2}\psi(\rho_t)\zeta_t^2\Big)dt + \tfrac{\partial\psi(\rho_t)}{\partial\rho_t}\sqrt{\psi(\rho_t)}\zeta_t^{3/2}dW_t^\rho \\ d\zeta_t d\psi(\rho_t) &= (\varsigma\sqrt{\zeta_t}dW_t^\zeta)(\tfrac{\partial\psi(\rho_t)}{\partial\rho_t}\sqrt{\psi(\rho_t)\zeta_t}dW_t^\rho) = \tfrac{\partial\psi(\rho_t)}{\partial\rho_t}\sqrt{\psi(\rho_t)}\varsigma\xi_2\zeta_t dt \end{aligned} \tag{4.38}$$

the stochastic differential equation for $\sigma^2_{\rho_t}$ is given by

$$\begin{aligned} d\sigma^2_{\rho_t} &= \Big(\kappa(\theta-\zeta_t)\psi(\rho_t) + (\varsigma\xi_2\sqrt{\psi(\rho_t)} + \beta(\bar\rho-\rho_t))\tfrac{\partial\psi(\rho_t)}{\partial\rho_t}\zeta_t + \tfrac{1}{2}\tfrac{\partial^2\psi(\rho_t)}{\partial\rho_t^2}\psi(\rho_t)\zeta_t^2\Big)dt \\ &+ \tfrac{\partial\psi(\rho_t)}{\partial\rho_t}\sqrt{\psi(\rho_t)}\zeta^{3/2}dW_t^\rho + \varsigma\sqrt{\zeta_t}\psi(\rho_t)dW_t^\zeta \end{aligned} \tag{4.39}$$

From equations (4.36) and (4.39), we obtain

$$\begin{aligned} d\langle\rho,\sigma^2_\rho\rangle_t &= \Big(\tfrac{\partial\psi(\rho_t)}{\partial\rho_t}\zeta_t + \xi_2\varsigma\psi^{1/2}(\rho_t)\Big)\psi(\rho_t)\zeta_t dt \\ \sqrt{d\langle\rho\rangle_t} &= \sqrt{\zeta_t\psi(\rho_t)dt} \\ \sqrt{d\langle\sigma^2_\rho\rangle_t} &= \sqrt{\Big((\tfrac{\partial\psi(\rho_t)}{\partial\rho_t})^2\zeta_t^2 + \varsigma^2\psi(\rho_t) + 2\xi_2\varsigma\tfrac{\partial\psi(\rho_t)}{\partial\rho_t}\zeta_t\psi^{1/2}(\rho_t)\Big)\psi(\rho_t)\zeta_t dt} \end{aligned} \tag{4.40}$$

Therefore, the leverage for CSV model 2 is given by

$$\begin{aligned} Leverage &= \frac{d\langle \rho, \sigma_\rho^2 \rangle_t}{\sqrt{d\langle \rho \rangle_t} \sqrt{d\langle \sigma_\rho^2 \rangle_t}} \\ &= \frac{\frac{\partial \psi(\rho_t)}{\partial \rho_t} \zeta_t + \xi_2 \varsigma \psi^{1/2}(\rho_t)}{\sqrt{(\frac{\partial \psi(\rho_t)}{\partial \rho_t})^2 \zeta_t^2 + \varsigma^2 \psi(\rho_t) + 2\xi_2 \varsigma \frac{\partial \psi(\rho_t)}{\partial \rho_t} \zeta_t \psi^{1/2}(\rho_t)}} \end{aligned} \tag{4.41}$$

See proof of lemma in Appendix.

4.4 Additional approach to model SCP

The Wishart Affine Stochastic Correlation (WASC) model is a new continuous time process that can be considered as a multivariate extension of the [Heston, 1993] model, with a more accurate correlation structure. The framework of this model was introduced in [Gourieroux and Sufana, 2003]. It relies on the following assumption.

Assumption 1 The evolution of asset returns is conditionally Gaussian while the stochastic variance -covariance matrix follows a Wishart process.

In formulas, we consider a n-dimensional risky asset S_t whose risk-neutral dynamics are given by

$$dS_t = \text{diag}[S_t](\mu dt + \sqrt{\Sigma_t} dZ_t) \tag{4.42}$$

where μ is the vector of returns and $Z_t \in \mathbb{R}^n$ is a vector Brownian motion. The variance-covariance matrix of the risky assets is denoted by $\sum_t$ which is assumed to satisfy the following dynamics:

$$d\Sigma_t = (\Omega\Omega^\top + M\Sigma_t + \Sigma_t M^\top)dt + \sqrt{\Sigma_t} dW_t Q + Q^\top (dW_t)^\top \sqrt{\Sigma_t} \tag{4.43}$$

with $\Omega, M, Q \in M_n$, Ω invertible, and $W_t \in M_n$ a matrix Brownian motion. In [Da Fonseca et al., 2014], authors assume that the above dynamics are inferred from observed asset price time series, hence the stochastic differential equation is written under the historical measure.

(4.43) characterizes the Wishart process introduced by [Bru, 1991]. Note that in terms of factor analysis, this model not only allows the stochastic evolution of principal components (eigenvalues of the covariance matrix) but also a stochastic evolution of the factor loadings of observed factors with respect to latent factors.

In order to ensure the strict positivity and the typical mean-reverting feature of the volatility, the matrix M is assumed to be negative semi-definite while Ω satisfies

$$\Omega\Omega^\top = \beta\Omega^\top\Omega \tag{4.44}$$

with the real parameter $\beta > n - 1$ (see [Bru, 1991]).

In full analogy with square-root process, the term $\Omega\Omega^\top$ is related to the expected long-term variance -covariance matrix Σ_∞ through the solution to the following linear equation:

$$\Omega\Omega^\top = M\Sigma_\infty + \Sigma_\infty M^\top \tag{4.45}$$

Moreover, Q is the volatility of the volatility matrix, and its parameters will be crucial in order to explain some stylized observed effects in equity markets.

If there is negative correlation between the noise driving the returns and the noise driving their variance, it is possible to approximately reproduce observed negative skewness within the [Heston, 1993] model. This is why [Da Fonseca et al., 2007] using the following assumption to prove the linear relationship among the noises:

Assumption 2 The Brownian motions of the asset returns and those driving the covariance matrix are linearly correlated.

[Da Fonseca et al., 2007] proved that Assumption 2 leads to the following relation:

$$dZ_t = dW_t\eta + \sqrt{1-\eta^\top\eta}dB_t \tag{4.46}$$

with $dZ_t = (dZ_1, dZ_2, ..., dZ_n)^\top$. B is a vector of independent Brownian motions orthogonal to W.

[Da Fonseca et al., 2007] and [Da Fonseca et al., 2014] show in the special case of two assets (n=2), for which the variance- covariance matrix is given by

$$\Sigma_t = \begin{bmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{12} & \Sigma_t^{22} \end{bmatrix} \tag{4.47}$$

(4.46) is written by:

$$\begin{aligned} dZ_t^1 &= \sqrt{1-(\eta_1^2+\eta_2^2)}dB_t^1 + (dW_t^{11}\eta_1 + dW_t^{12}\eta_2) \\ dZ_t^2 &= \sqrt{1-(\eta_1^2+\eta_2^2)}dB_t^2 + (dW_t^{21}\eta_1 + dW_t^{22}\eta_2) \end{aligned} \tag{4.48}$$

The correlation between assets' returns and their volatilities admit a closed form expression , highlighting the impact of the η parameters on its value and positivity:

$$\begin{aligned} corr(dlogS_1, d\Sigma^{11}) &= \frac{\eta_1 Q_{11}+\eta_2 Q_{21}}{\sqrt{Q_{11}^2+Q_{21}^2}} \\ corr(dlogS_2, d\Sigma^{22}) &= \frac{\eta_1 Q_{12}+\eta_2 Q_{22}}{\sqrt{Q_{12}^2+Q_{22}^2}} \end{aligned} \tag{4.49}$$

where we recall that $\sqrt{\Sigma^{11}}$ represents the volatility of the first asset. Therefore, the sign and magnitude of the leverage effects are determined by both the matrix Q and the vector η.

It is well known that in two asset case ($n = 2$).

$$\rho_t^{12} = \frac{\Sigma_t^{12}}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}} \tag{4.50}$$

Differentiating $\rho_t^{12}\sqrt{\Sigma_t^{11}\Sigma_t^{22}} = \Sigma_t^{12}$ we obtain

$$\sqrt{\Sigma_t^{11}\Sigma_t^{22}}d\rho_t^{12} + \rho_t^{12}d(\sqrt{\Sigma_t^{11}\Sigma_t^{22}}) + (.)dt = d\Sigma_t^{12} \tag{4.51}$$

and

$$\sqrt{\Sigma_t^{11}\Sigma_t^{22}}d\rho_t^{12} + \rho_t^{12}(1/2\sqrt{\frac{\Sigma_t^{22}}{\Sigma_t^{11}}}d\Sigma_t^{11} + 1/2\sqrt{\frac{\Sigma_t^{11}}{\Sigma_t^{22}}}d\Sigma_t^{22} + (.)dt = d\Sigma_t^{12} \tag{4.52}$$

then

$$d\rho_t^{12} = \frac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}d\Sigma_t^{12} - 1/2\rho_t^{12}(\frac{1}{\Sigma_t^{11}}d\Sigma_t^{11} + \frac{1}{\Sigma_t^{22}}d\Sigma_t^{22}) + (.)dt \tag{4.53}$$

$$d\rho_t^{12} = \frac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}(d\Sigma_t^{12} - \frac{\rho_t^{12}}{2\Sigma_t^{11}}d\Sigma_t^{11} - \frac{\rho_t^{12}}{2\Sigma_t^{22}}d\Sigma_t^{22} + (.)dt \tag{4.54}$$

By using the covariation among the Wishart elements we have

$$\begin{aligned} d\langle\rho^{12}\rangle_t = \; & \tfrac{1}{\Sigma_t^{11}\Sigma_t^{22}}[\Sigma_t^{11}(Q_{12}^2+Q_{22}^2) + 2\Sigma_t^{12}(Q_{11}Q_{12}+Q_{21}Q_{22}) + \Sigma_t^{22}(Q_{11}^2+Q_{21}^2)] \\ & +(\Sigma_t^{12})^2(\tfrac{Q_{11}^2+Q_{21}^2}{\Sigma_t^{11}} + \tfrac{Q_{12}^2+Q_{22}^2}{\Sigma_t^{22}} \\ & +2\tfrac{\Sigma_t^{12}}{\Sigma_t^{11}\Sigma_t^{22}}(Q_{11}Q_{12}+Q_{21}Q_{22})) - 2\tfrac{\Sigma_t^{12}}{\Sigma_t^{11}}(\Sigma_t^{11}(Q_{11}Q_{12}+Q_{21}Q_{22}) + \Sigma_t^{12}(Q_{11}^2+Q_{21}^2)) \\ & -2\tfrac{\Sigma_t^{12}}{\Sigma_t^{22}}(\Sigma_t^{22}(Q_{11}Q_{12}+Q_{21}Q_{22}) + \Sigma_t^{12}(Q_{12}^2+Q_{22}^2))dt \end{aligned} \tag{4.55}$$

which leads to

$$d\langle\rho^{12}\rangle_t = (1-(\rho_t^{12})^2)(\frac{Q_{12}^2+Q_{22}^2}{\Sigma_t^{22}} + \frac{Q_{11}^2+Q_{21}^2}{\Sigma_t^{11}} - 2\frac{\rho_t^{12}(Q_{11}Q_{12}+Q_{21}Q_{22})}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}})dt \tag{4.56}$$

Now let us compute the drift of the process ρ_t^{12}.

We differentiate both sides of the equality $\rho_t^{12} = \frac{\Sigma_t^{12}}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}$ and we consider the finite variation terms:

$$\begin{aligned} d\rho_t^{12} = \; & \tfrac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}d\Sigma_t^{12} + \Sigma_t^{12}d(\tfrac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}) + d\langle\Sigma_t^{12}, \tfrac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}\rangle_t \\ & = \tfrac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}(\Omega_{11}\Omega_{21} + \Omega_{12}\Omega_{22} + M_{21}\Sigma_t^{11} + M_{12}\Sigma_t^{22} + (M_{11}+M_{22})\Sigma_t^{12})dt \\ & +\Sigma_t^{12}[\tfrac{1}{\sqrt{\Sigma_t^{22}}}(-\tfrac{1}{2\sqrt{\Sigma_t^{11^3}}})(\Omega_{11}^2\Omega_{12}^2 + 2M_{11}\Sigma_t^{11} + 2M_{12}\Sigma_t^{12}) \\ & + \tfrac{1}{\sqrt{\Sigma_t^{11}}}(-\tfrac{1}{2\sqrt{\Sigma_t^{22^3}}})(\Omega_{21}^2\Omega_{22}^2 + 2M_{21}\Sigma_t^{12} + 2M_{22}\Sigma_t^{22}) \\ & + \tfrac{3}{8\sqrt{\Sigma_t^{11}\Sigma_t^{22}}(\Sigma_t^{11})^2}d\langle\Sigma^{11}\rangle_t + \tfrac{3}{8\sqrt{\Sigma_t^{11}\Sigma_t^{22}}(\Sigma_t^{22})^2}d\langle\Sigma^{22}\rangle_t \\ & + \tfrac{1}{4\sqrt{\Sigma_t^{11}\Sigma_t^{22^3}}}d\langle\Sigma^{11}\Sigma^{22}\rangle_t]dt + \tfrac{1}{\sqrt{\Sigma_t^{22}}}(-\tfrac{1}{2\sqrt{\Sigma_t^{11^3}}})d\langle\Sigma^{11}\Sigma^{12}\rangle_t \\ & + \tfrac{1}{\sqrt{\Sigma_t^{11}}}(-\tfrac{1}{2\sqrt{\Sigma_t^{22^3}}})d\langle\Sigma^{12}\Sigma^{22}\rangle_t + \sqrt{\tfrac{d\langle\rho^{12}\rangle_t}{dt}}dW \end{aligned} \tag{4.57}$$

Now we use the formulas of the covariation of the Wishart elements and we arrive to an expression which can be written as follows:

$$\begin{aligned} d\rho_t^{12} \; & = (A_t(\rho_t^{12})^2 + B_t\rho_t^{12} + C_t)dt \\ & + \sqrt{(1-(\rho_t^{12})^2)(\tfrac{Q_{12}^2+Q_{22}^2}{\Sigma_t^{22}} + \tfrac{Q_{11}^2+Q_{21}^2}{\Sigma_t^{11}} - 2\tfrac{\rho_t^{12}(Q_{11}Q_{12}+Q_{21}Q_{22})}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}})}dW \end{aligned} \tag{4.58}$$

where:

$$\begin{aligned} A_t \; & = \tfrac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}(Q_{11}Q_{12}+Q_{21}Q_{22}) - \sqrt{\tfrac{\Sigma_t^{22}}{\Sigma_t^{11}}}M_{12} - \sqrt{\tfrac{\Sigma_t^{11}}{\Sigma_t^{22}}}M_{21} \\ B_t \; & = -\tfrac{\Omega_{11}^2+\Omega_{12}^2}{2\Sigma_t^{11}} - \tfrac{\Omega_{21}^2+\Omega_{22}^2}{2\Sigma_t^{22}} + \tfrac{Q_{11}^2+Q_{21}^2}{2\Sigma_t^{11}} + \tfrac{Q_{12}^2+Q_{22}^2}{2\Sigma_t^{22}} < 0 \\ C_t \; & = \tfrac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}(\Omega_{11}\Omega_{21} + \Omega_{12}\Omega_{22} - 2(Q_{11}Q_{12}+Q_{21}Q_{22})) + \sqrt{\tfrac{\Sigma_t^{22}}{\Sigma_t^{11}}}M_{12} + \sqrt{\tfrac{\Sigma_t^{11}}{\Sigma_t^{22}}}M_{21} \end{aligned} \tag{4.59}$$

From the definition of $\Omega = \sqrt{\beta}Q^\top$ and the Gindikin condition we deduce that B_t is negative. As a by-product, we easily deduce the instantaneous covariation between the Wishart element Σ_t^{11} and the correlation process:

$$\begin{aligned} d\langle\rho^{12}, \Sigma^{11}\rangle_t &= \frac{1}{\sqrt{\Sigma_t^{11}\Sigma_t^{22}}}(d\langle\Sigma^{11}\Sigma^{11}\rangle_t - \frac{\Sigma_t^{12}}{2\Sigma_t^{22}}d\langle\Sigma^{22}, \Sigma^{22}\rangle_t) \\ &= 2\frac{\Sigma_t^{11}}{\Sigma_t^{22}}(1-(\rho_t^{12})^2)(Q_{11}Q_{12}+Q_{21}Q_{22})dt. \end{aligned} \tag{4.60}$$

[Da Fonseca et al., 2007] also gave the following proposition:

$$d\langle\rho^{12}, S^{(i)}\rangle_t = (\frac{\Sigma_t^{ii}}{\Sigma_t^{jj}}(1-(\rho^{12})^2)Tr[R_jQ])dt \; i, j = 1, 2. \tag{4.61}$$

where S is the vector of asset price.

4.5 Stochastic modelling of HIX

[Guillaume and Linders, 2015] presented some different plausible stochastic models for HIX by referring to [van Emmerich, 2006] and [Teng et al., 2013]. Their HIX models are obtained by combining some mean-reverting processes with some mapping functions. These functions can map the domain of a mean-reverting process to an unit interval. The such obtained Ito processes preserve the mean-reverting trend of the underlying process while satisfying the fundamental properties of the HIX index which has [0,1] as domain. Even though the Brownian motion is not a mean-reverting process, [Guillaume and Linders, 2015] combined it with a mapping function to take the advantage of an analytical expression for the conditional transition probability. [Guillaume and Linders, 2015] predicted HIX with parameters inferred from maximizing the conditional transition density. Another important result in [Guillaume and Linders, 2015] is that for a given underlying Ito process, the choice of mapping functions does not really matter. Readers can consult the numerical study in [Guillaume and Linders, 2015] for more information.

Let us consider an Ito process defined as (4.1) and a mapping function $g(x)$. $g(x)$ is twice continuously differentiable on $\mathbb{R}$. The Vasicek process and CIR process are defined as process (4.15) and process (4.6).

$$\begin{aligned} g_1(x) &:= \tanh(x) \\ g_2(x) &:= \tfrac{\tanh(x)+1}{2} \\ g_3(x) &:= 1-\exp(-x) \\ g_4(x) &:= \tfrac{1}{1+\exp(-x)} \end{aligned} \tag{4.62}$$

Combing mapping functions in (4.62) with two mean-reverting processes leads to four model specifications as shown in Table 4.3. The coefficients $\tilde{a}(t, X_t)$ and $\tilde{b}(t, X_t)$ are obtained by plugging formula (4.62) to formula (4.2)

Table 4.3: HIX model specifications

X_t	$g(X_t)$	$\tilde{a}(t, X_t)$	$\tilde{b}(t, X_t)$
Vasicek	$g_2(x)$	$\frac{\kappa}{2}\text{sech}^2(X_t)(\eta - X_t) - \frac{\varsigma^2}{2}\tanh(X_t)\text{sech}^2(X_t)$	$\frac{\varsigma}{2}\text{sech}^2(X_t)$
Vasicek	$g_4(x)$	$\kappa\frac{\exp(X_t)}{(\exp(X_t)+1)^2}(\eta - X\text{sech}_t) + \frac{\varsigma^2}{2}\frac{\exp(X_t)(1-\exp(X_t))}{(\exp(X_t)+1)^3}$	$\varsigma\frac{\exp(X_t)}{(\exp(X_t)+1)^2}$
CIR	$g_1(x)$	$\kappa\text{sech}^2(X_t)(\eta - X_t) - \eta^2\tanh(X_t)\text{sech}^2(X_t)X_t$	$\varsigma\text{sech}^2(X_t)\sqrt{X_t}$
CIR	$g_3(x)$	$\kappa\exp(-X_t)(\eta - X_t) - \frac{\varsigma^2}{2}\exp(-X_t)X_t$	$\varsigma\exp(-X_t)\sqrt{X_t}$

Chapter 5

Applications of CIX

In this chapter, we introduce some applications of depended indices. We first give an example of constructing covariance and correlation swaps for two risky assets. Then we perform an empirical study by applying our CSV models in Section 4.3 to S&P 500 Implied Correlation Indexes historical data. Finally, we propose two digital CIX options and price them via Monte Carlo method.

5.1 Covariance and correlation Swap for two risky assets [Salvi and Swishchuk, 2014]

Consider a model with two risky assets and a risk free bond. Assuming that the risky assts satisfy the following stochastic differential equations

$$\begin{aligned} dS_t^{(1)} &= S_t^{(1)}(\mu_t^{(1)}dt + \sigma^{(1)}(x_t)dw_t^{(1)}) \\ dS_t^{(2)} &= S_t^{(2)}(\mu_t^{(2)}dt + \sigma^{(2)}(x_t)dw_t^{(2)}) \end{aligned} \tag{5.1}$$

where $\mu^{(1)}, \mu^{(2)}$ are deterministic functions of time, and $(w_t^{(1)})_t, (w_t^{(2)})_t$ are Brownian motions with quadratic covariance given by $d < w_t^{(1)}, w_t^{(2)} >= \eta_t dt$. Note that $(w_t^{(1)})_t, (w_t^{(2)})_t$ are assumed independent of the Markov process x_t. A stochastic process x_t is called Markov if $P(x_{t_n} \leq x_n | x_{t_{n-1}}, ..., x_{t_1}) = P(x_{t_n} \leq x | x_{t_{n-1}})$ for every n and $t_1 < t_2 < ... < t_n$ (see [Papoulis, 1984]).

Covariance swaps and a simple Pricing model A covariance swap is a covariance forward contract on the realized covariance between two risky assets for which payoff at maturity is equal to

$$N(Cov_R(S^{(1)}, S^{(2)}) - K_{cov}) \tag{5.2}$$

where K_{cov} is a strike reference value, N is the notional amount and $Cov_R(S^{(1)}, S^{(2)})$ is the realized covariance of the two assets $S^{(1)}$ and $S^{(2)}$ defined by

$$Cov_R(S^{(1)}, S^{(2)}) = \frac{1}{T}[lnS_T^{(1)}, lnS_T^{(2)}] = \frac{1}{T}\int_0^T \eta_t \sigma^{(1)}(x_t)\sigma^{(2)}(x_t)dt \tag{5.3}$$

The value of a covariance swap for Markov-modulated stochastic volatility is

$$P_{cov}(x) = E\{e^{-rT}(Cov_R(S^{(1)}, S^{(2)}) - K_{cov})\} = e^{-rT}\{\frac{1}{T}\int_O^T \eta_t e^{tQ}[\sigma^{(1)}(x)\sigma^{(2)}(x)]dt - K_{cov}\} \quad (5.4)$$

where we assumed that $N = 1$.

Correlation swaps and a simple Pricing model A correlation swap is a forward contract on the correlation between the underlying assets $S^{(1)}, S^{(2)}$ for which payoff at maturity is equal to

$$N(Corr_R(S^{(1)}, S^{(2)}) - K_{corr}) \quad (5.5)$$

where K_{corr} is a strike reference level, N is the notional amount and $Corr_R(S^{(1)}, S^{(2)})$ is the realized correlation defined by

$$Corr_R(S^{(1)}, S^{(2)}) = \frac{Cov_R(S^{(1)}, S^{(2)})}{\sqrt{\sigma_R^{(1)^2}(x)}\sqrt{\sigma_R^{(2)^2}(x)}} \quad (5.6)$$

where the realized variance is given be

$$\sigma_R^{(i)^2}(x) = \frac{1}{T}\int_0^T (\sigma^{(i)}(x_t))^2 dt, \ i = 1, 2 \quad (5.7)$$

The price of the correlation swap is the expected present value of the payoff in the risk neutral world

$$P_{corr}(x) = E\{e^{-rT}(Corr_R(S^{(1)}, S^{(2)}) - K_{corr})\} \quad (5.8)$$

where we set $N = 1$ for simplicity.

5.2 Empirical study

In this section we perform an empirical analysis under CSV models with historical data of S&P 500 Implied Correlation Indices.

5.2.1 Data and properties

The historical data of S&P 500 Implied Correlation Indexes accessed from CBOE website are displayed in Figure 5.1. We see that KCJ2009 almost exceeds the upper bound 1 because of the 2008 crisis while other CIXs take values within $(0, 1)$. Higher CIX suggests a herd behaviour during a crisis, which has been discussed in [Dhaene and Vyncke, 2012]. CIX exceeded 0.5 from 2008 to 2016 but fall in 2017. The Table 5.1 displays five summary statistics for CIXs: median, mean, mode, excess kurtosis and skewness. Most CIXs have positive kurtosis indicating a fat-tailed distribution. The negative skewness indicates that the data distribution is left-skewed.

We first fit ARMA models to CIX series then perform an ARCH Lagrange multiplier Test with a lag of 12. The outputs of ARCH test for different CIXs are shown in Table 5.2. If p-value is less than 0.05, we reject the null hypothesis and detect ARCH effects. From Table 5.2, we see that all CIX series have ARCH effects except for KCJ 2012.

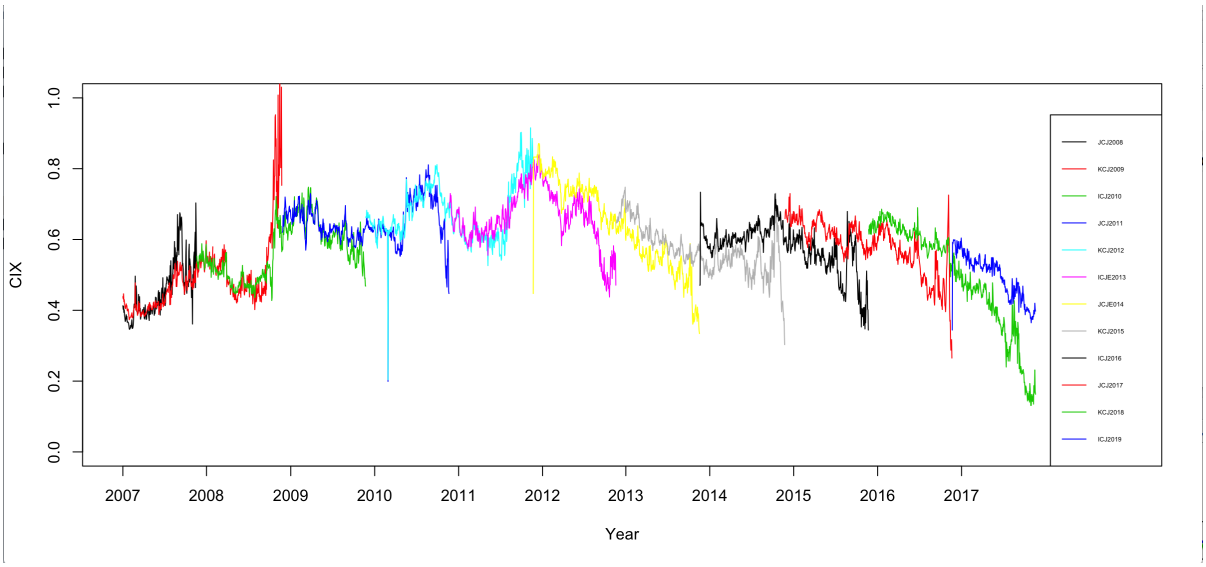

Figure 5.1: S& P 500 Implied Correlation Indexes between 2007-01-03 and 2017-11-17 released by CBOE. Symbols KCJ,ICJ and JCJ are cycled as time elapses but some of them are overlapped. KCJ 2009 = 105.93 on 2008-11-20 implies that the CIX on that day was 105.93. This calculation was based on individual options that expired at Jan.2009

.

CIX	Median	Mean	Mode	Kurtosis	Skewness
KCJ2009	0.470	0.489	0.401	8.426	2.484
ICJ2010	0.558	0.568	0.535	-1.011	0.254
JCJ2011	0.637	0.648	0.621	6.728	-0.716
KCJ2012	0.654	0.673	0.625	2.112	0.202
ICJE2013	0.661	0.664	0.628	0.108	-0.285
JCJ2014	0.640	0.640	0.833	-0.763	-0.248
KCJ2015	0.565	0.571	0.660	0.802	0.006
ICJ2016	0.592	0.576	0.639	1.563	-1.168
JCJ2017	0.594	0.581	0.642	1.522	-1.121
KCJ2018	0.532	0.494	0.460	-0.341	-0.816

Table 5.1: Table of 5 basic summary statistics for CIXs: median, mean, mode, excess kurtosis and skewness.

CIX	χ^2 **value**	**degree of freedom**	P **value**	**ARCH effects**
KCJ2009	277.01	12	$< 2.2 * 10^{-16}$	Y
ICJ2010	52.524	12	$5.006 * 10^{-7}$	Y
JCJ2011	29.056	12	0.003866	Y
KCJ2012	13.618	12	0.3258	N
ICJE2013	38.682	12	0.0001187	Y
JCJ2014	79.182	12	$5.91 * 10^{-12}$	Y
KCJ2015	183.56	12	$< 2.2 * 10^{-16}$	Y
ICJ2016	43.962	12	$1.55 * 10^{-5}$	Y
JCJ2017	52.332	12	$5.413 * 10^{-7}$	Y
KCJ2018	34.015	12	0.0006708	Y

Table 5.2: Table of ARCH effects for CIXs based on S&P 500 Implied Correlation Indexes from January 2009 using ArchTest() in R. Y represents ARCH effects while N suggests no ARCH effects.

5.2.2 Estimation

Next we estimate some parameters by using an autoregressive representation and the method of least squares. We are interested in the parameters driving the correlation excluding those driving the volatility. It should be noted that a full estimation exercise would be quite challenging due to the non-observability (hidden) nature of the stochastic volatility and the lack of closed-form expression for moments or Fourier transforms of the underlying. For our first CSV model, we redo an ARCH Test for $X_t = \text{artanh}(CIX_t)$. From Table 5.3, we see most of X_t series has ARCH effects.

Year	χ^2 value	degree of freedom	P value	ARCH effects
2009	118.98	12	$< 2.2 * 10^{-16}$	Y
2010	51.235	12	$8.465 * 10^{-7}$	Y
2011	44.033	12	$1.508 * 10^{-5}$	Y
2012	36.826	12	0.0002384	Y
2013	26.034	12	0.01061	Y
2014	55.965	12	$1.214 * 10^{-7}$	Y
2015	168.55	12	$< 2.2 * 10^{-16}$	Y
2016	31.31	12	0.001766	Y
2017	79.118	12	$6.079 * 10^{-12}$	Y
2018	15.597	12	0.2104	N

Table 5.3: Table of ARCH effects for $X_t = \text{artanh}(CIX_t)$ using ArchTest() in R. Y represents ARCH effects while N suggests no ARCH effects.

Due to the varying volatility of correlation, the disturbance variances in linear regression model is heteroscedastic. Although the disturbance process is unknown, it is similar to an autoregressive conditionally heteroscedastic (ARCH) process. Therefore, our analysis takes advantage of ARCH errors that are serial uncorrelated but have time-varying conditional variance. Since [Pantula, 1988] present the consistency of the ordinary least squares estimators for a p-th order autoregressive model with ARCH errors, we can use the method of least squares underlying a regression analysis to estimate the derived parameters even though the model has ARCH errors. In the following context, the varying volatility of correlation in each model is assumed to be constant for estimating the parameters of interest. Since we only consider positive volatilities, we acknowledge that there is bias in the simulation.

Estimating parameters on CSV model 1

Assuming the volatility of underlying process X_t is constant, we write the autoregressive representation in formula (4.20) as follows

$$X_{t_{i+1}} = \vartheta\eta\Delta t + (1 - \vartheta\Delta t)X_{t_i} + \sqrt{\upsilon}Z_{t_i} \quad i = 0, 1, ..., N \tag{5.9}$$

where $X_t = \text{artanh}(\text{CIX}_\text{t})$ and $\sqrt{\upsilon}$ is a constant volatility.

The formula (5.9) can be shown in a general form of Linear Regression:

$$\mathbf{Y} = \beta_0 \mathbf{1}_N + \beta_1 \boldsymbol{X}^{(\mathbf{1})} + \boldsymbol{\epsilon} \tag{5.10}$$

where $\mathbf{1}_N$ is a N-vector of ones, $\boldsymbol{\epsilon}$ is an error vector, and

$$\begin{aligned} \beta_0 &= \vartheta\eta\Delta t \\ \beta_1 &= 1 - \vartheta\Delta t \end{aligned} \tag{5.11}$$

$$\begin{aligned} \mathbf{Y} &= \left[X_{t_1}, X_{t_2}, \cdots, X_{t_N}\right]^\top \\ \boldsymbol{X}^{(\mathbf{1})} &= \left[X_{t_0}, X_{t_1}, \cdots, X_{t_{N-1}}\right]^\top \\ \boldsymbol{\epsilon} &= \left[\sqrt{\upsilon}Z_{t_0}, \sqrt{\upsilon}Z_{t_1}, \cdots, \sqrt{\upsilon}Z_{t_{N-1}}\right]^\top \end{aligned} \tag{5.12}$$

Table 5.4 displays that errors $\boldsymbol{\epsilon}$ are ARCH errors. Thus, we have

$$\begin{aligned} \hat{\vartheta} &= \tfrac{1-\hat{\beta_1}}{\Delta t} \\ \hat{\eta} &= \tfrac{\hat{\beta_0}}{\hat{\vartheta}\Delta t} = \tfrac{\hat{\beta_0}}{1-\hat{\beta_1}} \end{aligned} \tag{5.13}$$

We take advantage of the constant $\sqrt{\upsilon}$ to derive the estimated long-run variance $\hat{\theta}$ as the squared volatility υ follows a mean-reverting process with mean θ. To be specific, $\text{Var}(\boldsymbol{\epsilon})$, the squared residuals, is an approximation of $\theta\Delta t$ thus the estimator of long-run variance is $\hat{\theta} = \frac{\text{Var}(\boldsymbol{\epsilon})}{\Delta t}$.

The historical dataset shows the time-to-maturity T of CIXs is same equal to 2 years. The step size Δt is set to be 0.004. The estimation of parameters has been shown in Table 5.5.

Year	χ^2 **value**	**degree of freedom**	P **value**	**ARCH effects**
2009	313.04	12	$< 2.2 * 10^{-16}$	Y
2010	52.181	12	$5.759 * 10^{-7}$	Y
2011	161.87	12	$< 2.2 * 10^{-16}$	Y
2012	81.589	12	$2.052 * 10^{-12}$	Y
2013	25.144	12	0.01416	Y
2014	42.939	12	$2.312 * 10^{-5}$	Y
2015	210.04	12	$< 2.2 * 10^{-16}$	Y
2016	40.602	12	$5.71 * 10^{-5}$	Y
2017	68.757	12	$5.467 * 10^{-10}$	Y
2018	37.028	12	0.0002211	Y

Table 5.4: Table of ARCH effects for residuals in formula (5.10). Y represents ARCH effects while N suggests no ARCH effects.

CIX	$\hat{\beta_0}$	**Std. Error** ($\hat{\beta_0}$)	$\hat{\beta_1}$	**Std. Error** ($\hat{\beta_1}$)	$\hat{\vartheta}$	$\hat{\eta}$	$\hat{\theta}$
KCJ2009	0.311	0.02628	0.448	0.04116	138	0.563	18.496
ICJ2010	0.027	0.008525	0.959	0.012871	10.25	0.659	0.287
JCJ2011	0.091	0.01714	0.882	0.02184	29.5	0.771	0.610
KCJ2012	0.086	0.01701	0.898	0.02010	25.5	0.843	1.26
ICJE2013	0.028	0.01030	0.965	0.01248	8.75	0.8	0.400
JCJ2014	0.018	0.008168	0.977	0.010021	5.75	0.783	0.561
KCJ2015	0.025	0.009272	0.961	0.013979	9.75	0.641	0.245
ICJ2016	0.054	0.01241	0.918	0.01853	20.5	0.659	0.385
JCJ2017	0.027	0.01006	0.958	0.01479	10.5	0.647	0.304
KCJ2018	0.003	0.003909	0.992	0.006627	2	0.375	0.195

Table 5.5: Estimated parameters for CSV model 1 based on S&P 500 index from January 2009 using Multiple Linear Regression where T=2, $\Delta t = 0.004$

Estimating parameters on CSV model 2

To make $(h-g(\rho_{t_i}))(g(\rho_{t_i})-f)$ always positive, we define function $G(\rho_{t_i}) := \sqrt{(h-g(\rho_{t_i}))(g(\rho_{t_i})-f)}$ with $h = 1, f = -1$ and set the value of $g(\rho_{t_i})$:

$$g(\rho_{t_i}) := \begin{cases} -0.999 & \text{if } \rho_{t_i} \leq -1 \\ 0.999 & \text{if } \rho_{t_i} \geq 1 \\ \rho_{t_i} & \text{otherwise} \end{cases} \tag{5.14}$$

Then, the autoregressive representation in formula (4.25) is given by

$$\frac{\rho_{t_{i+1}}}{G(\rho_{t_i})} = (1-\beta\Delta t)\frac{g(\rho_{t_i})}{G(\rho_{t_i})} + \frac{\beta\bar{\rho}\Delta t}{G(\rho_{t_i})} + \sqrt{\zeta}Z_{t_i} \quad \text{for } i = 0, 1, ..., N \tag{5.15}$$

where $\sqrt{\zeta}$ is the constant volatility of correlation parameter. To be more precise, we rewrite the formula (5.15) as follows

$$\mathbf{Y} = \beta_1\mathbf{X^{(1)}} + \beta_2\mathbf{X^{(2)}} + \boldsymbol{\epsilon} \tag{5.16}$$

where $\boldsymbol{\epsilon}$ is an error vector, and

$$\begin{aligned} \beta_1 &= 1-\beta\Delta t \\ \beta_2 &= \beta\bar{\rho}\Delta t \end{aligned} \tag{5.17}$$

$$\begin{aligned} \mathbf{Y} &= \left[\frac{\rho_{t_1}}{G(\rho_{t_1})}, \frac{\rho_{t_2}}{G(\rho_{t_2})}, \cdots, \frac{\rho_{t_N}}{G(\rho_{t_N})}\right]^\top \\ \mathbf{X^{(1)}} &= \left[\frac{g(\rho_{t_0})}{G(\rho_{t_0})}, \frac{g(\rho_{t_1})}{G(\rho_{t_1})}, \cdots, \frac{g(\rho_{t_{N-1}})}{G(\rho_{t_{N-1}})}\right]^\top \\ \mathbf{X^{(2)}} &= \left[\frac{1}{G(\rho_{t_0})}, \frac{1}{G(\rho_{t_1})}, \cdots, \frac{1}{G(\rho_{t_{N-1}})}\right]^\top \\ \boldsymbol{\epsilon} &= \left[\sqrt{\zeta}Z_{t_0}, \sqrt{\zeta}Z_{t_1}, \cdots, \sqrt{\zeta}Z_{t_{N-1}}\right]^\top \end{aligned} \tag{5.18}$$

Thus,

$$\begin{aligned} \hat{\beta} &= (1-\hat{\beta_1})/\Delta t \\ \hat{\bar{\rho}} &= \hat{\beta_2}/(1-\hat{\beta_1}) \end{aligned} \tag{5.19}$$

Table 5.6 shows that $\boldsymbol{\epsilon}$ are ARCH errors. Now we apply the method of least squares to formula (5.16) to estimate the desired parameters. As the errors $\sqrt{\zeta}Z_t$ are iid $N(0, \zeta\Delta t)$ random variables, the ordinary least squares are linear unbiased consistent estimators of β and $\bar{\rho}$. Since $\text{Var}(\boldsymbol{\epsilon})$ is an approximation of $\theta\Delta t$, the estimator of long-run variance is given by $\hat{\theta} = \frac{\text{Var}(\boldsymbol{\epsilon})}{\Delta t}$.

We estimate the parameters of interest with same historical dataset. The time-to-maturity T equals to 2 years, and the step size Δt is set to be 0.004. The estimates of parameters has been shown in Table 5.7.

Year	χ^2 **value**	**degree of freedom**	P **value**	**ARCH effects**
2009	282.59	12	$< 2.2 * 10^{-16}$	Y
2010	49.743	12	$1.55 * 10^{-6}$	Y
2011	91.113	12	$3.009 * 10^{-14}$	Y
2012	86.133	12	$2.756 * 10^{-13}$	Y
2013	27.678	12	0.006163	Y
2014	34.283	12	0.0006085	Y
2015	205.44	12	$< 2.2 * 10^{-16}$	Y
2016	43.651	12	$1.751 * 10^{-5}$	Y
2017	61.978	12	$9.819 * 10^{-9}$	Y
2018	50.269	12	$1.253 * 10^{-6}$	Y

Table 5.6: Table of ARCH effects for residuals in formula (5.16). Y represents ARCH effects while N suggests no ARCH effects.

CIX	$\hat{\beta_0}$	**Std. Error** ($\hat{\beta_0}$)	$\hat{\beta_1}$	**Std. Error** ($\hat{\beta_1}$)	$\hat{\beta}$	$\hat{\bar{\rho}}$	$\hat{\theta}$
KCJ2009	0.551	0.01289	0.217	0.01138	-112.25	0.485	4.914
ICJ2010	0.960	0.012294	0.023	0.007177	10.25	0.561	0.177
JCJ2011	0.879	0.02357	0.078	0.01552	30.25	0.645	0.415
KCJ2012	0.890	0.01827	0.068	0.01274	25.25	0.673	0.575
ICJE2013	0.967	0.012345	0.022	0.008439	8.25	0.667	0.218
JCJ2014	0.981	0.008915	0.012	0.006097	4.75	0.632	0.256
KCJ2015	0.963	0.01366	0.021	0.00797	9.25	0.568	0.158
ICJ2016	0.922	0.01791	0.045	0.01051	19.5	0.577	0.241
JCJ2017	0.960	0.01481	0.023	0.00880	10	0.575	0.202
KCJ2018	0.994	0.006574	0.002	0.003528	1.5	0.333	0.146

Table 5.7: Estimated parameters for CSV model 2 based on S&P 500 index from January 2009 using Multiple Linear Regression where T=2, Δt = 0.004

5.2.3 Pricing

In this section, we design two European style digital CIX options with same time-to-maturity T, and price them based on our two CSV models (4.16) and (4.19) via Monte Carlo method. As we know, the correlation between two stocks is relative higher when the market goes down, see [Guillaume and Linders, 2015]. This characteristic phenomenon is verified in the S& P 500 Implied Correlation Indexes historical dataset, see Figure 5.1. CIX hit its upper bound 1 at the end of 2008, which coincides with the 2008 financial crisis. In fact, the exceptionally high correlation between two different stock prices will reduce the benefits of diversification (see [Linders and Schoutens, 2014]) so that investors will suffer more losses during crises. Nevertheless, investors can take advantage of this characteristic for hedging risk. In this thesis, we design two kinds of digital CIX options. These two digital CIX options use different values of CIX as a signal to identify a crisis. To be specific, if the signal value exceeds the pre-determined dangerous threshold, the financial market is regarded as in crisis. Holders are supposed to execute the option to earn the nonzero benefit. If the signal value does not exceed the dangerous threshold, there is no crisis and the option is not executed. In this scenario, option holders only lose the premium that is used to enter into the contract at time 0. We assume that the digital CIX option will bring L dollars benefit if the option is executed at maturity. The dangerous threshold of CIX is set to 0.9 as that is the minimum level only achieved in the 2008 financial crisis (see Figure 5.1). Next, we give the definition of these two kinds of digital CIX options:

Digital CIX option 1 The Digital CIX option 1 is designed for investors who worry about the financial crisis happened at maturity T. Thus the signal value of CIX is the value of CIX at time T. We introduce an indicator function as follows:

$$\mathbf{1}_{\{CIX_T \geq 0.9\}} := \begin{cases} 1 & \text{if } CIX_T \geq 0.9 \\ 0 & \text{otherwise} \end{cases} \tag{5.20}$$

where r is the risk-free rate and CIX_T is the value of CIX at maturity T.

With no arbitrage principle, the premium of digital CIX option 1 is:

$$\text{Price}_1 = \mathbb{E}(e^{-rT} \cdot L \cdot \mathbf{1}_{\{CIX_T \geq 0.9\}}) \tag{5.21}$$

Investors pay a premium Price_1 at time 0 to enter into the contract. Since $CIX_T \geq 0.9$ implies a crisis at time T, holders would execute the option to get L dollars benefit during crisis. $CIX_T < 0.9$ suggests no crisis at time T thus option holders will not execute the option. In this scenario, investors will have zero benefit and only lose the premium Price_1. According to no arbitrage principle, the price of digital option should be the expectation of the present value of payoff which is $L\mathbf{1}_{\{CIX_T \geq 0.9\}}$.

Digital CIX option 2 The second digital CIX option works for investors who worry about crisis that will arise over the life of the contract. Thus the digital CIX option 2 uses the maximum of CIX over [0,T] as the signal value to identify a crisis. An indicator function is defined as follows:

$$\mathbf{1}_{\{\overline{CIX}_T \geq 0.9\}} := \begin{cases} 1 & \text{if } \overline{CIX}_T \geq 0.9 \\ 0 & \text{otherwise} \end{cases} \tag{5.22}$$

where $\overline{CIX}_T = \max_{0\le t\le T} CIX_t$ is the maximum correlation parameter over the period $[0, T]$.

With no arbitrage principle, the premium of Digital CIX option 2 is given by:

$$\text{Price}_2 = \mathbb{E}(e^{-rT} \cdot L \cdot \mathbf{1}_{\overline{\{CIX_T \ge 0.9\}}}) \tag{5.23}$$

If $CIX_t, 0 \le t \le T$ reaches the dangerous threshold 0.9, option holders would execute the option to obtain L dollars benefit at time T. Otherwise, there is no crisis during the life of option. The Digital CIX option 2 will not be executed, and investors will lose the premium Price_2. Following no arbitrage principle, the price of option should be the expectation of the present value of payoff which is $L\mathbf{1}_{\overline{\{CIX_T \ge 0.9\}}}$.

To compute CIX option prices, we set the risk-free rate $r = 0.02$, the benefit $L = 1000$ and time-to-maturity $T = 2$. The default values of parameters are shown in Table 5.8. The default values of ς, η and θ in CSV model 1, and the defualt values of β, $\bar{\rho}$ and θ in CSV model 2 are estimated from same data set -JCJ 2014, see Section 5.2.2. Further we perform a sensitive analysis on Price_1 and Price_2 with our two CSV models by Monte Carlo simulation.

Model	Parameter	Value
Advanced model 1	Initial squared volatility	$\upsilon_0 = 0.56$
	Speed of reversion of the squared volatility parameter	$\kappa = 2$
	Long term mean level of the squared volatility parameter	$\theta = 0.56$
	Volatility of the squared volatility parameter	$\varsigma = 0.1$
	Initial value of the underlying process	$X_0 = 0.783$
	Long term mean level of the underlying process	$\eta = 0.783$
	Speed of reversion of the underlying process	$\vartheta = 5.75$
	Correlation of W_t^X and W_t^υ	$\xi_1 = -0.5$
	Risk free rate	$r = 0.02$
	Time to maturity	$T = 2$
	Number of simulation steps	$N = 500$
	Number of trajectories	$M = 10000$
	Benefit (dollars)	$L = 1000$
Advanced model 2	Initial squared volatility	$\zeta_0 = 0.256$
	Speed of reversion of the squared volatility parameter	$\kappa = 2$
	Long term mean level of the squared volatility parameter	$\theta = 0.256$
	Volatility of the squared volatility parameter	$\varsigma = 0.15$
	Initial value of correlation process	$\rho_0 = 0.632$
	Speed of reversion of the correlation parameter	$\beta = 4.75$
	Long term mean level of the correlation parameter	$\bar{\rho} = 0.632$
	Bounds of correlation parameter	$h = 1, f = -1$
	Correlation of W_t^ρ and W_t^ζ	$\xi_2 = -0.5$
	Risk free rate	$r = 0.02$
	Time to maturity	$T = 2$
	Number of simulation steps	$N = 500$
	Number of trajectories	$M = 10000$
	Benefit (dollars)	$L = 1000$

Table 5.8: Default parameters for pricing of CSV model 1 & 2

Pricing on CSV model 1

Recall that the level of CIX is affected by two groups of parameters in our CSV model 1, see formula (4.16). One is embedded in the mean-reverting process X_t consisting of parameters ϑ,

η and ξ_1. The other is associated with the stochastic volatility including parameters κ, θ and ς. Since we are interested in the independent effect of these parameters on prices, we vary each parameter in turn. Apart form that, we always keep $\upsilon_0 = \theta$ for varying θ and $X_0 = \eta$ for varying η to identify mean reversion in diagrams, as we did in Section 4.3.2.

Price_1 and Price_2 with CSV model 1 are plotted with varying κ, ς and ξ_1 in Figure 5.2 respectively. Since the second Digital CIX options provide higher safety, it is reasonable to observe that Price_2 is significantly greater than Price_1. From the left hand sides of Figure 5.2, we observe that κ has less effect on Price_1 and Price_2. As κ increases from 2 to 10, the Price_2 increases 7%. However, ς (the volatility of volatility) has strong effect on prices. As ς increases from 0.01 to 0.25, Price_1 falls 44% and Price_2 falls 34%. These distinct differences between pricing with non-zero SV parameters and that without SV parameters suggest the importance of stochastic volatilities in modelling. Otherwise, the arbitrage opportunity will emerge in trading CIX options.

Figure 5.3 shows Price_1 and Price_2 on the y-axis, with the parameters of interest θ, ϑ and η on the x-axis. Since θ and η are long term mean levels of υ_t and X_t, they can have significant impacts on option prices. More precisely, option price stays around zero when parameters θ and η start growing. Then the price jumps into an obvious value when θ exceeds 0.4, or η exceeds 0.6. Note that Price_2 are greater than Price_1. We also see that Price_1 and Price_2 obviously drop to zero when ϑ exceeds 8.

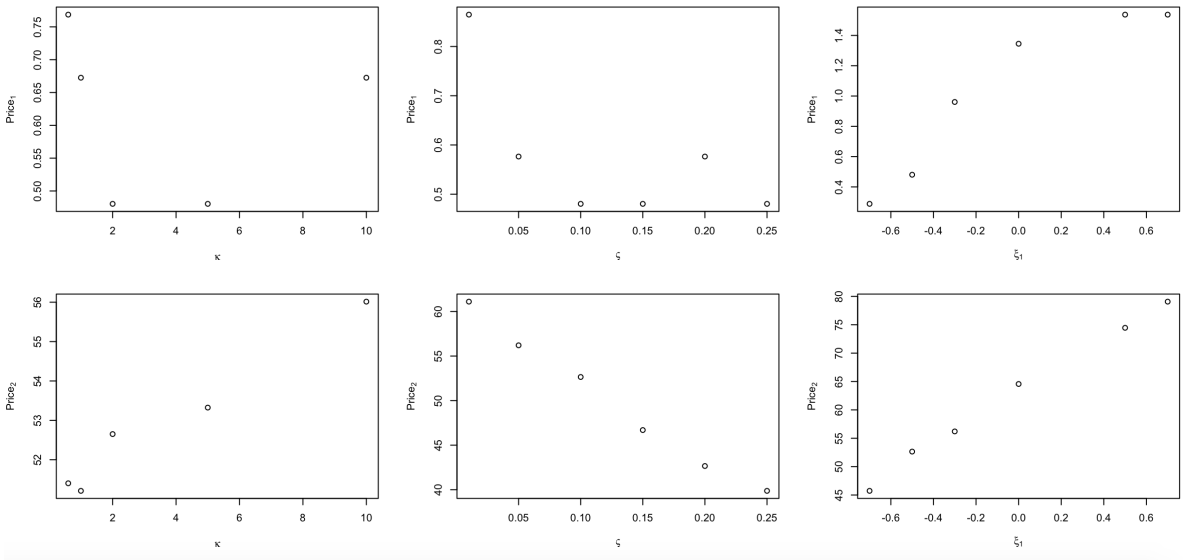

Figure 5.2: Price_1 (upper row) and Price_2 (lower row) on CSV model 1 with three varying parameters: the speed of reversion of volatility parameter κ (left), the volatility of volatility parameter ς (middle), and the correlation between two Brownian motions ξ_1(right)

.

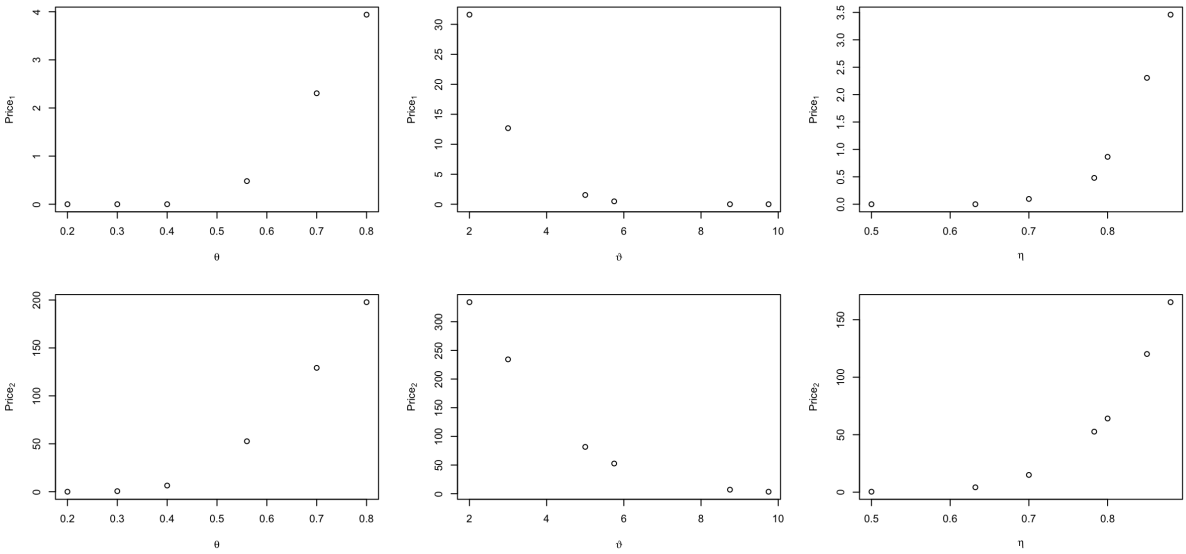

Figure 5.3: Price_1 (upper row) and Price_2 (lower row) on CSV model 1 with three varying parameters: long term mean level of squared volatility θ (left), speed of reversion of underlying process ϑ (middle), and long term mean level of underlying process η (right)

.

Pricing on CSV model 2

In CSV model 2, we study the impact of parameters κ, θ, ξ_2, ς, β and $\bar{\rho}$ on prices, see formula (4.25). Parameters κ, θ and ξ_2 have direct influences on ρ_t while ς, β and $\bar{\rho}$ affect ρ_t through stochastic volatilities $\sqrt{\zeta}$.

Figures 5.4 and 5.5 show the Price_1 (upper row) and Price_2 (lower row) on the y-axis with parameters on the x-axis. All other things being equal, Price_2 is always greater than Price_1 due to its higher safety. From the left panel in Figure 5.4, we see that both prices are influenced by κ. As $\kappa = 10$ increases from 2 to 10, Price_1 rises 11 %, and Price_2 rises 16%. The middle panel in Figure 5.4 shows that the parameter ς has negative effect on prices. As ς increases from 0.01 to 0.25, Price_1 falls 44% while Price_2 falls 31% . From the right panels, we see that Price_1 and Price_2 have growth trending as ξ_2 increases from -0.7 to 0.7.

In addition, the long term mean level of squared volatility (θ) and the long term mean level of correlation process ($\bar{\rho}$) have significantly positive influences on prices. From Figure 5.5, we observe that both prices are nearly zero when θ and $\bar{\rho}$ start increasing but prices rise sharply at $\theta = 0.4$, or at $\bar{\rho} = 0.7$. The middle panel in Figure 5.5 suggests that prices significantly drop as β increases from 2 to 5, but stay around zero when β exceeds 5.

Compared to prices in CSV model 1, CSV model 2 presents triple prices. Recall that we only have three free parameters in simulation: the speed of reversion of the squared volatility parameter κ, the volatility of the squared volatility parameter ς and the correlation between two Brownian motions ξ_1 (ξ_2). From Figures 5.2 and 5.4, we see that these free parameters do not lead to such huge magnitude (300%). Changes on those parameters only cause slight differences in prices. For unexpected differences in prices between these two models, we will study this in future research.

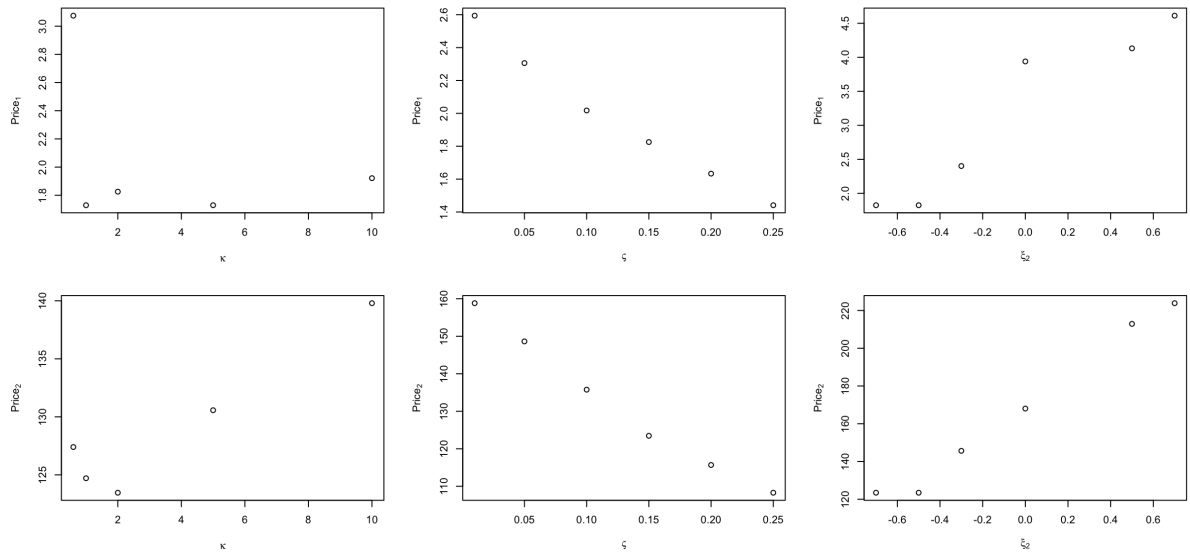

Figure 5.4: Price_1 (upper row) and Price_2 (lower row) on CSV model 2 with three varying parameters: the speed of reversion of the squared volatility parameter κ (left), the volatility of the squared volatility parameter ς (middle), and the correlation between two Brownian motions ξ_2 (right)

.

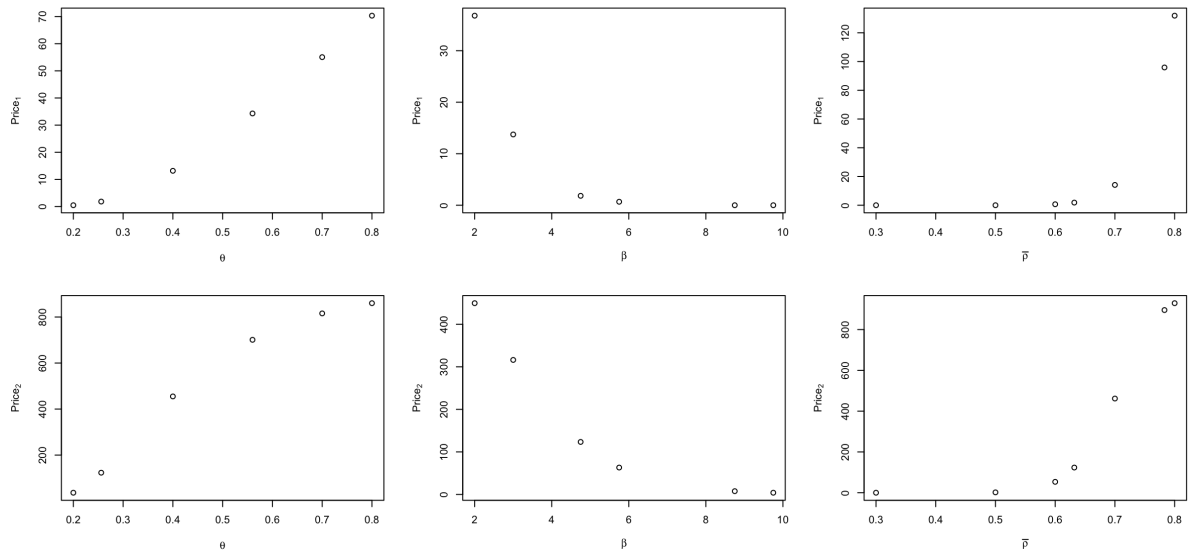

Figure 5.5: Price_1 (upper row) and Price_2 (lower row) on CSV model 2 with three varying parameters: the long term mean level of squared volatility θ (left), the speed of reversion of the correlation parameter β (middle), and the long term mean level of the correlation parameter $\bar{\rho}$ (right)

.

Chapter 6

Conclusions

In this thesis, we have provided definitions, studied properties, and crafted new stochastic models for two dependence indices: the implied correlation index and the herd behaviour index (HIX). We modelled the basic implied correlation index (CIX) by considering stochastic volatility models. We displayed the definition of CIX (see [Skintzi and Refenes, 2005]), and developed two Correlation Stochastic Volatility (CSV) models for CIX, of which the volatility follows a CIR process. The first CSV model (4.16) consists of a mean-reverting process and a proper mapping function. The second CSV model (4.19) is an extension of general Jacobi process. In both models, CIX takes value in (-1,1) and satisfies [van Emmerich, 2006]'s properties.

Then we presented a definition and properties of HIX (see [Dhaene and Vyncke, 2012]). Based on [Guillaume and Linders, 2015]'s work, we modelled HIX by combining a mean-reverting process with some particular mapping functions (i.e $\tanh(x)$). The mapping function will map the domain of the mean-reverting process to $[0, 1]$ such that HIX always takes value in $[0, 1]$.

In the empirical study, we analyzed the historical data of CBOE S&P 500 CIX. The ARCH Test shows that most of the CIX series has heteroscedasticity. Parameters are estimated via Monte Carlo method. For a given series of CIX, we found that the long term mean level of squared volatility parameter in the first CSV model was similar to that in the second CSV model. As we only consider positive volatilities, we acknowledge the bias in simulation.

Finally, we proposed two kinds of digital CIX options and priced them via Monte Carlo method. The first digital option pricing used the value of CIX at maturity. The second digital option pricing used the maximum CIX over the life of the option. As the second digital option reflected a higher risk aversion, all other things being equal, the prices of the second digital option should be greater than the prices of the first digital option. Our analysis in Section 5.2.3 has confirmed this point. Moreover, we see that option prices are increasing functions of the long term mean level (θ) and decreasing functions of the volatility of volatility parameter (ϑ). Since parameters κ, θ, and ς have significant influence (i.e. 60%) on option pricing, it is reasonable to use stochastic volatility models to model CIX. In addition, prices in our two CSV models have slight differences with respect to same values of stochastic volatility parameters κ, ς and ξ. The unexpected differences of prices between our two CSV models will be studied in the future.

Bibliography

[Bossu, 2005] Bossu, S. (2005). Arbitrage pricing of equity correlation waps. jpmorgan equity derivatives. Technical report, Working paper.

[Bru, 1991] Bru, M.-F. (1991). Wishart processes. *Journal of Theoretical Probability*, 4(4):725–751.

[Chen and Vanmaele, 2008] Chen, X., D. G. D. J. and Vanmaele, M. (2008). Static super-replicating strategies for a class of exotic options. *Insurance: Mathematics and Economics*, 42(3):1067–1085.

[Da Fonseca et al., 2014] Da Fonseca, J., Grasselli, M., and Ielpo, F. (2014). Estimating the wishart affine stochastic correlation model using the empirical characteristic function. *Studies in Nonlinear Dynamics & Econometrics*, 18(3):253–289.

[Da Fonseca et al., 2007] Da Fonseca, J., Grasselli, M., and Tebaldi, C. (2007). Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 10(2):151–180.

[Dhaene and Vyncke, 2002a] Dhaene, Jan, D. M. G. M. J. K. R. and Vyncke, D. (2002a). The concept of comonotonicity in actuarial science and finance: applications. *Insurance: Mathematics and Economics*, 31(2):133–161.

[Dhaene and Vyncke, 2002b] Dhaene, J., D. M. G. M. K. R. and Vyncke, D. (2002b). The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, 31(1):3–33.

[Dhaene and Vyncke, 2012] Dhaene, J., L. D. S. W. and Vyncke, D. (2012). The herd behavior index: A new measure for the implied degree of co-movement in stock markets. *Insurance: Mathematics and economics*, 50(3):357–370.

[Engle, 2002] Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.

[Exchange, 2009] Exchange, C. B. O. (2009). Cboe s&p 500 implied correlation index. Technical report, Working Paper.

[Fenzl and Pelzmann, 2012] Fenzl, T. and Pelzmann, L. (2012). Psychological and social forces behind aggregate financial market behavior. *Journal of Behavioral Finance*, 13(1):56–65.

[Gouriéroux, 2006] Gouriéroux, C. (2006). Continuous time wishart process for stochastic risk. *Econometric Reviews*, 25(2-3):177–217.

[Gourieroux and Sufana, 2003] Gourieroux, C. and Sufana, R. (2003). Wishart quadratic term structure models.

[Guillaume and Linders, 2015] Guillaume, F. and Linders, D. (2015). Stochastic modelling of herd behaviour indices. *Quantitative Finance*, 15(12):1963–1977.

[Heston, 1993] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343.

[Hobson* et al., 2005] Hobson*, D., Laurence, P., and Wang, T.-H. (2005). Static-arbitrage upper bounds for the prices of basket options. *Quantitative finance*, 5(4):329–342.

[Kim and Ahn, 2013] Kim, C., C. Y. L. W. and Ahn, J. (2013). Analyzing herd behavior in global stock markets: An intercontinental comparison. *arXiv preprint arXiv:1308.3966*.

[Linders and Schoutens, 2014] Linders, D. and Schoutens, W. (2014). A framework for robust measurement of implied correlation. *Journal of Computational and Applied Mathematics*, 271:39–52.

[Linders and Vanmaele, 2012] Linders, D., D. J. H. H. and Vanmaele, M. (2012). Index options: a model-free approach.

[Lord et al., 2010] Lord, R., Koekkoek, R., and Dijk, D. V. (2010). A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194.

[Ma, 2009] Ma, J. (2009). A stochastic correlation model with mean reversion for pricing multi-asset options. *Asia-Pacific Financial Markets*, 16(2):97–109.

[McNeil et al., 2015] McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools*. Princeton university press.

[Pantula, 1988] Pantula, S. G. (1988). Estimation of autoregressive models with arch errors. *Sankhyā: The Indian Journal of Statistics, Series B*, pages 119–138.

[Papoulis, 1984] Papoulis, A. (1984). Brownian movement and markoff processes. *Probability, random variables, and stochastic processes*, pages 515–553.

[Salvi and Swishchuk, 2014] Salvi, G. and Swishchuk, A. V. (2014). Covariance and correlation swaps for financial markets with markov-modulated volatilities. *International Journal of Theoretical and Applied Finance*, 17(01):1450006.

[Skintzi and Refenes, 2005] Skintzi, V. D. and Refenes, A.-P. N. (2005). Implied correlation index: A new measure of diversification. *Journal of Futures Markets*, 25(2):171–197.

[Teng et al., 2013] Teng, L., van Emmerich, C., Ehrhardt, M., and Gunther, M. (2013). A general approach for stochastic correlation using hyperbolic functions. *Preprint, Bergische Universität Wuppertal*.

[Teng et al., 2016] Teng, L., Van Emmerich, C., Ehrhardt, M., and Günther, M. (2016). A versatile approach for stochastic correlation using hyperbolic functions. *International Journal of Computer Mathematics*, 93(3):524–539.

[van Emmerich, 2006] van Emmerich, C. (2006). Modelling correlation as a stochastic process. *Preprint*, 6(03).

[Wilmott,] Wilmott, P. Derivatives, the theory and practice of financial engineering, 1998. *John Wiley&Sons, Chichester*.

[Zhou, 2013] Zhou, H. (2013). On the predictive power of the implied correlation index. *Working Paper*.

Appendix A

Proofs

A.1 Proof of Lemma 1

At first, we derive the general stochastic differential equation for ρ_t by Ito's lemma

$$\begin{aligned} d\rho_t &= df(X_t) \\ &= \left(\frac{\partial f(X_t)}{\partial X_t} \vartheta(\eta - X_t) + \frac{1}{2} \frac{\partial^2 f(X_t)}{\partial X_t^2} \upsilon_t \right) dt + \frac{\partial f(X_t)}{\partial X_t} \sqrt{\upsilon_t} dW_t^X \end{aligned}$$

Since $f(X_t) = \tanh(X_t)$, we have

$$\begin{aligned} \frac{\partial f(X_t)}{\partial X_t} &= \operatorname{sech}^2(X_t) = 1 - \rho_t^2 \\ \frac{\partial^2 f(X_t)}{\partial X_t^2} &= -2 \tanh(X_t) \operatorname{sech}^2(X_t) = -2\rho_t(1 - \rho_t^2) \end{aligned}$$

Hence, $d\rho_t$ can be written in the form

$$d\rho_t = \left((1 - \rho_t^2)\vartheta(\eta - \operatorname{artanh}(\rho_t)) - \rho_t(1 - \rho_t^2)\upsilon_t \right) dt + (1 - \rho_t^2) \sqrt{\upsilon_t} dW_t^X$$

Since $m(X_t) := (\frac{\partial f(X_t)}{\partial X_t})^2$, we derive $\upsilon_t dm(X_t)$, $m(X_t)d\upsilon_t$ and $d\upsilon_t dm(X_t)$ respectively. By Ito's lemma,

$$\begin{aligned} dm(X_t) &= \frac{\partial m(X_t)}{\partial X_t} dX_t + \frac{1}{2} \frac{\partial^2 m(X_t)}{\partial X_t^2} \upsilon_t dt \\ &= \left(\frac{\partial m(X_t)}{\partial X_t} \vartheta(\eta - X_t) + \frac{1}{2} \frac{\partial^2 m(X_t)}{\partial X_t^2} \upsilon_t \right) dt + \frac{\partial m(X_t)}{\partial X_t} \sqrt{\upsilon_t} dW_t^X \end{aligned}$$

Then

$$\begin{aligned}
\upsilon_t dm(X_t) &= \left(\frac{\partial m(X_t)}{\partial X_t}\vartheta(\eta - X_t)\upsilon_t + \frac{1}{2}\frac{\partial^2 m(X_t)}{\partial X_t^2}\upsilon_t^2\right)dt + \frac{\partial m(X_t)}{\partial X_t}\upsilon_t^{3/2}dW_t^X \\
m(X_t)d\upsilon_t &= \kappa(\theta - \upsilon_t)m(X_t)dt + \varsigma\sqrt{\upsilon_t}m(X_t)dW_t^\upsilon
\end{aligned}$$

Because the correlation between W_t^X and W_t^υ is ξ_1, we have

$$\begin{aligned}
d\upsilon_t dm(X_t) &= (\kappa(\theta - \upsilon_t)dt + \varsigma\sqrt{\upsilon_t}dW_t^\upsilon)\left((\frac{\partial m(X_t)}{\partial X_t}\vartheta(\eta - X_t) + \frac{1}{2}\frac{\partial^2 m(X_t)}{\partial X_t^2}\upsilon_t)dt + \frac{\partial m(X_t)}{\partial X_t}\sqrt{\upsilon_t}dW_t^X\right) \\
&= \frac{\partial m(X_t)}{\partial X_t}\varsigma\xi_1\upsilon_t dt
\end{aligned}$$

According to the definition of quadric variance, we have

$$\begin{aligned}
d\langle\rho\rangle_t &= \upsilon_t(1-\rho_t^2)^2dt \\
d\langle\sigma_\rho^2\rangle_t &= (16\rho_t^2\upsilon_t^2 + \varsigma^2 - 8\xi_1\varsigma\rho_t\upsilon_t)(1-\rho_t^2)^4\upsilon_t dt \\
d\langle\rho,\sigma_\rho^2\rangle_t &= (1-\rho_t^2)\sqrt{\upsilon_t}(-4\rho_t(1-\rho_t^2)^2\upsilon_t^{3/2})dW_t^X dW_t^X + \left((1-\rho_t^2)\sqrt{\upsilon_t}\varsigma\sqrt{\upsilon_t}(1-\rho_t^2)^2\right)dW_t^X dW_t^\upsilon \\
&= (-4\rho_t\upsilon_t + \varsigma\xi_1)(1-\rho_t^2)^3\upsilon_t dt
\end{aligned}$$

A.2 Proof of Lemma 2

From formula (4.25), we have

$$\psi(\rho_t)d\zeta_t = \kappa(\theta - \zeta_t)\psi(\rho_t)dt + \varsigma$$

By Ito's lemma,

$$\begin{aligned}
d\psi(\rho_t) &= \frac{\partial\psi(\rho_t)}{\partial\rho_t}d\rho_t + \frac{1}{2}\frac{\partial^2\psi(\rho_t)}{\partial\rho_t^2}\psi(\rho_t)\zeta_t dt \\
&= \left(\beta(\bar\rho - \rho_t)\frac{\partial\psi(\rho_t)}{\partial\rho_t} + \frac{1}{2}\frac{\partial^2\psi(\rho_t)}{\partial\rho_t^2}\psi(\rho_t)\zeta_t\right)dt + \frac{\partial\psi(\rho_t)}{\partial\rho_t}\sqrt{\psi(\rho_t)\zeta_t}dW_t^\rho
\end{aligned}$$

Then

$$\zeta_t d\psi(\rho_t) = \left(\beta(\bar\rho - \rho_t)\frac{\partial\psi(\rho_t)}{\partial\rho_t}\zeta_t + \frac{1}{2}\frac{\partial^2\psi(\rho_t)}{\partial\rho_t^2}\psi(\rho_t)\zeta_t^2\right)dt + \frac{\partial\psi(\rho_t)}{\partial\rho_t}\sqrt{\psi(\rho_t)}\zeta_t^{3/2}dW_t^\rho$$

Since the correlation between W_t^ρ and W_t^ζ is ξ_2, we obtain

$$d\zeta_t d\psi(\rho_t) = (\varsigma\sqrt{\zeta_t}dW_t^\zeta)(\frac{\partial\psi(\rho_t)}{\partial\rho_t}\sqrt{\psi(\rho_t)\zeta_t}dW_t^\rho) = \frac{\partial\psi(\rho_t)}{\partial\rho_t}\sqrt{\psi(\rho_t)}\varsigma\xi_2\zeta_t dt$$

Curriculum Vitae

Name:	Lin Fang
Education and Degrees:	Western University London, ON, CA 2016 - 2018 MSc. Sichuan University Chengdu, PRC 2012 - 2016 BSc.
Related Work Experience:	Teaching Assistant/Researching Assistant Western University 2016 - 2018