

A Study Of Stock Volatility In The Context Of Factor Volatility Models For Large Datasets: Factor Analysis And Forecasting

Lui, Sze Wai Silvia

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author

For additional information about this publication click this link.

<http://qmro.qmul.ac.uk/jspui/handle/123456789/1848>

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk



A Study Of Stock Volatility In The Context
Of Factor Volatility Models For Large
Datasets: Factor Analysis And Forecasting

by

Sze Wai Silvia Lui

A thesis submitted for the degree of Doctor of Philosophy
(Ph.D.) in Economics

University of London
Queen Mary
Department of Economics

August 2007

Author's Declaration

I wish to declare this thesis, titled as "A Study Of Stock Volatility In The Context Of Factor Volatility Models For Large Datasets: Factor Analysis and Forecasting", submitted to the University of London in pursuance of the degree of Doctor of Philosophy (Ph.D.) in Economics is my own work.

Signed

A handwritten signature in black ink, appearing to read 'Sze Wai Silvia Lui', with a small flourish at the end.

Sze Wai Silvia Lui

Abstract

This thesis is a study of stock volatility adopting two factor volatility models for large datasets: the orthogonal GARCH model and the stochastic volatility factor model. An application is made to the constituent stocks of five Asian indexes. Factor analysis and volatility forecasting exercise are carried out.

Chapter 1 is an empirical application of the orthogonal GARCH model to Asian stock returns. Correlation analysis, eigenvalue and eigenvector analysis and several diagnostic tests are carried out. Our results show that using large number of principal components cannot guarantee an improvement in capturing dynamics. Moreover, GARCH(1,1) is the appropriate specification for the principal components of stock returns of some datasets. An empirical example of how GARCH analysis of all series in the entire dataset can be summarised by a univariate GARCH analysis of the first principal component is also provided.

Chapter 2 is a factor analysis using stochastic volatility factor models. In contrast to the first part of the study, common factors are estimated from large datasets of Asian stock volatilities via principal components. Correlation analysis of stock volatilities is performed. Examinations of the dynamics of factor estimates and their explanatory power are also carried out. Our results

confirm that large dataset with many cross-sectional series from the same category may not always be desirable for factor analysis. Evidence of long memory is found in the first principal component of some datasets but not all of them.

Chapter 3 is a volatility forecasting exercise. In-sample analysis is implemented using the stochastic volatility factor models and the orthogonal GARCH model. Moreover, we propose an extension of a local-factor model to a multi-factor model. Testing of factor significance is scrutinised. A comparison of forecasting performance shows that the stochastic volatility factor models outperform the orthogonal GARCH model in forecasting Asian volatilities.

Acknowledgment

I would like to thank my parents for their love and support. They have been the biggest motivation for the completion of my Ph.D. study.

I would also like to thank Dr. Marika Karanassou, for her kind support and advice since my master degree. I also owe much to Dr. Andrea Cipollini who has spent time talking with me and making helpful suggestions.

Above all, I would like to express my deepest gratitude to my supervisor Professor George Kapetanios. This thesis could not have been accomplished without his invaluable guidance and continuous help in my research. It has been my greatest pleasure to learn from him.

Many thanks to Giles for helping me with the proof-reading. Special thanks go to Iolanda and Zeenat, who always stand by my side, sharing with me all the ups and down throughout my study. The days in room W316 will always be memorable.

List of Tables

Chapter 1

Table 1:	Descriptive statistics of the mean return series for the five indexes	73
Table 2:	Summary statistics of average correlations	74
Table 3:	Cumulative explained variation of the first 20 principal components from the five datasets	75
Table 4:	Distributional statistics of the first principal components	76
Table 5:	Test for unit root in first principal components	76
Table 6:	GARCH(1,1) estimation outputs for the first principal components	77
Table 7:	Descriptive statistics on the standardised residuals of GARCH(1,1) of the first principal components	78
Table 8:	LM test for serial correlation in GARCH(1,1) standardised residuals of principal components	78
Table 9:	ARCH LM test for heteroscedasticity in GARCH(1,1) standardised residuals of principal components	78
Table 10:	LM test results for serial correlation and heteroscedasticity of standardised residuals	79
Table 11:	Diagnostic test results on return residuals in the five datasets	80

Chapter 2

Table 1:	Descriptive statistics of average correlations of daily volatilities	136
Table 2:	Cumulative R^2 for the first 10 factor estimates of the five datasets	137
Table 3:	Breusch-Godfrey LM test statistics and Engle's ARCH LM test statistics	138
Table 4:	AR representations of the first factor estimates of the five datasets	139
Table 5:	ARFI(1,d) specification for the dominant factors of KOS200, NIK225 and NIK500	140

Chapter 3

Table 1:	Cumulative R^2 of KOS200 and NIK225 factors	204
Table 2:	No. of times (out of total no. of stocks) a factor is proved significant	205
Table 3:	Average adjusted R^2 of FARs including different factors	206
Table 4:	In-sample ARFI(1,d) estimation results for KOS200 factors	207
Table 5:	In-sample ARFI(1,d) estimation results for NIK225 factors	207

Table 6:	Serial correlation test on measurement errors of idiosyncratic volatilities	208
Table 7:	In-sample GARCH(1,1) estimation of first principal Component	209
Table 8:	Diagnostic test results on return residuals of OGARCH with only the first principal component	210
Table 9:	Forecasting performance comparison	211

List of Figures

Chapter 1		
Figure 1:	Time plots of mean returns	81
Figure 2:	Plots of first principal components from the five datasets of stock returns	82
Figure 3:	Factor weights on first principal component of KOS200	83
Figure 4:	Factor weights on first principal component of NIK225	84
Figure 5:	Factor weights on first principal component of NIK500	85
Figure 6:	Factor weights on first principal component of HSCI	86
Figure 7:	Factor weights on first principal component of SING	87
Figure 8:	Histogram of KOS200 first principal component	88
Figure 9:	Histogram of NIK225 first principal component	89
Figure 10:	Histogram of NIK500 first principal component	90
Figure 11:	Histogram of HSCI first principal component	91
Figure 12:	Histogram of SING first principal component	92
Figure 13:	Histogram of GARCH(1,1) standardised residuals of KOS200 principal component	93
Figure 14:	Histogram of GARCH(1,1) standardised residuals of NIK225 principal component	94
Figure 15:	Histogram of GARCH(1,1) standardised residuals of NIK500 principal component	95

Figure 16:	Histogram of GARCH(1,1) standardised residuals of HSCI principal component	96
Figure 17:	Histogram of GARCH(1,1) standardised residuals of SING principal component	97
Chapter 2		
Figure 1:	Plots of the first factors	141
Figure 2:	Autocorrelations of KOS200 first factor	142
Figure 3:	Autocorrelations of HSCI first factor	143
Figure 4:	Autocorrelations of NIK225 first factor	144
Figure 5:	Autocorrelations of NIK500 first factor	145
Figure 6:	Autocorrelations of SING first factor	146
Figure 7:	Autocorrelations of 1 st differenced NIK500 first factor	147
Figure 8:	Autocorrelations of 1 st differenced NIK225 first factor	148
Figure 9:	Autocorrelations of 1 st differenced KOS200 first factor	149
Chapter 3		
Figure 1:	In-sample KOS200 dominant factor	212
Figure 2:	In-sample NIK225 dominant factor	213
Figure 3:	Autocorrelations of in-sample KOS200 dominant factor	214
Figure 4:	Autocorrelations of in-sample KOS200 second factor	215
Figure 5:	Autocorrelations of in-sample KOS200 third factor	216

Figure 6:	Autocorrelations of in-sample KOS200 fourth factor	217
Figure 7:	Autocorrelations of in-sample KOS200 fifth factor	218
Figure 8:	Autocorrelations of in-sample NIK225 dominant factor	219
Figure 9:	Autocorrelations of in-sample NIK225 second factor	220
Figure 10:	Autocorrelations of in-sample NIK225 third factor	221
Figure 11:	Autocorrelations of in-sample NIK225 fourth factor	222
Figure 12:	Histogram of hyperparameters in measurement equations of KOS200 idiosyncratic volatilities of SSVF	223
Figure 13:	Histogram of hyperparameters in transition equations of KOS200 idiosyncratic volatilities of SSVF	224
Figure 14:	Histogram of hyperparameters in measurement equations of KOS200 idiosyncratic volatilities of MSVF	225
Figure 15:	Histogram of hyperparameters in transition equations of KOS200 idiosyncratic volatilities of MSVF	226
Figure 16:	Histogram of hyperparameters in measurement equations of NIK225 idiosyncratic volatilities of SSVF	227
Figure 17:	Histogram of hyperparameters in transition equations of NIK225 idiosyncratic volatilities of SSVF	228
Figure 18:	Histogram of hyperparameters in measurement equations of NIK225 idiosyncratic volatilities of MSVF	229
Figure 19:	Histogram of hyperparameters in transition equations of NIK225 idiosyncratic volatilities of MSVF	230

Contents

Introduction	15
1. Some Stylised Facts of Financial Volatility	16
2. Modelling Volatility -- The Basic Building Block of GARCH models and Stochastic Volatility Models	18
3. The Asian Stock Markets -- Some Backgrounds	24
Chapter 1: An Application of the Orthogonal GARCH Model to Large Datasets of Asian Stock Indexes Constituents	32
1. Introduction	32
2. The Model	34
2.1 From Univariate GARCH to Orthogonal GARCH – An Overview	35
2.2 The Orthogonal GARCH Model	44
3. Empirical Results	49
3.1 Data and Preliminary Statistical Analysis	49
3.2 Analysis of Average Correlations	52
3.3 Eigenvector and Eigenvalue Analysis	54
3.4 GARCH Analysis of the First Principal Components	58
3.5 Diagnostic Tests on Residuals	65
4. Concluding Remarks	70
5. Tables and Figures	73

Chapter 2: Factor Analysis of Stock Volatility Using Stochastic Volatility

Factor Model: Evidence from Five Asian Stock Indexes	98
1. Introduction	98
2. The Model	102
2.1 Stochastic Volatility -- An Overview	103
2.2 Stochastic Volatility Factor Model	108
2.3 Estimation of Common Factors: Principal Components Method	113
2.4 Investigation of Long Memory in Common Factors	117
3. Empirical Analysis	121
3.1 Correlation Analysis	122
3.2 Explanatory Power of Factor Estimates	124
3.3 Dynamics of Factor Estimates	129
3.3.1. Estimation of Autoregressive Models of Common Factors	130
3.3.2. Long Memory in Common Factors	132
4. Concluding Remarks	133
5. Tables and Figures	136

Chapter 3: Forecasting Stock Volatility: Stochastic Volatility Factor

Models versus Orthogonal GARCH Model	150
1. Introduction	150
1.1. Orthogonal GARCH Model and Stochastic Volatility Factor Model – A Comparison	152

2. The Models	156
2.1. Stochastic Volatility Factor Models	156
2.1.1. The Single Local-Factor Specification -- A Revision	156
2.1.2. The Multi-factor Specifications -- An Extension	163
2.1.3. Testing for Factor Significance	169
2.2. Orthogonal GARCH Model	173
3. Forecasting Methodology	176
3.1. Forecasting with the Single Local-Factor Specification	176
3.2. Forecasting with the Multi-Factor Specifications	179
3.3. Forecasting with Orthogonal GARCH Model	181
4. Forecast Evaluation	182
5. Empirical Results	183
5.1. In-sample Analysis using Stochastic Volatility Factor Models	184
5.2. In-sample Analysis using Orthogonal GARCH Model	192
5.3. Out-of-sample Forecasting	194
6. Concluding Remarks	201
7. Tables and Figures	204
Conclusion	231
References	235

Introduction

Study of financial volatility has long been under the interest of financial researchers and international investors. Developments of volatility models, modifications of existing modelling and estimation methodologies, as well as improvements of forecasting procedure are ongoing and will continue to be under a great deal of attention. Research in this topic was initially aim to improve the modelling method in order to account for time-varying properties of financial volatility and its observed stylised facts. Different univariate volatility models have thus been proposed. However, univariate models only allows for the study of single financial time series. When the interest is to consider linkages and comovements of multiple financial time series, univariate models become insufficient. Evolution of the study thus moved from focusing on one financial asset or one financial market, to the consideration of the relationships between different financial assets and linkages between different financial markets. This has triggered the development of multivariate modelling technique. Proposals of different multivariate models and estimation procedures have undoubtedly nourished the study of financial time series. However, the fact that when many time series are modelled simultaneously has made modelling and estimation procedures complicated. Recent development in the topic allows one to use

common factors to represent the entire system of data series and the use of factor volatility models are advocated.

The aim of this thesis is to provide an empirical factor analysis and an evaluation of forecasting performance of two factor volatility models, the orthogonal GARCH model of Alexander (2001a, b) and the stochastic volatility factor model of Cipollini and Kapetanios (2005), with applications to Asian stock volatilities. These two models stem from the two main categories of volatility models – the GARCH family models and the stochastic volatility models. We believe our factor analysis could provide further insight into an understanding of the comovements in, the modelling and forecasting of Asian stock volatilities.

In the remainder of this introductory chapter, we will discuss some commonly observed stylised facts in financial volatility that have given indication to the developments and modifications of volatility models. We will also look at the basic building block of the two major types of volatility models. Some backgrounds of the Asian stock markets will then be presented. Finally, we will outline the organisation of this thesis.

1 Some Stylised Facts of Financial Volatility

Prior to the invention of volatility models with time-varying conditional variance, the Autoregressive Moving Average (ARMA) type models had been commonly used for the study of time series data. However, a major shortcoming of the ARMA models is the assumption of constant variance over time, which

has been proved unrealistic in many studies of financial time series. The need to account for the time-varying properties of volatility has initiated the invention of various types of volatility models. The large amount of empirical studies in the topic has revealed some stylised facts of financial volatility. This has in turn, provided indications to further developments and modifications of different methods of volatility modelling.

The design of different volatility models aims to capture the following well-known stylised facts in financial volatility. First of all, distribution of financial time series normally has fatter tails than the normal distribution. It is commonly found that kurtosis of financial time series is above 3, which is the standardised fourth moment of normal distribution. It means large positive and negative returns occur more often than expected. Moreover, volatility clustering has been commonly observed. It means there are alternate periods of large movements in financial prices and periods when the prices have hardly moved. Another interesting feature of financial prices is that negative shocks seem to have more pronounced effect on volatility than positive shocks would have on it. A negative relationship between price movement and volatility can be observed from many stocks. The nature of declining volatility with increasing returns and the increasing volatility with decreasing returns is known as leverage effect. Volatility models with asymmetry have been proposed to account for this feature.

In addition to the above characteristics, developments of volatility models in the recent decades also consider long memory in financial volatility. Persistence is generally observed in financial volatility, especially in high frequency

financial data. Evidence can be seen from plots of autocorrelations from which significant correlations exist even at long lags. Conditional variance equations are evident with near unit root rather than exact unit root. It means autocorrelations of shocks exhibit a slow hyperbolic decay rather than an exponential decay. The objective to capture persistence in volatility has promoted the development of long memory volatility models. For example, the long memory GARCH models (see for example, Baillie, Bollerslev, Mikkelsen (1996)), and the long memory stochastic volatility models (see for example, Harvey (1993)) have been proposed. Finally, movements in one financial market always induces movements in another financial markets. Stock returns in several markets change contemporaneously. This is an evidence of cross-markets relationships. It is therefore more appropriate to examine cross-markets financial time series simultaneously. This has addressed the importance of using multivariate models to study comovements in volatility in different financial markets and different financial assets.

2 Modelling Volatility – The Basic Building Block of GARCH Models and Stochastic Volatility Models

This thesis concerns factor analysis and forecasting using factor volatility models. The factor volatility models we adopt are the advanced developments of the two comparable volatility models, the GARCH model and the stochastic volatility models. We present here a brief explanation of the basic set-up of these two

models. A thorough review of the GARCH and stochastic volatility models can be found in Knight and Satchell (2007). Reviews of GARCH models can be found in Franses and van Dijk (2000). Bauwens, Laurent and Rombouts (2006) also provides a survey of multivariate GARCH models. Furthermore, the survey papers by Poon and Granger (2003) and Andersen, Bollerslev, Christoffersen and Diebold (2006) discuss the various volatility models used for forecasting.

The basic building block for modelling volatility is to consider the shocks in the mean of a financial time series, ε_t , as a product of a random variable, z_t , and a factor, σ_t^2 . Where z_t is an independently identically distributed random variable with mean zero and unit variance, and σ_t^2 is the time-varying variance, and $t = 1, \dots, T$. That is,

$$\varepsilon_t = z_t \sigma_t \tag{1}$$

The task is then to find a process to model the variance, σ_t^2 , so that the stylised facts of financial time series can be well-captured. The variety of volatility models that have been proposed for this purpose can be classified into two main approaches. The first approach refers to modelling the variance using past observations of the shocks via autoregressive conditional heteroscedasticity, that is the ARCH models, first proposed by Engle (1982). Further generalisation of this model by Bollerslev (1986) allows the past observations of the conditional variance impact on the current conditional variance, that is the GARCH models. Consider the following GARCH(1,1) model,

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 \tag{2}$$

where $\beta_0 > 0$, $\beta_1 > 0$, and $\beta_2 > 0$ to ensure non-negative σ_t^2 . The sum of β_1 and β_2 has to be less than one to ensure covariance stationarity. The unconditional fourth moment of ε_t is finite if $(\beta_1 + \beta_2)^2 + 2\beta_0^2 < 1$ (see Bollerslev (1986)). The kurtosis of ε_t is greater than 3 if z_t also follows a normal distribution, which means its distribution has a fatter tails than the normal distribution.

The most distinguished attribute of the ARCH and GARCH models are their capability to capture volatility clustering. Moreover, the standard GARCH model is linear. Current volatility is modelled by an ARMA process with past shocks and past volatilities. This set-up has favoured the estimation procedure as σ_t^2 can be deduced by using the past information of the shocks. This has made maximum likelihood estimation tractable. Nevertheless, there are still some drawbacks of the standard GARCH model. First of all, normal conditional densities, which are commonly used for the functional form of the standard GARCH model, is insufficient to account for the fat tails in the unconditional distributions of financial prices and returns. Evidence of excess kurtosis not being captured under normality assumption is found, see for example, Baillie and Bollerslev (1989). This has led to the use of some non-normal distribution in GARCH modelling. Secondly, the impact of a shock on volatility depends only on its magnitude in standard GARCH model and the sign of the shock is irrelevant. However, asymmetry is generally observed in financial time series that means a negative shock has larger impact on volatility than a positive shock with the same magnitude would have on it. Thirdly, the non-negativity constraints on the parameters in GARCH models may cause

difficulties in estimation procedures (see Rabemananjara and Zakoain (1993)). Under the non-negativity constraint, past shocks always have a positive impact on volatility at current period. Larger size of the shock gives larger impact on current volatility, regardless of its sign. Thus, non-linearity in volatility cannot be captured by the model. Finally, standard GARCH model fails to capture the near unit root property in conditional variance.

The deficiencies of standard GARCH models have led to further modifications and developments in the literature, including the exponential GARCH model and various extensions of GARCH model that consider threshold effect. These models are capable to capture the leverage effect in financial time series. Furthermore, there are also developments which aims to account for persistence in conditional variance, for example, the integrated GARCH model and the Fractionally integrated GARCH model. All these developments aim to seek for modifications of the conditional variance equation so that the commonly observed stylised facts can be well-captured. In the first chapter of this thesis, an overview of the GARCH volatility models will be presented.

An alternative class of volatility models which have also received considerable attention is the stochastic volatility. In the stochastic volatility models, the variance, σ_t^2 in equation (1) above is treated as an unobserved variable and it is modelled by a stochastic process, for example in its simplest form, is an autoregression

$$h_t = \alpha_0 + \alpha_1 h_{t-1} + v_t \tag{3}$$

where $h_t = \ln(\sigma_t^2)$. v_t is independently normally distributed with mean

0 and variance σ_v^2 . v_t and z_t are uncorrelated. The conditional variance, h_t , is strictly stationary with mean μ_h equals $\frac{\alpha_0}{1-\alpha_1}$ and variance σ_h^2 equals $\frac{\sigma_v^2}{1-\alpha_1^2}$ if $|\alpha_1| < 1$. Moreover, ε_t in equation (1) is a white noise process due to the independence of v_t and z_t . h_t is not deterministic conditional upon the past history, which is opposite to the GARCH model.

The set-up of the stochastic volatility models can overcome some drawbacks in the standard GARCH model. The above autoregression is not set up for the unobserved variance σ_t^2 , but for the logarithm of it, that is, h_t . Therefore, σ_t^2 is always positive. In addition, if z_t is normally distributed, the kurtosis of ε_t is $3 \exp(\sigma_h^2)$, which is greater than the normal value, 3^1 . In contrast to the standard GARCH model, the stochastic volatility model is more capable to account for the excess kurtosis in financial time series as the term $\exp(\sigma_h^2)$ can take any value. Furthermore, if we relax the assumption that z_t and v_t are independent and allow for $Cov(z_t, v_t) < 0$, asymmetry in financial returns can be accounted for (see Knight and Satchell (2007), Hull and White (1987), and Engle (1982)).

Although the stochastic volatility model allows more flexibility than the standard GARCH model in accounting for some of the stylised facts in financial volatility, the popularity of GARCH models in financial applications has never been adversely affected. The reason for this is that in contrast to the standard GARCH model, estimation procedure for the stochastic volatility model is more complicated. Different estimation methods for the models have been examined. These procedures include generalised method of moments approach

¹For derivation of the kurtosis coefficient of the stochastic volatility model, see Franses and van Dijk (2000).

(see Melino and Turnbull (1990), and Andersen and Sørensen (1996)), quasi-maximum likelihood via Kalman filter by Harvey, Ruiz and Shephard (1994), and Bayesian method (see Jaquier, Nicholas and Rossi (1994)). Even though there has not been an agreement on which method is the appropriate one, these progressions have enriched the modelling of stochastic volatility over time. The above set-up of stochastic volatility model in equation (3) is the simplest form of the model. Further developments of the model have been proposed in order to capture more stylised facts of financial volatility. A brief overview of the literature on stochastic volatility models will be presented in a later chapter of this thesis.

It is worth mentioning that although the stochastic volatility models may seem to be more appealing than the GARCH models, it is difficult to determine upon which of them are the winning models from their practical performance. The decision upon which of the two models is used for analysis should depend on the nature of the data we are dealing with and a model selection procedure should be implemented prior to further analysis and forecasting exercises.

Notice that our discussion of the two models so far has limited to the univariate set-ups. As mentioned at the beginning of this introductory chapter, univariate volatility model limits one's interest in a single financial time series. The need to examine or study different financial time series or different financial markets simultaneously promotes the use and evolution of multivariate volatility models. Although different multivariate versions of GARCH model and stochastic volatility model have been proposed to approach the issue, estimation methods involved are always complicated when the number of time

series considered gets very large. This necessitates the simplification of the modelling procedure. The idea of using a few factors to summarise comovements among a large system of data series and thus simplify the modelling and estimation procedure has become popular. Later in the chapters of this thesis, we will discuss the multivariate versions of the two volatility models and the incorporation of common factors into the models.

3 The Asian Stock Markets – Some Backgrounds

Volatility factor models provide a platform to study large datasets of financial assets, the comovements among them and the linkages among financial markets with ease. By adopting the factor volatility models, our empirical study aims to analyse the common factors underlying the Asian stocks and forecasting Asian stock volatilities. Before we move on to the core of this thesis, we present here some background of the Asian stock markets.

Stock markets in Asia have attracted a great deal of attention from researchers, academics and international investors in recent decades due to their sound record of trade and investment. High returns from the investment in the Asian stock markets in comparison to other developed markets have contributed to their popularity. However, the high returns always accompanied by high volatility. This has provided an interesting topic for financial researchers and an incentive for the international investors to further explore the behaviour of these markets. Efforts have been put on studying for the intra-regional linkages among those markets, as well as the international linkages between

them and other financial markets around the world. Other researchers have attempted to study the variation in Asian stock returns, modelling and forecasting Asian stock volatility. Implications and conclusions drawn from these studies have provided crucial information on security pricing strategies and global hedging strategies.

The positions in the world's rankings of stock markets taken by some South-east Asian stock markets reveal their important roles in international investment. In 1995, the Japanese stock market was ranked the second, after the US market, with a total market capitalisation of US\$366.7 billion. Hong Kong stock market was ranked the 9th with total market capitalisation of US\$30.4 billion. South Korea and Singapore were ranked the 15th and 17th, with market capitalisation of US\$18.2 billion and US\$14.8 billions, respectively. Although the markets have been impacted by events like the 1997 financial crisis, their importance remained. The total market capitalisation of the Japanese stock market fell to US\$315.7 billion in year 2000. But it remained at the 2nd place, after the US and followed by the UK. The Hong Kong market has gone down to the 10th largest but its market capitalisation doubled between 1995 and 2000, to US\$62.3 billion. The Singaporean market has also moved down to a lower position, to the 22nd place but with an increase of capitalisation to US\$15.3. Whereas the South Korean market was ranked the 20th, with a fall in capitalisation to US\$17.2 billion during the period². The significant international position of the Southeast Asian stock markets continue even until recent years. Tokyo Stock Exchange of Japan was ranked the 2nd with market capi-

²Figures detail from Li (2006). "The Hong Kong Economy: Recovery and Restructuring". Source: Standard and Poor's Stock Market Factbook, 1996 and 2001.

talisation of US\$20.7 trillion as at May 2007. The Hong Kong Stock Exchange and the Korea Stock Exchange were ranked the 7th and the 16th with market capitalisation of US\$1.92 trillion and US\$1 trillion, respectively³.

The role of Asian stock markets is undoubtedly significant, the linkages among those markets and their relationships with the other world stock markets are also tight. These linkages have been revealed through different empirical literatures on volatility transmission, causal relationship and financial contagion, and also through some daily news. These linkages are even more obvious in the situation of significant events. We list out the following three events in the past decade as examples to illustrate these linkages.

The 1997 Financial Crisis⁴

It is no doubt that the 1997 Asian financial crisis can be viewed as the most notable shock in the financial markets of Southeast Asia in the past decade. Issues of the crisis have been under attention and debate for many years. Some economists understood the crisis as a "bubble burst" after years of economic growth, exchange rate stability and prosperity within the region. Whereas some others take a different view and argue that this is purely a consequence of speculative attack. No matter what has been the actual cause of the crisis, the impact of the event on the Asian currency markets and stock markets, as well as some overseas markets was immense. The initial currency crisis,

³Ranking by market capitalisation as at May 2007. Source: World Federation of Exchanges, June 2007.

⁴Source of information and figures:

Chronology of Asian currency crisis and its global contagion, Nouriel Roubini's Global Macroeconomic and Financial Policy Site, <http://pages.stern.nyu.edu/~nroubini/asia/>
1997 East Asian financial crisis, Wikipedia, the free encyclopedia, http://en.wikipedia.org/wiki/Asian_financial_crisis

which was incurred by speculative attack in Thailand, has propagated rapidly across the countries in Southeast Asia. It has then evolved quickly into stock market clashes and economic crises. Following the Thai and the Indonesian stock market clashes in June and August, the Hong Kong stock market also suffered from its heaviest slump ever between 20 to 23 October 1997. The clash was even more severe than the one in 1987. Hang Seng index plummeted by more than 20%. After regaining 718 points on 24 October 1997, the index lost another 5.8%. South Korean stock market also fell sharply by 6.9% on 7 November 1997 and plunged by another 7.2% on 24 November. The shocks in stock markets also spread to some Latin American markets, most notably the Brazilian and Argentinean markets, and they also suffered from heavy losses. The US market experienced its largest point loss ever in October 1997, the Dow Jones Industrial Average suffered by 7% loss and the New York Stock Exchange suspended trading for a short time. The Asian stock markets experienced the losses again in November 1997. The NIKKEI 225 index of Japan dropped to its lowest level in more than two years and the market faced major sell-off after that.

Loss incurred and damages caused to the currency and stock markets in Southeast Asia were enormous. Crisis countries attempted to stabilise their financial system by requesting for international financial aids. The International Monetary Fund, the World Bank and the Asian Development Bank have taken vital roles in securing financial stability in the crisis region. Some governments in the crisis region have also exercised some domestic emergency procedures to restore stability and investors' confidence, such as the purchase of shares and

futures contracts in its local market by the Hong Kong government.

The Dot-com Bubble⁵

The bursting of the dot-com bubble can be viewed as the beginning of the early 2000s recession, which has affected mostly the western countries. The dot-com boom covered the period 1995 and 2001 and stock markets in the Western countries enjoyed a surge due to the rapid growth in the IT and related industries. However, the Asian markets started to take pleasure from the boom since 1999. After the painful period of the 1997 financial crisis, Asian firms were eager to search for revenue that could be earned at low cost. The dot-com boom has offered them with new hopes to make easy money in a short period by just launching new businesses which have shared the common label ".com". An example of how the Asian market benefited from the boom can be seen from a Hong Kong internet venture, Tom.com that launched its Initial Public Offering on NASDAQ and its share price increased by five times by the end of its first trading day.

However, the boom came to an end when the US technology NASDAQ composite index reached its peak at 5048.62 on 10 March 2000, exceeded a double of its value a year before. The bubble bursted shortly after that. On 13 March, NASDAQ opened about 4% point lower and the index lost almost 9%

⁵Source of information and figures:

"Dot-com bubble", Wikipedia, the free encyclopedia.
http://en.wikipedia.org/wiki/Techology_bubble/

"Asia falls as techs swoon", CNN Money, 4 April 2000.
<http://money.cnn.com/2000/04/04/markets/asiawrap/>

"Tokyo avoids Asian slump", CNN Money, 24 April 2000.
<http://money.cnn.com/2000/04/24/markets/asiawarp>

"From boom to bust and back again", ZDNet Asia, 29 Novemeber 2005.
<http://news.zdnet.co.uk/internet/0,1000000097,3923883,22.htm>

in three days time. The Asian stock markets were inevitably affected by this burst of bubble. The Tom.com suspended trading six months after its launch of IPO on NASDAQ. On 4 April 2000, NIKKEI index fell by 0.6% with prices of Japanese internet stocks plunged after the volatile NASDAQ. The Singapore Strait Times Index also tumbled by 0.6% on the same day and a 1.47% fall on 24 April due to a selling of financial and technology stocks. Whereas the Korea Stock Exchange Composite (KOSPI) Index and the Malaysian Kuala Lumpur Stock Exchange Composite (KLSE) Index fell by 1.9% and 0.5% on 4 April, respectively. The KOSPI index fell another 2.55% on 24 April due to the decline in telecommunication and technology stocks, while the KLSE index fell by 1.23%. The dot-com bubble burst has also caused many job losses and bankruptcies in the IT sector in the Asian countries.

The September 11 Terrorists Attacks⁶

The September 11 attacks have caused a sharp plummet in global stock markets. The New York Stock Exchange, American Stock Exchange and the NASDAQ were closed between 11 and 17 September. The Dow Jones Industrial Average experienced its largest one-day point decline with 7.1% fall and the largest one-week point decline with 14.3% fall when the stock markets reopened.

⁶Source of information and figures:

"Gloom lifts from stunned markets", BBC News, 12 September 2001.
<http://news.bbc.co.uk/1/hi/business/153916.stm>

"Stock prices across Asia plummet in wake of attacks in the US", International Herald Tribune. http://www.ihl.com/articles/2001/09/13/t15_8.php

"Asian markets open in shock", CNN Money. 11 September.
http://money.cnn.com/2001/09/11/markets/markets_asia/

"World economic effects arising from the September 11, 2001 attacks", Wikipedia, the free encyclopedia.
http://en.wikipedia.org/wiki/World_economic_effects_arising_from_the_September_11%2C_2001_attac

"September 11, 2001 attacks", Wikipedia, the free encyclopedia.
http://en.wikipedia.org/wiki/September_11_2001#Economic_aftermath

The Asian stock markets were also in turmoil. Following the attacks on the 11th, NIKKEI 225 fell by 6.6% the next day, while the Hang Seng Index fell by 7% soon after the market was open and fell by 10% shortly after that. The South Korean stock market had a three hour delay in opening and the KOSPI index fell by 12%. The Singaporean shares have also tumbled by 7.4%. Whereas the stock markets in Taiwan and Malaysia were closed.

Although this thesis does not focus on the study of the impact of extreme events on Asian stock volatility, the aforementioned events and the facts given at the beginning of this section both reveal the close relationships between the Asian markets and the financial markets around the world and the significance of those markets in international investment. Application of the stochastic volatility factor model of Cipollini and Kapetanios (2005) to Asian volatilities for factor analysis and volatility forecasting is original in this thesis. We hope our empirical application can make a contribution to the study of Asian stock volatilities through factor analysis and thus provide further insight into the understanding of the common movements of Asian stock volatilities.

4 Outline of the thesis

This thesis is a study of stock volatility adopting two factor volatility models for large datasets: the orthogonal GARCH model and the stochastic volatility factor model. An application is made to the constituent stocks of five Asian indexes. Factor analysis and volatility forecasting exercise are carried out. The study consists of three major parts. The first part of the study is an empirical

application of the orthogonal GARCH model to Asian stock returns. We focus on investigating how well the model can perform when applying to Asian stock returns. In particular, several diagnostic tests are carried out to see if the model with one principal component or more than one principal component is more desirable; and whether the commonly used GARCH(1,1) is the appropriate specification for the principal components of stock returns. Correlation analysis, and eigenvalue and eigenvector analysis are also considered.

The second part of the thesis is a factor analysis carried out using stochastic volatility factor models. In contrast to the first part of the study, common factors are estimated from large datasets of Asian stock volatilities via principal components of Stock and Watson (2002a, b). Correlation analysis of stock volatility is performed. Examination of the dynamics of factor estimates and their explanatory power is also carried out. Moreover, we investigate long memory in common factors of stock volatilities and analyse the impact of the size of dataset on explanatory power of factor estimates. We compare several results obtained here with those from the first part.

The last part of the study is an exercise of volatility forecasting. In-sample dynamics of idiosyncratic volatility are investigated in the context of the stochastic volatility factor model. Moreover, we propose an extension of a local-factor model to a multi-factor model that is believed to be more applicable for empirical situation. Testing of factor significance and a comparison of forecasting performance between the factor models and the orthogonal GARCH model are scrutinised.

Chapter 1

An Application of the Orthogonal GARCH Model to Large Datasets of Asian Stock Indexes Constituents

1 Introduction

The invention of univariate ARCH and GARCH models by Engle (1982) and Bollerslev (1986), and the further generalisation in both univariate and multivariate settings have provided financial researchers with means to carry out more realistic modelling and forecasting of financial volatility. Multivariate GARCH models are widely used for the simultaneous study of the relations between the volatilities and covariances of various financial assets and between various financial markets. However, the deficiency of multivariate GARCH models in estimation, interpretation, as well as computation of large variance-covariance matrices have led to further developments. These developments can be classified into two streams. In one stream, the shortcomings of the multivariate GARCH models has encouraged academics to seek modifications of

the volatility models, which is an alternative of GARCH, that is the stochastic volatility models. For example, the multivariate stochastic volatility model by Harvey, Ruiz and Shephard (1994) has resolved the problems encountered in multivariate GARCH modelling. Further development of this model is the stochastic volatility factor model by Cipollini and Kapetanios (2005).

In the other stream, the drawbacks have encouraged further development in the GARCH modelling itself – the use of latent factor in GARCH modelling. In this chapter, we will focus on an improvement in multivariate GARCH model that is ideal for datasets with large dimensions and at the same time, able to resolve difficulties encountered in multivariate GARCH models – the orthogonal GARCH model by Alexander (2001a, b). Orthogonal GARCH model provides a breakthrough to the conventional GARCH modelling by incorporating principal component analysis into GARCH analysis. A few principal components are used to represent the entire system of data series. Reduction of dimensionality is thus achieved. Only univariate GARCH estimation is required for the construction of volatilities and covariances, computation is then simplified. Moreover, positive definiteness of variance-covariance matrix can be achieved.

In our empirical study, we carry out an application of the orthogonal GARCH model to large datasets of constituent stock returns of five Asian stock indexes. We carry out an analysis to examine the average correlations in the Asian stock returns and to investigate the relationship between overall correlations among the return series in a dataset and the explanatory power of the principal components. We then examine the distributional characteristics of the principal components that represent the systems of Asian stock returns. Uni-

ivariate GARCH analysis is carried out to check if GARCH(1,1) is appropriate specifications for Asian stock return principal components. Our analysis also provides empirical examples on how GARCH analysis of the dataset of our Asian returns can be summarised by univariate GARCH analysis of their first principal component. Finally, diagnostic tests are scrutinised to check if the first principal component is able to capture enough dynamics in the datasets of Asian stock returns. We also provide empirical evidence to show that using large number of principal components in modelling may allow one to benefit from larger amount of total variation being explained, but it cannot guarantee an improvement in capturing dynamics in the datasets, especially when series in the datasets are not very highly correlated.

Organisation of this chapter is as follow. Section 2.1 gives a brief overview of how univariate GARCH model is developed to orthogonal GARCH model. The orthogonal GARCH model is discussed in section 2.2. Empirical analysis is described in section 3. Concluding remark is in section 4.

2 The Model

In this section, we will first give an overview of how the univariate GARCH model has been extended to the various specifications that account for complicated time properties of financial time series. We will discuss the progressions of univariate GARCH models and their development into multivariate GARCH models. Advantages and drawbacks of these models are discussed. We will also review further advances into models involving latent factors in view of the

shortcomings of the multivariate GARCH models. The orthogonal GARCH model is discussed in section 2.2.

2.1 From Univariate GARCH to Orthogonal GARCH – An Overview

Time-varying volatility, also known as heteroscedasticity, is one of the most well-known features in financial returns. Ever since the introduction of the univariate Autoregressive conditionally heteroscedasticity (ARCH) models by Engle (1982), and the further generalization of the model, GARCH, by Bollerslev (1986), the GARCH family models have been used extensively for volatility modelling and forecasting. The standard GARCH models are popular for the fact that they are not only able to capture heteroscedasticity. They are also able to capture some other common characteristics found in financial time series, including excess kurtosis and thick tailedness. And also volatility clustering, that is, periods of large returns alternate with periods of small returns.

In standard univariate GARCH modelling, time varying volatility is captured by allowing the conditional volatility, σ_t^2 , to be a function of past squared observations of unexpected returns and past variances. In a GARCH(1,1) representation, conditional volatility evolves over time via an ARMA(1,1) with the autoregressive term being the past volatility and the moving average term being the past squared shocks to the time series. Conditional volatility is driven by the same shocks as its conditional mean. Maximum likelihood estimation of univariate GARCH model is straight forward due to the fact that although

the conditional volatility is unobserved, it can be computed by using the information of past shocks.

Although the standard GARCH models can capture some stylised facts of financial time series, they are far too simple to explain some more complicated time properties of financial volatility which have also been empirically proven. In the simple GARCH model, the impact of a shock on volatility depends on its magnitude. The sign of the shock is not taken into account. However, it is empirically observed that a negative shock in the current period causes higher conditional volatility in the next period than a positive shock would. This non-symmetrical dependence is not captured by standard GARCH models.

The exponential GARCH (EGARCH) model introduced by Nelson (1991) is the earliest improvement which aims to capture the asymmetric effects of a shock on conditional volatility. In the EGARCH model, conditional volatility is specified in logarithmic form. There is therefore, no need to impose estimation restrictions on the parameters to ensure nonnegativity of the conditional volatility. Other models that consider asymmetric impact of shocks are for example, Threshold GARCH (TGARCH), the GJR-GARCH introduced by Glosten, Jagannathan and Runkle (1993), Quadratic GARCH (QGARCH) of Sentana (1995), and the Logistic Smooth Transition GARCH (LST-GARCH) discussed by González-Rivera (1998), etc.

Apart from the leverage effect, another empirically proven time property of financial time series that has attracted a lot of attention is volatility persistence. The impact of a shock on conditional volatility seems to have very long memory. Covariance stationarity in a standard GARCH(1,1) model exists if and only

if the persistence parameters sum to unity. It implies that the conditional volatility is persistent. Engle and Bollerslev (1986) refer the model as the Integrated GARCH (IGARCH) model when the parameters sum to one. In the IGARCH model, conditional variance is a decaying function of the shocks in past and current periods. However, Nelson (1990) argues that the sum-to-unity constraint can cause very parsimonious representation of the distribution of a financial return. In fact, a shortcoming of IGARCH model is that the autocorrelations of the shock is not well-defined.

Although the autocorrelations approximation for an IGARCH model given by Ding and Granger (1996) shows exponentially decaying autocorrelations of the shocks, it is in fact commonly observed that the autocorrelations decay is slower than what is assumed by the IGARCH model. When dealing with financial time series with high frequency, one is more likely to obtain the sum of the persistence parameters very close to one. In fact, a hyperbolic decay is a more realistic pattern of autocorrelations (see Ding, Granger and Engle (1993)). This characteristic is regarded as long memory in financial time series. Baillie, Bollerslev and Mikkelsen (1996) propose a further development of the GARCH model to incorporate long memory, that is, the Fractionally Integrated GARCH (FIGARCH(p,d,q)) models. In the FIGARCH model, the fractional differencing operator has value lies between 0 and 0.5 and autocorrelations of the shocks exhibit hyperbolic decay. The model by Davidson (2004) also considered long memory in financial time series.

The various GARCH models that have been discussed so far are all in univariate settings, in which the impact of shocks on the volatility of a single

asset is studied. In reality, the arrival of news and shocks impacts on various financial assets simultaneously, and thus volatilities of different financial assets move together over time. Multivariate modelling of conditional volatility is therefore more suitable for studying the real financial situations. An extension of univariate GARCH models to multivariate GARCH models allows one to study the relationships among volatilities of different financial assets and different financial markets. Moreover, when the formation of covariances among different financial assets are needed, for example in portfolio construction, multivariate GARCH models play a crucial role to model time-varying conditional covariances of these assets.

Unlike the univariate GARCH model, conditional variances and covariances in multivariate GARCH model is given by a matrix, H_t . Parameterisation for this conditional variance-covariance matrix need to be specified in order to construct the model. However, this is not as straight forward as in univariate GARCH modelling. Various parameterisation of this conditional variance-covariance matrix have been proposed, and this has lead to a board literature on multivariate GARCH models.

A generalisation of matrix H_t is proposed by Bollerslev, Engle and Wooldridge (1988), in which the vec transformation is applied. The $vech(\bullet)$ operator stack the lower portion of the matrix H_t such that $vech(\bullet)$ contains all unique elements of H_t . Engle and Kroner (1995) thus call this model as the VEC representation. The VEC model is a flexible setting as it allows all elements in the conditional variance-covariance matrix to be a linear function of the lagged squared shocks, cross-products of the shocks, and the lagged values of

the elements in the conditional variance-covariance matrix. However, the VEC model has two main drawbacks. First, it is difficult to ensure positive definiteness of matrix H_t without imposing strong constraints on the parameters. Second, the number of estimated parameters explodes as the number of data series increase. Another specification discussed by Engle and Kroner (1995) is the BEKK model, which is acronym from the joint work on multivariate simultaneous generalised ARCH by Baba, Engle, Kraft and Kroner (1990). In the BEKK model, parameters in the generating process of the conditional variance-covariance matrix are expressed in quadratic forms. In this parameterisation, positive definiteness of H_t can be obtained without the need to impose constraint on the parameters. However, the problem of too many estimated parameters remains. Because of this reason, the VEC and the BEKK models are not widely used for datasets consist of many series.

An alternative multivariate GARCH model that has also guaranteed positive definiteness of the conditional variance-covariance matrix is the conditional correlation models proposed by Bollerslev (1990). This specification can be regarded as a nonlinear combination of univariate GARCH models. Conditional correlations between the shocks of different series in a dataset are assumed constant over time. The conditional covariances are proportional to the product of their conditional standard deviations. The number of parameters can be reduced by imposing these constraints. Individual conditional volatilities are assumed to have a univariate GARCH(1,1) representations. However, positive definiteness of matrix H_t can be guaranteed if and only if all individual conditional variances modelled by the univariate GARCH models are positive and

the correlation matrix is positive definite. Moreover, the time-invariant conditional correlations assumption seems to be unrealistic. An alternative model to the conditional correlation model is the dynamic conditional correlation (DCC) model proposed by Engle (2002), which assumes time-dependent conditional correlation matrix. However, the major disadvantage of the DCC model is that the parameters are assumed to be scalar, so all the conditional correlations obey the same dynamics. Moreover, although estimation of the DCC model can still be implemented via a two step approach when a large number of series is considered, the restriction of common dynamics become stronger and it is more difficult to deal with.

It is undoubtedly that GARCH models have provided a lot of advantages in volatility modelling and forecasting. Multivariate GARCH models allow one to generate large variance-covariance matrix. In particular, when modelling term structure, the models are able to give mean-reverting term structure of volatility and correlation with simple analytic form, making the use of them become popular. However, we can also see the shortcomings in various multivariate GARCH models from the above discussion. When a large system of series is considered and the need for a large covariance matrix is required, e.g. for the purpose of pricing and hedging, the use of multivariate GARCH models becomes problematic. In general, multivariate GARCH models all expose to similar disadvantages when number of series in a dataset increases – too many parameters to be estimated, flat likelihood function, and convergence problems arise in the estimation process. Moreover, using GARCH models to compute a large positive semi-definite covariance matrix becomes more complex when the

number of series included in a system increases. Parameterization of multivariate GARCH is normally required for computation and it is hard to implement.

The above shortcomings of multivariate GARCH models, together with the increasing interest in modelling datasets with large dimensions have led to further progressions in the literature that involves using latent factors to summarise the comovements among a large number of data series by a few common factors. Early work of multivariate GARCH models that involve retrieving latent factor that have GARCH property can be seen from Diebold and Nerlove (1989), King, Sentana and Wadhvani (1994) and Dungey, Martin and Pagan (2000). However, these models are difficult to deal with when apply to large datasets. Alexander (2001a, 2001b) (see also Alexander and Chibumba (1997)) has shown how problems in estimation and the creation of large positive definite variance-covariance matrix can be resolved by introducing orthogonal factors and principal component analysis into multivariate GARCH modelling. This type of GARCH models is known as principal component GARCH or according to Alexander (2001a, 2001b), the orthogonal GARCH model (see also Ding (1994), and Engle, Ng and Rotschild (1990)). The principal component GARCH model and the orthogonal GARCH model are very similar, except the latter suggests the use of less principal components to represent the entire system of data series in order to ensure stable correlation and robust volatility estimates. The idea behind these models is to diagonalised the multivariate problem, and thus only univariate GARCH estimation is involved. In the orthogonal GARCH model of Alexander (2001a, b), orthogonal factorisation via principal component analysis is exercised. Principal components

extracted from the correlation matrix of a system of assets are used to represent the entire system. The time-varying variances of the principal components are then estimated using univariate GARCH models. Variance-covariance matrix of the entire system is thus approximated as the volatilities of the principal components time the squared of factor weights.

Orthogonal GARCH model has a number of advantages over standard GARCH models. One of the most important advantages is that only the first few principal components representing the system of data series are used for calculating GARCH variances. It makes computation relatively simple and convergence problems in the optimisation process can thus be avoided. Moreover, the variance-covariance matrix is always positive semi-definite, i.e. all eigenvalues are nonnegative when number of principal components used are less than the number of series in the dataset. And it is positive definite when the number of principal components used equals to the number of data series. Mean-reverting term structure of volatility and correlation is still ensured. Furthermore, since only the first few principal components are used for the estimation of GARCH variances that represents all data series in a system, correlation estimates become more stable and less likely to be affected by those variation caused by the noise in the system.

Although orthogonal GARCH model can resolve the major drawbacks that have commonly exist in multivariate GARCH models, it still has its own disadvantage. For orthogonal GARCH model to perform well, the few chosen principal components are required to have accounted for sufficiently large amount of total variation in a dataset. Whether the principal components are powerful

enough depends on how correlated the series in a dataset are. The higher the correlations among the series in the dataset, the stronger the explanatory power of the principal components. In her orthogonal GARCH analysis of crude oil futures data and the UK zero coupon yield data, Alexander (2001a) has shown that the first few principal components from the latter dataset explain a less amount of total variation than those from the former dataset. It is because some data series included in the zero coupon yield dataset have low correlations with the rest of the series in the dataset, as oppose to the crude oil futures dataset. Since variance-covariance matrix of the entire dataset is computed as an approximation of the product of the conditional volatilities of the principal components and the squared factor weights. Weak dependence among data series results in weak approximation and thus causing orthogonal GARCH model not so favourable. The problem of orthogonal GARCH model with weakly correlated series has further led to some recent developments, see for example the Generalised orthogonal GARCH model, by van der Weide (2002), but these models are less straight forward to implement.

In the paper by Engle and Colacito (2006), the empirical performance of five multivariate GARCH models, including the orthogonal GARCH model. They apply tests to stock and bonds, and also to highly correlate assets. The tests they have considered, including a comparison of volatility and the Diebold and Mariano (1995) approach, show no evidence that the orthogonal GARCH model is better than the alternative GARCH models they concern.

Our empirical applications in this chapter will be implemented in the context of orthogonal GARCH model of Alexander (2001a, b). The model will be

further discussed in the next section.

2.2 The Orthogonal GARCH model

Orthogonal GARCH model have a lot of advantages over the classical GARCH models. First of all, orthogonal GARCH solves the problem of parameter estimation that exist in multivariate GARCH models. The model not only allows handy computation of the variance-covariance matrices. It also allows forecasting exercise to be implemented with ease. When volatility forecast on the return series is required, it can be done by multiplying the GARCH(1,1) forecast of the few chosen principal components by the squared values of the factor weights. We will leave a forecasting exercise to the last chapter of this thesis. In this section, we provide a discussion of the orthogonal GARCH model of Alexander (2001a, b).

Suppose there are N stock return series in a dataset, Y is a $T \times N$ matrix of daily stock returns. Let $y_{i,t}$ denote the daily return of stock i at time t , $i = 1, \dots, N$. Normalising the data series gives

$$x_{i,t} = \frac{y_{i,t} - \mu_i}{\sigma_i} \quad (1)$$

Where μ_i and σ_i are the mean and standard deviation of y_i , x_i and y_i are both $T \times 1$ vectors. Normalisation of the stock returns prior to analysis is necessary since the results of the principal component analysis are sensitive to re-scaling of data. Matrix X represents the $T \times N$ matrix of normalised dataset and each column of this matrix has mean equals to zero and unit variance. That

is, each stock return series becomes a $T \times 1$ standard normal vector.

Matrix $X'X$ is a $N \times N$ symmetric matrix of unconditional correlations between stock return series in X , with diagonal elements of ones. Extracting the eigenvectors from the matrix $X'X$ gives us the factor weights. These factor weights are contained in matrix W , which is an orthogonal matrix. Λ denotes a diagonal matrix of the eigenvalues of the unconditional correlation matrix $X'X$. Columns in W are arranged to descending order according to the size of the corresponding eigenvalues.

If all N extracted principal components are considered, then the matrix containing all the principal components, P , is given by

$$P = XW \quad (2)$$

The k th principal component can be written as

$$p_k = w_{1k}x_1 + w_{2k}x_2 + \cdots + w_{Nk}x_N \quad (3)$$

where $k = 1, \dots, N$. We can see from equation (3) that each principal component is a linear combination of the column of X with factor weights given by the elements in W . The first principal component explains the majority of the total variation in X and the principal components are unconditionally uncorrelated with each other. Since W is orthogonal, so $W' = W^{-1}$ and $P'P = \Lambda$. Inverting equation (2) gives us

$$X = PW' \quad (4)$$

In the case when only K principal components are chosen for modelling, where $K < N$, equation (4) becomes

$$X = PW' + \Xi \quad (5)$$

Matrix P contains the chosen K principal components becomes a $T \times K$ matrix and W' is a $K \times N$ matrix. PW' represents the part of the stock returns that is explained by the K principal components. Alternatively, it can be interpreted as the common component of the stock returns with weights given by the columns in W . Ξ is the part of the stock returns that is not captured by the common component. Let λ_k be the k th eigenvalue, $k = 1, \dots, K$, the proportion of variation in X that is explained by the k th principal component is given by $\frac{\lambda_k}{N}$ and $\sum_{k=1}^K \frac{\lambda_k}{N} = 1$.

Principal component analysis provides good results when it is applied to a dataset containing series that are highly correlated. The majority of the common movements in the dataset can be explained by only a few principal components extracted from the correlation matrix $X'X$. However, when asset series contained in a dataset have quite unique characteristics, they tend to respond to unexpected market events quite differently. This implies lower assets correlation. In such case, the term Ξ will be relatively large.

The conditional variance-covariance matrix of the dataset of stock returns, V , is calculated as

$$V = W\Lambda W' + V_\epsilon \quad (6)$$

M denotes the conditional variance-covariance matrix of the chosen K principal components. It gives the conditional variances and covariances of the common movement in the original dataset of stock returns. V contains the GARCH volatilities of the series in the original dataset. V_ϵ denotes covariance matrix of the errors. Notice that M is a diagonal matrix with the GARCH variances of the K chosen principal components on the diagonal. That is, $M = \text{diag}(\text{Var}(p_1), \dots, \text{Var}(p_k))$, where $\text{Var}(p_k)$ is the variance of the k th principal components. Approximation of variance of X is given by

$$V \approx WMW' \quad (7)$$

Accuracy of this approximation depends on the amount of total variation in the dataset that is explained by the K chosen principal components. In other words, this approximation will be more accurate if the principal components chosen, normally the first few of them, are capable of accounting for a sufficient amount of total variation in the dataset. It is because the size of the term Ξ in equation (5) will become smaller and thus, V_ϵ in equation (6) will also become smaller. From here, we can see how orthogonal GARCH model allows the $N \times N$ GARCH variance-covariance matrix to be generated by only K univariate GARCH models. Multiplying the conditional variance-covariance matrix of the chosen principal components by the squared values of the factor weights give us the approximation of the conditional variance-covariance matrix of the original dataset of returns.

The diagonal matrix of conditional variances of principal components, M , is estimated by univariate GARCH models via maximum likelihood estimation.

If we assume each principal component follows a GARCH(1,1), the conditional variance of the k th principal component at time t is given by

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 \quad (8)$$

where $\beta_0 > 0$, $\beta_1, \beta_2 \geq 0$. β_1 represents the market reaction parameter, it measures the intensity of reaction of volatility to unexpected market return in the last period, ε_{t-1}^2 . β_2 represents the volatility persistence, it measures the persistence in volatility. Both parameters, β_1 and β_2 , should sum to less than 1 to ensure convergence.

We can see from above how reduction in dimensionality and ease of computation can be achieved in the calculation of the GARCH variance-covariance matrix of the original dataset of returns using the orthogonal GARCH model. Moreover, matrix V is positive definite if all N extracted principal components are used. When $K < N$, which is usually the case, matrix V will still be positive semi-definite.

In orthogonal GARCH modelling, the choice of the value K is crucial. When dealing with a highly correlated system of data series, it would be preferable to choose K much less than N . It is because in such case, extraneous noise embedded in those relatively less representative principal components is excluded from modelling, and thus resulting in more stable correlation estimates and more robust volatility estimates (see Alexander (2001a)). However, keeping all principal components in the estimation can allow one to always benefit from obtaining a positive definite variance-covariance matrix. Reducing the number of principal components used in modelling can only guarantee a posi-

tive semi-definite variance-covariance matrix. When dealing with a dataset in which series are less correlated, one may want to keep all principal components so that a sufficient amount of total variation in the dataset can be captured and to ensure approximation as in equation (7) to be more accurate. But on the other hand, this may lead to the danger of having unstable correlation estimates.

The above dilemma has revealed the major drawback of orthogonal GARCH model, The model works well when series in the dataset being analysed are highly correlated. When the dataset is made up of data series with low correlation, orthogonality condition partly breaks down, leading to weak performance of orthogonal GARCH analysis. This will be proved in the diagnostic test on the return residuals later in this chapter.

3 Empirical Results

3.1 Data and Preliminary Statistical Analysis

Daily data of constituent stocks of five Asian indexes are obtained from Datastream. The five indexes are NIKKEI 225 (NIK225) and NIKKEI 500 (NIK500) of Japan; Heng Seng Composite Index (HSCI) of Hong Kong; Korean Stock Exchange Composite Index 200 (KOS200) of South Korea; and Stock Exchange of Singapore All Share Index (SING) of Singapore. The reason for investigating both the NIKKEI 225 and NIKKEI 500 of Japan is for us to investigate if there is any effect of the size of a dataset on the explanatory power of a first principal component. The sample ranges from 3 January 2000 to 30 July 2004,

for a total of 1194 daily return observations. Daily returns on each constituent stock is calculated as

$$y_{i,t} = \ln(pr_{i,t}) - \ln(pr_{i,t-1}) \quad (9)$$

where $y_{i,t}$ denotes return on constituent stock i at time t , $pr_{i,t}$ and $pr_{i,t-1}$ denote prices of the constituent stock at time t and time $t - 1$, respectively. For each index, only the constituent stocks that have data available throughout the whole sample period are considered. This leads us to have 217 stocks for NIK225, 481 stocks for NIK500, 176 stocks for KOS200, 227 stocks for SING and 161 stocks for HSCI. We exclude the periods when the markets are closed from the dataset, the number of observations then becomes 1128 for both NIK225 and NIK500, 1121 for KOS200, 1185 for SING and 1128 for HSCI.

Notice that the preliminary statistical analysis in this subsection is carried out prior to normalisation of the dataset, we first compute a mean return over all N constituent stocks in a dataset at each period in the entire sample to get a $T \times 1$ vector of mean returns. The statistics on the mean returns reported in table 1 is the summary statistics for this $T \times 1$ vector of mean returns. The reason for not carrying out the analysis using the actual index series is because the indexes are weighted indexes¹. To compute the mean return series, we give equal weight to all stocks in each dataset and thus do not need to worry about different weighting systems being involved in the analysis.

¹NIKKEI index is a price-weighted index, HSCI is a capitalisation-weighted index and KOSPI is a market capitalisation-weighted index.

Figure 1 displays the time plot of mean returns of the constituent stocks for the five indexes. We can see from these plots that the mean return series appear stationary and they all fluctuate around their long term averages as what one would expect. Two negative extremes can be seen from almost all of the plots around April 2000 and September 2001, these coincide with the burst of Dot-com bubble and the September 11 terrorists attacks. Table 1 summarises descriptive statistics of the mean return series. Mean returns of SING constituents and KOS200 constituents are negative on average during the sample period. Average Mean returns on the constituents of HSCI is higher than that on the constituents of the other four indexes. Mean returns on KOS200 constituents has the highest variance, followed by NIK225, HSCI and NIK500. Whereas mean return on SING constituents has the lowest. It indicates the returns on South Korean stocks are the most volatile on average, but Singaporean stock returns are the least volatile on average. The values of standard deviation tell us that KOS200 mean returns seem to have deviated most largely from its sample mean comparing with the other three markets during the sampling period. Whereas the mean returns of SING has the least deviation from its sample mean. With regards to the distributional characteristics, it can be seen clearly from the skewness and kurtosis coefficients that none of the mean return series are normally distributed and this is also confirmed by the Jarque-Bera test statistics as the null hypothesis of normality is rejected at 5%.

3.2 Analysis of Average Correlations

Following the preliminary analysis, we now move on to examine the correlations among the series in our five datasets. The datasets of stock returns are normalised before further analysis. Since each dataset consists of more than a hundred, and some of them more than two hundreds, stock return series, it is difficult to report the correlation matrices. We therefore report summary statistics on average correlations. For each stock i in a dataset, we compute the correlations between this stock with the rest of the $N - 1$ series to get a vector of correlations with dimensions equal $(N - 1) \times 1$. We then take the average of these $N - 1$ correlation coefficients to get an average correlation between stock i and the remaining stocks in the system. We do it for $i = 1, \dots, N$ and get a vector of $N \times 1$ average correlations. Table 2 shows some summary statistics of the average correlations for our five datasets.

First of all, all of the means of the average correlations are positive during the sample period. This indicates that stock return series in each of the four markets are positively correlated in general, that is, they tend to response to changes in market condition in the same direction. By looking at the values of maximum and minimum average correlations, we can see that in fact the South Korean, Hong Kong and Japanese markets all have stock returns in their own markets that are positively correlated on average. Except for Singapore, we see evidence of negative correlation among its stock returns.

The most highly correlated datasets of stock returns appear to be NIK225 and KOS200. We expect these high correlations among the stock returns in these two datasets can make contribution to good results of principal compo-

ment analysis. And thus, the first principal components from these two datasets will be powerful in accounting total variation. Comparing the two Japanese datasets, we can see that NIK225 stock returns are more correlated than stock returns in the NIK500 dataset in general. Therefore, the first principal component from the NIK225 dataset is expected to be more powerful than the one from the NIK500 dataset.

KOS200 stock returns have the largest values of standard deviation in their average correlations, followed by HSCI. The smallest value of standard deviation of average correlation appears to be the SING stock returns. The two Japanese datasets of stock returns have comparable standard deviations of average correlation, with the one of NIK225 slightly higher. It means there is evidence of more observations of average correlation deviate from the mean average correlations in general in the South Korean and Hong Kong markets. This indicates the level of correlations among stock return series varies considerably in these two markets. On the contrary, the level of correlations among the Singaporean return series do not vary that much in general.

From equation (5), we can see that the stock returns matrix X is divided into two components if only K chosen principal components are used to represent comovements in the entire system. PW' is the common component of stock returns, in which the common variation within the dataset is summarised by the chosen K principal components, with weight given by W' . The idiosyncratic component of stock returns Ξ , is the remaining part of the stock returns that is not captured by the common component. Equation (5) tells us that if the chosen K principal components are able to explain a large amount of total

variation in the dataset, then the term Ξ will have small magnitude. Therefore, principal components extracted from a highly correlated system should be powerful enough to ensure a small Ξ . Whereas for a less correlated system, Ξ will be relatively large if K is not large enough.

3.3 Eigenvector and Eigenvalue Analysis

In this subsection we apply principal component analysis to the $N \times N$ symmetric unconditional correlation matrix, $X'X$, to carry out an analysis on eigenvectors and eigenvalues. As explained in section 3.2, eigenvectors from the correlation matrix give us the factor weights, whereas eigenvalues of the correlation matrix give us an indication of the explanatory power of the principal components. The proportion of total variation in data matrix X explained by the k th principal component is given by $\frac{\lambda_k}{N}$, where λ_k is the eigenvalue corresponds to the k th eigenvector of matrix $X'X$.

Table 3 shows the cumulative explained variation computed for the first 20 principal components extracted from each of the five correlation matrices. We can see from the table that the first principal component always manage to explain a large proportion of the total variation in the dataset. Proportions of total variation explained by additional principal components are always decreasing. As we have seen from table 2, Singaporean stock returns have the weakest correlations on average comparing with those in the other three markets. This low correlation has reflected in the explanatory power of its first principal component as it can explain only 8% of the total variation. The first principal components of KOS200 and NIK225 stock returns are the strongest

as they can explain about 29% and 28% of total variation, respectively. Explanatory power of the HSCI first principal components is ranked the third as it captures 23% of the total common movement in its stock returns. Although NIK500 first principal component can also explain about 22%, it is relatively less powerful than the one of NIK225 dataset. It is because bringing in extra stocks to constitute the NIK500 dataset has lowered the overall correlations in the dataset and thus causing lower explanatory power of the first principal component extracted from the dataset. In the empirical analysis of chapter 2, a more detail explanation on why larger dataset with more cross-sectional series is not always desirable for factor analysis will be given. In fact, given that the number of series included in the NIK500 system are at least the double of the others, its explanatory power is not weak.

One may argue that 20% of the total variation being explained by the first principal component can hardly be regarded as a large amount. Consider the fact that our datasets are systems of individual stock returns, which have more idiosyncratic properties than other financial time series. They are thus less correlated when comparing with some other system of highly correlated financial assets. In her application to the WTI crude oil futures data on all monthly maturity from 1-month to 12-month, Alexander (2001) has shown that the first three principal components can already explain 99.8% of total variation with the first one alone explains 95.9%. WTI crude oil futures data is a highly correlated system which favours the use of principal component analysis. The results from her study, together with what we have found from our datasets, confirm that a dataset with assets that have sufficiently large correlation is desirable

for application of principal component analysis. It is because asset that has very idiosyncratic properties tend to corrupt the volatilities and correlations of other assets in the dataset. However, this does not demote the implication of our application to the returns on stock index constituents. It is because consider the large number of stocks included in our datasets and also consider the idiosyncratic characteristics of equities in general, 20% of the total is already quite an appealing result.

Findings in the overall explanatory power of the first 20 principal components from the five datasets are quite consistent with the results we have obtained regarding the explanatory power of the first principal components. It can be seen that around 50% of the total variation in KOS200 and NIK225 stock returns, and around 46% and 40% of total variation in HSCI and NIK500 stock returns can be explained by using all first 20 principal components. But only 30% of comovements in the SING stock returns can be explained by using 20 principal components. It means that the Singaporean stock return principal components have weak ability in capturing variation in general.

Although a larger amount of explained variation can be resulted using more principal components to model the data system, there is no guarantee that more dynamics in a datasets can be captured if more principal components are used. This can be seen later from the results of diagnostic tests on return residuals we discuss later in this chapter.

We are aware that the low correlations among stock returns in our datasets may not be favourable for orthogonal GARCH modelling and volatility prediction, especially if there are not enough principal components chosen to repre-

sent the datasets. However, the aim of this chapter is just to perform a factor analysis in the context of orthogonal GARCH, rather than to evaluate the performance of orthogonal GARCH. We will carry out an evaluation of using orthogonal GARCH models for forecasting volatility in the last chapter of this thesis.

Having investigated the explanatory power of the principal components, we now move on to analyse the factor weights on the principal components. We can see from Table 3 that the fact that the first principal component can alone account for the largest amount of variation is applicable to all of the datasets. We therefore concentrate our analysis here on the first principal components. We compute the factor weights of equation (3), that is the eigenvectors from the matrix $X'X$ and these eigenvectors are contained in matrix W in equation (2). Figures 3 to 7 then plot the factor weights on the first principal components from our five datasets. In principal component analysis, the first eigenvector that corresponds to the largest eigenvalue takes values less than one in a data matrix with full rank. Values in the largest eigenvector are larger and more similar if the series in the dataset are more highly correlated. This can be confirmed by plots of the factor weights on the five first principal components. Constituent returns on the Singaporean stocks have the lowest correlation on average comparing with those in the other four indexes. It can be seen clearly that the factor weights on the first principal component of the Singaporean stock returns vary quite a lot in values when comparing with the factor weights on the first principal components of the other four datasets of stock returns. Return series in a data system that have higher correlations tend to have higher

degree of common movement, or put it differently, they tend to respond in a more similar way to changes in market condition. This can be reflected in the values of the factor weights on their first principal component. And it is because of this reason, the first principal component from the correlation matrix tend to explain the majority of the variation in the dataset as this is confirmed by results in table 3.

3.4 GARCH Analysis of the First Principal Components

In this subsection, we report the results on the analysis of the first principal components obtained from our five Asian datasets. We first examine their distributions and to check if they are stationary. We then analyse the GARCH(1,1) representation of the five first principal components. We assume conditional mean equation follows a strict white noise process and the GARCH(1,1) of the principal component has the following representation²:

$$p_t = \alpha_0 + \varepsilon_t \quad (10)$$

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 \quad (11)$$

Plots of the first principal components from the five systems of stock returns are shown in figure 2. It is not surprising to observe that the behaviour of the first principal components of KOS200, NIK225, NIK500 and HSCI datasets

²The conditional variance equation of principal component, equation (11), ignores the possibility that the conditional variance of principal components may have long memory. To account for persistence, equation (11) can be replaced by a long memory GARCH model, e.g. a FIGARCH representation.

mimic the behaviour of their mean stock returns counterparts shown in figure 1. It is because the first principal components from these datasets manage to capture a reasonable amount of total variation, they are thus able to represent their datasets in the same way that the mean return series can do. Especially, less noise is embedded in this principal components than in the mean return series. However, an exceptional case can be seen from the KOS200 first principal component. The plot of its first principal component does not have the same pattern as the plot of its mean return series. It is because the Singaporean first principal component fails to capture enough variation in the dataset, and thus fail to represent the overall behaviour of the dataset. In the case when a dominant principal component is representative enough to summarise the co-movements and the behaviour of the all stock returns in a dataset, a GARCH analysis of this principal component can give indication on the GARCH analysis of the entire dataset.

Figures 8 to 12 plot the histograms of the five first principal components. Table 4 reports the distributional statistics of these first principal components. All five principal components are leptokurtic, meaning the tails of the distributions of these series are fatter than normal distribution. That is, large positive and negative values in a principal component series occur more often. Moreover, all series appear to have slight negative skewness, indicating the left tail in each distribution is slightly fatter than the right tail. This implies there are more large negative values than large positive values in a principal component series. The non-normal underlying distribution of the first principal components are confirmed by the Jarque-Bera normality test. There is clear

rejection of the null hypothesis of normality at all levels for the five principal components from the five datasets of returns. Having observed these findings, we can conclude that the first principal components from the South Korean, Hong Kong and Japanese returns have the distributional characteristics that can commonly be observed in empirical stock return series, and at the same time these principal components are able to represent the behaviour of their entire system of returns. Although the Singaporean principal component also exhibits the commonly observed distributional characteristics of stock returns, it does not mimic the overall behaviour of the entire system.

Table 5 shows results of Augmented Dickey-Fuller test and Phillips-Perrons test. The 10% MacKinnon critical value for both tests is about -2.568. The first principal component is the trend component and represents comovements of all return series in the dataset. If the series is covariance stationary, then shocks to the series are temporary. This implies shocks that affect the comovements of all series in the dataset are non-permanent. Both its finite variance and autocovariance are also time-invariant. Effects of shocks on the trend component faded away over time. However, if the series follows a random walk, effects of temporary shocks are then permanent. The Augmented Dickey-Fuller specification for computing the test statistic includes a constant but not a trend, and we consider a lag order of 5. Whereas 6 lags truncations are chosen for the Phillips-Perrons test. Test statistics from both tests are less than the 10% critical value, leading to rejection of unit root. We can conclude that the first principal components for all five datasets are stationary, and thus impacts of shocks on the comovements of all series are non-permanent.

Since most empirical studies on orthogonal GARCH model suggest that an GARCH(p,q) with both p and q higher than 1 is rarely needed, our analysis focuses on the use of GARCH(1,1) representations for modelling the first principal components. GARCH model is famous for not only capable of describing volatility clustering, but also for capturing many other features in financial time series, for example excess kurtosis and fat-tailedness. We also try to examine whether GARCH(1,1) is appropriate specifications for the first principal components extracted from the five systems of stock return series. Univariate GARCH(1,1) as in equation (10) and (11) is estimated for the first principal component from each dataset. Estimation results are shown in Table 6. Due to the non-normality nature found in the distributions of all our first principal components, we suspect the residuals may also not be conditionally normally distributed, the reported standard errors computed are thus the quasi-maximum likelihood standard errors by Bollerslev and Wooldridge (1992).³

β_1 in equation (11) measures the market reaction, it tells how intensely volatility at current period reacts to unexpected market returns in the last period. Whereas β_2 measures the persistence in volatility. Both parameters must have a sum to a value less than one in order to ensure a finite unconditional variance, and β_0 must be larger than zero. We can see that all parameter estimates in the conditional variance equations for the principal components

³ARCH parameter estimates will still be consistent if the mean and variance equations are correctly specified when assumption of conditional normality is no longer valid. However, estimates of the covariance matrix will not be consistent and resulting in incorrect standard errors being computed. Quasi-maximum likelihood covariance and standard errors computed using the method described by Bollerslev and Wooldridge (1992) are robust to conditional non-normality.

of all five datasets satisfy these conditions and they are significant at 5 % significance level. Notice that the sum of β_1 and β_2 is quite close to one and this is what usually observed in high frequency financial data. We can also see from the values of β_2 that all principal components from the five markets appear to have quite high volatility persistence, with the one for South Korean market to be the highest. It can be seen from the values of β_1 that the Singaporean stock volatility seems to response to last period's unexpected return most intensely as its value of this market reaction parameter being the largest among the five. Level of intensity of market reaction for Japanese and Hong Kong markets are not very different as the values of their β_1 are quite similar. There are no big difference in the values of β_1 and β_2 between the two Japanese principal components, as stocks in NIK225 is a subset of NIK500.

We also carry out diagnostic tests to see if the GARCH(1,1) is an appropriate specification for the first principal components of the Asian stock returns. In particular, we look at whether univariate GARCH(1,1) is enough to capture the dynamics of the first principal components. Some statistics on the standardised residuals are presented in table 7. It is clearly shown that standardised residuals of the principal components of all five datasets are negatively skewed with those of KOS200 and HSCI principal components being the most. Distributions of standardised residuals of these two principal components deviate quite largely from normality. Although those of the two Japanese datasets do not strictly follow normality, the level of deviation is less severe comparing with the others. This can also be seen from the histograms of the GARCH(1,1) standardised residuals for all five principal components in figures 13 to 17.

Breusch-Godfrey LM test for serial correlation and an ARCH LM test in standardised residuals are also performed. We include 5 lags in the auxiliary equation for computing the test statistics for both tests. Table 8 shows the results of serial correlation LM test, the LM test statistics and the corresponding probability values are reported. Evidence of serial correlation is found in the GARCH(1,1) standardised residuals of KOS200, HSCI and SING first principal components at 5% significance level. Whereas no sign of remaining serial correlations is detected in those of the two Japanese datasets. However, test on heteroscedasticity shows different results. Table 8 reports ARCH LM test statistics and the corresponding probability values. At 5% significance level, the GARCH(1,1) is insufficient to capture all the heteroscedasticity present in the first principal components of the Japanese stock returns. Whereas the model is sufficient to capture those in the KOS200, HSCI and SING datasets. To sum up, although GARCH(1,1) is an appropriate specification for the conditional volatility of the South Korean, Hong Kong and Singaporean principal components, their conditional means have time properties that require a more complicated process than a strict white noise. GARCH(1,1) is not sufficient in capturing the dynamics of the conditional volatilities of the two Japanese principal components.

We believe remaining serial correlation found in the standardised residuals indicates that strict white noise is not an appropriate specification for the conditional mean equation. It also indicates that comovements in the stock returns of a market may have an autoregressive specification. Our findings suggest that comovements in the Korean, Hong Kong and Singaporean stock

returns at current period may depend on their comovements in the local market of past periods. This claim can be proved by increasing the order of lags in the mean equation to see if it can help to soak up the remaining serial correlation. Similarly, unexplained ARCH effect indicates conditional variance in the trend component has more heteroscedasticity than a GARCH(1,1) can explain. It indicates variance in the comovements of stock returns not only depends on the variance of the comovements in the last period, but also on variance of the comovements in more remote past. To illustrate our claim, we try to examine if there is any improvement in diagnostic test results by increasing the order of lags in the conditional mean for the South Korean, Hong Kong and Singaporean principal component. But for the Japanese principal components, we only increase the order of lags in their conditional variance equations. Table 10 shows the LM test results on both heteroscedasticity and serial correlation for three sets of specifications. For the first principal component of the KOS200 and HSCI, the statistics are computed for AR(1)-GARCH(1,1) standardised residuals. However, for the SING dataset, we find out that its principal component needs an AR(3) in the mean equation to show some improvement. For the first principal component of the NIK225 and NIK500, statistics are computed for GARCH(2,2) standardised residuals. We can see that there is improvement in the explanation of dynamics for all datasets at 5% level. Improvements can also be confirmed by comparing the values of Akaike Information criterion (AIC) reported in table 6 and 10.

Since we are working with daily observations of the first principal component that summarise the comovements of the returns on the constituent

stocks of a stock index, these findings provide interesting empirical implications. Throughout our sampling period, we find evidence to show current common movements in South Korean and Hong Kong stock returns depend on the comovements in their local market in the last period. For the South Korean market, comovements in returns today depends on comovements in the past three days. The same does not apply to the Japanese market. Common movement in current Japanese returns is independent from the comovements in the past period. Moreover, current conditional volatilities in stock return comovements of Hong Kong, South Korean, Singapore all depends on the unexpected comovements in local market return, and also on the fluctuation in stock return comovements in the last period. Whereas comovements of current conditional volatility depend on the unexpected comovements and the volatility in the last two days in the Japanese market.

From here, we can observe empirically the beauty of orthogonal GARCH modelling: applying only a univariate GARCH analysis of the dominant principal component that represents mostly the behaviour of the entire dataset of financial system allows us to get a picture of GARCH analysis of the entire system, provided the dominant principal component is strong enough to represent the entire system. Time-consuming computation can thus be simplified.

3.5 Diagnostic Tests on Residuals

We know when a larger number of principal components are used for modelling a system of data series, a larger proportion of the total variation in the system can be explained. However, using a large number of principal components for

modelling is not always desirable. First of all, notice that the idea of orthogonal GARCH model and any factor models that involve the use of principal component analysis is aiming to reduce dimensionality and to encourage convenient estimation procedure. That is, we try to summarise the comovements among series in a dataset using K principal components, rather than concerning all N series in the dataset, with K always being much smaller than N . The adoption of a large number of principal components in modelling contradicts the framework of those modelling techniques. For example, if the dataset consists of more than a hundred series and one is to use a large number of principal components in a orthogonal GARCH model, then there may be a need to estimate more than a hundred univariate GARCH models of the principal components. Therefore, the amount of work and the complication involved in the analysis will not be a lot less than the use of multivariate GARCH models.

Secondly, Alexander (2001a) points out that when only the first few principal components that represent the entire system are used for estimation of GARCH volatilities, correlation estimates become more stable and less likely to be affected by those variation caused by the noise in the system and volatilities estimates are more robust. In other words, we can avoid the stability of correlation estimates being corrupted by the extra noise from those relatively less important or less representative principal components. Finally, the other reason why using a large number of principal components may not be desirable relates to the amount of dynamics that can be captured by the chosen principal components. Although the use of large number of principal components can guarantee large proportion of total variation in a dataset being captured, there

is no guarantee that one can benefit from an improvement in diagnostic test results by modelling the dataset with such a large number of principal components. With the above understandings in mind, it is essential to decide upon the appropriate number of principal components to be chosen for orthogonal GARCH modelling. Decision should be made in the sense that a reasonable amount of total variation in the dataset can be accounted for and the dynamics in the dataset can be captured, while the number of principal components involved in modelling is not too large.

In this subsection, we carry out a diagnostic tests on the return residuals when the returns are modelling by K chosen principal components. We want to check whether using a large number of principal components can be rewarding in terms of the explanation of the temporal features of a dataset of stock return series. This will also give an implication on whether large amount of variation being explained can guarantee a better diagnostic test results.

Residuals in stock returns, Ξ , are computed as in equation (5). That is, after the common component is computed by multiplying the chosen K principal components with the matrix containing the corresponding factor weight, we subtract it from the matrix of return series to obtain the residuals at level. Consider the explanatory power of the principal components as shown in table 3, we consider two cases. In the first case, we test on the residuals of all N return series in a dataset when only the first principal component is used for modelling. In the second case, we test on the residuals of all N return series in a dataset when the first 20 principal components are used. The choice of these two numbers is not arbitrary. For all of our five datasets, except SING, the first

principal component alone can already explain over 20% of the total variation. With reference to the large number of series that each dataset contains, the explanatory power of these first principal components is quite appealing. The reason for considering the second case is because for most of our datasets, the first 20 principal components can account for over 40% of total variation, except for SING. For NIK500, they can even account for more than half of the total variation.

In order to detect if there exists any remaining serial correlation and heteroscedasticity not being captured by the principal components, we carry out the Breusch-Godfrey LM test for serial correlation and an ARCH LM test to on return residuals of all N series in each dataset. We include 6 lags in the auxiliary equations of both tests and the level of significance is equal to 1%. Since the number of return series, N , contained in each datasets is more than a hundred, this involves estimating more than N auxiliary equations and the computation of the N corresponding probability values. To report the test results in table 10, we calculate the number of stocks that show no remaining serial correlation and no remaining ARCH effect when $K = 1$ principal component, and when $K = 20$ principal components are used.

First of all, consider the case when only the first principal component is used, only the return residuals of more than 50% or about 50% of the total number of stocks in KOS200, NIK225 and NIK500 show no remaining serial correlation not explained by the model. However, in view of heteroscedasticity, the majority of the return residuals in all five datasets have unexplained ARCH effect. This finding has suggested that modelling the return series with only

the first principal component is not appropriate.

In the case when the first 20 principal components are used for modelling, we can see that only the return residuals of over half of the total number of stocks in KOS200 show no remaining serial correlation whereas the majority of the return residuals of all other four datasets show significant serial correlation not being captured. For test on heteroscedasticity, we have the same finding as in the case when only the first principal component is used, that is, the majority of the return residuals in all five datasets have unexplained ARCH effect.

Comparing the two cases, we can see that slight improvement in capturing both of the two temporal features when larger number of principal components are used only present in the return residuals of KOS200. However, the number of KOS200 stocks that show no remaining serial correlation and ARCH effect is not a lot larger when 20 principal components are used. For the other four datasets of return residuals, improvement is found only in capturing ARCH effect. The number of stocks in these datasets, which have no remaining ARCH effect in their residuals, is larger when 20 principal components are used. Nevertheless, less number of their stocks have residuals that shows serial correlation when only the first principal component is used.

To sum up, four out of five cases show that using large number of principal components to represent the datasets cannot guarantee an improvement in capturing dynamics. Only one case shows that slight improvement can be obtained with a large number of additional principal components are used. We believe the reason behind is that these extra principal components that

are relatively weak in representing the behaviour of system have contained extra noise. Therefore, when they are used for modelling, the problem of misspecification is worsened. These findings imply that although we can benefit from having a larger proportion of total variation in a dataset be accounted for when more principal components are used for modelling, we cannot always be better off in terms of capturing the dynamics of our datasets by modelling with more principal components.

4 Concluding Remarks

In this chapter, we have applied the orthogonal GARCH model into an analysis of the constituent stock returns of five Asian indexes. Results of our correlation analysis reveal that the South Korean stock return series are the most highly correlated on average but those in the Singaporean market are the least. Higher correlations among stock return series contribute to stronger explanatory power of the first principal component and this claim is once again empirically confirmed by our study here. Moreover, the lowest correlations among the Singaporean stock returns on average is also reflected in the values of the factor weights on its first principal component as they vary quite a lot when comparing with the factor weights on the first principal components of the other four datasets of stock returns.

Examination of the distributional characteristics of the first principal components extracted from the five datasets of returns reveals that first principal components from the South Korean, Hong Kong and Japanese returns have the distributional characteristics that can be commonly observed in empirical

stock return series, and at the same time these principal components are able to mimic the overall behaviour of their entire system of returns. Although the Singaporean first principal component also exhibits the commonly observed distributional characteristics of stock returns, it does not mimic the overall behaviour of the entire system.

Results from the univariate GARCH(1,1) analysis on the first principal component shows that although GARCH(1,1) is an appropriate specification for the conditional volatilities of the South Korean, Hong Kong and Singaporean principal components, their conditional means have time properties that require a more complicated process than a strict white noise. GARCH(1,1) is not sufficient in capturing the dynamics of the conditional volatilities of the Japanese principal components. We have also shown how GARCH analysis of the entire dataset of our Asian returns can be summarised by an analysis of univariate GARCH analysis of their first principal components⁴.

Finally, diagnostic tests are scrutinised to check if a first principal component is able to capture enough dynamics in the datasets of Asian stock returns. Our findings show that modelling the return series of all five datasets with only the first principal component is not appropriate. However, only slight improvement in capturing dynamics in South Korean stock returns is shown when 20 principal components are used. For the other four datasets, using more principal components for modelling has worsened the problem of misspecification.

⁴The idea of applying factor analysis into GARCH analysis allows reduction in dimensionality which in turn overcome the common drawbacks multivariate GARCH models suffer from when dealing with large dataset. However, if the entire system of data series can be summarised by only K principal components, with K very small. One may considered estimating these principal components via a multivariate GARCH, provided K is small enough to ensure those common drawbacks do not kick in.

Our empirical evidence shows that using large number of principal components in modelling may allow one to benefit from larger amount of total variation being explained, but it cannot guarantee an improvement in capturing dynamics in the datasets, especially when series in the datasets are not very highly correlated.

The GARCH analysis carried out in this chapter allows us to gain a basic understanding of the characteristics of the principal components. These basic understanding will help us to build insight into how the model can be used for forecasting purpose. Furthermore, the above findings reveal the fact that orthogonal GARCH model works well for highly correlated system of data series. When the dataset is made up of data series with low correlations, orthogonality condition partly breaks down, leading to weak performance of orthogonal GARCH analysis. In the next chapter, we will carry out an empirical factor analysis on stock volatilities using a comparable factor volatility model, the stochastic volatility factor model of Cipollini and Kapetanios (2005). Notice that the basic building blocks of these two factor volatility models are the two main categories of the commonly used volatility models for financial time series analysis, that is, the GARCH family model, and the stochastic volatility model. In view of the shortcoming of the orthogonal GARCH model, we believe the stochastic volatility factor model provides a more flexible framework than the orthogonal GARCH model both in terms of modelling volatility and in terms of modelling the idiosyncratic part of the dataset which is not captured by the common component. Forecasting performance evaluation of the two factor volatility models will be carried out in chapter 3.

5 Tables and Figures

Table 1:					
Descriptive statistics of the mean return series for the five indexes					
	SING	HSCI	KOS200	NIK225	NIK500
No. of periods	1185	1128	1121	1128	1128
No. of stocks	227	161	176	217	481
Mean	-0.00037	0.00027	-0.00012	0.00005	0.00006
Median	-0.00081	0.00078	0.00063	0.00035	0.0006
Maximum	0.0397	0.0537	0.0808	0.0662	0.0544
Minimum	-0.0526	-0.1028	-0.1357	-0.0764	-0.0729
Std. Dev.	0.00915	0.01432	0.01935	0.01435	0.01274
Variance	0.000084	0.000205	0.000374	0.000206	0.000162
Skewness	-0.1321	-0.8678	-0.8486	-0.1858	-0.4005
Kurtosis	6.1624	7.7122	7.6298	4.8508	5.1817
J-B test	497.24 [0.000]	1176.8 [0.000]	1135.7 [0.000]	167.49 [0.000]	253.89 [0.000]
Note: probabilities are reported in brackets.					

Table 2: Summary statistics of average correlations					
	KOS200	NIK225	NIK500	HSCI	SING
No. of stocks	176	217	481	161	227
Max. of average correlations	0.3856	0.3627	0.3345	0.3253	0.1367
Min. of average correlations	0.0672	0.0845	0.0552	0.0544	-0.0128
Median of average correlations	0.2787	0.2808	0.2206	0.2158	0.0632
Mean of average correlations	0.2670	0.2723	0.2134	0.2129	0.0609
S.D. of average correlations	0.0672	0.0488	0.0452	0.0617	0.0300

Table 3: Cumulative explained variation of the first 20 principal components from the five datasets					
PCs	KOS200	NIK225	NIK500	HSCI	SING
1	0.2883	0.2846	0.2250	0.2359	0.0818
2	0.3118	0.3363	0.2676	0.2794	0.1564
3	0.3316	0.3610	0.2870	0.3001	0.1697
4	0.3508	0.3829	0.3024	0.3148	0.1814
5	0.3647	0.3972	0.3138	0.3282	0.1912
6	0.3762	0.4090	0.3243	0.3408	0.2001
7	0.3874	0.4198	0.3331	0.3522	0.2083
8	0.3976	0.4304	0.3407	0.3627	0.2166
9	0.4076	0.4395	0.3478	0.3725	0.2247
10	0.4168	0.4484	0.3543	0.3820	0.2327
11	0.4256	0.4566	0.3603	0.3913	0.2405
12	0.4342	0.4645	0.3661	0.4003	0.2482
13	0.4426	0.4717	0.3718	0.4093	0.2557
14	0.4509	0.4789	0.3770	0.4181	0.2632
15	0.4588	0.4858	0.3821	0.4268	0.2707
16	0.4665	0.4927	0.3871	0.4353	0.2779
17	0.4739	0.4994	0.3920	0.4437	0.2849
18	0.4813	0.5059	0.3967	0.4519	0.2919
19	0.4886	0.5124	0.4014	0.4600	0.2989
20	0.4956	0.5187	0.4060	0.4680	0.3058

Table 4: Distributional statistics of the first principal components					
	KOS200	NIK225	NIK500	HSCI	SING
Skewness	-0.8138	-0.1426	-0.3371	-0.7621	-0.6643
Kurtosis	7.5563	4.9648	5.0887	7.3829	8.5988
Jarque-Bera statistics	1093.4	185.27	226.42	1004.9	1643.9
p-values	0.000	0.000	0.000	0.000	0.000

Table 5: Test for unit root in first principal components					
	KOS200	NIK225	NIK500	HSCI	SING
ADF test statistics	-13.514	-13.884	-14.212	-13.256	-12.224
PP test statistics	-32.167	-33.648	-32.011	-30.118	-32.118
Note: 10% MacKinnon critical values = -2.568					

Table 6: GARCH(1,1) estimation outputs for the first principal components

		α_0	β_0	β_1	β_2
KOS200	Coefficient	0.2760 (0.1630)	1.639 (0.5178)	0.1069 (0.0232)	0.8679 (0.0242)
	z-statistic	1.693 {0.0905}	3.165 {0.0016}	4.595 {0.000}	35.833 {0.000}
	AIC	6.65			
NIK225	Coefficient	0.1251 (0.2113)	4.281 (1.521)	0.0882 (0.030)	0.8438 (0.0433)
	z-statistic	0.5923 {0.5536}	2.815 {0.0049}	2.937 {0.0033}	19.448 {0.000}
	AIC	6.97			
NIK500	Coefficient	1.770 (0.2791)	8.179 (3.209)	0.0813 (0.0277)	0.8446 (0.0458)
	z-statistic	0.6343 {0.5259}	2.548 {0.010}	2.935 {0.003}	18.412 {0.000}
	AIC	7.48			
HSCI	Coefficient	0.1874 (0.1687)	3.1854 (1.424)	0.0831 (0.0218)	0.8318 (0.045)
	z-statistic	1.111 {0.2664}	2.236 {0.025}	3.802 {0.000}	18.355 {0.000}
	AIC	6.41			
SING	Coefficient	-0.0103 (0.1039)	1.1146 (0.4324)	0.1268 (0.0329)	0.8157 (0.0427)
	z-statistic	-0.9918 {0.9210}	2.577 {0.010}	3.848 {0.000}	19.081 {0.000}
	AIC	5.64			

Note: Standard errors are reported in parentheses. Probabilities are reported in curly brackets.

Table 7: Descriptive statistics on standardised residuals of GARCH(1,1) of first principal components

	KOS200	NIK225	NIK500	HSCI	SING
Mean	-0.054	-0.023	-0.026	-0.037	-0.001
Std. Dev.	0.999	1.000	1.000	0.999	1.000
Skewness	-1.169	-0.191	-0.370	-0.927	-0.278
Kurtosis	9.251	4.058	4.563	8.417	6.058

Table 8: LM test for serial correlation in GARCH(1,1) standardised residuals of principal components

	KOS200	NIK225	NIK500	HSCI	SING
LM test statistics	13.15	4.580	10.55	20.41	47.29
P-values	{0.022}	{0.469}	{0.060}	{0.001}	{0.000}

Table 9: ARCH LM test for heteroscedasticity in GARCH(1,1) standardised residuals of principal components

	KOS200	NIK225	NIK500	HSCI	SING
LM test statistics	1.297	16.79	15.03	2.941	6.206
P-values	{0.935}	{0.004}	{0.010}	{0.709}	{0.286}

Table 10: LM test results for serial correlation and heteroscedasticity of standardised residuals		
Results of standardised residuals of AR(1)-GARCH(1,1)		
	KOS200	HSCI
Serial correlation LM test	6.169	1.672
P-values	{0.290}	{0.892}
ARCH LM test	1.091	2.549
P-values	{0.954}	{0.769}
AIC	6.64	6.39
Results of standardised residuals of AR(3)-GARCH(1,1)		
	SING	
Serial correlation LM test	8.874	
P-values	{0.114}	
ARCH LM test	7.176	
P-values	{0.207}	
AIC	5.61	
Results of standardised residuals of GARCH(2,2)		
	NIK225	NIK500
Serial correlation LM test	4.090	10.47
P-values	{0.536}	{0.07}
ARCH LM test	9.079	5.335
P-values	{0.106}	{0.376}
AIC	6.89	7.46

Table 11: Diagnostic test results on return residuals in the five datasets		
K=1		
	No autocorrelation	No ARCH effect
KOS200	99	21
NIK225	109	58
NIK500	246	96
HSCI	76	59
SING	75	71
K=20		
	No autocorrelation	No ARCH effect
KOS200	104	30
NIK225	81	71
NIK500	188	144
HSCI	69	68
SING	65	85
Note: Number of stocks are reported in this table.		

Figure 1: Time plots of mean returns

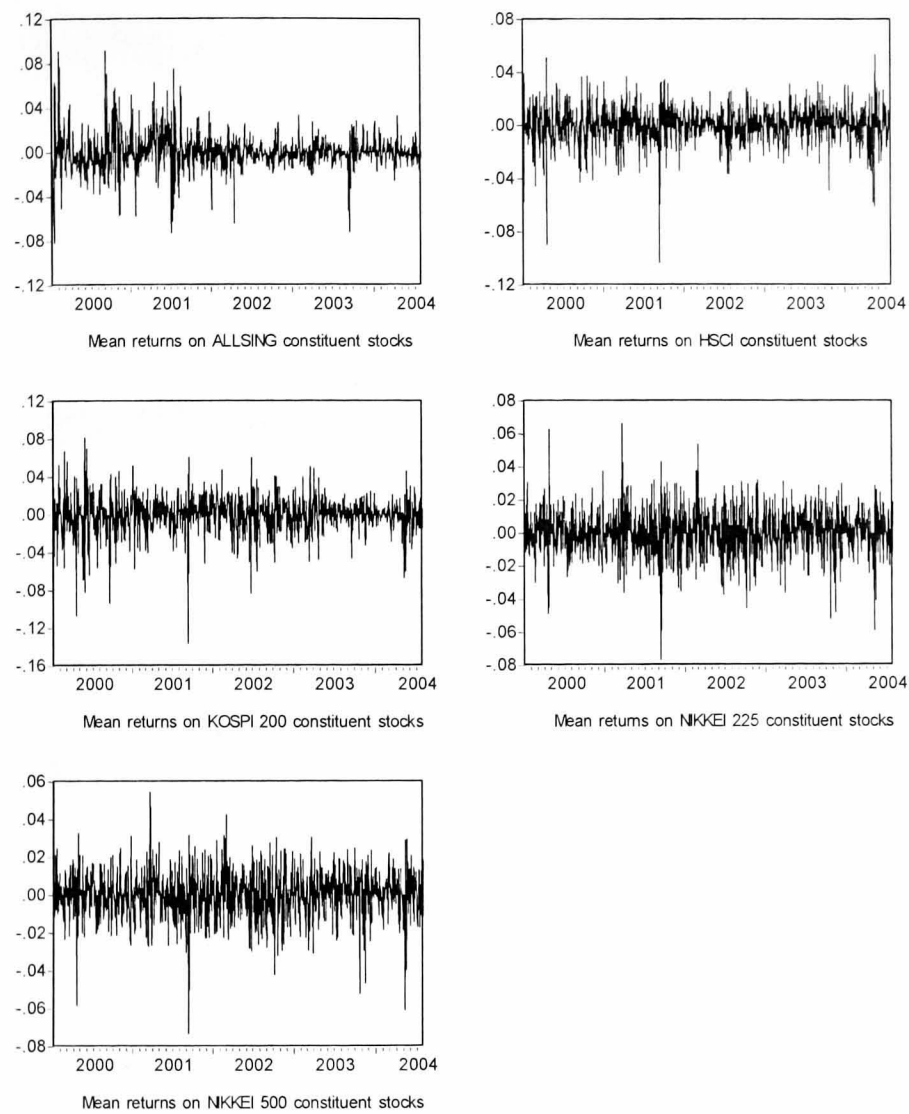


Figure 1: Time plots of mean returns

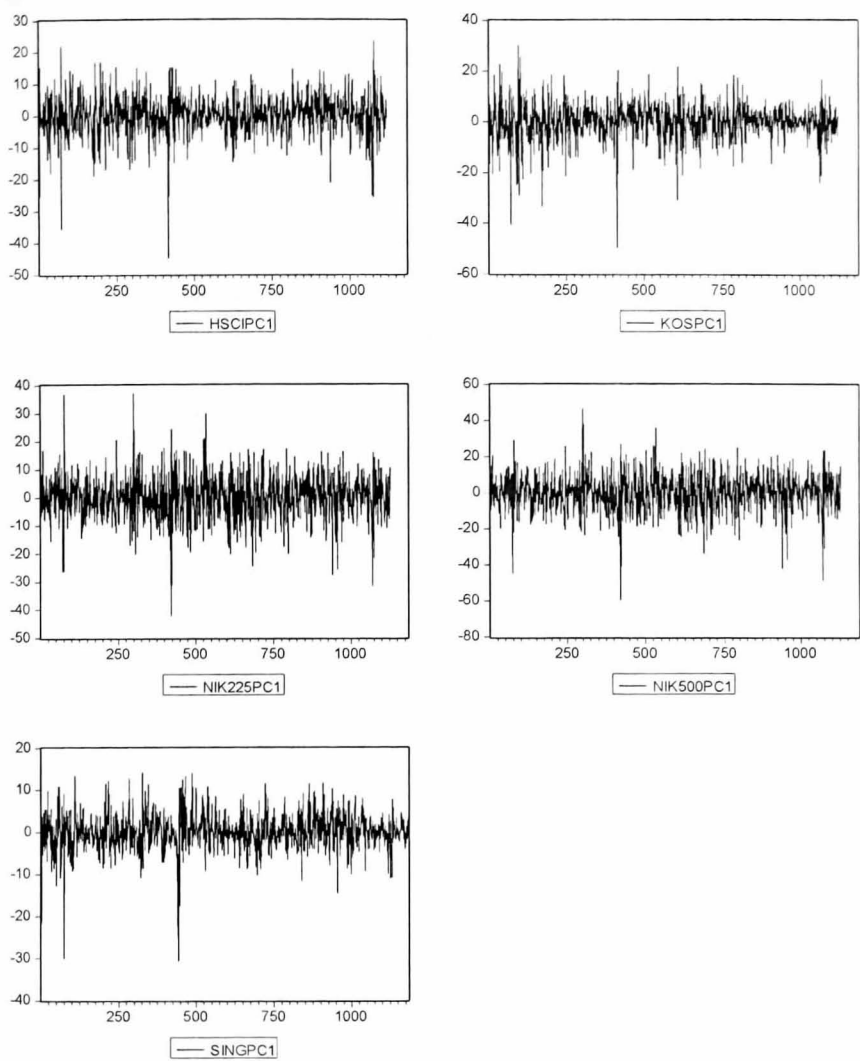


Figure 2: Plots of first principal components from the five datasets of stock returns

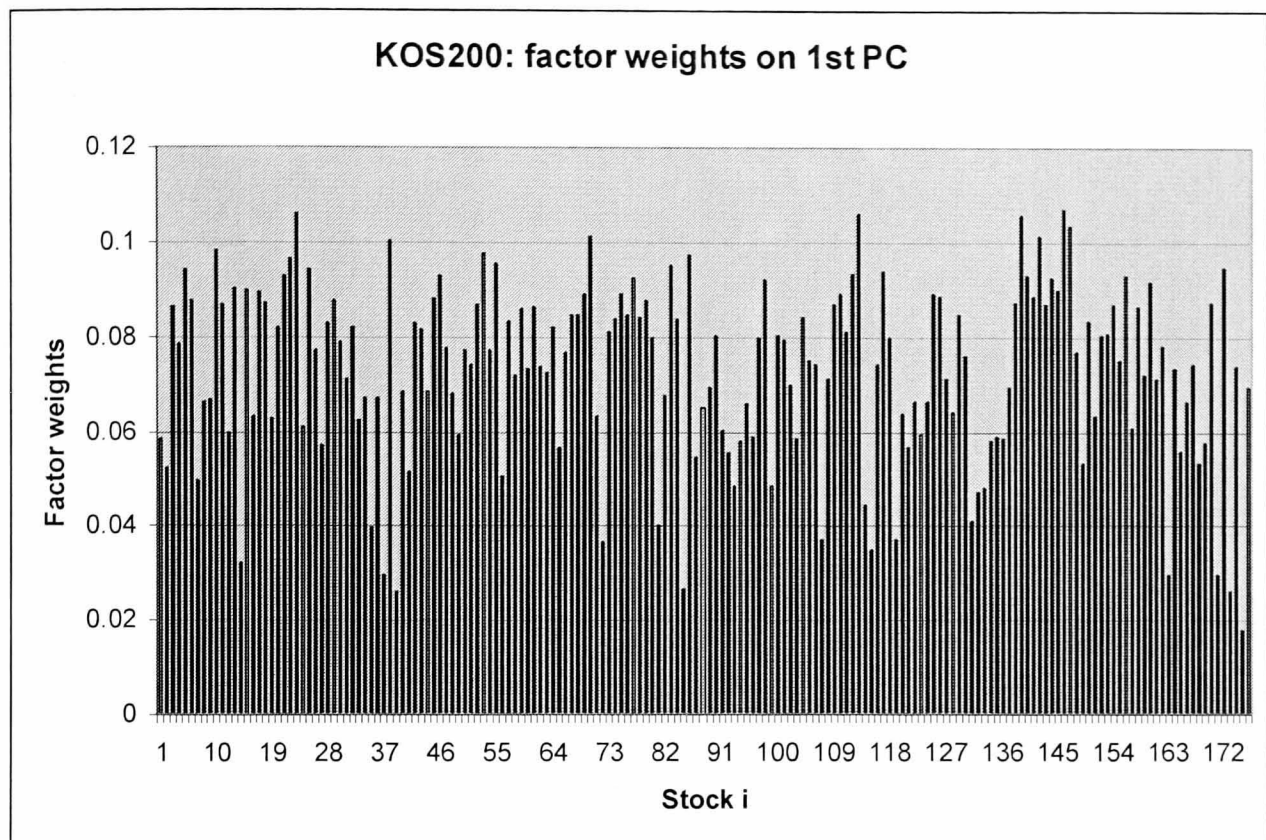


Figure 3: Factor weights on first principal component of KOS200 (Key: Factor weights plotted here is the values of the eigenvector that corresponds to the largest eigenvalue of unconditional correlation matrix $X'X$ of normalised returns of KOSPI 200 constituents. That is, vector $w_{i,k}$ in equation (3), where $k = 1$ and $i = 1, \dots, N$)

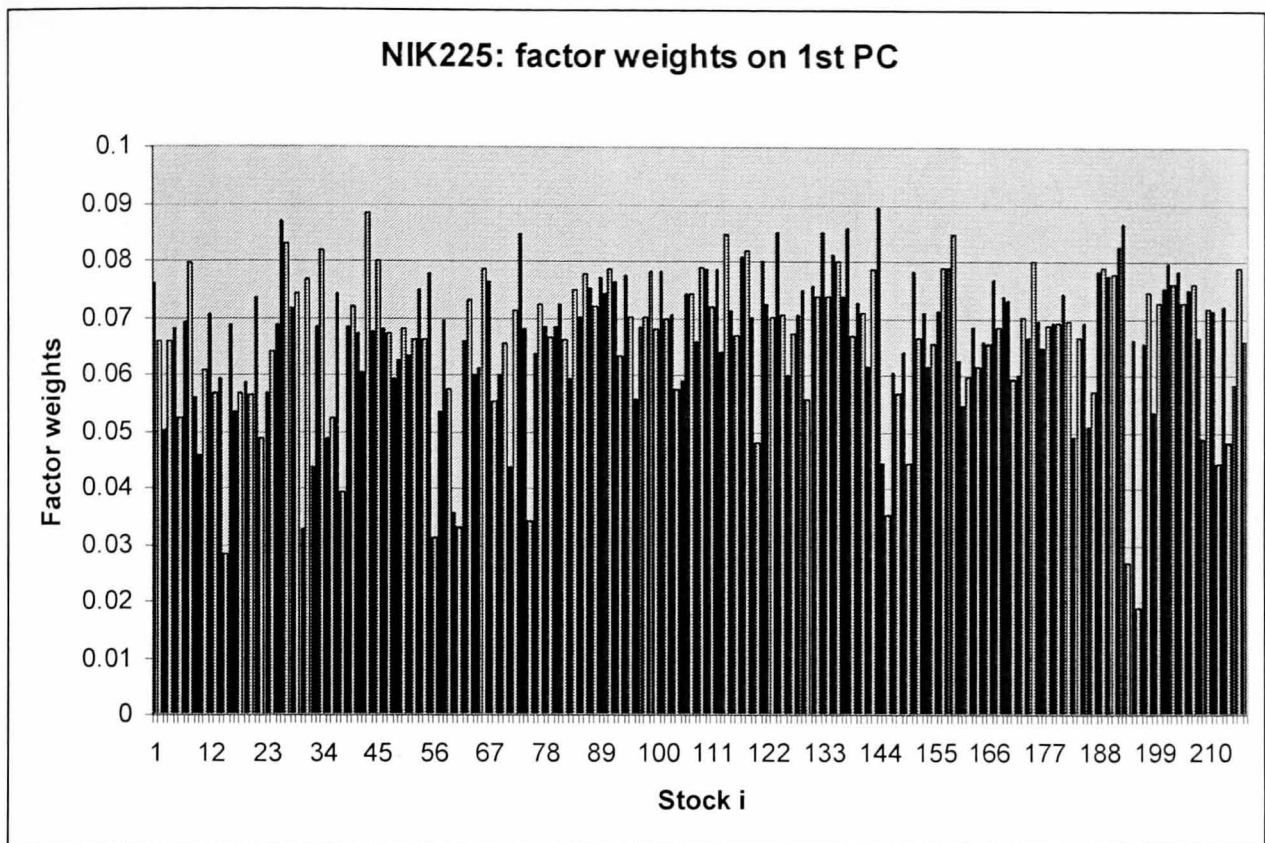


Figure 4: Factor weights on first principal components of NIK225 (Key: Factor weights plotted here is the values of the eigenvector that corresponds to the largest eigenvalue of unconditional correlation matrix $X'X$ of normalised returns of NIKKEI 225 constituents. That is, vector $w_{i,k}$ in equation (3), where $k = 1$ and $i = 1, \dots, N$)

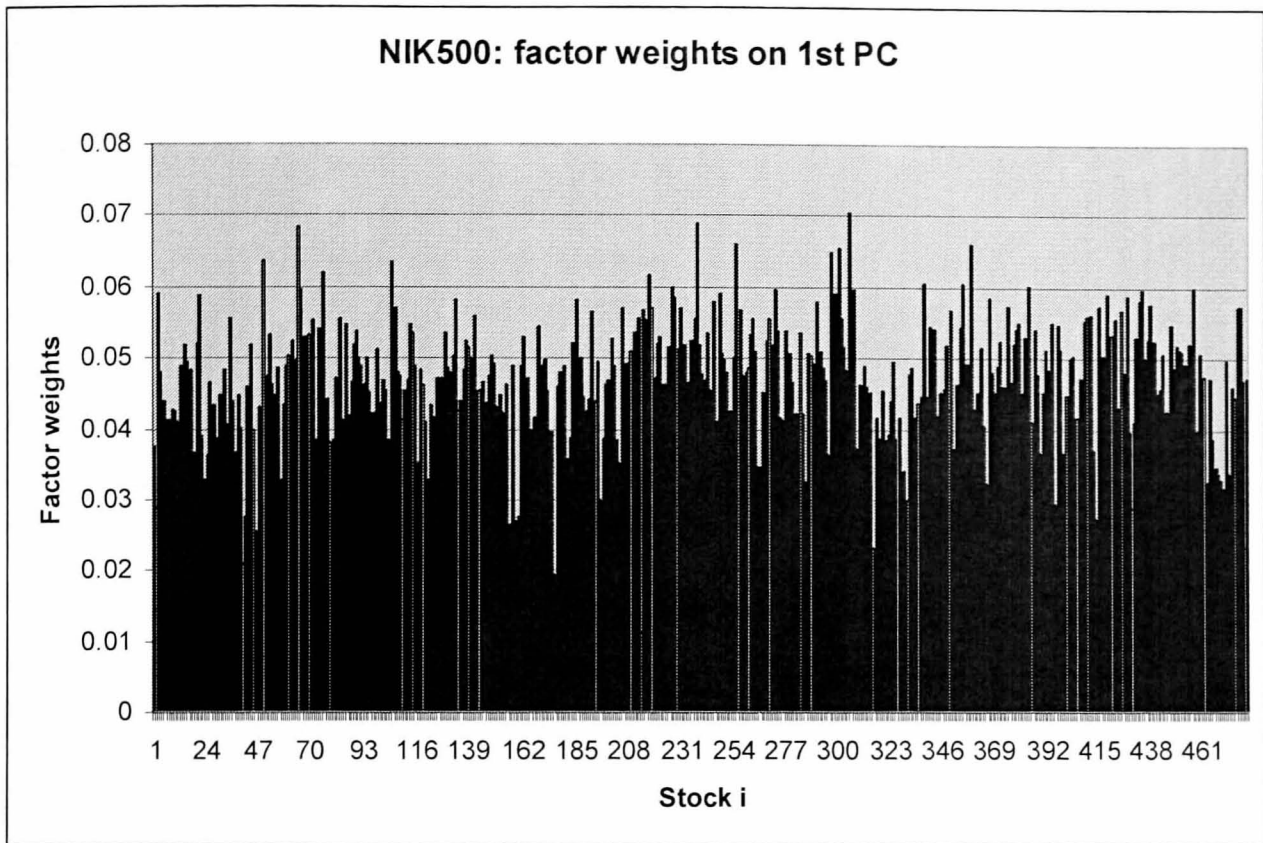


Figure 5: Factor weights on first principal components of NIK500 (Key: Factor weights plotted here is the values of the eigenvector that corresponds to the largest eigenvalue of unconditional correlation matrix $X'X$ of normalised returns of NIKKEI 500 constituents. That is, vector $w_{i,k}$ in equation (3), where $k = 1$ and $i = 1, \dots, N$)

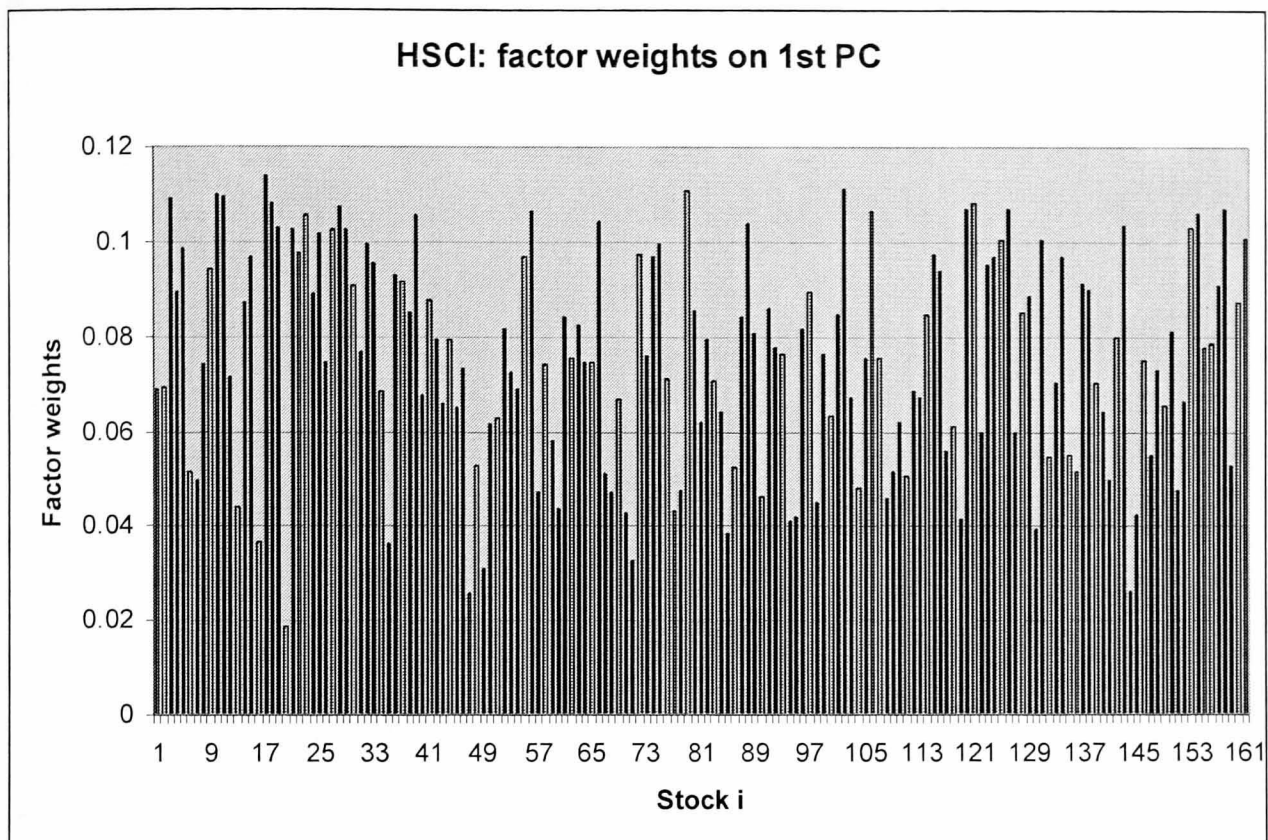


Figure 6: Factor weights on first principal components of HSCI (Key: Factor weights plotted here is the values of the eigenvector that corresponds to the largest eigenvalue of unconditional correlation matrix $X'X$ of normalised returns of HSCI constituents. That is, vector $w_{i,k}$ in equation (3), where $k = 1$ and $i = 1, \dots, N$)

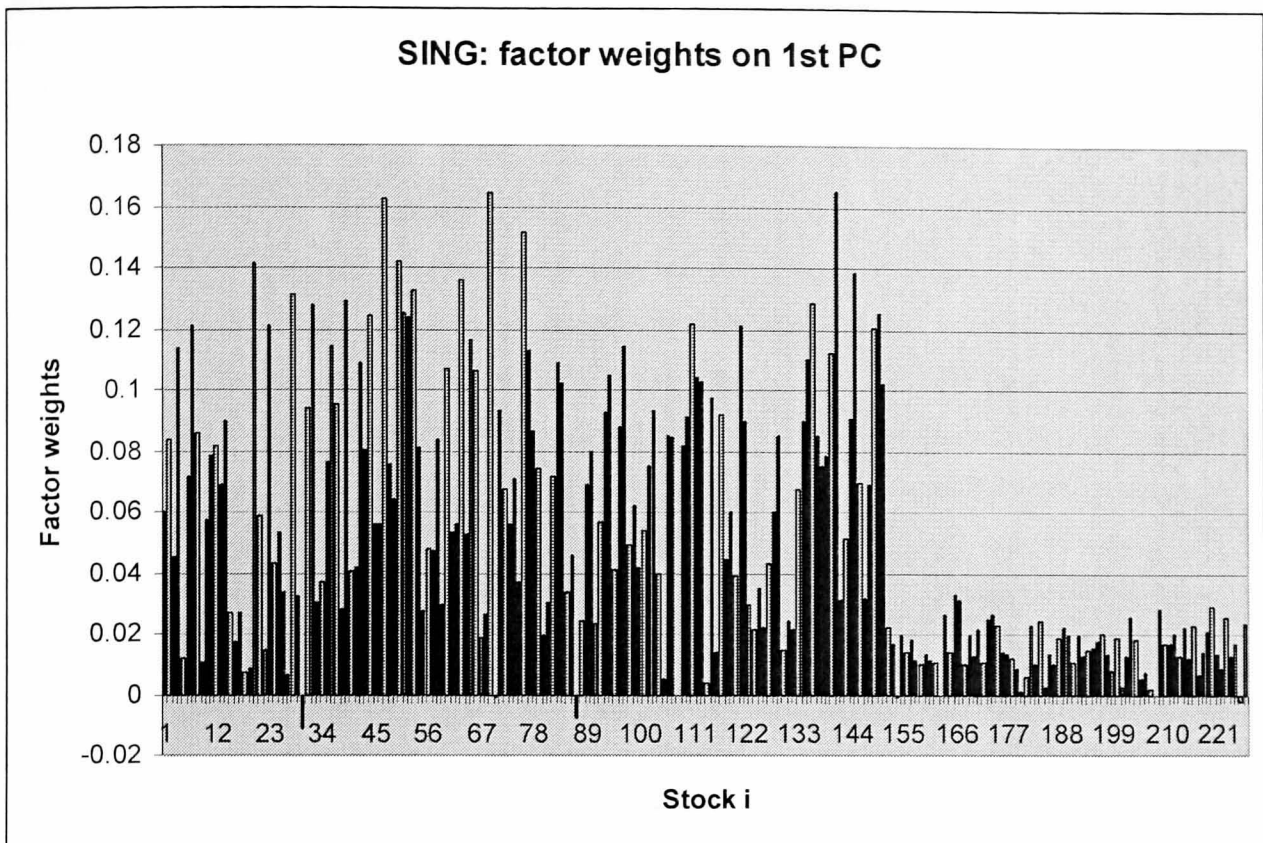


Figure 7: Factor weights on first principal components of SING (Key: Factor weights plotted here is the values of the eigenvector that corresponds to the largest eigenvalue of unconditional correlation matrix $X'X$ of normalised returns of Singapore All Share Index constituents . That is, vector $w_{i,k}$ in equation (3), where $k = 1$ and $i = 1, \dots, N$)

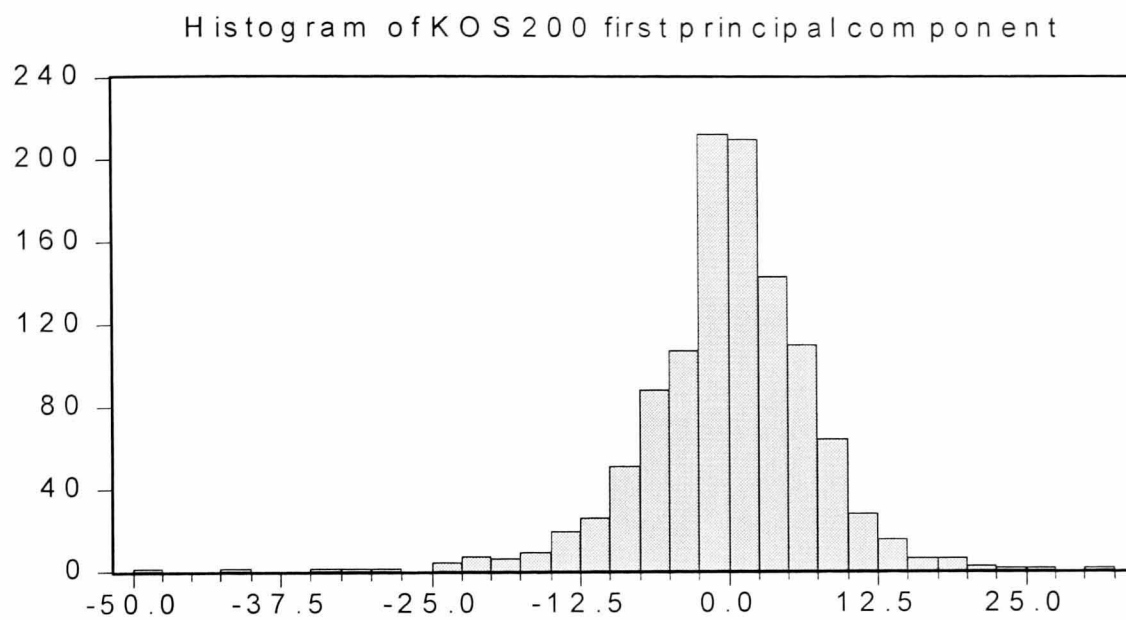


Figure 8: Histogram of KOS200 first principal component

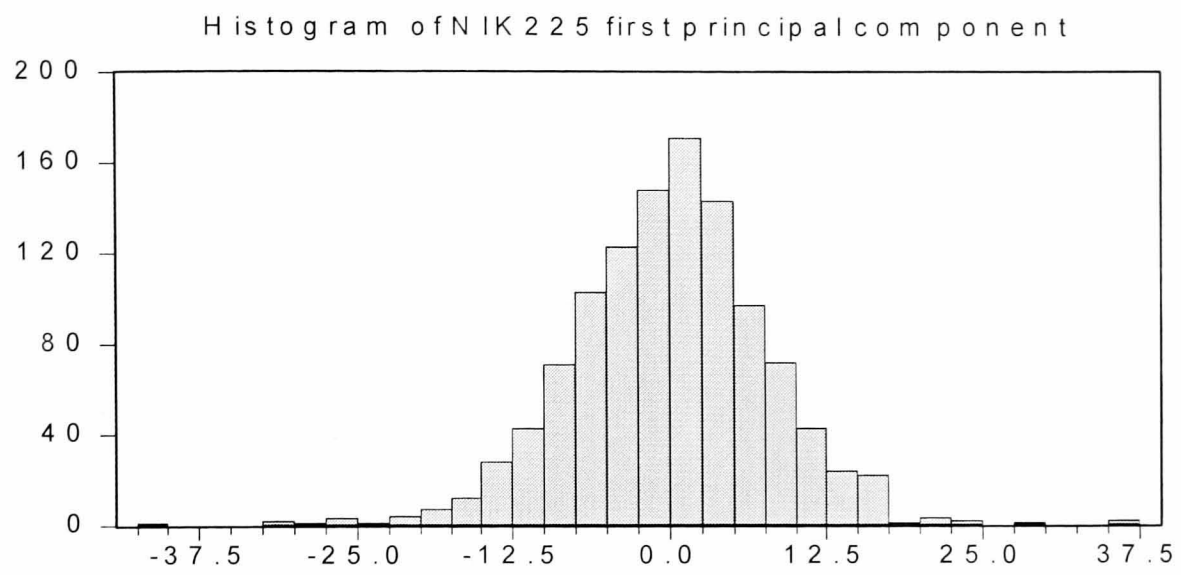


Figure 9: Histogram of NIK225 first principal component

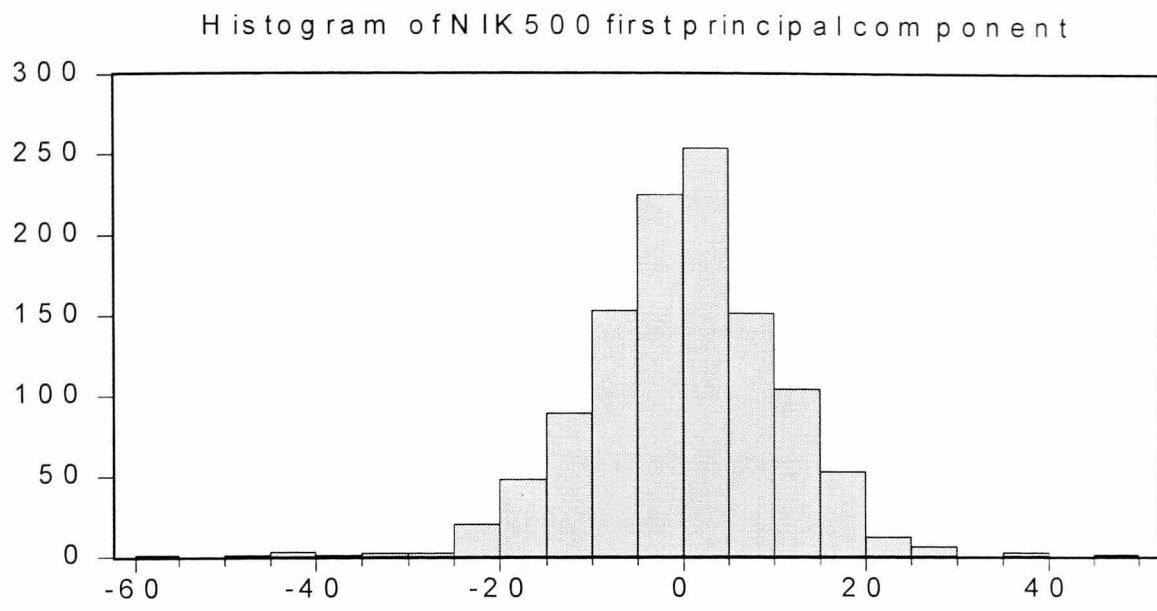


Figure 10: Histogram of NIK500 first principal component

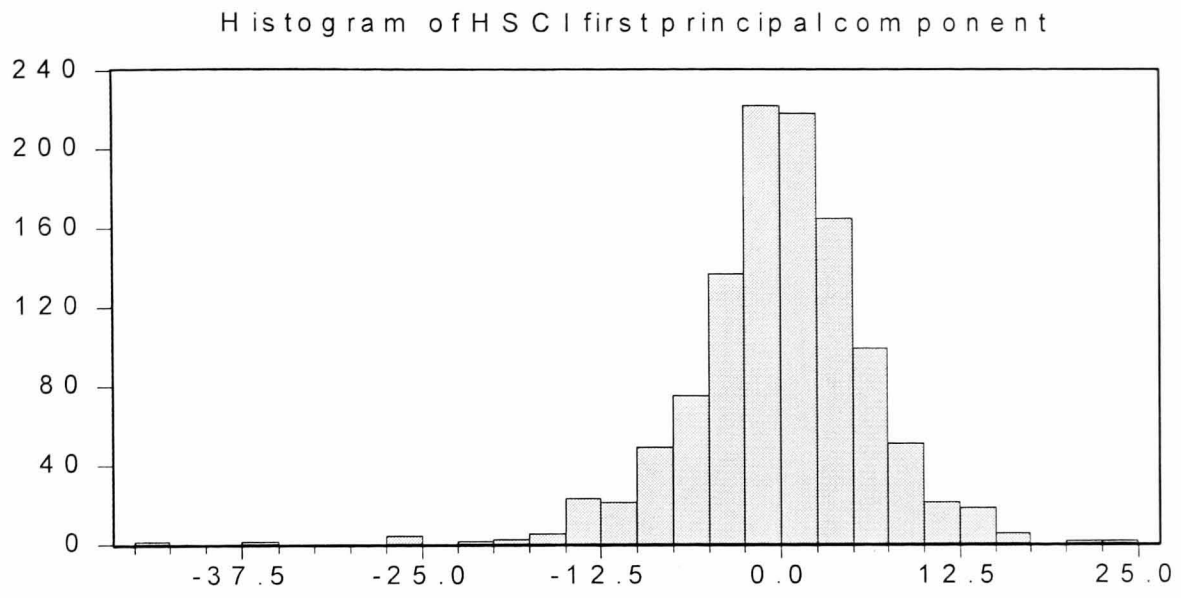


Figure 11: Histogram of HSCI first principal component

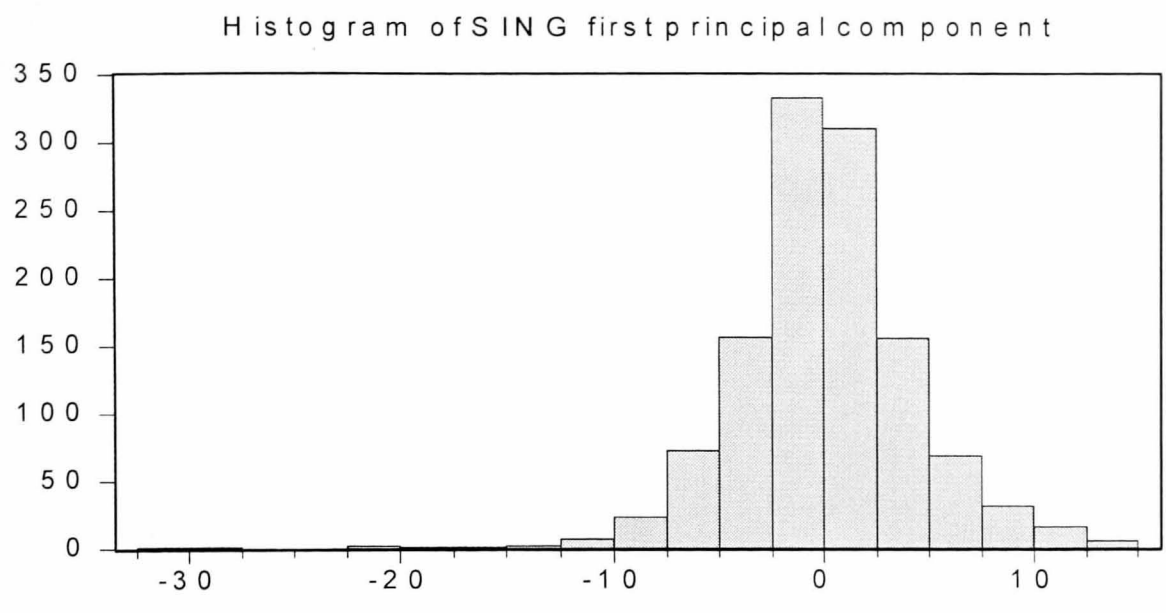


Figure 12: Histogram of SING first principal component

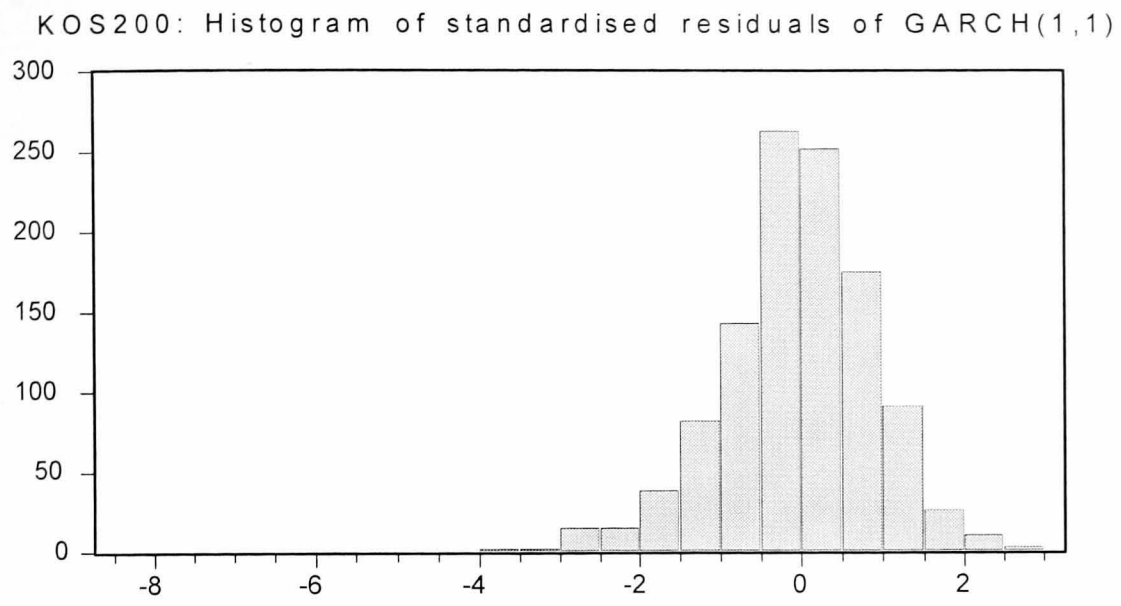


Figure 13: Histogram of GARCH(1,1) standardised residuals of KOS200 principal component

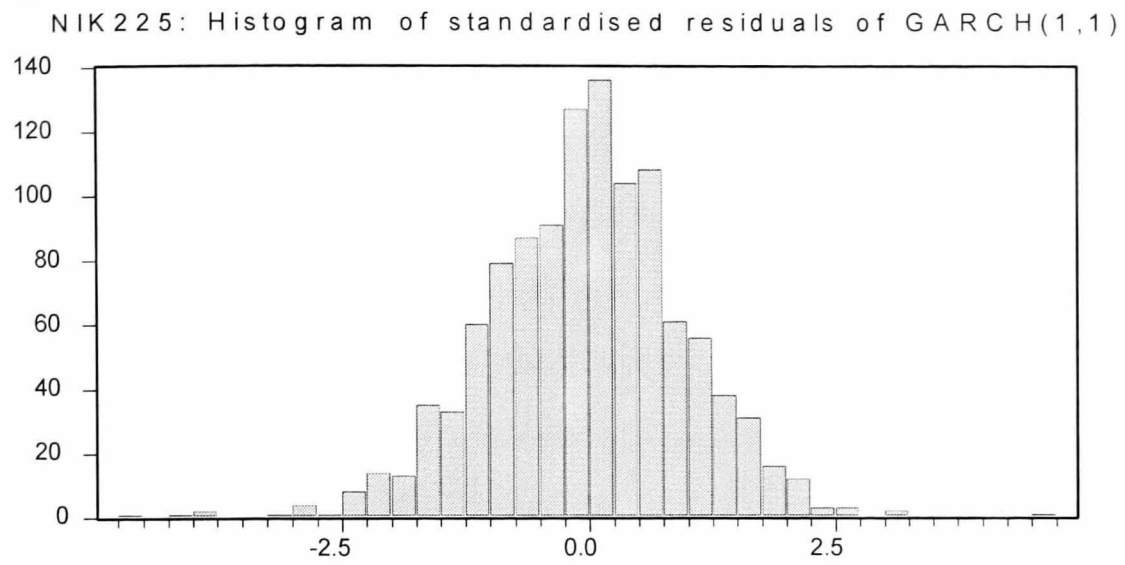


Figure 14: Histogram of GARCH(1,1) standardised residuals of NIK225
principal component

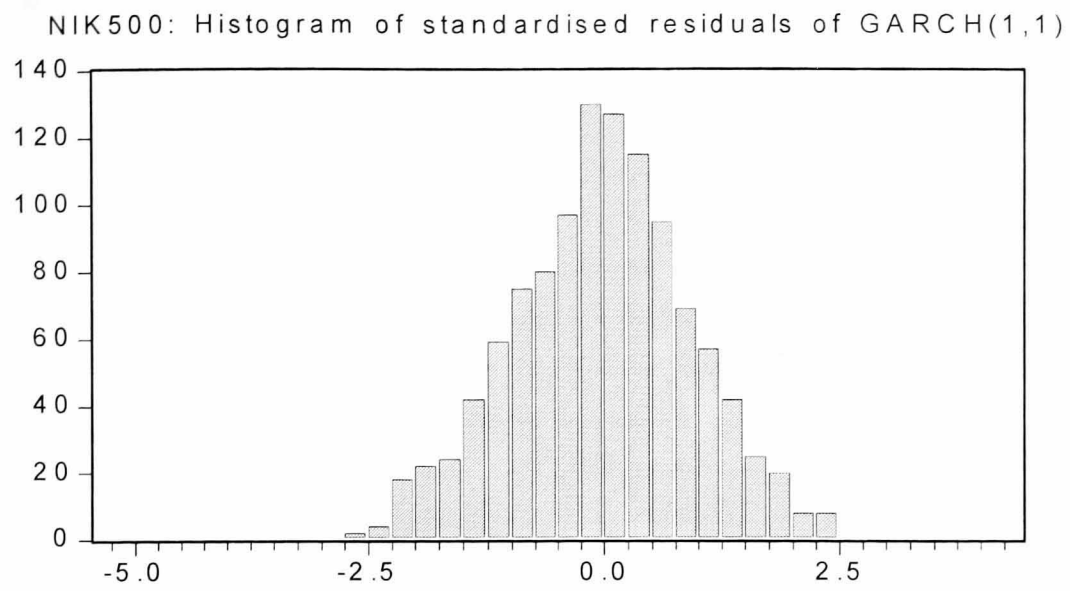


Figure 15: Histogram of GARCH(1,1) standardised residuals of NIK500 principal component

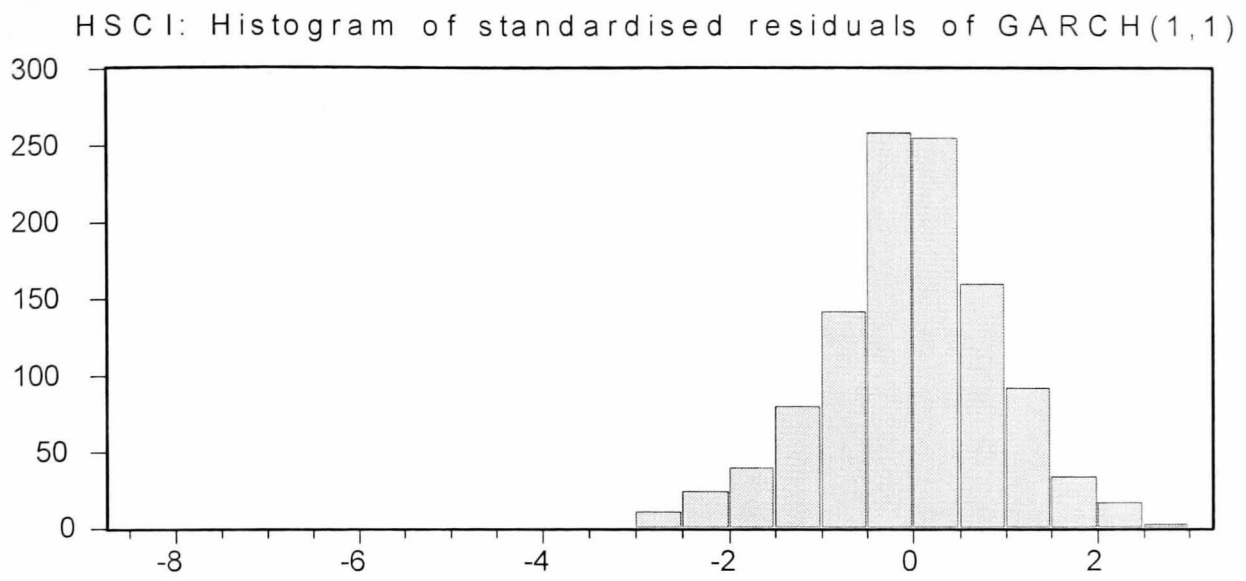


Figure 16: Histogram of GARCH(1,1) standardised residuals of HSCI principal component

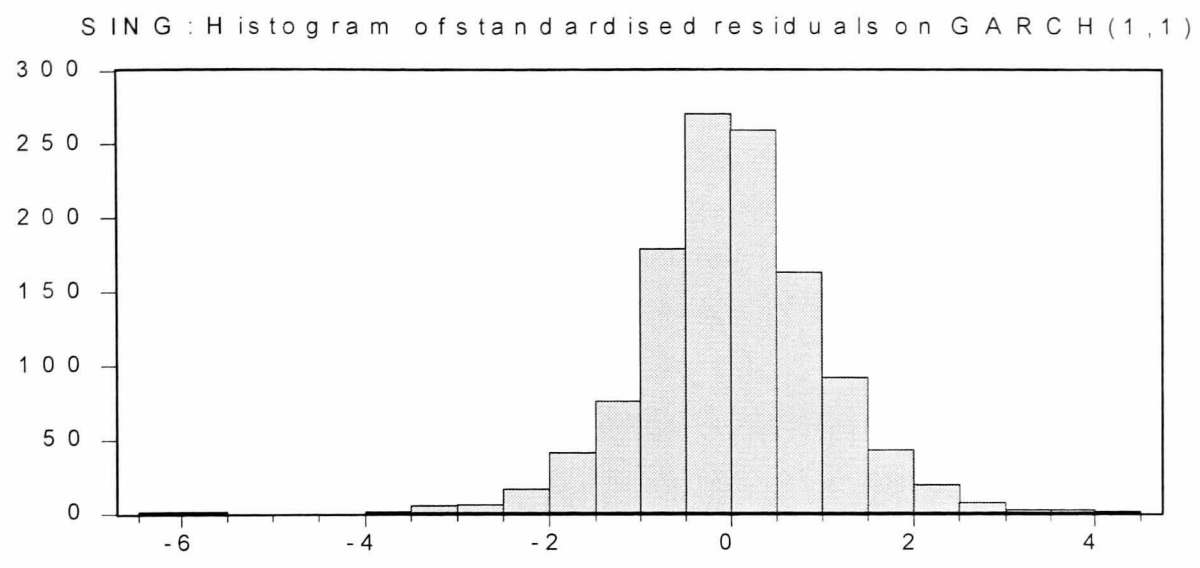


Figure 17: Histogram of GARCH(1,1) standardised residuals of SING principal component

Chapter 2

Factor Analysis of Stock Volatility Using Stochastic Volatility Factor Model: Evidence from Five Asian Stock Indexes

1 Introduction

In the previous chapter, our analysis has been carried out in the context of the orthogonal GARCH model of Alexander (2001a, b). In this chapter, our analysis will be carried out in the context of a comparable model, the stochastic volatility factor model of Cipollini and Kapetanios (2005). In contrast to the orthogonal GARCH analysis, common factors in stochastic volatility factor model are estimated from large datasets of Asian stock volatilities via principal components of Stock and Watson (2002a). We will perform a correlation analysis of stock volatilities and an examination of the dynamics of common volatility factor estimates. The explanatory power of common factors is also investigated. In particular, we will analyse the impact of the size of dataset on explanatory power of factor estimates. Moreover, we will detect if long memory

exists in common factors of stock volatilities.

Studies of the underlying forces that cause fluctuations in financial time series have long been under the interest of empirical researchers. Empirical studies in this literature tend to take two different approaches. Some literatures study a set of pre-defined global and local macroeconomic variables that are believed to have made a contribution to the fluctuation in a particular stock market as a proxy for common factors. For example, Bilson, Brailford and Hooper (2001) employ a multifactor model to select common explanatory factors for emerging markets from a set of global risk variable and local economic variables. This type of studies relies on observable significance of the "proxy common factors". However, this may be quite restrictive in the sense that if fluctuations in the financial time series are actually driven by some unobserved common forces, then these forces will be neglected in the analysis. Another approach employs state-space representation. Common factors are defined as some unobserved components in the state-space setting. Stock and Watson (1998, 2002a) call these common factors "diffusion indexes". These unobserved common factors summarise the information from a large group of driving forces that account for variation in a dataset. This group of driving forces may be generated by both macroeconomic variables and some other unobserved forces. No pre-definition of the common factors is made therefore, state-space factor analysis is relatively more flexible.

The seminal paper by Harvey, Ruiz and Shephard (1994) is influential in the stochastic volatility literature. In view of the complications in estimation and interpretation of the multivariate GARCH models due to large number of

parameters and the need to impose constraints, they propose a multivariate stochastic volatility model. They have also made suggestion to incorporate common factors, which is assumed to follow a multivariate random walk, into the model. The model is estimated by quasi-maximum likelihood in a state-space approach.

However, Cipollini and Kapetanios (2005) argue that although state-space approach is powerful and intuitive, but it is not computationally tractable for datasets with very large dimensions. Moreover, allowing the common factor to follow only a multivariate random walk is too restrictive in the sense that more complicated dynamics in the factors cannot be accounted for. They have therefore proposed a stochastic volatility factor model for large datasets by extending the model of Harvey et. al. (1994). Redefining the factor vector and by making no prior assumption to the form of the process that underlies the common factors, their model allows more complicated temporal features of common factors other than non-stationarity being captured. Moreover, their suggestion of using principal components method of Stock and Watson (2002a) for extracting common factors has simplified computation with large datasets. Their studies and that of Bai (2003) show principal components estimation provides consistent estimates in large datasets. Bai (2003) also studies the asymptotic properties of this estimator. All these are remarkable developments in factor analysis.

The stochastic volatility factor model is comparable to the orthogonal GARCH model. Both of them belong to the type of volatility models with latent factors. The former is an advanced development of stochastic volatility model that

involves using latent factors to model common volatility. Whereas the latter is an advanced development of the GARCH model and involved using latent factors to summarise comovements in a dataset of financial asset and compute volatilities of the entire system using GARCH volatilities of these latent factors.

In this chapter, we will apply the stochastic volatility factor model to analyse the Asian stock volatilities. We adopt the approach that Cipollini and Kapetanios (2004) use in their paper for our analysis. Applying their stochastic volatility factor model and using principal components for factor estimation, we analyse the common factor in the constituent stock volatilities of the Japanese, South Korean, Hong Kong and Singaporean indexes. The outline of this chapter is as follows. Section 2.1 presents a brief overview of stochastic volatility literature. The remainder of section 2 outlines the methodology used for the analysis in this chapter. The extension of the original multivariate stochastic volatility model into the stochastic volatility factor model is reviewed in section 2.2. Principle components method of Stock and Watson (2002a) is also discussed in section 2.3. In section 2.4, we discuss the fractionally integrated autoregressive moving average (ARFIMA) representation of the common factors. Section 3 displays our empirical results. We first implement a correlation analysis to look at the correlation among constituent stock volatility series of the Asian indexes. We then examine the estimated factor series for all five datasets to see if there is any evidence of heteroscedasticity and serially correlations. Explanatory power of the set of factors for each dataset of constituent stock volatilities are then evaluated. In particular, cumulative R^2 of the dominant factors of the two Japanese datasets of volatilities are compared in order

to see how size of a dataset impacts on the explanatory power of a factor. Followed by the investigation of the dynamics of the factor series, we focus on the examination of the long memory nature in the factor estimates. Conclusion is made in section 4 and tables and graphs are presented at the end of this chapter.

2 The Model

As mentioned in the introductory chapter of this thesis. The stochastic volatility models are comparable to the GARCH type models. The main difference between the two models is that an unobserved shock to the return variance is explicitly included in the volatility dynamics in the former. The variance has thus become a latent process and it is unobserved even when we have full knowledge of all the observable past information. This has made the use of standard maximum likelihood estimation of the parameters in the stochastic volatility model infeasible. Despite the fact that the estimation of stochastic volatility model is less straight forward when comparing with the standard GARCH models, the former have naturally overcome some drawbacks of the latter. Moreover, increasing popularity of the study of realised volatility and the use of high frequency financial data has enriched the development of the stochastic volatility model due to their close and natural linkage to the continuous-time version of the model. All of these have contributed to the vast and rapid growth of the stochastic volatility literature. In this section, we present a brief overview of the stochastic volatility literatures.

2.1 Stochastic Volatility – An Overview

In the early progressions of the stochastic volatility literature, most of the econometricians focused on the discrete time versions of the model. Whereas the financial mathematicians and financial econometricians focused more on the continuous time versions of model, which were mainly used to deal with option pricing. Recall equation (3) in the introductory chapter, Taylor (1986) is the first to introduce this logarithmic stochastic volatility model by using an autoregression to represent the unobserved logarithmic volatility, h_t . The disturbance term in this autoregression and the shocks in the return equation are assumed to be uncorrelated. The normality assumption on the disturbance in the autoregression of the logarithmic conditional volatility has been widely adopted. This model was then known as the stochastic volatility model. As opposed to the standard GARCH model, the simplest form of the stochastic volatility model has a more flexible set-up that allows the model to capture leverage effect or asymmetry that the standard GARCH model fails to account for. This can be done by allowing a negative correlation between the shocks in the autoregression of the logarithmic volatility and the shocks in the return. In this sense, the stochastic volatility model is analogous to the EGARCH model. Moreover, the model is more capable of explaining excess kurtosis found in financial time series than the standard GARCH model.

Persistence is another well-known feature of financial volatility. By allowing the logarithmic conditional volatility to follow a random walk process, persistence in financial volatility can be captured. This specification is analogous to the IGARCH model. There has been considerable empirical evidence to show

that autocorrelations in financial time series has a faster rate of decay at short lags but a slower rate of decay at long lags. Example can be seen from the study of Standard and Poor 500 index by Ding, Granger and Engle (1993), in which the autocorrelation of fractional moments of return series has a very slow decay. The study of Psaradakis and Sola (1995) on stock volatility using 22 UK stocks via FIGARCH model also provides evidence of long memory. Their results confirm the commonly observed fact that shocks die out at a slow hyperbolic decay. In the long memory stochastic volatility model by Harvey (1998), the logarithmic conditional volatility, h_t , is generated by a fractional noise. The model is covariance stationary if the differencing operator is less than 0.5. In the cases when the operator takes the values of 1 and 0, the process reduces to a random walk and a white noise, respectively. The long memory stochastic volatility model has a hyperbolic decay. The study carried out by Harvey (1998) has shown that the autocorrelation function of the long memory process of h_t has a much slower decay as compared to the AR(1) representation of h_t . Breidt, Crato and de Lima (1993) also suggest a long memory stochastic volatility model. They construct the model by introducing an ARFIMA process into a standard stochastic volatility set-up. Their application have shown the model performs better than other volatility models.

Modern development of the stochastic volatility literatures lie in the context of realised volatility and the application of the model into high frequency financial data due to their close linkage with the continuous time version of the model. See for example, Andersen, Bollerslev, Diebold and Labys (2001) and Barndorff-Nielsen and Shephard (2002).

Although the construction of stochastic volatility model favours its flexibility in capturing stylised facts of volatility than the standard GARCH model, its complication in estimation has also been an issue. The construction of GARCH model has made available its likelihood function but this is not the case for the stochastic volatility model. An extensive amount of stochastic volatility literatures focus on the estimation and inference of the models.

Some literatures on the inference of stochastic volatility models initiate estimators based on moments of the model that can be easily computed. For example, Taylor (1982) uses the method of moments for estimation. GMM approach has also been discussed, see for example, Andersen and Sørensen (1996). The approach suggested in the seminal paper by Harvey, Ruiz and Shephard (1994) has been significant among the literature of the stochastic volatility model. They suggest to estimate the model using quasi-maximum likelihood via Kalman filter. By adopting their method, the parameter estimates, filtered estimates and smoothed estimates of the conditional volatility process can all be obtained.

Other literatures on the estimation of stochastic volatility models advocate the use of simulation-based inference. For example, Danielsson and Richard (1993) propose an accelerated Gaussian importance sampler. Jacquier, Polson and Rossi (1994) develop the Bayesian framework and using Markov Chain Monte Carlo (MCMC) algorithm for the simulation from the posterior distribution. Their sampling experiments show Bayesian estimators have better performance than the methods of moments and the quasi-maximum likelihood estimators. More discussions on the development of MCMC methods for the

stochastic volatility models have also been carried out, see for example, Andersen (1994a) and Kim, Shephard and Chib (1998). Apart from the use of MCMC method, there is another simulation based approach so-called the indirect inference approach or the efficient methods of moments (EMM), see for example, Gouriéroux, Monfort and Renault (1993). EMM is a simulation-based moment matching procedure. The matched moments are called the score generator. Efficiency in the parameter estimates of the stochastic volatility model as in maximum likelihood estimation can be obtained if the score generator gives a good approximation of the distribution of the data (see Gallant, Hsieh and Tauchen (1997) for application of the method).

So far the part of the above discussion regarding the modifications of the stochastic volatility models has been limited to the univariate stochastic volatility models, the need to study several financial time series simultaneously has nourished the extension to a multivariate set-up. Harvey, Ruiz and Shephard (1994) propose the discrete time multivariate stochastic volatility model in which the logarithmic unobserved variance is modelled by a random walk. They estimate the model via state-space approach. Chib, Nardari and Harvey (2002) have suggested Bayesian estimation of this model via MCMC method. However, Cipollini and Kapetanios (2005) argue that estimation via state-space approach is powerful but not ideal when the dimensions of dataset is large due to computational constraints. Whereas the Bayesian estimation via MCMC requires knowledge in the computational algorithm and involve selection of several parameters (e.g. the prior density), this may be contentious. Moreover, Harvey, Ruiz and Shephard (1994) also discussed the incorporation of latent

factors into the model by allowing the common factors to follow multivariate random walk. This suggestion has nurtured further extension of the model into the stochastic volatility factor model by Cipollini and Kapetanios (2005). In contrast to the suggestion by Harvey et. al. (1994), they have redefined the dynamics of the factors and thus allow more complex dynamics rather than a random walk to be captured. In particular, they advocate the use of principal components of Stock and Watson (2002a) to estimate their factor model. This has simplified both the modelling and estimation procedure. The studies in both their paper and in Bai (2003) show this estimators are consistent.

Our empirical study in this chapter is to apply the stochastic volatility factor model into large datasets of constituent stock volatilities of the five Asian stock indexes considered in chapter 1. We adopt the approach and the factor model of Cipollini and Kapetanios (2005) to carry out empirical analysis. They extend the multivariate stochastic volatility model of Harvey, Ruiz and Shephard (1994) to allow more complex temporal features to be captured in the process that underlie the common factors by redefining the factor vector. Their application of the model is made to the constituent stock volatilities of Standard and Poor 100 and 500 (S&P 100 and S&P 500) indexes. Our empirical application here aims to examine the dynamics of the common factors estimated via principal components. We will compare some of the findings here with the those obtained in chapter 1 when common factors are extracted from datasets of returns. Moreover, we will also provide empirical evidence to show that large datasets may not always be desirable for factor analysis as claimed by Boivin and Ng (2003). In the next subsection, we will go through the progression

of the original multivariate stochastic volatility factor model to the stochastic volatility factor model. Estimation method and the underlying process of the common factors will also be explored.

2.2 Stochastic Volatility Factor Model

It is well known that financial time series are featured with heteroscedasticity and serial correlation. The ARCH and GARCH family models by Engle (1982) and Bollerslev (1986), and further modifications of those models that feature with time-varying volatility and serial correlation are widely used in financial econometric analysis. However, Harvey, Ruiz and Shephard (1994) argue that the multivariate version of GARCH family models may not be convenient to estimate and interpret due to the possibility of large number of parameters. The maximum likelihood estimation of this multivariate version requires imposition of restrictions. They have therefore, proposed an alternative approach, which is to model volatility as an unobserved stochastic process. The logarithm of this process is then a linear stochastic process like an AR process. They call this kind of model a stochastic volatility model or stochastic variance model. Let's first look at the basic setting of this model.

Consider first a univariate time series, y_t

$$y_t = u_t \sigma_t \tag{1}$$

where $t = 1, \dots, T$. y_t is the product of a Gaussian white noise process

with zero mean and variance equals to one. σ_t is the standard deviation and h_t denotes $\ln(\sigma_t^2)$. Harvey et al. (1994) suggest that h_t can be modelled by a simple AR process, an ARMA process or it can follow a random walk. Now suppose there are N constituent stocks included in a particular stock index. Consider the following generalization to a multivariate time series of N constituent stock returns y_t . For stock i at time t ,

$$y_{i,t} = u_{i,t}(\exp(h_{i,t}))^{1/2} \quad (2)$$

where $i = 1, \dots, N$, $t = 1, \dots, T$. $y_t = (y_{1,t}, \dots, y_{N,t})'$; and $u_t = (u_{1,t}, \dots, u_{N,t})'$ is a multivariate normal vector of disturbance with mean zero and variance-covariance matrix Σ , which has diagonal elements of ones and off-diagonal elements of $\rho_{i,t}$'s. Squaring both sides of equation (2) and taking natural logarithm gives us

$$\ln(y_{i,t}^2) = E(\ln(u_{i,t}^2)) + h_{i,t} + \ln(u_{i,t}^2) - E(\ln(u_{i,t}^2)) \quad (3)$$

Denotes $\ln(y_{i,t}^2)$ by $y_{i,t}^*$, $E(\ln(u_{i,t}^2))$ by a_i and $\ln(u_{i,t}^2) - E(\ln(u_{i,t}^2))$ by $\zeta_{i,t}$. Equation (3) can be written as

$$y_{i,t}^* = a_i + h_{i,t} + \zeta_{i,t} \quad (4)$$

Harvey et al. suggest that logarithm of the unobserved variance, $h_{i,t}$ can be modelled by a multivariate random walk.

$$h_{i,t} = h_{i,t-1} + \epsilon_{i,t} \quad (5)$$

Alternatively, one can also incorporate common factors into the stochastic volatility model and to treat equation (4) and (5) as the measurement and transition equations in a state-space representation. Allowing the $N \times 1$ vector of unobserved variance to incorporate a vector of K common factors f_t , i.e. $h_t = \theta f_t$, where θ is a $N \times K$ matrix of factor loadings with $K \leq N$, equation (4) and (5) becomes

$$y_t^* = a + \theta f_t + \zeta_t \quad (6)$$

$$f_t = f_{t-1} + \nu_t \quad (7)$$

$y_t^* = (y_{1,t}^*, \dots, y_{N,t}^*)$ is a $N \times 1$ vector of transformed data using standard logarithmic transformation. This is used as a volatility proxy. f_t is a $K \times 1$

vector of common factors and it follows a multivariate random walk. θ is a $N \times K$ matrix. ζ_t and ν_t are uncorrelated.

However, Cipollini and Kapetanios (2005) argue that this setting of the model can be quite restrictive, since it only allows the factors to follow a random walk. Therefore, if the true underlying process of the common factors has more complicated dynamic natures, for example, long memory, the state-space version of the stochastic volatility model becomes inadequate. Their extension to the above stochastic volatility model allows one to capture more temporal properties of the data. They propose that the unobserved variance, h_t should have a common component and a disturbance

$$h_{i,t} = \theta'_i f_t + \eta_{i,t} \quad (8)$$

Substitute equation (8) into equation (4) gives

$$y_{i,t}^* = a_i + \theta'_i f_t + \omega_{i,t} \quad (9)$$

where $\omega_{i,t} = \eta_{i,t} + \zeta_{i,t}$ is an idiosyncratic error term. From equation (9), we can see that stochastic volatility of stock i , $y_{i,t}^*$, has two components. A common component represented by $\theta'_i f_t$, and an idiosyncratic component represented by $\omega_{i,t}$. The former tells us how much of the volatility in stock i is due to

market condition as the common factor f_t summarises the unobserved forces that cause changes in market condition. The latter tells us how much of the volatility in stock i is caused by forces that are unique to the stock itself. The analysis in this chapter restricts to the common component as our aim is to examine the dynamics and temporal features of this common comovements of the stocks in a market. The extension as in equation (9) is more flexible than the original stochastic volatility model setting as they do not make any specific assumption about the form of the process that underlies the common factors. Therefore, once factors are estimated, then by careful examination of the estimates, one can seek for an appropriate specification that best captures the temporal features of the series. In their paper, they have found persistence in the autocorrelations of their factor estimates of S&P 100 and S&P 500 constituent stock volatilities. This suggests that the underlying specification for their estimated factor series should fall in the category of long memory models. We will also see later from the empirical results in this paper, constituent stock volatilities of some Asian indexes have factor estimates that also appear to have long memory.

There are several assumptions on the factor, f_t , the factor loading and the error term, $\omega_{i,t}$ in order for us to obtain consistent factor estimates using principal components method by Stock and Watson (2002a, 2000b). The stochastic volatility factor model of Cipollini and Kapetanios (2005) is also flexible in the way that one can use principal components estimation to obtain consistent factor estimates without making further assumptions additional to the estimation method. We will go through the basic mechanism of principal

components estimation method and the relevant assumptions in the following subsection.

2.3 Estimation of Common Factors: Principal Components Method

Conventional factor analysis focuses on small datasets. The analysis requires some restrictive assumptions to hold, and the use of maximum likelihood estimation. Bai (2003) discusses some limitations of classical factor analysis due to those restrictive assumptions. He points out that the assumption of fixed number of cross-section dimension (N), which is required to be smaller than the time dimension (T) is unrealistic since the number of series included in the dataset is much larger than the number of time series observations in economic datasets. The assumption of idiosyncratic innovations being *i.i.d.* across time and across cross-section is too strong for economic time series. Moreover, maximum likelihood estimation is not feasible for estimating factor model with a very large number of series in the dataset.

Stock and Watson (2002a) suggest the method of principal components for estimating factors in an approximate dynamic factor model with large dataset. Principal components method involves eigenvalue decomposition of sample variance-covariance matrix. It is simple to use and asymptotically equivalent to the maximum likelihood estimation. In their paper, Stock and Watson (2002b) study the finite sample properties of principal component estimator. They show that under rather general assumptions, the factor estimates of an approximate factor model obtained by using this method are consistent,

even if idiosyncratic innovations are serially and cross-sectionally correlated. Bai (2003) also shows that the necessary conditions for ensuring consistency are asymptotic orthogonality and asymptotic homoscedasticity in idiosyncratic innovations.¹ Consistency in factor estimates can be obtained even in the presence of serial correlations and heteroscedasticity. Cipollini and Kapetanios (2005) claim that the results of Bai (2003) on consistency of factor estimates estimated by principal components can still be valid if the underlying processes of the factor estimates are stationary long memory ARFIMA(p,d,q) with finite fourth moment.

Consider the following approximate dynamic factor model. $x_t = (x_{1,t}, \dots, x_{N,t})$ is a N -dimensional vector of multivariate time series. At time t for series i

$$x_{i,t} = \lambda'_i f_t + \xi_{i,t} \quad (10)$$

f_t is a K -dimensional vector of common factors with $t = 1, \dots, T$. λ'_i is the i th row of matrix Λ , which is a matrix of factor loadings. $x_{i,t}$ is the element in the t th row and i th column of a $T \times N$ data matrix X . $\xi_{i,t}$ is the i th element of $\xi_t = (\xi_{1,t}, \dots, \xi_{N,t})$, which is a vector of idiosyncratic innovations.

¹Consider the factor model in equation [10], Bai (2003) calls the restrictions

$$N^{-1} \sum_{i=1}^N \xi_{i,t} \xi_{i,s} \rightarrow 0, \text{ for } t \neq s$$

$$\text{and } N^{-1} \sum_{i=1}^N \xi_{i,t} \rightarrow \sigma^2, \text{ for all } t \text{ as } N \text{ tends to } \infty$$

asymptotic orthogonality and asymptotic homoscedasticity, respectively.

In contrast to a strict factor model which assumes idiosyncratic innovations to be *i.i.d.*, these ξ_t here are allowed to be weakly time and cross-sectionally dependent. Estimation of factors and factor loadings by principal components method requires us to minimise the following objective function with respect to f and Λ .

$$V(f, \Lambda) = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \lambda'_i f_t)^2 \quad (11)$$

This is analogous to minimising the variance of the idiosyncratic innovations $\xi_{i,t}$. Estimating f_t requires the eigenvectors of the matrix XX' . Minimising the above objective function with respect to the factor is equivalent to maximising the matrix trace of $f'(XX')f$, subject to the restriction on $T^{-1}(f'f)$ being orthogonal, i.e. $T^{-1}(f'f) = I_k$, where $f = (f_1, \dots, f_K)$. In other words, the consistent estimate of f is given by the K largest eigenvectors of matrix XX' . The matrix of factor loadings, $\Lambda = (f'f)^{-1} f'X = \frac{f'X}{T}$. Several mild assumptions on the factors f_t , the factor loadings Λ and the innovations ξ_t , are required for the estimation to provide consistent estimates. Those assumptions are outlined as follows.²

1. For the factor loadings, $(\Lambda'\Lambda/N) \rightarrow I_k$ and $\|\lambda_i\| \leq \bar{\lambda} < \infty$
2. For the factors, $f'f$ has finite unconditional second moment, that is, $E(f'f)^2 < \infty$. The factors are also allowed to be serially correlated, so $E(f'f) = \Sigma_f$, where Σ_f is a diagonal matrix with diagonal elements

²For detailed discussion of the assumptions, see Stock and Watson (2002a, b).

$\rho_{i,i} > \rho_{j,j} > 0$, for $i < j$. Moreover, Σ_f is the probability limit of $T^{-1}(\sum_{t=1}^T f' f)$

3. The innovations, $\xi_{i,t}$ are uncorrelated with the factors f_t . Moreover, $\xi_{i,t}$ has zero unconditional mean and they are assumed to be serially correlated, so $E(\xi_t' \xi_{t+s})/N = \gamma_{N,t}(s)$ and $\sup_N \sum_{s=-\infty}^{\infty} |\gamma_{N,t}(s)|$ has finite limit as $N \rightarrow \infty$. Cross-sectional correlations in innovations are also allowed, so $E(\xi_{i,t} \xi_{j,t}) = \tau_{ij,t}$ and $\sup_t N^{-1} \sum_{i=1}^N \sum_{j=1}^N |\tau_{ij,t}|$ has finite limit as $N \rightarrow \infty$. The size of fourth moment is limited as well, so $\sup_{t,s} N^{-1} \sum_{i=1}^N \sum_{j=1}^N |cov(\xi_{i,s} \xi_{i,t}, \xi_{j,s} \xi_{j,t})|$ has finite limit as $N \rightarrow \infty$.

Cipollini and Kapetanios (2005) advocate the use of principal components method to estimate the factors from their extended stochastic volatility factor model as in equation (9). They point out that as long as the assumptions on the innovations and the factors are valid for $\omega_{i,t}$ and f_t in equation (9), then this estimation method can provide consistent estimates. Estimation of factors becomes very straightforward in this context. We can take the following four-step approach in estimating the common factors. Suppose there are N constituent stocks in a particular stock index, then for stock i

1. First compute the daily return on this stock, demean this return series and denote it as $y_{i,t}$
2. Transform the return series by standard logarithm transformation and name it as $y_{i,t}^*$ i.e. $y_{i,t}^* = \ln(y_{i,t}^2)$. This is the volatility proxy.

3. Demean the transformed returns, denote it as $\tilde{y}_{i,t}$. The stochastic volatility factor model in equation (9) can then be written as

$$\tilde{y}_{i,t} = a_i + \theta'_i f_t + \omega_{i,t} \quad (12)$$

4. Apply principal components method to matrix $\tilde{Y}\tilde{Y}'$, that is, to extract the K largest eigenvectors from matrix $\tilde{Y}\tilde{Y}'$. Where \tilde{Y} is the $T \times N$ matrix of demeaned and transformed constituent stock returns, which is thus a matrix of stochastic volatilities. The element in the t th row and i th column of matrix \tilde{Y} is $\tilde{y}_{i,t}$. The extracted K largest eigenvectors are the K estimated factor series. Grouping these estimated series together gives us a $T \times K$ matrix of factor estimates F , with the k th factor series in its k th column, $f_{k,t} = (f_{k,1}, \dots, f_{k,T})'$.

2.4 Investigation of Long Memory in Common Factors

A large amount of existing empirical studies show that financial time series exhibit persistence of shocks and slow hyperbolic rate of decay in autocorrelations. For example the paper by Baillie, Bollerslev, and Mikkelsen (1996) on daily nominal percentage returns on Deutschmark-US dollar exchange rate in the framework of FIGARCH model; and Davidson's (2004) study on dollar exchange rates for Korean Won, Indonesian Rupiah and Taiwan Dollar. Wright (1999) use log-periodogram regression to find evidence for positive long memory in some emerging market stock returns.

We expect the factor estimates of our Asian constituent stock volatilities have long memory. Our investigation starts off by examining the autocorrelations of our factor estimates and the first difference of them. We then move on to specify our factor estimates in an Autoregressive Fractionally Integrated Moving Average (ARFIMA(p,d,q)) representation. Hosking (1981), Granger (1980, 1981), and Granger and Joyeux(1980) propose fractional differencing i.e. fractionally integrated I(d) time series. An I(d) time series shows long memory and thus have slowly decaying impulse-response weight. For simplicity, the moving average terms are dropped out in our analysis. An ARFI(p,d) process of the estimated factor series, $f_{i,t}$ takes the following form³

$$\Phi(L)(1 - L)^d(f_{i,t} - \mu) = e_t \quad (13)$$

where $\Phi(L) = 1 - \sum_{i=1}^p \phi_i L^i$ are polynomial of lag operator of order p and the roots lie outside the unit circle. e_t is a white noise process with variance σ^2 . d is the fractional differencing parameter. For $-0.5 < d < 0.5$, the process is covariance stationary. f_t exhibits long memory and its autocorrelations have a hyperbolic decay for $0 < d < 0.5$. However, if $-0.5 < d < 0$, f_t is said to have immediate memory. The fractional difference is defined as

$$(1 - L)^d = \sum_{j=0}^{\infty} \pi_j(d) L^j \quad (14)$$

³Our analysis here concern modelling each factor series using univariate *ARFIMA*(p, d, q) model. However, future work on using more than one factor series for the model may consider a multivariate *ARFIMA*(p, d, q) specification for the factors.

and

$$\pi_j(d) = (-1)^k \frac{\Gamma(d+1)}{\Gamma(j+1)\Gamma(d-j+1)} \quad (15)$$

where $\Gamma(\bullet)$ denotes the gamma function. Conventional methods of estimating ARFIMA(p,d,q) process rely on maximum likelihood estimation. Whittle's (1951) Approximate Frequency Domain Maximum Likelihood Estimator (AFDMLE) is in the context of frequency domain. In the context of time domain, ARFIMA model can be estimated by Exact Maximum Likelihood Estimation (EML) (see Sowell (1992)), Modified Profile Likelihood (MPL) (see Cox and Reid (1987)), or Approximate Maximum Likelihood Estimator (AMLE) of Beran (1995). In this paper, we adopt Beran's (1995) method. This estimation method bases on minimising the sum of squared naive residuals (see also the Conditional Sum of Square estimator (CSS) of Chung and Baillie (1993), and Beveridge and Oickle (1993)). One of the remarkable advantages of this estimation method is that it is applicable for non-stationary ARFIMA process with $d > 0.5$, without the need of prior differencing⁴. Since our data have already been demeaned, we can drop the term μ in the above ARFI(p,d) process. The

⁴EML and MPL require imposing $-1 < d < 0.5$. Moreover, AMLE (or CSS) is easier to extend and thus can be used to estimate *ARFIMA* process with conditional heteroskedasticity (See Chung (1998)). CSS is preferred to the other methods due to its computational efficiency.

AMLE method of Beran minimizes the sum of squared residuals in equation (13). That is,

$$S(\beta) = \sum_{t=2}^n e_t^2 \quad (16)$$

where $\beta = (d, \phi_1, \phi_2, \dots, \phi_p)$ is a vector of estimated parameters. We evaluate the value of our estimated fractional differencing parameters to see if our factor estimates of Asian stock volatilities have long memory.⁵

Furthermore, heteroscedasticity is another common feature that is found in the financial time series. Baillie, Bollerslev and Mikkelsen (1996) incorporate long memory fractional differencing into GARCH and propose the FIGARCH model, with the assumption that the fractionally differencing parameter lies between 0 and 1. We believe our findings about dynamics of common factors of Asian stock volatilities are similar to those in the FIGARCH literatures.

The next section presents our empirical results. We start our analysis off by analysing the correlations among the constituent stock volatility series in Singaporean, Hong Kong, South Korean and Japanese indexes. We then use

⁵We use here the value of fractional differencing parameter, d , estimated by Beran's (1995) AMLE and the pattern of autocorrelation plot as indication of whether a factor exhibits long memory. Some tests for long memory, for example the Lobato and Robinson (1998) test, can also be implemented for the detection of persistence. See also Robinson (1995) for his proposal of a log-periodogram semiparametric estimator for d , which can be used to estimate d in the cases of long memory (when d lies between 0 and 0.5), short memory (when d equals 0) or negative memory (when d lies between -0.5 and 0) under the condition that standard normal approximation holds.

principal components to estimate the common factors from those datasets of stock volatilities. Next, we move on to look at the explanatory power of the factor estimates and followed by an examination of the dynamics of those factor estimates. Autoregressive (AR) representation is to be fitted to the dominant factor of each dataset. Autocorrelations of the factor estimates are to be checked carefully to see if there is any evidence of long memory nature. Finally, ARFIMA process is fitted to the dominant factor series that shows persistence in autocorrelations in order to confirm their long memory nature. Dominant factors of Korean and Japanese stock volatilities are found to have autocorrelations that show slow hyperbolic decay and evidence of long memory. Whereas factor estimates of Singaporean and Hong Kong shows low explanatory power and thus lead us to conclude that the stochastic volatility factor specification being an inappropriate underlying model of them.

3 Empirical Analysis

The five Asian indexes considered for our empirical analysis in chapter 1 are reconsidered here. Daily observations of constituent stocks of five Asian indices are obtained from Datastream. The five indexes are NIKKEI 225 (NIK225) and NIKKEI 500 (NIK500) of Japan; Heng Seng Composite Index (HSCI) of Hong Kong; Korean Stock Exchange Composite Index 200 (KOS200) of South Korea; and Stock Exchange of Singapore All Share Index (SING) of Singapore. The reason for investigating both the NIKKEI 225 and NIKKEI 500 of Japan is for us to look at the impact of the size of dataset on the explanatory power of a dominant factor. Daily returns on each constituent stocks i is calculated as the

difference between the logarithm of constituent stock price at time t and time $t - 1$. Daily volatility is computed as the logarithm of squared daily return.

As in chapter 1, our analysis considers only stocks that have data available throughout the entire sample period. The number of stocks in each datasets are thus 217 stocks for NIK225, 481 stocks for NIK500, 176 stocks for KOS200, 227 stocks for SING and 161 stocks for HSCI. We look at the same sample period as in chapter 1, which is from 3 January 2000 to 30 July 2004, for a total of 1194 daily observations of stock volatility. Once again, we exclude the periods when the markets are closed from the dataset, the number of observations then becomes 1128 for both NIK225 and NIK500, 1121 for KOS200, 1185 for SING and 1128 for HSCI. For the preliminary statistical analysis of the daily return, one can refer to table 1 of chapter 1. In the next subsection, we present a correlation analysis of the constituent stock volatilities of our five Asian indexes.

3.1 Correlation Analysis

As a first step of our analysis in this chapter, we carry out a correlation analysis of our datasets. In contrast to the correlation analysis carried out in chapter 1 in which datasets of stock returns are concerned, here we look at the correlation among the series in our five datasets of stock volatilities. Once again, since the number of series in each dataset is too large for us to report the correlation matrix, we compute average correlations in the same manner as in chapter 1. For each stock i in a dataset, we compute the correlations of its volatility with the rest of the $N - 1$ volatility series in the dataset to get a vector of correlations

with dimensions equal $(N - 1) \times 1$. Then we take the average of these $N - 1$ volatility correlation coefficients to get an average correlation between volatility of stock i and the volatility of the remaining stocks in the system. We do it for $i = 1, \dots, N$ to obtain a $N \times 1$ vector of average correlations. Table 1 shows some summary statistics of the average correlations of the Asian daily returns.

We can see from table 1 that volatility series in KOS200 and NIK225 appear to be the most correlated on average when comparing with the other three datasets, the former has mean average correlation of about 0.09 whereas the latter has mean average correlation of about 0.07. NIK500 volatility series are less correlated than the NIK225 volatility series on average, and its mean average correlation is 0.05. Concerning the fact that NIK500 index contain a wider range of stocks for which some them are relatively less related, this finding is reasonable. There is a weak correlation among the SING volatility series on average, with mean average correlations of 0.05 only. Based on these results, we expect the first factor extracted from the South Korean stock volatilities will be the most powerful one in terms of the amount of variance in the dataset being captured, but the one from the Singaporean stock volatilities will be the least powerful. From the maximum and minimum values of the average correlations in the five datasets, we can see that the volatilities in KOS200, NIK225 and HSCI are all positively correlated on average. However, evidence of negative correlation among volatility series is found in NIK500 and SING datasets as the minimum values of their average correlations are negative.

The HSCI and KOS200 stock volatilities have the largest values of standard deviation in their average correlations, both have the values around 0.024. The

smallest value of standard deviation of average correlations appear to be the NIK500 stock volatilities. It means there are more observations of average correlation that have deviated from the mean average correlations in general among the Hong Kong and South Korean stock volatilities. This indicates the level of correlations among stock volatility series varies quite a lot. However, correlations among the NIKKEI 500 stock volatility series do not vary that much in general.

The results obtained here is quite different from the results obtained from the correlation analysis of stock returns in chapter 1. We can see that the summary statistics on average correlations we present in table 1 of this chapter are much smaller in value than those presented in table 2 of chapter 1, except statistics on standard deviation. These results simply reveal the fact that there are more correlations among stock return series than among the volatility series. It is because individual volatility series has more variation than individual return series, and thus it is less correlated with other volatility series in the dataset which also exhibit large amount of variation.

3.2 Explanatory Power of Factor Estimates

By implementing the four-step approach described in section 2.3, the first 10 factors are extracted from the dataset of constituent stock volatilities of each of the five indexes.

Table 2 shows the cumulative R^2 of these factors.⁶ Some clarifications need

⁶As we will see later the cumulative R^2 , computed via Factor Augmented Regression (FAR), is used here to give us an indication of the explanatory power of factor estimates in terms of average goodness of fit. However, if we use some other procedures, then K , the

to be made before we move on to the interpretation of these statistics. In contrast to the statistics shown in table 3 of chapter 1, the cumulative R^2 here is computed from the factor model. They are the average coefficients of determination when K factors are included in the factor model. Factor Augmented Regression (FAR) is formed for the computation of these statistics.⁷ That is, for each stock, we regress its volatility proxy on the factor estimates. We do this cumulatively for $k = 1, \dots, K$. We then calculate the average R^2 across all \tilde{y}_t and report it cumulatively for the first K factors and we set $K = 10$ here. These statistics shows the average goodness of fit of the model when K factors are used. Whereas in table 3 of chapter 1, each value of the cumulative explained variation is computed by adding up the value $\frac{\lambda_k}{N}$ for $k = 1, \dots, K$, where λ_k is the eigenvalue corresponds to the k th eigenvector. If all N extracted principal components are considered, then $\sum_{k=1}^N \frac{\lambda_k}{N} = 1$. Therefore, the difference between each value with its preceding value in that table gives the amount of total variation explained by that principal component, i.e. $\frac{\lambda_k}{N}$. And $\frac{\lambda_k}{N}$ shows the relative importance of the k th principal component. Although the statistics in these two tables are computed in different ways, they are very heavily related due the fact that the conventional principal component analysis carried out in chapter 1 and the principal components method of Stock and Watson (2002a) applied in this chapter are statistical equivalent but the latter is more robust.

optimal number of factors to be included in the factor models for our dataset can be found. For example, Bai and Ng (2002) develop a modification of standard Akaike Information Criteria as a decision rule of finding an appropriate number of factors. Whereas Kapetanios (2004) proposes a new method to estimate the number of factors in a factor model by adding factors to a series until no neglected factor structure is detected in the residuals from the factor analysis. His method shows better performance than the Bai and Ng (2002) approach via Monte Carlo studies.

⁷For details of FAR, see equation (16) of Chapter 3.

So, comparison between the two sets of statistics can still be made for this reason.

We can see from the table that for KOS200, NIK225 and NIK500, the first factors have the highest values of cumulative R^2 . The first factor of KOS200 is the most powerful one among the five, 10.1% of variation for a total number of 176 stocks can be explained when it is included in the factor model. When the first factor of NIK225 is used in the factor model, about 8.1% of the variation can be explained for a total of 217 stocks. While for the first factor of NIK500, about 6.1% of the variation can be explained in the dataset of 481 stocks. Although they do not seem as powerful as the first factor of KOS200, they are quite good in general. We can then conclude that the first factor of KOS200, and that of NIK225 and NIK500 are the dominant factors for these datasets. Considering the first 10 factors from these three datasets, we can see additional factors only have marginal contribution to the explanatory power of the set of factors. These findings are quite similar to the results of the explanatory power of principal components we obtained in chapter 1. We also found large proportion of total variation being explained by the first principal components from the constituent stock returns of these three indexes.

However, the same story does not apply to the factors of SING and HSCI stock volatilities. The first factor of SING stock volatilities can only explain 5.3% of the variation for a dataset of 227 stocks when it is included in the factor model. This amount is quite low. An interesting finding appear to be the low explanatory power of the first factor from the HSCI dataset of volatilities as the model has the value of R^2 equals to 1.7% only when it is included in the

factor model. However, we can see from table 2 that when its first three factors are used for modelling, the cumulative R^2 increases a lot, with the third factor gives the largest contribution to the average fit. These results imply that for these two datasets, modelling the volatilities with only the first factor does not provide a good fit on average.

Notice that although we find high average correlations in the HSCI volatilities, but this does not mean modelling with only the first factor will result in good fit of the factor model. High average correlations does not mean good fit of the factor model. It only suggests possible factor structure in the dataset and the first factor embeds the majority of the variance. It is because according to the definition of principal components, the first factor is the linear combination that has maximum variance subject to a normalisation. Our finding shows the factor model may not be an appropriate specification for both HSCI and SING in our case.

Moreover, high average correlation in a dataset suggests possible factor structure. This applies to our analysis no matter it is carry out via conventional principal component analysis as in chapter 1 or via the Stock and Watson (2002a) method as in chapter 2. High correlation implies powerful principal components in terms of accounting for variation. We expect the proportions of total variation explained by the principal components are in descending order, meaning the first principal component explains the most due to the fact that majority of the variance from the dataset is in the first factor. This is proved in the analysis in chapter 1.

In addition, attention should be paid to the amount of explained variation

for the first factor of the NIK225 dataset and that of the NIK500 dataset. We can see that having more stocks in the dataset does not strengthen, but weaken the explanatory power of the dominant factor. Boivin and Ng (2003) explain why using more series to estimate factors may not be desirable in factor analysis. Their studies show the resulting factor estimates from a dataset with more series added to it are only useful if the errors of factor estimates are *i.i.d.* It is because factor estimates may appear to have innovations that are heteroscedastic and cross-sectionally correlated. If more series from the same category are added to the dataset, average size of their common component will become smaller. Number of correlated innovations will increase as more series from the same category are included. As a result, the correlations among innovations may be too large for the factor estimates to remain consistent, making larger dataset not advantageous to factor analysis.⁸ In our case, the dominant factors of NIK225 and NIK550 are highly correlated with correlation coefficient equals 0.9527 and we have also seen that the errors of our factor estimates are serially correlated. Stocks in NIK225 are more representative than stocks in NIK500 and they are believed to have made a large contribution to the fluctuation in the Japanese market. NIK500 dataset contains both stocks in NIK225 and other stocks that are relatively less "important" in explaining the fluctuation. Including these relative less "important" stocks may generate noise and therefore, reduce the average size of common components, making the innovations more time and cross-sectionally dependent, and thus lower

⁸Principal components estimation method by Stock and Watson (2002) shows factors estimates are asymptotically consistent if innovations are stationary, factor loadings are trivial and idiosyncratic errors are weakly serial and cross-sectional correlated. See also Bai and Ng (2002) and Boivin and Ng (2003)

the explanatory power of the factor estimates. Our results here confirm the suggestion of Boivin and Ng (2003) and consistent with the findings in the Cipollini and Kapetanios (2005) paper.

Our findings here shows that factor model may not be an appropriate specification for both SING and HSCI constituent stock volatilities. Therefore, follow-on examination of long memory in factor dynamics will be centered on the dominant factors of KOS200, NIK225 and NIK500. We will also not carry out the examination of long memory on the third factor of HSCI. It is because although including it into the model gives a better fit, less variance from the dataset is embedded in the third factor than in the first factor. Since our analysis ignored the first HSCI factor, we also ignored the third HSCI factor.

3.3 Dynamics of Factor Estimates

The first factor estimates from all five datasets of stock volatilities are plotted in Figure 1. If we compare these plots with the plots of the first principal components extracted from the datasets of returns in the last chapter, it is not difficult to see that the factors from the stock volatilities exhibit larger fluctuations than the principal components extracted from the stock returns.

As a first step to understand the properties of factor estimates, the estimated first factor for each dataset is regressed on a constant term plus *i.i.d.* disturbances to form a strict white noise process. We then test for serial correlation in the residuals and squared residuals. This is the same as breaking our factor estimates into a deterministic component (a constant) and a stochastic

component (the disturbance), this allows us to carry out residual tests, which in turn give us the idea of the temporal properties of our factor estimates. Table 3 shows the computed Breusch-Godfrey serial correlation LM test statistics and Engle's ARCH LM test statistics, and the probabilities corresponding to these statistics⁹. Clear rejections of no serial correlation are found for the residuals and squared residuals for all five factors at 5%. This indicates the estimated factor series have residuals and squared residuals that are serially correlated and they are clearly not *i.i.d.*. It also suggests the model underlying those common factors should incorporate these properties.

3.3.1 Estimation of Autoregressive Models of Common Factors

In the above section, we have already seen two important dynamic properties of the factor estimates for all of our five datasets, i.e. serial correlation in residuals and squared residuals. We now take a further step to examine more deeply into their dynamic characteristics. Based on the findings in section 3.2, we center our analysis on the dominant factors of KOS200, NIK225 and NIK500 datasets. The first factor of SING and that of HSCI are also investigated for comparison purpose. AR(p) representation of the following form is fitted into each of these estimated factor series

⁹5 lags are included in the auxiliary equations of both Breusch-Godfrey LM test and Engle's ARCH LM test. Auxiliary equation of Breusch-Godfrey LM test is

$$e_t = \alpha_0 + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \alpha_3 e_{t-3} + \alpha_4 e_{t-4} + \alpha_5 e_{t-5}$$

Auxiliary equation of Engle's ARCH LM test is

$$e_t^2 = \beta_0 + \beta_1 e_{t-1}^2 + \beta_2 e_{t-2}^2 + \beta_3 e_{t-3}^2 + \beta_4 e_{t-4}^2 + \beta_5 e_{t-5}^2$$

Both LM statistics have asymptotic χ^2 distribution under null hypothesis of no serial correlation and no ARCH effect, respectively.

$$f_{k,t} = \alpha_1 f_{k,t-1} + \alpha_2 f_{k,t-2} + \dots + \alpha_p f_{k,t-p} \quad (17)$$

where $f_{k,t}$ denotes factor k at time t , and $(\alpha_1, \dots, \alpha_p)$ denote the *AR* coefficients. Table 4 reports the estimated *AR* coefficients, *t*-statistics and the corresponding probabilities for the factor estimates of the five datasets¹⁰. Number of lags included in the above *AR* representation of the first factor series for each of the five datasets are chosen according to the Akaike Information Criterion (AIC) and the Schwartz Bayesian Criterion (SBC). Besides, *AR*(p) processes with highly insignificant coefficients are avoided. These selection criteria lead us to choose *AR* representations with 7 lags for the dominant factors of KOS200, NIK225 and NIK500; but only 3 lags and 5 lags for SING and HSCI, respectively. It can be seen clearly that those factor estimates that appear to have stronger explanatory power are found to sustain an *AR* representation with longer lags. Empirical studies of financial time series have shown that volatility in stock return has extremely long memory. If a dominant factor is capable of representing the behaviour of the entire system of stock volatilities, it can also show persistence, and thus sustain a higher order *AR* process. The first factor of HSCI apparently does not persist. However, the dominant factors of KOS200, NIK225 and NIK500 seem to have obtained this characteristic. In the next section, we move on to examine the long memory nature of the factor estimates. We focus our analysis on the dominant factors of KOS200, NIK225

¹⁰The autoregression in equation (17) is known as Factor Augmented Regression (FAR). Using the statistical inference of principal component factor estimates developed by Bai and Ng (2006), hypothesis testing can be carried out. Further discussion on this issue will be presented in chapter 3 of this thesis.

and NIK500 datasets. We concentrate out the first factor of SING and that of HSCI due to the fact that they have low explanatory power since it implies that the factor models may be inappropriate specifications for them.

3.3.2 Long Memory in Common Factors

Long memory nature of the dominant factors of these three datasets is further confirmed by the persistence exists in their autocorrelations. Figures 2 to 6 graph the autocorrelations of the factor estimates up to 200 lags. Time series with long memory should have autocorrelations that are persistently significant at long lags with a hyperbolic decay. When they are differenced, they appear to have the characteristics of alternating positive and negative autocorrelations out to long lags, which indicates the series has been over-differenced. Autocorrelations of the dominant factors of KOS200, NIK225 and NIK500 shows persistence even up to 200 lags and they also exhibit a slow rate of decay. This is apparently an evidence of persistence. Autocorrelations of the first factor of HSCI however, does not show hyperbolic decay. Figures 7 to 9 plot the autocorrelations of the first differenced dominant factors of KOS200, NIK225 and NIK500 datasets. Their autocorrelations are alternating positively and negatively even up to 200 lags, meaning the estimated factor series are over-differenced.

Having observed this nature, we then move on to fit a Autoregressive Fractionally Integrated Moving Average (ARFIMA(p,d,q)) process without the moving average terms to the series of the dominant factors of KOS200, NIK225 and NIK500 only and we estimate the process by AMLE. The number of lags

included in the ARFI process as in equation (13) of a dominant factor is once again determined by the AIC, which tell us that ARFI(1,d) should be chosen for all three dominant factors. Table 5 reports the estimated parameters, the standard errors and the computed 95% confidence interval of \hat{d} for the ARFI specifications for these dominant factors. We can see from the results that the estimated \hat{d} 's lie between 0 and 0.5 for all three dominant factors. This is an indication of hyperbolic decay in their autocorrelations. Computed 95% confidence intervals of our estimated \hat{d} also lies within this range, showing evidence that the true d also lie between 0 and 0.5. Moreover, these values also lies between -0.5 and 0.5, meaning the ARFI(1,d) processes of these factor estimates are covariance stationary.

There is quite strong evidence to show that the underlying process of dominant factors of KOS200, NIK225 and NIK500 constituent stock volatilities should be the one that characterised with long memory. This finding is similar to the FIGARCH literatures. A dominant factor represents the common component of a dataset of stochastic volatilities. Long memory in this common component means persistence can be found in all constituent stock volatilities in general.

4 Concluding Remarks

In this chapter, we have applied principal components estimation to the stochastic volatility factor model to study the constituent stock volatilities of five Asian stock indexes. We have found evidence to show that the first factor of

KOSPI 200, NIKKEI 225 and NIKKEI 500 can account for a lot of variation in the dataset when they are used for modelling volatilities. But factor model may not be an appropriate specification for the HSCI and Singapore All Share Index stock volatilities. We have also provided empirical evidence to confirm that datasets with more series from the same category are not always desirable for factor analysis. Long memory characteristic is also detected in these factors. These findings provide an insight into empirical studies of common factors that contribute to Asian stock volatilities.

It is worth mentioning that some studies show some long memory in financial time series may be caused by neglected structural breaks in the series, see for example, Granger and Hyung (2004). It may thus be interesting to further investigate the presence of structural breaks in common factors.

It is obvious that the stochastic volatility factor model of Cipollini and Kapetanios (2004) is a single local-factor specification. Our analysis here also restricts to this local-factor settings. Given the empirical evidence of volatility transmission, it may be interesting to look at how common factor in stock volatilities extracted from one market impacts on stock volatilities in another market. An extension from their single local-factor model to a multi-factor model allows one to investigate on spillover effects. Forecasting exercise can also be carried out by using both a single local-factor and a multi-factor specification of stochastic volatility factor model. One should notice that this factor model is favourable for large dataset with many cross-sectional series. Further empirical in-sample and out-of-sample comparison between this factor model and other

factor volatility models that are also designed for large datasets, for example the orthogonal GARCH model discussed in chapter 1, may bring insight into appropriate model selection for large financial dataset in the context of factor analysis.

The empirical analysis in this chapter and chapter 1 is on an in-sample basis. In the next chapter, we will focus on evaluating the empirical performance of the orthogonal GARCH model and the stochastic volatility factor models in predicting volatility. We will also discuss an extension of the single local-factor model into a multi-factor model. Testing of factor significance using the asymptotic results of principal component factor estimates developed by Bai and Ng (2006) will be performed in the selection of factors in the multi-factor models.

5 Tables and Figures

	SING	HSCI	KOS200	NIK225	NIK500
Max. of average correlations	0.0904	0.1185	0.1429	0.1168	0.1037
Min. of average correlations	-0.0136	0.0048	0.0266	0.0350	-0.0267
Median of average correlations	0.0511	0.0676	0.0904	0.0717	0.0545
Mean of average correlations	0.040	0.0662	0.0896	0.0734	0.0547
S.D. of average correlations	0.0174	0.0243	0.0239	0.0171	0.0150

Table 2: Cumulative R^2 for the first 10 factor estimates of the five datasets					
No. of factors	SING	HSCI	KOS200	NIK225	NIK500
1	0.053	0.017	0.101	0.081	0.061
2	0.109	0.032	0.115	0.096	0.071
3	0.124	0.092	0.122	0.111	0.083
4	0.135	0.101	0.131	0.121	0.088
5	0.141	0.112	0.143	0.133	0.094
6	0.152	0.123	0.152	0.141	0.101
7	0.159	0.134	0.162	0.149	0.105
8	0.165	0.143	0.171	0.156	0.111
9	0.171	0.152	0.181	0.163	0.115
10	0.177	0.162	0.189	0.170	0.120

Table 3: Breusch-Godfrey LM test statistics and Engle's ARCH LM test statistics

	SING	HSCI	KOS200	NIK225	NIK500
B-G LM test	483.08 [0.000]	32.74 [0.000]	305.06 [0.000]	210.14 [0.000]	284.21 [0.000]
ARCH LM test	81.78 [0.000]	21.52 [0.001]	82.99 [0.000]	77.74 [0.000]	88.18 [0.000]

Note: Both LM statistics are asymptotically χ^2 distributed with 5 degree of freedom. Probabilities are reported in brackets.

Table 4:					
AR representations of the first factor estimates of the five datasets					
AR coeff.	SING	HSCI	KOS200	NIK225	NIK500
α_1	0.334 {11.59}	0.113 {3.786}	0.149 {4.970}	0.076 {2.561}	0.119 {3.991}
α_2	0.119 {3.914}	-0.05 {-1.683}	0.235 {7.779}	0.161 {5.369}	0.187 {6.238}
α_3	0.126 {4.164}	0.119 {4.001}	0.118 {3.804}	0.150 {4.981}	0.176 {5.798}
α_4	0.072 {2.382}		0.061 {1.959}	0.044 {1.452}	0.025 {0.840}
α_5	0.151 {5.264}		0.044 {1.446}	0.131 {4.360}	0.106 {3.478}
α_6			0.088 {2.926}	0.050 {1.677}	0.039 {1.288}
α_7			0.062 {2.071}	0.079 {2.653}	0.087 {2.926}
Max. abs. eigenvalue	0.918	0.50	0.920	0.908	0.918
Note: t-statistics are reported in curly parentheses.					

Table 5: ARFI(1,d) specifications for the dominant factors of KOSPI, NIK225 and NIK500			
	KOS200	NIK225	NIK550
$\hat{\phi}_1$	0.5177	0.3181	0.4181
\hat{d}	0.2677	0.1492	0.1831
Standard error of $\hat{\phi}_1$	0.0370	0.0499	0.0452
Standard error of \hat{d}	0.0411	0.0493	0.0489
95% C.I. of \hat{d}	[0.1871, 0.3493]	[0.0514, 0.2470]	[0.0873, 0.2789]
Standard error of residuals	0.0256	0.0295	0.0287

Figure 2: Plots of the first factors

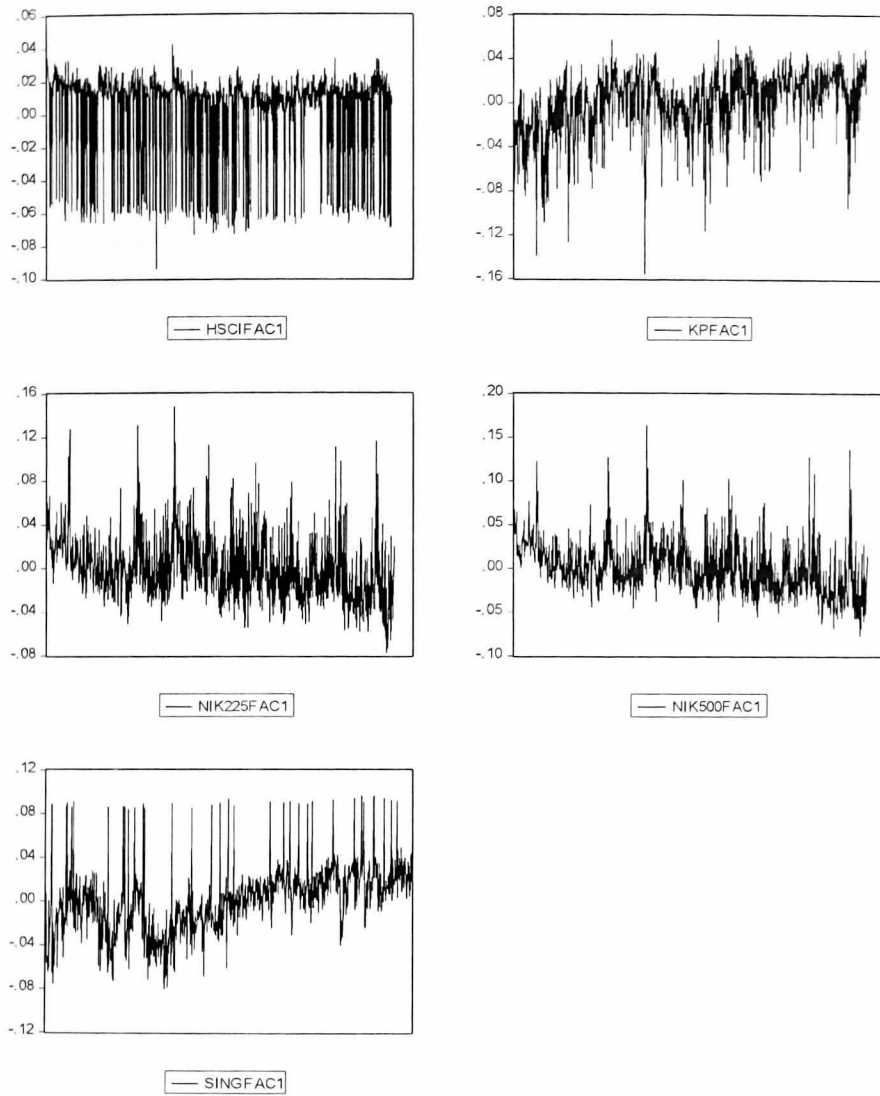


Figure 1: Plots of the first factors

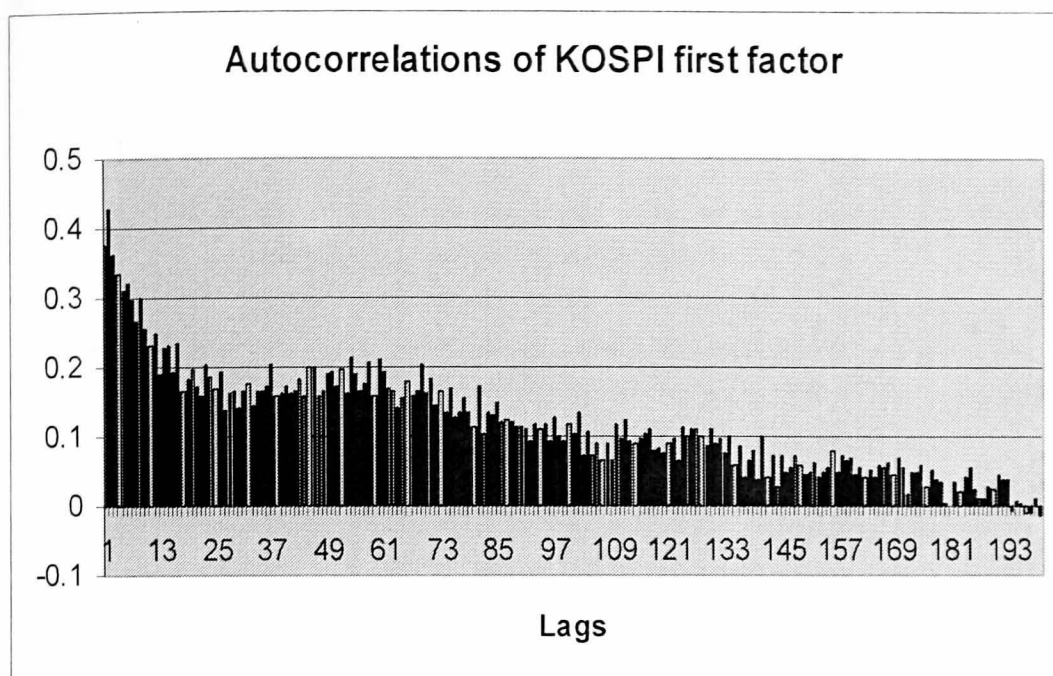


Figure 2: Autocorrelations of KOS200 first factor

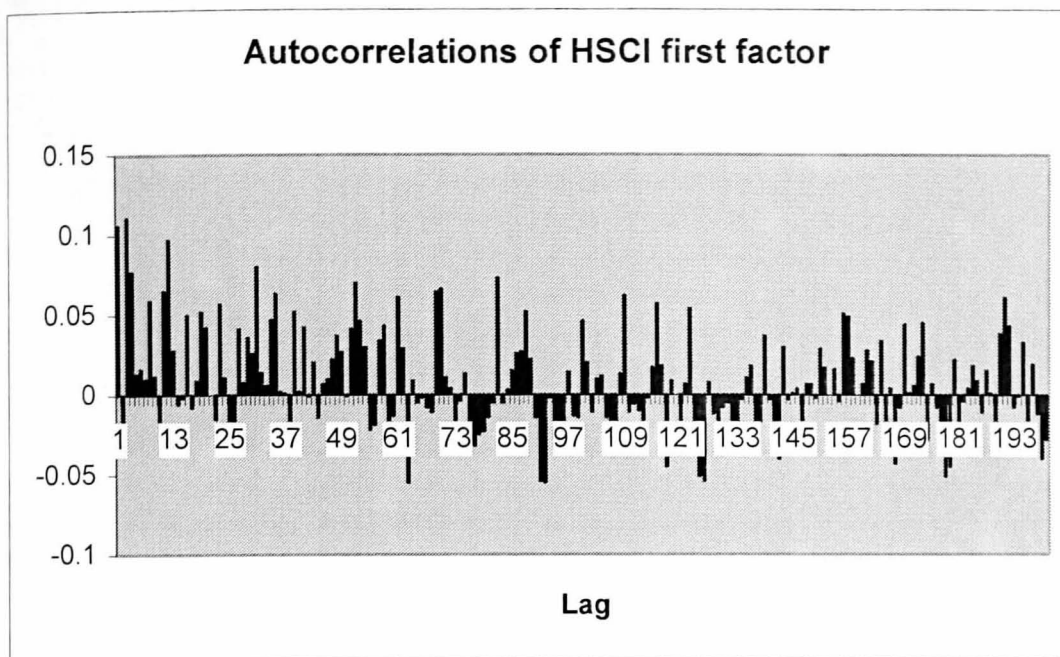


Figure 3: Autocorrelations of HSCI first factor

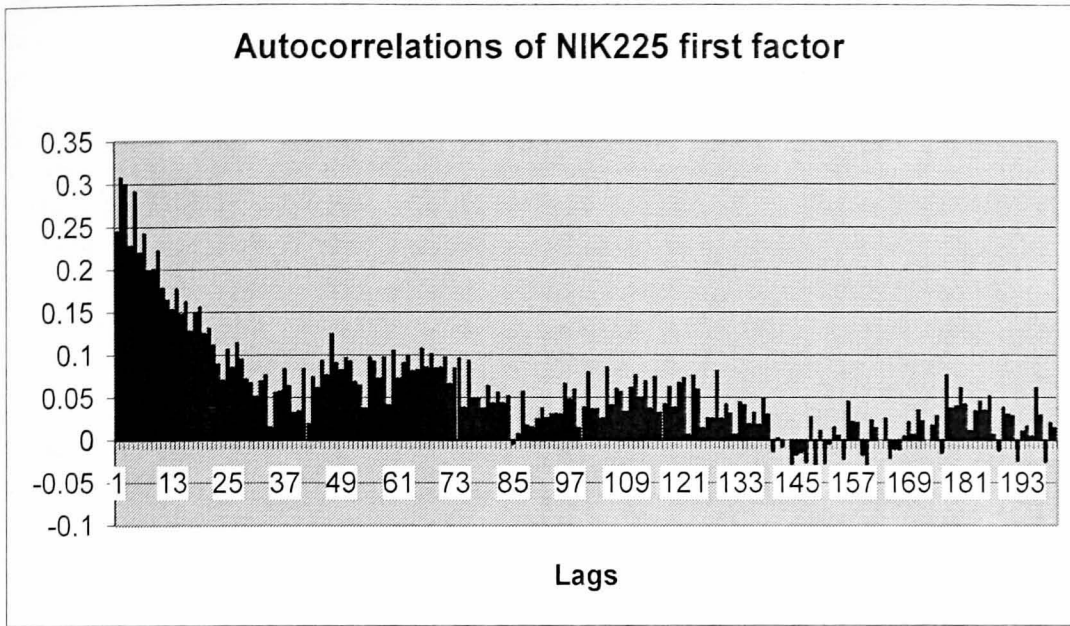


Figure 4: Autocorrelations of NIK225 first factor

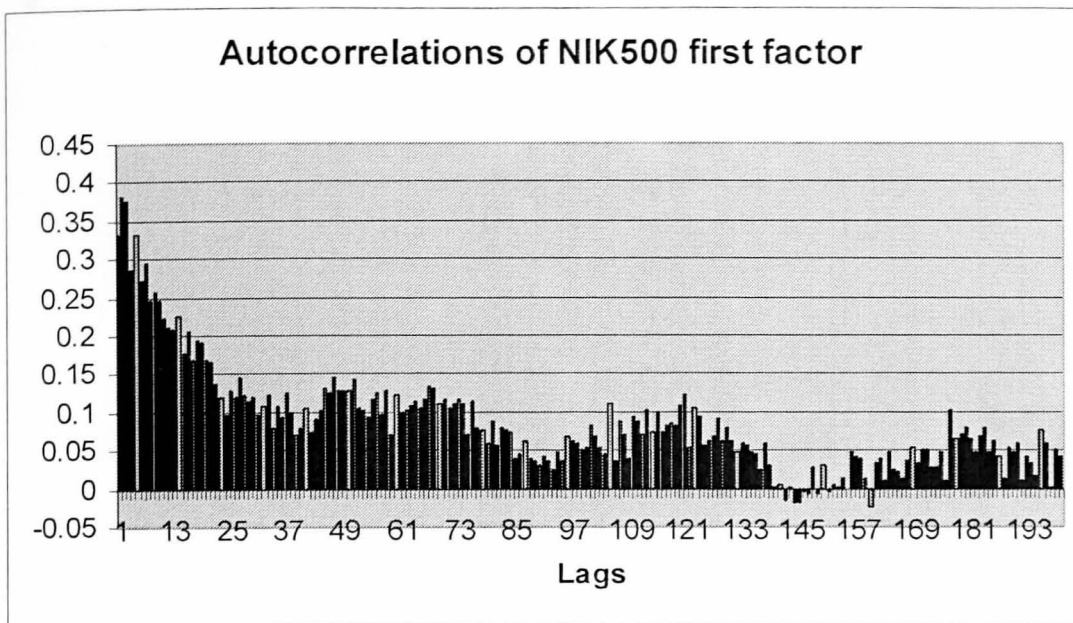


Figure 5: Autocorrelations of NIK500 first factor

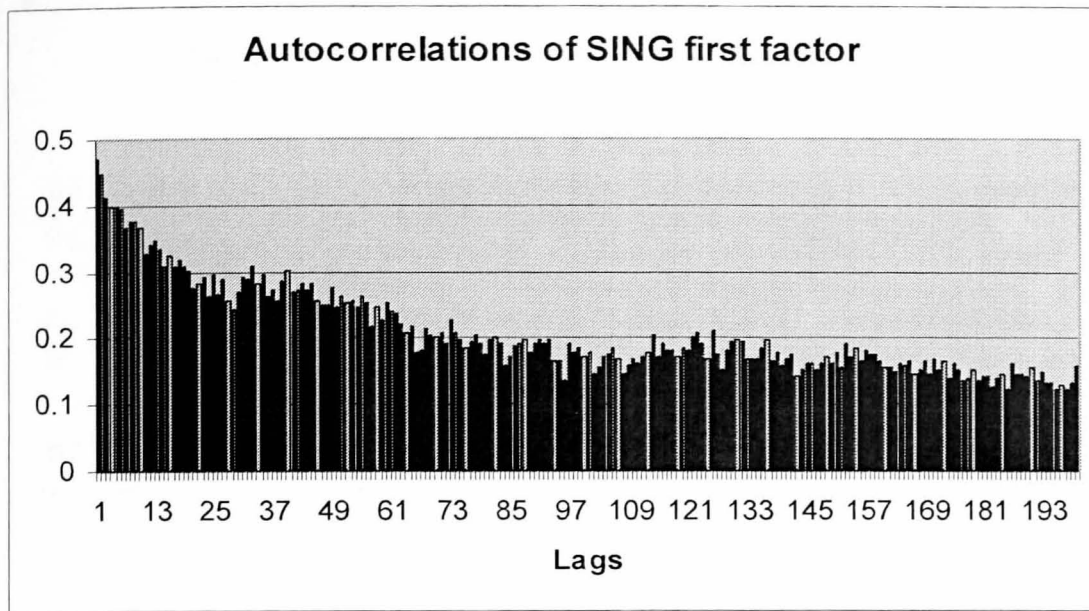


Figure 6: Autocorrelations of SING first factor

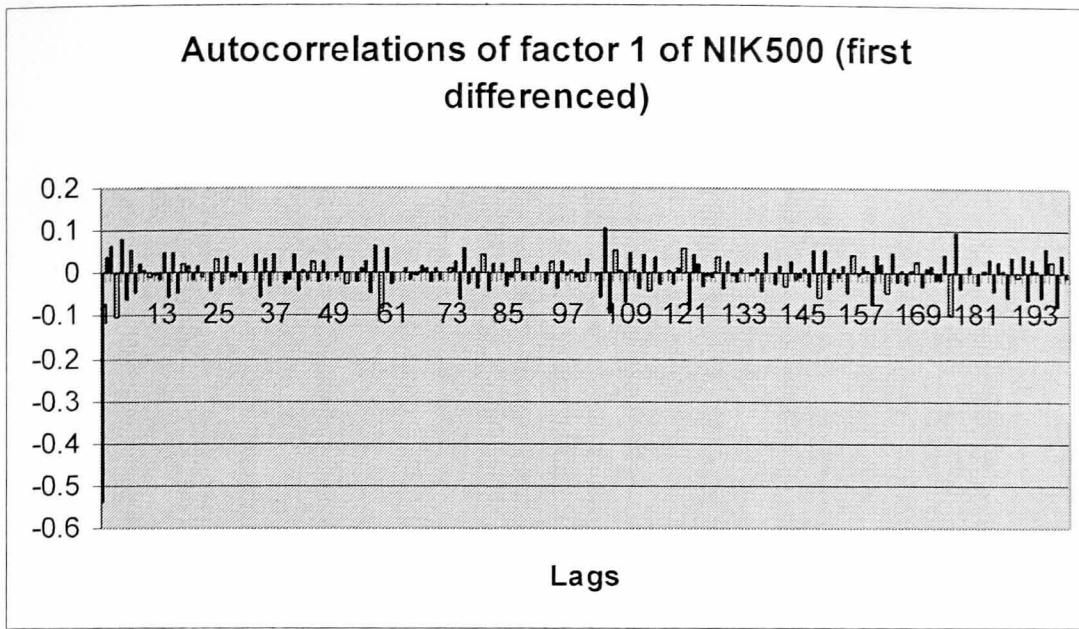


Figure 7: Autocorrelations of 1st differenced NIK500 first factor

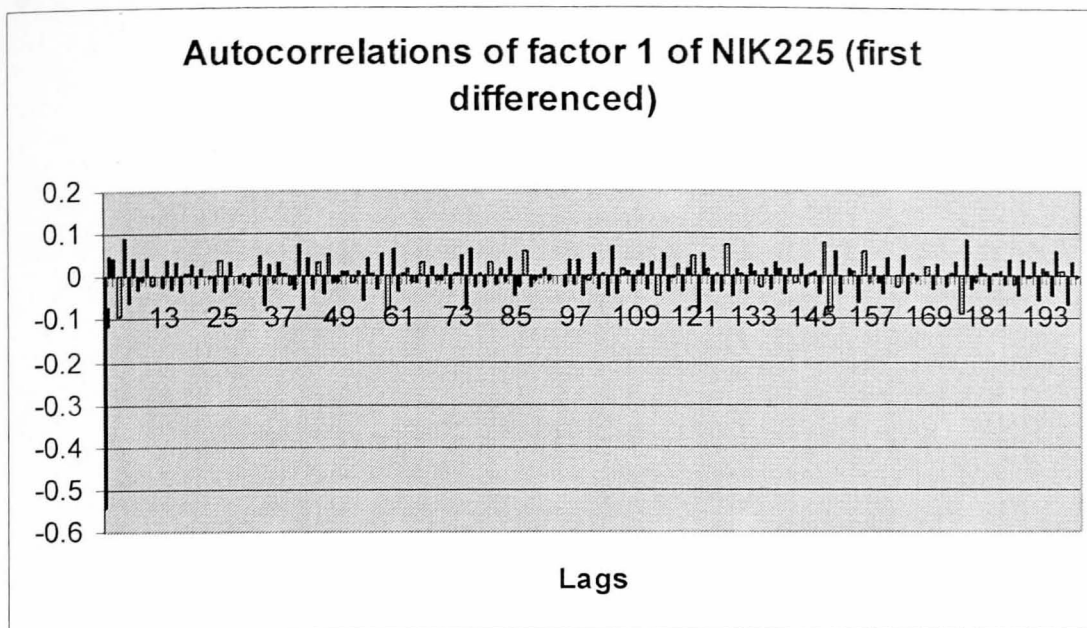


Figure 8: Autocorrelation of 1st differenced NIK225 first factor

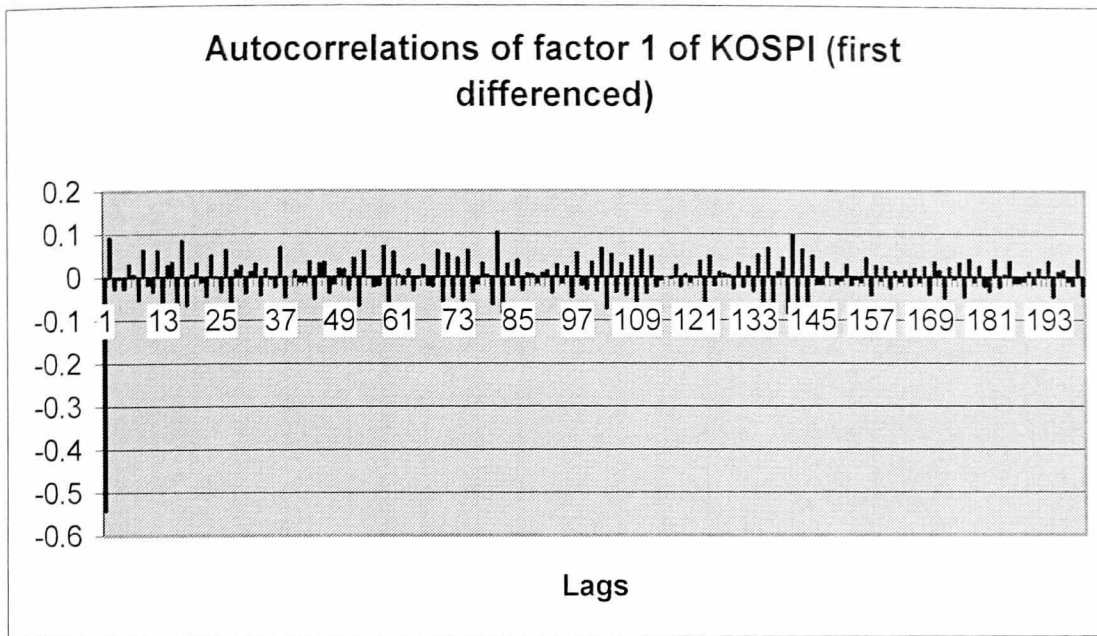


Figure 9: Autocorrelations of 1st differenced KOS200 first factor

Chapter 3

Forecasting Stock Volatility: Stochastic Volatility Factor Models versus Orthogonal GARCH Model

1 Introduction

Forecasting volatility is a crucial exercise in the study of financial and macroeconomic time series. Various volatility models have been proposed and the topic has attracted a lot of attentions from not only academics, policy makers and financial researchers, but also international investors. According to a survey paper by Poon and Granger (2003), at the time of the production of their survey, at least 93 papers regarding volatility forecasting have been published. this number does not include the number of working papers that have already been produced at that time. We can also see an extensive amount of research in volatility forecasting continue to be carried out after then.

As aforementioned in the previous chapters, volatility models can be classified into two main categories: the stochastic volatility models and the GARCH family models. The two types of models differ from each other in different aspects. First of all, in terms of specifying the conditional volatility, time varying volatility is captured by allowing the conditional volatility to be a function of past squared unexpected returns and past variances in GARCH models. Conditional volatility is driven by the same shocks as its conditional mean. It evolves over time via an autoregressive moving average process with the autoregressive terms being the past volatility and the moving average term being the past squared shocks to the time series. In stochastic volatility modelling, an unobserved variance component is included in the model and a linear stochastic process is used for modelling of this unobserved variance component. Current volatility is subject to an additional contemporaneous shock, which has an impact on the logarithm of current volatility.

Secondly, comparing the univariate GARCH(1,1) model with the univariate stochastic volatility model, the former has been criticised for not capable of capturing excess kurtosis in financial time series completely. Finally, maximum likelihood estimation of univariate GARCH model is straight forward. Conditional distribution of the shock on past history can be derived easily in univariate GARCH model. Therefore, the likelihood function can be constructed with ease. However, estimation of stochastic volatility model is less straight forward. It is because the conditional distribution of the shock on past history can not be described explicitly, due to the fact that the conditional volatility is unobserved and past information on the time series cannot help

in the derivation. Therefore, making standard maximum likelihood estimation not favourable for the estimation of stochastic volatility model. Various estimation procedures have thus been developed. These procedures have been outlined in chapter 2.

Further developments of the two types of models involve more complex specifications of the conditional variance and also an extension of the models into multivariate settings. However, in facing with large datasets, the conventional multivariate versions of both models encounter several constraints. In the past two chapters, we have already looked at the advanced generalisations of these two types of models – the orthogonal GARCH model of Alexander (2001a, b) and the stochastic volatility factor model of Cipollini and Kapetanios (2005). They are designed to allow for effective estimation and computation when working with large datasets and at the same time resolve constraints faced by conventional multivariate GARCH and stochastic volatility models. In the next subsection, we provide a comparison of the two factor volatility models in several aspects.

1.1 Orthogonal GARCH Model and Stochastic Volatility Factor Model – A Comparison

A remarkable attribute of both the orthogonal GARCH and the stochastic volatility factor models is the ease of factors or principal components estimation. Principal component analysis is the basic framework of both models. Comovements within a large dataset are summarised by a few principal components (in orthogonal GARCH) or common factors (in stochastic volatility factor

model) and thus reduction in dimensionality can be easily achieved. Both models are ideal for applications to and analysis with large datasets that contain a lot of cross-sectional time series. However, although the two models are both ideal for large datasets and involve the adoption of principal components. They are clearly different in various aspects.

First of all, in terms of the estimation of principal components, orthogonal GARCH model is constructed by extracting principal components from a system of financial time series. In our analysis, this is a system of stock return series. Common variation in the dataset is thus represented by a few principal components. In contrast to the orthogonal GARCH model, the stochastic volatility model is constructed by applying principal components to a dataset of financial volatilities. In our analysis, this is a system of stock volatilities. Common factors are estimated via principal components method of Stock and Watson (2002a).

Secondly, in terms of modelling volatility, an univariate GARCH(1,1) is the underlying process of the conditional volatility of each of the extracted principal components in orthogonal GARCH model. Volatility and covariances of the series in the original system of financial time series are approximated by the variance-covariance matrix of the chosen principal components multiplied by the squared of the factor weights. Accuracy of this approximation depends on the level of correlations among the data series. Whereas in stochastic volatility factor model, common factors summarise comovements in the system of volatilities and are modelled by long memory processes.

Finally, the models are different in terms of accounting for idiosyncratic

characteristics of the individual financial time series in a dataset. In orthogonal GARCH modelling, although the squared of the idiosyncratic part of each chosen principal components in its conditional mean goes into its conditional variance equation in the univariate GARCH(1,1). The idiosyncratic part of the original system which is not approximated by the chosen principal components are left out of the model. This has made the performance of orthogonal GARCH model highly depends on how good the principal components are in capturing variation in the dataset. In the case of weakly correlated system, a large number of principal components will be needed for accounting a large amount of total variation. However, using a large number of principal components is not desirable. It is because this will result in unstable correlation estimates and the volatility estimates will also be less robust since extraneous noise in the less important principal components gets into the model. As a result, causing problem of misspecification as we have already seen from the empirical analysis in chapter 1. When a small number of principal components are used, the volatility approximation becomes weak. As a result, it will weaken the power of the orthogonal GARCH model.

In contrast, stochastic volatility factor model does not have this shortcoming. It is because the idiosyncratic part of individual stochastic volatility that is not captured or approximated by the common component is modelled by a state-space representation. So even though the series in a dataset is not very highly correlated and thus not a lot of common variation can be captured by the factors included in the model, the "larger" idiosyncratic part will still be taken into account for modelling. When the stochastic volatility factor model is used

for forecasting volatility, idiosyncratic volatility forecast has also made contribution to the formation of an overall volatility forecast for individual financial asset in the dataset. This ability of capturing idiosyncratic characteristics has made the stochastic volatility factor model relatively flexible.

The aim of this chapter is to provide an evaluation of the volatility forecasting performance of the orthogonal GARCH model and the stochastic volatility factor models. In the consideration of volatility transmission, we also propose an extension of the stochastic volatility factor model to several multi-factor specifications which allow overseas factors to play roles in the determination of local market stock volatilities. We also recommend a procedure in selecting an appropriate multi-factor specification for local stock volatilities.

This chapter is organised as follow. Section 2 provides a discussion of the orthogonal GARCH model and the stochastic volatility factor model. We propose an extension of the single local-factor stochastic volatility model of Cipollini and Kapetanios (2005) into a multi-factor specification to allow for information transmission of common component in stochastic volatility among stock markets. Various multi-factor specifications are presented. We also discussed an application of the asymptotic results developed by Bai and Ng (2006) for testing factor significance. This can provide an indication of the choice of an appropriate multi-factor specification for a local market. Forecasting methodology is discussed in section 3. Section 4 provides an outline of the statistics adopted in our evaluation of out-of-sample forecasting performance of the factor volatility models. Empirical results are analysed in section 5. We examine both the in-sample and out-of-sample performance of the factor volatility models.

Section 6 concludes.

2 The Models

This section provides a review of the forecasting models we use for forecasting Asian stock volatilities. As well as revising the stochastic volatility factor model, we will look at the specification of the idiosyncratic part of individual stock volatility in the stochastic volatility factor model as suggested by Cipollini and Kapetanios (2005). To further develop the factor model, we will propose an extension of the original stochastic volatility factor model that contains only a local factor into a multi-factor specification that is believed to be more applicable to empirical situation. Moreover, we also consider how to determine which factors should enter into the multi-factor specification for a particular market.

2.1 Stochastic Volatility Factor Models

2.1.1 The Single Local-Factor Specification – A Revision

In the paper by Harvey, Ruiz and Shephard (1994), they propose a multivariate stochastic volatility model. This model is a better alternative to the multivariate GARCH models for modelling financial time series featured with time-varying volatility. Consider there are N stocks in a dataset. Multivariate stochastic volatility model of daily stock returns takes the following form

$$y_{i,t} = u_{i,t}(\exp(h_{i,t}))^{1/2} \quad (1)$$

where $y_{i,t}$ denotes the daily return of stock i at time t , and $i = 1, \dots, N$ and $t = 1, \dots, T$. So, y_t is a N -dimensional vector of stock return, i.e. $y_t = (y_{1,t}, \dots, y_{N,t})'$. $u_t = (u_{1,t}, \dots, u_{N,t})'$ is a multivariate normal vector of disturbances with zero mean and variance-covariance matrix Σ . Σ has diagonal elements of ones and off-diagonal elements $\rho_{i,t}$. Applying logarithmic transformation to equation (1), it becomes

$$\tilde{y}_{i,t} = a_i + h_{i,t} + \zeta_{i,t} \quad (2)$$

where $\tilde{y}_{i,t}$ denotes $\ln(y_{i,t}^2)$, a_i denotes $E(\ln(u_{i,t}^2))$, $\zeta_{i,t}$ denotes $\ln(u_{i,t}^2) - E(\ln(u_{i,t}^2))$. Harvey et. al. (1994) make two suggestions to model the logarithm of the unobserved variance, $h_{i,t}$. They suggest that it can either be modelled by a multivariate random walk, or to incorporate common factors into the stochastic volatility model, i.e. the $N \times 1$ vector of unobserved variance, $h_t = \theta f_t$, where f_t is a K -dimensional vector of common factor, $f_t = (f_{1,t}, \dots, f_{K,t})'$. They estimate this model using Kalman filter and assume normality for u_t . However, both suggestions have limitations on the nature of the underlying processes of the unobserved variance and the common factors. Thus, if common factors have more complex temporal features, their model will become insufficient. Moreover, estimation via Kalman filter is not ideal when dealing with datasets with very large dimensions.

Cipollini and Kapetanios (2005) propose a generalisation to the stochastic volatility model, the stochastic volatility factor model, that improves the shortcomings of the state-space version mentioned above. In their factor model, unobserved variance h_t features a common component and a disturbance that

is unique to each individual stock. That is,

$$h_{i,t} = \theta'_i f_t + \eta_{i,t} \quad (3)$$

substituting equation (3) into (2) gives us the stochastic volatility factor model

$$\tilde{y}_{i,t} = a_i + \theta'_i f_t + \omega_{i,t} \quad (4)$$

where $\omega_{i,t} = \eta_{i,t} + \zeta_{i,t}$ is the idiosyncratic volatility of stock i . It can be seen from equation (4) that the logarithmic of squared daily return is used as a proxy for daily volatility. Principal components of Stock and Watson (2002a) is suggested for factor estimation of this model. Moreover, we can also see from equation (4) that stochastic volatility of individual stock is made up of two components, a common component $\theta'_i f_t$, and an idiosyncratic component, $\omega_{i,t}$.

Common component constitutes the part of stock volatility that has been caused by some common driving forces which induce fluctuations in returns on all stocks in the dataset. In other words, it is the contribution to individual stock volatility made by market fluctuations. The idiosyncratic component, on the other hand, summarises the part of stochastic volatility that is mainly caused by some other factors that is unique to individual stocks. Comparing the original stochastic volatility model with the stochastic volatility factor model, it is not difficult to see that the latter is far more flexible than the former as it allows both the factor estimates and the idiosyncratic shocks to be driven by more complicated underlying processes that best describe their

empirical temporal properties. Thus, the dynamics of financial volatility can be best captured. Moreover, no extra assumptions, in addition to those required by principal component estimation, are required for the factor model to provide consistent factor estimates. All these advantages make the stochastic volatility factor model easy to use, more applicable to empirical situation and very tractable.

It is worth mentioning that in the studies of the Standard and Poor stock volatilities by Cipollini and Kapetanios (2005), the common factor in their factor model is the dominant factor of the local US market. Being the world leading stock market, it is not surprising to find that the US local dominant factor contains sufficient information to provide a good forecast for stock volatilities in its own market. However, when our interest is to concern stock markets around the world, whether a single local dominant factor contains enough information for modelling stock volatilities in a local market is uncertain. We will leave a further discussion of this issue in the next section. We now move on to look at the specifications of the common factor and the idiosyncratic volatility.

The Common Factors Dynamics of common factors have already been discussed in the last chapter. Let us have a revision here. We have already seen from equation (4) that individual stock volatility contains the common component that involves the common factor, and the idiosyncratic component that represents the part of fluctuations in return that is unique to the stock. To obtain the common factors from stock volatilities, Cipollini and Kapetanios (2005) suggest the principal components method of Stock and Watson (2002a). Extending the results of Bai (2003), they point out that this estimation method

can still provide consistent factor estimates if the factor process is a stationary long memory Fractionally Integrated Autoregressive Moving Average process (ARFIMA(p,d,q)) with shocks that have finite fourth moment. Their Monte Carlo analysis shows that this factor estimation method performs well in estimating their factor model. An appropriate specification for the factor estimates can be obtained by carefully examining the time properties of the factor estimates. It is well-known that financial time series have long memory and this is evident by many existing empirical literatures. If common factor is able to represent comovements in stock volatilities then the common factor should also have long memory. This claim is first proved by Cipollini and Kapetanios (2005). They suggest factor estimates follows an ARFIMA(p,d,q) process that takes the following form

$$\Phi(L)(1 - L)^d(f_t - \mu) = \Psi(L)\epsilon_t \quad (5)$$

where $\Phi(L)$ and $\Psi(L)$ are lag polynomials and the roots of $\Phi(L)$ lie outside the unit circle. ϵ_t is a white noise process with variance σ^2 . The fractional difference is defined as

$$(1 - L)^d = \sum_{j=0}^{\infty} \pi_j(d) L^j \quad (6)$$

and

$$\pi_j(d) = (-1)^j \frac{\Gamma(d+1)}{\Gamma(j+1)\Gamma(d-j+1)} \quad (7)$$

where $\Gamma(\cdot)$ denotes the gamma function. The ARFIMA(p,d,q) process is

covariance stationary when the value of the fractional differencing operator d lies between -0.5 and 0.5 . For $0 < d < 0.5$, f_t exhibits long memory and its autocorrelations show persistence. For simplicity, we drop out the moving average terms in equation (5) in our empirical analysis, and thus we assume the common factor follow only an ARFI(p,d) process. We adopt the Approximate Maximum Likelihood estimation (AMLE) of Beran (1995) (see also Conditional sum of Square estimator (CSS) by Chung and Baillie (1993)) to estimate ARFI(p,d) process of the factor.

Notice that persistence does exist in a dominant factor, that is the first factor that corresponds to the highest eigenvalue of matrix YY' , where Y is a $T \times N$ matrix of stock volatilities. This can also be seen from our analysis of Asian factors in the previous chapter. However, persistence may not always exist in additional factors. This can be seen later in our empirical analysis. When long memory is absent from an additional factor that is also proved significant in explaining stock volatility, an autoregressive process (AR(p)) is recommended.

The Idiosyncratic Stochastic Volatility Our discussion so far has only focused on the dynamics and the underlying process of the common factor and little have been said about the idiosyncratic volatility, $\omega_{i,t}$. In order to construct the factor model, a specification of the idiosyncratic volatility is required. The idiosyncratic stochastic volatility of a particular stock i is obtained by subtracting the common component from the stochastic volatility. Cipollini and Kapetanios (2005) recommend an univariate state-space model as the underlying process of this idiosyncratic component. In a state-space representation,

idiosyncratic volatility of stock i is assumed to be driven by some unobserved forces that are summarised by the state vector. That is,

$$\omega_{i,t} = \gamma_i \eta_{i,t} + \zeta_{i,t} \quad (8)$$

$$\eta_{i,t} = \lambda_i \eta_{i,t-1} + \kappa_{i,t} \quad (9)$$

where $\zeta_{i,t} \sim N(0, q_i)$, $\kappa_{i,t} \sim N(0, 1)$, and $\eta_{i,0} \sim N(\alpha_0, \rho_0)$. $\eta_{i,0}$ is the initial state of stock i and it has mean equals α_0 and variance equals ρ_0 .

Equation (8) is known as the measurement equation. $\omega_{i,t}$ is the idiosyncratic volatility of stock i at time t . $\eta_{i,t}$ is known as the state. $\zeta_{i,t}$ is serially uncorrelated disturbances with mean zero and variance q_i . $\eta_{i,t}$ is unobservable in general. It is generated by a first-order Markov process as in equation (9), which is known as the transition equation. $\kappa_{i,t}$ is the serially uncorrelated disturbance with zero mean and we assume it has unit variance, i.e. $E(\kappa_{i,t}^2) = 1$. We also assume that $\alpha_0 = E(\eta_{i,0}) = 0$ and $\rho_0 = E(\eta_{i,0}^2) = \frac{1}{1+\lambda_i^2}$. γ_i , λ_i , and q_i are the hyperparameters of the above univariate state-space model, which are estimated via prediction error decomposition by Gaussian maximum likelihood using Kalman filter. Harvey, Ruiz and Shephard (1994) show in their paper that Gaussian maximum likelihood can provide consistent estimates.¹

Notice that in their study of Standard and Poor 500 constituent stock volatilities, Cipollini and Kapetanios (2005) use only the dominant factor extracted from this dataset to forecast US stock volatilities. Therefore, their

¹Further discussion on normality assumption in state-space model can be found in, chapter 3 of Harvey (1990).

model in the form of equation (4) is actually a single local-factor model specification. Their findings show that the volatility forecasts by the single local-factor stochastic volatility factor model outperforms the forecasts produced by some other time series models. Since the US market is the world leading market, its local fluctuations have a lot more significant impact on other markets in the world than the fluctuations in those markets have on it. Therefore, it is sensible to regard the US local dominant factor as the global barometer of stock volatility and thus to consider its local factor being strong enough in explaining stock volatilities of its own market. However, whether the single local dominant factor of other stock markets in the world also contains enough information to produce accurate volatility forecasts for their own markets remains a question. Moreover, local-factor model only allows the study of comovements within the dataset of volatilities in a local market, it does not allow for study of linkages among different markets. With these concerns in mind, we propose an extension of the single local-factor model to a multi-factor specification, namely multi-factor stochastic volatility factor model. We suggest various specifications of this multi-factor model for the investigation of the impacts of overseas factors on local market stock volatilities.

2.1.2 The Multi-factor Specifications (MSVF) – An Extension

A lot of empirical studies provide evidence of volatility transmission and interdependence among financial markets. Some studies show significant stock volatility spillovers from US and Japanese markets to some Asian markets. (see e.g. Ng (2000), and Miyakoshi (2003)). Other studies show volatility transmission and causality among some Asian markets (see e.g. In, Kim, Yooh and

Viney (2001), and So, Lam and Li (1997)). Linkages among Asian stock markets and financial markets around the world can also be seen from the impacts of extreme events on stock markets. Based on these empirical evidence, we believe if there exists volatility spillovers, then dominant factor in a foreign market that summarises the common movement in its stock volatilities should impact on the stock volatilities of a local market. Therefore, by introducing the common factors of the foreign market into the factor volatility model for this local market should increase the model's goodness of fit and thus, strengthen the forecasting power of the model. This brings about an idea of extending a single local-factor model into a multi-factor model. The remaining question we are facing now is what factors, other than the local factor, should be included in the volatility model of a local market?

Although our empirical analysis in this chapter focuses on only the Japanese and Korean markets, the discussion of the multi-factor models here applies to all stock markets in general. According to existing empirical evidence, we believe that apart from a local dominant factor, three other types of factors should also play essential roles in explaining stock return volatility. These factors are, first of all, a US factors that summarise the common driving forces of stock volatility in the US market. Regional leading factors that describe the common fluctuation in stock returns of a leading market of the region. Finally, regional factors that are extracted using data of constituent stocks of all representative indexes of the markets in the same region, except the regional leading factor. We also consider the situation that more than one local factor may be needed for modelling and forecasting local market stock volatilities.

Our forecasting exercise in this chapter has been carried out for only the South Korean and Japanese markets. When the analysis will be further extended to cover the entire Asian region, we suggest the followings: to obtain the three types of common factors explained above, principal component estimation will be applied to extract the US factors from a dataset of US stock volatilities. Regional leading factors are represented by the factors of Japanese market and these factors are extracted from the Japanese stock volatilities. The regional factors are to be extracted from the constituent stock volatilities of all representative indexes of the markets in the Asian region, except those in the regional leading index. The regional factors describe the common variation in stock volatilities in the region except the regional leading market. Stock returns in the regional leading market are not included in the dataset for extracting the regional factor because we want to see how much comovements in stock volatilities of the entire region as a whole can contribute to the volatility forecasts for a local market, without the impact of the stock volatilities of the leading market.² Moreover, since we are introducing the common factors from a leading market as regional leading factors, we have already singled out and emphasized the importance of stock return fluctuations in this market in determining local market stock volatilities.

Following the above argument, we propose an extension to the original single local-factor stochastic volatility factor model, that is, the multi-factor models. Consider the following multi-factor stochastic volatility factor model

²This can of course be verified to include regional leading market stock returns in the dataset if one is more interested in examining the contribution of common variations in stock returns in all markets of the region.

$$\tilde{y}_t^L = a + \theta f_t^M + \omega_t \quad (10)$$

where $\tilde{y}_t^L = (\tilde{y}_{1,t}, \dots, \tilde{y}_{N,t})'$ is a $N \times 1$ vector of local market stock volatility. N denotes the number of constituent stocks of a particular index. θ is a $N \times R$ matrix of factor loadings. ω_t is a $N \times 1$ vector of idiosyncratic volatilities, $\omega_t = (\omega_{1,t}, \dots, \omega_{N,t})'$, and this is obtained after the common component is removed from the stochastic volatilities. f_t^M is a R -dimensional vector of factors. Notice that the dimension of vector f_t^M (and also, θ) depends on the number of factors we include in our multi-factor model.

Let f_t^L , f_t^{US} , f_t^R , and f_t^{RL} denote $K^{(j)}$ -dimensional vectors of local factors, US factors, regional factors, and regional leading factors, respectively. With $K^{(j)} < R$, $j = L, US, R, RL$, and the value of $K^{(j)}$ depends on how many of each types of these factors are proved significant in modelling local market volatility. We will discuss the testing of factor significance later in this chapter. θ^L , θ^{US} , θ^R and θ^{RL} are the matrices of their corresponding factor loadings. We consider the following 5 specifications of a multi-factor model in equation (10).

Modification 1: Local - Regional Leading factor model (L-RL)

Stock volatilities of a local market depends on its local factors, and the factors from a leading market of the region.

$f_t^M = (f_t^L, f_t^{RL})'$ and $\theta = (\theta^L, \theta^{RL})$ i.e.

$$\tilde{y}_t^L = a + \theta^L f_t^L + \theta^{RL} f_t^{RL} + \omega_t \quad (11)$$

Modification 2: Local - US factor model (L-US)

Factors from the local market and the US market are used to explain local market stock volatilities.

$f_t^M = (f_t^L, f_t^{US})'$ and $\theta = (\theta^L, \theta^{US})$. So,

$$\tilde{y}_t^L = a + \theta^L f_t^L + \theta^{US} f_t^{US} + \omega_t \quad (12)$$

Modification 3: Local - Regional Leading - US factor model (L-RL-US)

This model assumes the local factors, together with the regional leading factors and the US factors, are all crucial determinants of local stock volatilities.

$f_t^M = (f_t^L, f_t^{RL}, f_t^{US})'$ and $\theta = (\theta^L, \theta^{RL}, \theta^{US})$. The model becomes

$$\tilde{y}_t^L = a + \theta^L f_t^L + \theta^{RL} f_t^{RL} + \theta^{US} f_t^{US} + \omega_t \quad (13)$$

Modification 4: Local-Regional-US factor model (L-R-US)

This model considers the local factors, the regional factors, and the US factors being significant in explaining local volatilities. The impact of the regional leading market is ruled out.

In this case, $f_t^M = (f_t^L, f_t^R, f_t^{US})'$ and $\theta = (\theta^L, \theta^R, \theta^{US})$.

$$\tilde{y}_t^L = a + \theta^L f_t^L + \theta^R f_t^R + \theta^{US} f_t^{US} + \omega_t \quad (14)$$

Modification 5: Local-Regional Leading-Regional-US factor model (L-RL-R-US)

In this specification, all common factors are concerned.

So, $f_t = (f_t^L, f_t^{RL}, f_t^R f_t^{US})'$ and $\theta = (\theta^L, \theta^{RL}, \theta^R, \theta^{US})$.

$$\tilde{y}_t^L = a + \theta^L f_t^L + \theta^{RL} f_t^{RL} + \theta^R f_t^R + \theta^{US} f_t^{US} + \omega_t \quad (15)$$

There is a reason for not simply to take a model that includes all US, regional, regional leading factors, but to concern different specifications of the multi-factor model. Although some studies of financial contagion have shown intra-regional contagion effects in the Asian financial markets e.g. Masih and Masih (1999), there are also some studies which show no support for contagion among some markets during some periods e.g. Khalid and Kawai (2003). Therefore, we do not want to rule out the possibility that a factor may be an important determinant of stock volatilities in some markets, but not in some other markets. Notice that these specifications can be further extended by including lagged factors. By doing so, it will allow us to study for causal relationships among international stock markets.

Estimation of the multi-factor model is straight forward and this is in the same manner as the estimation of the single local-factor model. We estimate each type of the common factors using principal components. Once we obtain the factor estimates, we remove them from the stochastic volatility of each stock in the dataset and what we are left with is the idiosyncratic stochastic volatility.

Similar to the single local-factor model, all common factors in the multi-factor stochastic volatility factor model may have a long memory underlying process. We model all the factors that exhibit long memory with ARFIMA(p,d,q)

processes. For those common factors that do not have long memory, we model them with an AR(p) process. Idiosyncratic component is the part of stochastic volatility that is unique to each individual stock. Analogous to the local-factor volatility model, this is caused by some driving force with unknown form. And thus, it will also be modelled by a state-space representation.

2.1.3 Testing for Factor Significance

When our interest is to examine how overseas factors impact on stock volatilities in a local market, and thus intend to use a multi-factor model, the next task is to decide on an appropriate specification. That is, which of the specifications in equation (11) to (15) should be chosen? The questions we need to answer here, are (1) are the chosen common factors significant in explaining local market stock volatilities? (2) Is the model including the selected factors strong enough to explain local market volatilities and thus to be used for forecasting? It is important to know how significant those factors are because this will give us an idea of whether our dataset of stock volatilities support a single local-factor or a multi-factor stochastic volatility factor models. If it is the latter that is needed, then how many of the local factors, regional leading factors and US factors should be included.

Answer to the first question can be easily drawn by carrying out test for significance on common factors. However, no existing literatures about factor models seem to have carried out a significance test on common factor. The reason behind is that factors are estimated rather than observed, statistical test for significance on factors cannot be carried out without some well-developed

statistical inferences on factor estimates and the parameters in those factor models. Until recently, Bai and Ng (2006) have made a remarkable contribution to this issue by carrying out a detailed study on statistical inference of principal components factor estimates via the use of factor-augmented regression.

Bai and Ng (2006) show that under some general assumptions, least square estimates of the above factor-augmented regression are asymptotically normal and \sqrt{T} consistent if $\frac{\sqrt{T}}{N}$ tends to 0. They also show that in the setting of factor-augmented regression for a given T , a large N , the number of series that are used for factor estimation, enables precise factor estimation. Thus, estimation errors can be ignored and the cost of having to estimate the factor is negligible. Moreover, consistency of parameter estimates is not affected by the fact that factor is estimated rather than observed as both T and N tend to infinity (see also Bai and Ng (2002)). These results still apply under the conditions of heteroscedasticity and cross-sectional dependence in the idiosyncratic shocks. Given these results, we can carry out significance tests on our factor estimates in the context of factor-augmented regression. We apply the inferences they have developed and we use least squares parameter estimates of factor-augmented regression to compute the test statistics.

Suppose there are N stocks in a dataset, consider the following factor-augmented regression

$$\tilde{y}_{i,t} = c + \alpha' f_t + \beta' M_t + e_{i,t} \quad (16)$$

where $i = 1, \dots, N$ and $t = 1, \dots, T$. $\tilde{y}_{i,t}$ denotes volatility of stock i at time t that is computed by applying standard logarithmic transformation

to daily return of stock i . f_t is a K -dimensional vector of common factors, $f_t = (f_{1,t}, \dots, f_{K,t})$. Common factors contained in matrix f_t is estimated from the $(T \times N)$ dataset of stochastic volatilities, \tilde{Y} , using principal components. c is a constant. α and β are vectors of least squares estimated coefficients of the common factors, f_t , and M_t is a set of other observable variables. $e_{i,t}$ is the disturbance. Notice that in our empirical study, matrix M_t is not considered.

In the Monte Carlo study by Bai and Ng (2006), confidence intervals of estimated conditional mean is computed using their Cross-section and heteroscedastic autocorrelation consistent (CS-HAC) covariance matrix estimator. Factor-augmented regression is used as a forecasting model with different chosen combinations of N and T . Their results show that when idiosyncratic errors are heteroscedastic and cross-sectionally correlated, the coverage rate for conditional mean is the highest when $N = 100$ and $T = 400$. While the coverage rate for forecasting variable is found the highest when $N = 100$ and $T = 400$, and when $N = 100$ and $T = 200$. High coverage rate suggests more robust confidence interval in general. When computing confidence interval, error variance in the prediction of conditional mean is needed. This variance has two parts – asymptotic variance of factor estimates and asymptotic variance of parameter estimates. Factor estimation error will be small due to precise factor estimation if N is sufficiently large. Variance of factor and error variance in prediction of conditional mean will then be small. As a results, narrower confidence interval will be found and high coverage rate is resulted over repeated sample. To sum up, when the conditions (1) $\frac{\sqrt{T}}{N}$ tends to 0, and (2) large N are met, robustness of confidence interval is ensured. This allows high coverage rate which in turn,

indicates consistent parameter and precise factor estimates.

As we will see later in our empirical results in section 5 that the $\frac{\sqrt{T}}{N}$ ratios of our Korean and Japanese datasets are quite close to the one they have chosen and we have a larger N . Consistency in parameter estimates and precise estimation of factor should also be obtained in our study. The asymptotic results they develop should thus be applicable to our analysis and allow us to perform test for significance on factor estimates. We use their results in this chapter for the determination of the number of common factors to be included in our multi-factor model. The following procedure will be taken for the selection of common factors. We first fit a factor-augmented regression into the stochastic volatility of every stock in each dataset, that is, we will estimate N factor-augmented regressions in total via least squares method. By applying their asymptotic theory, we compute t-statistics to test for factors significance in every one of these N regressions. We then move on to check how many times a common factor is found significant out of N models. This will give us an idea of how powerful this factor is in general.

In answering the second question stated at the beginning of this subsection, we believe the statistics that can provide us with the ideas of goodness of fit will be good indicators for selecting an appropriate specification. In order to determine which specification is appropriate for a local market, we estimate all five specifications using data of this market in the estimation sample. Then for each specification, we compute the adjusted R^2 for each stock in the dataset of a local market. We then move on to compare the average adjusted R^2 over all stocks between the five specifications and the one of the local-factor model.

The model that gives us highest average adjusted R^2 is preferable. Comparing adjusted R^2 allows us to pick the specification that is most appropriate to the local market of interest, at the same time account for loss in degree of freedom due to the addition of extra factors into the model. Another advantage of doing this is there will be less time consumption for a forecasting exercise. Instead of forecasting volatility by everyone of the above specifications as shown in equations (11) to (15) for every local market we consider, we use only the one which best describes stocks volatility of that market. We then compare it with the forecasting performance of a local factor model and that of the other volatility models. In our exercise, we compare their performance with the orthogonal GARCH model. The set-up of the orthogonal GARCH model is reviewed in the next subsection.

2.2 Orthogonal GARCH model

Orthogonal GARCH model is a comparable factor volatility model to the stochastic volatility factor models. We have once looked at the model in chapter 1 and a comparison between the two models have been given at the beginning of this chapter. We now have a revision of the orthogonal GARCH model here.

Let X represents the $T \times N$ matrix of normalised dataset of stock return series and each column of this matrix has mean 0 and variance 1. Matrix $X'X$ is a $N \times N$ symmetric matrix of unconditional correlations between data series in X , with diagonal elements of ones. Eigenvectors extracted from matrix $X'X$ gives the normalised factor weights, which are contained in matrix W . Each principal component, p_j , $j = 1, \dots, K$, is a linear combination of the column

of X with weight given by the corresponding column in W . Λ denotes the diagonal matrix of eigenvalues of the unconditional correlation matrix $X'X$.

Columns of W is arranged to descending order according to the magnitude of the corresponding eigenvalues. Thus, the first principal component explains the majority of the total variation in X . Moreover, the principal components are uncorrelated with each other. If the entire normalised dataset of stock return is represented by all N principal components, then the data matrix has the following form:

$$X = PW' \quad (17)$$

where P , X are both $T \times N$ matrices, but W is $N \times N$. Notice that P and W are orthogonal matrices. $W' = W^{-1}$ and $P'P = \Lambda$. However, if only K principal components are chosen to represent the dataset, which means when comovements in stock returns are only approximated by the K chosen principal components, the above representation will include a disturbance term and becomes

$$X = PW' + \Xi \quad (18)$$

Matrices X and Ξ both have dimensions equal to $T \times N$. Whereas P becomes a $T \times K$ matrix and W' is $K \times N$. PW' represents the common components of stock returns. Ξ is the disturbances, it represents the amount of stock returns that is not explained by the common components. The time-varying covariance matrix of X , when only K chosen principal components are used, can be calculated as

$$V = WMW' + V_\epsilon \quad (19)$$

where V_ϵ denotes covariance matrix of the errors, M denotes the time-varying covariance matrix of K principal components and it is a diagonal matrix. V denotes the time-varying conditional variance-covariance matrix of the dataset of stock returns. Approximation of the variance of X is given by

$$V \approx WMW' \quad (20)$$

The diagonal matrix M of covariances of principal components is estimated by univariate GARCH models. Assuming the conditional volatility of each of the K principal follows a GARCH(1,1) in the following form,

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 \quad (21)$$

where $\beta_0 > 0$, $\beta_1, \beta_2 \geq 0$. β_1 represents the market reaction parameter, it measures the intensity of reaction of volatility to the unexpected market return in the last period, that is, ε_{t-1}^2 . β_2 measures the persistence in volatility. Both parameters should sum to less than 1 to ensure convergence. Parameters in the GARCH representations of the principal components are estimated using maximum likelihood estimation. The conditional volatilities for the orthogonal GARCH model are computed as the GARCH(1,1) conditional variance estimates of principal components from the stock returns times the squared value of the corresponding factor weights as in equation (20).

Notice that the major differences between stochastic volatility factor models

and the orthogonal GARCH model is first, the application of principal component analysis to the two models. In the stochastic volatility factor model, principal components are extracted from the dataset of stock volatilities. Whereas in the orthogonal GARCH model, principal components are extracted from the dataset of stock returns. Moreover, it is not difficult to see that the stochastic volatility factor model is a more flexible specification than the orthogonal GARCH model in the sense that it allows the idiosyncratic part of the stochastic volatility to be modelled by a state-space representation. However, the idiosyncratic parts of all stock returns that are not approximated by the model are simply ignored in the orthogonal GARCH model. This has made the stochastic volatility factor models more advantageous in modelling datasets that consist of series with low correlations.

3 Forecasting Methodology

3.1 Forecasting With The Single Local-Factor Specifications

We adopt the method proposed by Cipollini and Kapetanios (2005) to produce our forecasts for Asian stock volatilities using their single local-factor stochastic volatility factor model. Recursive forecasting scheme is adopted to produce one-step ahead volatility forecast for each constituent stocks in a dataset. It means to produce each point forecast, the common factor, the factor loadings as well as the idiosyncratic shocks are to be re-estimated. The forecast is produced by using the stochastic volatility factor model in the form of equation (4). One-

step ahead volatility forecast is thus computed as

$$\widehat{\tilde{y}}_{i,t+1|t} = a_i + \theta'_i \widehat{f}_{t+1|t}^L + \widehat{\omega}_{i,t+1|t} \quad (22)$$

where $t = T, \dots, S$. T is the last period of the estimation sample and S is the last period of the entire sample, which is also the last period of the forecast sample. $\widehat{\tilde{y}}_{i,t+1|t}$ denotes the one-step ahead forecast for the volatility of stock i made at time t . $\widehat{\omega}_{i,t+1|t}$ denotes the one-step ahead forecast of idiosyncratic volatility, and $\widehat{f}_{t+1|t}^L$ denotes the one-step ahead forecast for the local dominant factor estimate. Equation (22) is a local-factor stochastic volatility forecast since only the dominant factor that is extract from the dataset of a local market is considered. We can see that the forecasting procedure using the stochastic volatility factor model is very different from the conventional forecasting exercise using a volatility model that does not involves common factor. The overall volatility forecast on stock i consists of a forecast using estimated data rather than empirical or observed data. Worries about whether an accurate forecast can be obtained may arise. It is because the factor estimation error may deteriorate the fitness of stochastic volatility factor model and thus it may contribute to the disparity between the actual volatility and forecasted volatility. However, the findings by Bai (2003) shows, although error is incurred due to the fact that a factor is an estimated series rather than observed, its size is tiny and can thus be neglected if \sqrt{T}/N tends to zero. Based on the findings of Bai (2003), Cipollini and Kapetanios (2005) point out that a factor model which involves estimated series as an explanatory variable can still provide good forecast for the stochastic volatility.

Equation (22) shows that the overall volatility forecast is produced by combining a forecast for the local dominant factor and a forecast of the idiosyncratic shocks. We now explain how the final forecast is constructed using the following five-step approach. Let $t = 1, \dots, T$ be the estimation sample, and $t = T + 1, \dots, S$ be the forecast evaluation sample. Therefore, the number of periods in the forecast evaluation sample is thus $S - T$. In the last period of the estimation sample, period T :

1. Estimate the dominant local factor from a dataset of stock volatilities using principal component method. Let f_t^L denotes the dominant local factor. Consistent estimate of f_t^L is given by the largest eigenvectors of matrix $\tilde{Y}\tilde{Y}'$, where \tilde{Y} is the matrix of volatility proxies. The factor loadings is computed as, $\Lambda = (f'f)^{-1} f'\tilde{Y} = \frac{f'\tilde{Y}}{T}$. New estimates of the factor, factor loadings and idiosyncratic shocks are then obtained.
2. Estimate the ARFIMA(p,d,q) process shown in equation (5), without the moving average terms of the local dominant factor using AMLE. The estimated parameters of this ARFI(p,d) are used to form a one-step ahead forecast of the local dominant factor for the next period, i.e. $\widehat{f_{T+1|T}^L}$
3. Then for each stock i in the dataset, we fit a univariate state-space model into its idiosyncratic volatility, $\omega_{i,T}$, and estimate it via prediction error decomposition by maximum likelihood using Kalman filter. A one-step ahead forecast for the idiosyncratic shock for the next period, that is $\widehat{\omega_{i,T+1|T}}$, is then obtained.
4. Finally, combine the forecasts produced in step 2 and 3, together with

the factor loadings estimated for period T , we can produce a forecast of volatility for stock i .

$$\widehat{\tilde{y}}_{i,T+1|T} = a_i + \theta_i \widehat{f}_{T+1|T}^L + \omega_{i,T+1|T} \quad (23)$$

5. Repeat steps 1 to 4 for the remaining dates $t = T + 1$ to S , until a $S - T$ dimensional vector of volatility forecasts for stock i is obtained.

Forecasting volatility using a local factor model clearly ignores the fact that fluctuations in other international financial markets can have predictive power towards stock fluctuations in a local market at current period. It assumes the common volatility in the stocks of the local market is explained solely by a local dominant factor and thus it can provide sufficient information in predicting future volatility. However, existing studies on transmission of shocks, causality and financial contagion provide us with lots of empirical evidence of how fluctuations in one market can impact on another one. Therefore, one may argue that using a local factor model to forecast volatility is not general enough. Based on this argument, we also consider forecasting volatility using the multi-factor models.

3.2 Forecasting With The Multi-Factor Specification

If additional factors other than the local factors are proved to be essential determinants of stock volatilities in a local market, volatility forecasts can then be achieved by using an appropriate specification of multi-factor stochastic volatility factor model. Forecasting exercise is done in a similar manner as

when the single local-factor stochastic volatility factor model is used. The only difference is that for every period in the forecast evaluation sample, we forecast not only the local factors in the first step, but we also produce forecasts on additional factors included in the chosen multi-factor model using either ARFI(p,d) or AR(p), depending on the time properties of the common factors. We then forecast idiosyncratic stochastic volatility for every stock using univariate state-space model. An overall volatility forecast for an individual stock is formed by combining these forecasts. Forecasts constructed by the five specifications of the multi-factor model as in equation (11) to (15) are as follow

L-RL factor model:

$$\widehat{\tilde{y}}_{t+1|t}^L = a + \theta^L \widehat{f}_{t+1|t}^L + \theta^{RL} \widehat{f}_{t+1|t}^{RL} + \widehat{\omega}_{t+1|t} \quad (24)$$

L-US factor model:

$$\widehat{\tilde{y}}_{t+1|t}^L = a + \theta^L \widehat{f}_{t+1|t}^L + \theta^{US} \widehat{f}_{t+1|t}^{US} + \widehat{\omega}_{t+1|t} \quad (25)$$

L-RL-US factor model:

$$\widehat{\tilde{y}}_{t+1|t}^L = a + \theta^L \widehat{f}_{t+1|t}^L + \theta^{RL} \widehat{f}_{t+1|t}^{RL} + \theta^{US} \widehat{f}_{t+1|t}^{US} + \widehat{\omega}_{t+1|t} \quad (26)$$

L-R-US factor model:

$$\widehat{\tilde{y}}_{t+1|t}^L = a + \theta^L \widehat{f}_{t+1|t}^L + \theta^R \widehat{f}_{t+1|t}^R + \theta^{US} \widehat{f}_{t+1|t}^{US} + \widehat{\omega}_{t+1|t} \quad (27)$$

L-RL-R-US factor model:

$$\widehat{\tilde{y}}_{t+1|t}^L = a + \theta^L \widehat{f}_{t+1|t}^L + \theta^{RL} \widehat{f}_{t+1|t}^{RL} + \theta^R \widehat{f}_{t+1|t}^R + \theta^{US} \widehat{f}_{t+1|t}^{US} + \widehat{\omega}_{t+1|t} \quad (28)$$

3.3 Forecasting with Orthogonal GARCH model

In the orthogonal GARCH model, comovements in a dataset of stock returns is approximated by K principal components and each principal component is assumed to follow an univariate GARCH(1,1). Conditional variance of principal components can then be computed easily by using the univariate GARCH(1,1). One-step ahead GARCH(1,1) volatility forecast for each principal component is computed as

$$\widehat{\sigma}_{t+1|t}^2 = \beta_0 + \beta_1 \varepsilon_t^2 + \beta_2 \sigma_t^2 \quad (29)$$

According to equation (20), volatility forecast is approximated by

$$\widehat{V}_{t+1|t} \approx W \widehat{M}_{t+1|t} W' \quad (30)$$

where $\widehat{M}_{t+1|t}$ is the matrix of the one-step ahead forecasts of the covariance matrix of all K principal components. $\widehat{V}_{t+1|t}$ is the matrix of one-step ahead forecasts of conditional volatilities of the dataset of stock returns. In other words, volatility forecast is calculated as the GARCH(1,1) volatility forecast of principal components multiplied by the squared of factor weights.

We can see that a forecast for the idiosyncratic volatility, V_ε , in equation (19) does not exist in equation (30). Therefore, the accuracy of the forecast depends on a good approximation as shown in equation (20).

4 Forecast Evaluation

In order to compare volatility forecasting performance using the various models discussed in section 2, we compute several forecast evaluation statistics. Let $\widehat{y_{i,t+1}|t}$ be the one-step ahead volatility forecast for stock i by a particular forecasting model at time t , $y_{i,t+1}$ be the volatility realisation at time t . T denotes the last period in the estimation sample, and S denotes the last period in the entire sampling period. So, $S - T = m$ is the number of periods in the forecast evaluation sample. We compute the following evaluation criteria.

Mean Absolute Error (MAE)

$$MAE = \frac{1}{m} \sum_{j=1}^m |y_{i,T+j} - y_{i,T+j|\widehat{T+j-1}}| \quad (31)$$

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{m} \sum_{j=1}^m \left| \frac{y_{i,T+j} - y_{i,T+j|\widehat{T+j-1}}}{y_{i,T+j}} \right| \quad (32)$$

Mean Square Error (MSE)

$$MSE = \frac{1}{m} \sum_{j=1}^m (y_{i,T+j} - y_{i,T+j|\widehat{T+j-1}})^2 \quad (33)$$

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{m} \sum_{j=1}^m (y_{i,T+j} - y_{i,T+j|\widehat{T+j-1}})^2} \quad (34)$$

U-Coefficient (Theil Inequality Coefficient)

$$U = \sqrt{\frac{\frac{1}{m} \sum_{j=1}^m (y_{i,T+j} - \widehat{y_{i,T+j}|T+j-1})^2}{\frac{1}{m} \sum_{j=1}^m y_{i,T+j}^2}} \quad (35)$$

Theil-Coefficient

$$THEIL = \frac{\sqrt{\frac{1}{m} \sum_{j=1}^m (y_{i,T+j} - \widehat{y_{i,T+j}|T+j-1})^2}}{\sqrt{\frac{1}{m} \sum_{j=1}^m y_{i,T+j}^2} + \sqrt{\frac{1}{m} \sum_{j=1}^m \widehat{y_{i,T+j}|T+j-1}^2}} \quad (36)$$

The above criteria are commonly used statistics for forecast evaluation. MAE, MSE, RMSE and U-coefficients are not bounded and the values of the RMSE, MSE and MAE depends on the values of the scale of the dependent variable. U-coefficient is also known as the standardised root mean squared forecasting error, as we can see from equation (35) that the mean squared error in the denominator is standardised by the sum of the squared actual realisation. The larger the values of these 4 statistics, the worse is the forecasting performance of a model. Whereas small values of them indicate good performance with values equal to zero meaning perfect forecast. MAPE and Theil-coefficient are scale invariant. The values of Theil-coefficient lies between zero and one.

5 Empirical Results

In our forecasting exercise, we produce one-step ahead stock volatility forecasts for the constituents of two Asian indexes. Daily returns on constituent stocks

of NIKKEI 225 of Japan (NIK225) and Korean Stock Exchange Composite Index 200 of South Korea (KOS200) are obtained from Datastream. We also use data on constituent stocks of Standard and Poor 500 (SP500) from the U.S. to produce the US factor for the analysis using multi-factor stochastic volatility factor model. The entire sample covers the period 3 January 1995 to 30 July 2004, leading to 2498 daily observations of stock return. There is a need to have the same number of observations on daily return in all datasets for the analysis using the multi-factor models. Due to the fact that the two markets have different closed market periods, we therefore interpolate returns for the closed market periods by approximating it with the arithmetic mean of the returns on the previous and next trading days. Moreover, we only include the stocks that have data available throughout the entire sample period. This has lead to 136 stocks, 211 stocks and 428 stocks to be included in the KOS200, NIK225 and SP500 datasets, respectively.

5.1 In-sample Analysis using Stochastic Volatility Factor Models

Estimation sample starts on 3 January 1995 and ends on 6 June 2003. This constitutes 2198 daily return observations. The remaining 300 periods are taken as the forecast evaluation sample and this is corresponding to the period 9 June 2003 to 30 July 2004. Let $pr_{i,t}$ denotes daily stock price of stock i at time t , daily return is computed as $\ln(pr_{i,t}) - \ln(pr_{i,t-1})$. We start off by first demeaning the daily returns of the constituents and name it as $y_{i,t}$. For the analysis using the single local-factor model and the multi-factor model, daily

return is transformed by standard logarithmic transformation, that is, $\ln(y_{i,t}^2)$. We then demean the transformed returns and denote it as $\widetilde{y}_{i,t}$. This is a proxy for stochastic volatility of stock i at time t .

To carry out an in-sample analysis using stochastic volatility factor model, we estimate the common factors of stock volatilities via the principal components. The Monte Carlo study by Cipollini and Kapetanios (2005) shows that this estimation method performs well when applying to their model. Principal components method involves eigenvalue decomposition of sample variance-covariance matrix. Computation is simple and the estimation method is asymptotically equivalent to maximum likelihood estimation. Several assumptions are required for consistent factor estimates. A thorough discussion can be found in Stock and Watson (2002b) (see also Cipollini and Kapetanios (2005)). Estimation of factors and factor loadings by this method requires minimising the following objective function³

$$V(f, \Theta) = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (\widetilde{y}_{i,t} - \theta'_i f_t)^2 \quad (37)$$

This is analogous to minimising the variance of idiosyncratic volatility $\omega_{i,t}$ in equation (4). Estimation of the 10 largest factors requires extracting the 10 largest eigenvectors from the matrix $\widetilde{Y}\widetilde{Y}'$, where \widetilde{Y} is the $T \times N$ matrix of demeaned transformed constituent stock returns, that is, volatility proxies. $\widetilde{y}_{i,t}$ is the element in the t th row and i th column of matrix \widetilde{Y} . The eigenvectors we

³Minimising objective function (18) with respect to the factor is equivalent to maximising matrix trace of $f'(XX')f$, subject to the restriction on $T^{-1}(f'f)$ being orthogonal, where $f = (f_1, \dots, f_T)$. Several mild assumptions on the factors, factor loadings, and innovations are required for consistent estimates being produced. For details, see Stock and Watson (2002b) and Cipollini and Kapetanios (2005).

have extracted are the first 10 estimated factor series.

Table 1 shows the cumulative R^2 for the first 10 factors extracted from each dataset. These statistics are computed in the same manner as those ones on table 2 of chapter 2. They are the average coefficients of determination when K factors are included in the factor model. Factor-augmented regressions in the form of equation (16) without a set of other observable variables are formed for the computation of these statistics. That is, for each stock, we regress its volatility proxy on the factor estimates. We do this cumulatively for $k = 1, \dots, K$. We then calculate the average R^2 across all \tilde{y}_t and report it cumulatively for the first K factors and again, we set $K = 10$ here. We can see that the first factors of both datasets explain most of the variation during the estimation period. The first factors from the two datasets are quite powerful which explain about 15% and 11% of the total variation in KOS200 and NIK225 stock volatilities, respectively. Additional factors only cause marginal increase in explanatory power. This has led us to conclude that the first factor estimates from each of the markets can be regarded as their local dominant factors.

In order to determine which specifications of the multi-factor stochastic volatility factor model are appropriate for the KOS200 and NIK225 datasets, we carry out significance test using the factor augmented regression in the form of equation (16) based on the asymptotic results developed by Bai and Ng (2006). As we know from empirical experience that only the first few factors are always powerful, we therefore focuses our tests on the first five local and overseas factors. We carry out significance tests at 10% level. In order to determine which factors should go into the multi-factor model of a local market,

we take the factors that are proved significant to explain the volatility of at least 50% of the stocks in a local market (that is, 68 out of 136 stocks in KOS200 and 105 out of 211 stocks in NIK225), and at the same time they are proved to be helpful in explaining the total variation in the dataset. Table 2 reports the number of times a factor show significance out of N factor-augmented regressions. Table 3 displays the average adjusted R^2 of the factor-augmented regressions with only local factors, local and regional factors, and local and US factors. Notice that the regional factors in the table are the Japanese ones if South Korea is considered as the local market, and vice versa.

We can see from table 2 that the first factor of both datasets is significant for all N factor-augmented regressions. Although the same does not apply for additional local factors, they are still significant for a subset of the stocks in the two datasets. For KOS200, local factors 2 to 5 are significant in explaining the volatility of more than 50% of the stocks. Whereas local factors 2 to 4, but not 5 are proved significant for accounting volatility in the majority of stocks in NIK225. One interesting finding is that none of the regional factor and the US factors are proved significant for at least 50% of the stocks for both datasets. This suggests that there seems to be no strong volatility transmission in general either from the US market to the two Asian markets, or between the two Asian markets throughout our estimation period. Comparing statistics of adjusted R^2 in table 3, we can see that having additional local factors show improvement in accounting for total variation in both datasets, as an increase in the value of adjusted R^2 by about 0.04 is observed for both datasets. However, we cannot benefit from having additional overseas factors. It indicates that the inclusion

of only the first 5 and 4 local factors of KOS200 and NIK225, respectively, in the factor-augmented regressions can help explaining more variation than the inclusion of overseas factors. Based on these statistics and test results, we conclude that multi-local factor specification with 5 and 4 local factors are chosen for KOS200 and NIK225 datasets, respectively.

Figure 1 plots the first local factor of KOS200, and figure 2 plots the first local factors of NIK225. We can see that both estimated factor series revolve around zero. Notice that this factor summarises the majority of the comovements in a dataset of stock volatilities. Figures 3 to 7 graph the sample autocorrelations of the first five local factors of the Korean market up to 500 lags. Whereas figures 8 to 11 display those of the first four local factors of the Japanese markets. Persistence and slow hyperbolic decay can be observed in sample autocorrelations of the first 5 local factors of KOS200 up to 500 lags, except for factor 2, which shows alternating positive and negative autocorrelations. It seems that long memory exists on all first 5 local factors of KOS200 except the second local factor. Similar results are obtained for NIK225 factors. We find that persistence is shown in the autocorrelations of the first 3 local NIK225 factors but not the 4th factor. These findings show that long memory is present in the first, second and third NIK225 factors, and also in the first, third, fourth and fifth KOS200 factors. It indicates that long memory process may be the appropriate specifications for these common factors. These findings also tell us that persistence exists in dominant factors. However, there may or may not be persistence shown in factors with relatively less importance. Dominant factors should embed most of the variance of the dataset, which is an

outcome of the principal component estimation.

To confirm this proposition, we also estimate AR(p) models for each of these local factors from the two datasets and pick the order of lag based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Our finding shows that first, second, third KOS200 factors, and the first, third, fourth and fifth NIK255 factors all support AR(p) representation with long lags (9 lags or more). However, the second KOS200 factor and the fourth NIK225 factor only need AR(4) and AR(6), respectively.

For the factors that exhibit persistence, we fit into each of them a long memory process, ARFIMA(p,d) process in equation (5), with the moving average terms dropped out for simplicity. Order of p is once again chosen by comparing AIC and BIC. Both information criteria suggest ARFI(1,d) is the most suitable model for those factors. We estimate the long memory process using AMLE, which bases on minimising the sum of squared naive residuals. ARFI(1,d) estimation results are shown in table 4 and 5. We can see from these tables that the 1st, 3rd, 4th and 5th KOS200 factors, together with the 1st, 2nd, 3rd NIK225 factors all have estimated fractional differencing parameters, \hat{d} , lie between 0 and 0.5. This is a clear evidence of long memory in those factors. Furthermore, since all estimated \hat{d} also lie between -0.5 and 0.5, this indicates the processes are also covariance stationary. This has led us to conclude that the appropriate underlying processes for those Japanese and Korean factors are ARFI(1,d), except for 2nd KOS200 and 4th NIK225 factors. These two factors should be modelled by an AR(4) and an AR(6), respectively.

Having analysed the dynamics of the common factors, we move on to ex-

amine the idiosyncratic stochastic volatilities. Both single local-factor model that contains only the local dominant factor, and multi-factor model that contains the significant local factors are concerned. Idiosyncratic volatility for each stock, $\omega_{i,t}$ in a dataset is obtained by removing the common component, $\theta'_i f_t$ from its stochastic volatility $\tilde{y}_{i,t}$.

We fit a univariate state-space model as in equation (8) and (9) into $\omega_{i,t}$ for every stock i . Hyperparameters, γ_i , λ_i , and q_i are estimated via prediction error decomposition by maximum likelihood using Kalman filter. Since our datasets have large number of constituents, our analysis involves estimation and reporting results for more than a hundred univariate models. We report our findings using histograms. Figures 12 to 15 are the histograms of the hyperparameters of single local-factor and multi-factor models for KOS200 idiosyncratic state-space models. Figures 16 to 19 are the histograms of the hyperparameters of single local-factor and multi-factor models for NIK225 idiosyncratic state-space models. It can be seen from the histograms that the majority of the idiosyncratic volatilities in the two markets have estimated hyperparameters in the transition equation, λ_i , between 0.8 and 1. These large values of estimated coefficient in the transition equation indicate persistence in the remaining idiosyncratic volatilities of the constituents of our two indexes.

We also check for remaining serial correlations in the measurement errors of idiosyncratic volatilities. This is done by first obtaining smoothed estimates of the states, $\eta_{i,t}$ in equation (9), that is denoted as $\hat{\eta}_{i,t}$. Then we subtract $\hat{\eta}_{i,t}$ from $\omega_{i,t}$ to obtain measurement errors $\hat{\zeta}_{i,t}$. We perform Lagrange Multiplier (LM) test to test for serial correlation up to 5 lags in measurement errors at 1%

significance level for every constituent stock volatility in KOS200 and NIK225 indexes. The null hypothesis of no serial correlations up to 6 lags is rejected if the probability value corresponding to the computed LM test statistic is less than the size of the test equals 0.01. Due to the large number of stock volatility series in each dataset, we reports in table 6 the number of stocks that shows no signs in serial correlation of measurement errors and squared measurement errors when single local-factor and multi-factor models are used.

For the single-local factor model, about 17% of the KOS200 stocks and 10% of NIK225 stocks are found to have no serial correlation in residuals. For the multi-factor model, no serial correlation is found in 15% of the KOS200 stocks and 21% of the NIK225 stocks. We also check for remaining serial correlation in squared measurement errors of idiosyncratic volatilities. LM test statistics tell us that about 77% of KOS200 stocks and 54% of NIK225 stocks show no sign of remaining correlation in squared residuals that is not explained by the single-local factor model. In the multi-factor model case, no remaining serial correlation in squared residuals is found in 80% and 67% of KOS200 and NIK225 stocks, respectively. Two things can be seen from these results. First of all, the factor models do not seem to capture enough serial correlation in the measurement error for the majority of the stock volatilities. Secondly, there is hardly improvement by moving from single-local factor to multi-factor models for KOS200 stock volatilities. Although the multi-factor model seems to be able to capture more serial correlation in squared residual, the extra 3% improvement only translates into 4 stocks. For NIK225 stock volatilities, it seems there is improvement in moving from single-local factor model to a

multi-factor model but there is still sign of serial correlation in the majority of the stocks in the multi-factor model case. We suspect that using a fixed number of local factors to model the volatilities of all stocks in the multi-factor case may not be appropriate. As we can see evidence from table 2 that individual stocks may need different number of local factors in modelling their volatility. Thus, further improvement can be achieved if individual stock volatilities are modelled using the appropriate number of factors that are proved significant in explaining them.

5.2 In-sample Analysis using Orthogonal GARCH Model

In-sample analysis is also performed in the context of orthogonal GARCH model. The first principal component extracted from KOS200 and that from NIK225 datasets of normalised return series are shown to have explained 31% and 29% of total variation, respectively. This amount of explained variation is even more than the amount of variation explained by the chosen number of local volatility factors in the multi-factor stochastic volatility factor models for both datasets. Our orthogonal GARCH analysis will then only use the first principal component for both datasets. A GARCH(1,1) specification as in equation (21) is used for modelling the first principal components extracted from the two datasets of returns. In the conditional mean equation, we regress the principal component on a constant, α_0 , only. In-sample estimation of a GARCH(1,1) is presented in table 7.

β_1 measures the market reaction, it tells how intensely volatility at current

period reacts to unexpected market returns in the last period. Whereas β_2 measures the persistence in volatility. Both parameters must have sum up to a value less than one in order to ensure a finite unconditional variance, and β_0 must be larger than zero. We can see that all parameter estimates in the conditional variance equations for the principal components of both datasets are significant at 5 % significance level. All $\beta_0 > 0$ and $\beta_1 + \beta_2$ sum to less than one for both markets. Notice that this sum is quite close to one and this is what usually observed in high frequency financial data. We can also see from the values of β_2 that principal components from the two markets appear to have quite high volatility persistence, with the South Korean stock volatilities to be more persist than the Japanese ones in general. From the values of β_1 , the Japanese stock volatilities seem to response to unexpected return in the last period more intensely than the South Korean market as it has higher value of the market reaction parameter.

A $T \times N$ normalised stock return matrix is computed for both datasets, that is, matrix X in equation (18). Diagnostic tests are carried out on return residuals, that is matrix Ξ in equation (18) when the return series are modelled by using only the first principal component. Table 8 reports the results of Breusch-Godfrey LM test for serial correlation and ARCH LM test. 6 lags are included in the auxiliary equations of both tests and level of significance is chosen to be 1%. With reference to the large number of stock return series in each dataset, we report the number of stocks that shows no significant serial correlation and no significant ARCH effect in their returns. Evidence of misspecification in general can be seen clearly from the results as less than

10% of stock returns have residuals that have no sign of ARCH effect for both indexes. No signs of serial correlation only found in 23% and 32% of stocks return residuals in KOS200 and NIK225, respectively.

5.3 Out-of-sample Forecasting

Our forecasting exercise is performed for the last 300 periods of the whole sample to produce 1-step ahead point forecasts using the single local-factor and multi-factor stochastic volatility factor models, and the orthogonal GARCH model. Recursive forecasting procedure is adopted for which our models are estimated with more data as forecasting move forward in time. Following the procedure and using the forecasting equations explained in section 3, we construct volatility forecast $\widehat{\tilde{y}}_{i,t+1|t}$ for every period in the forecast evaluation sample. For both single local-factor and multi-factor models, an overall 1-step ahead volatility forecast for each stock is formed by combining the 1-step ahead factor forecast(s) and the 1-step ahead idiosyncratic volatility forecast. However, attention should be paid in computing the final forecasts. Since the data we use for estimation and forecasting are demeaned, we therefore, need to add the removed mean back to the demeaned volatility forecast to obtain the final volatility forecast, $\widehat{\tilde{y}}_{i,t+1|t}$. So, $\widehat{\tilde{y}}_{i,t+1|t}$ is forecasted volatility with the removed mean added back to it. The forecasting exercise is performed for every period in the forecast sample until a vector of volatility forecasts is achieved for each of stock.

Since we have large datasets that contains more than a hundred stocks with more than a thousand observations. Thus there is a huge the number

of univariate state-space models for idiosyncratic volatilities need to be re-estimated along the forecast evaluation period. There are 40800 individual state-space models for KOS200 dataset and 63300 models for NIK225 dataset. We aware there may be a possibility that maximum likelihood estimation does not converge for some of the idiosyncratic models. And if this happens, the relevant periods will be excluded from the final results. However, we do not find such case in our analysis.⁴

Orthogonal GARCH forecasts are obtained by first performing a one-step ahead forecasts of a GARCH(1,1) of principal component extracted from the normalised dataset of stock returns. Then, we multiply this one-step ahead forecast by the squared of the factor weights to obtain the one-step ahead out of sample forecast of the volatility of each series. As described in equation (29) and (30).

To forecast the Asian volatilities using the stochastic volatility factor models, we consider two specifications, the single local-factor model for both datasets and the multi-factor models with 5 and 4 local factors for KOS200 and NIK225, respectively. However, considering the large number of stocks included in our datasets, it will involve enormous amount of work if we compute forecast for each individual stocks with different number of local factors in its multi-factor model. We believe although the overall forecast on volatility of all stocks are computed by using the multi-factor models with the same number of local factors, our results can still provide enough insight into how good overall forecast

⁴It is not surprising to find all idiosyncratic models converge in our analysis for all three datasets. The analysis of Cipolini and Kapetanios (2005) involves estimation of 43800 individual state-space model and they find only 8 models that do not converge.

performance can be achieved by using the single-factor and multi-factor models.

Due to the large number of stocks in the two Asian datasets, we compute the average of the statistics explained in section 4 for forecast performance comparison. Volatility forecasts by univariate GARCH(1,1) is also performed as a benchmark comparison. Moreover, the realisation of the volatility proxy computed as $\ln(\widetilde{y}_{i,t})$. We take logarithm of the forecasts computed by the orthogonal GARCH and the GARCH models to compare them with this proxy. Our result is reported in Table 9.⁵ First of all, we can see that overall performance of univariate GARCH(1,1) forecasts are the worst. This is not a surprising finding as it is well-known that univariate GARCH does not capture some complex temporal features commonly found in financial time series.

We can see for KOS200, both single local-factor and multi-factor stochastic volatility factor models outperform the orthogonal GARCH model in forecasting volatility when comparing all evaluation statistics, except for average MAPE, that shows the multi-factor model is comparable to the orthogonal GARCH model. However, the majority of these average statistics still tell us that the multi-factor model is performing slightly better than the orthogonal GARCH model. For NIK225, 4 out of 6 reported evaluation statistics suggest that the orthogonal GARCH model does not provide better volatility forecast than the single local-factor and multi-factor stochastic volatility factor models

⁵Tests on predictive ability can also be carried out to evaluate the accuracy of forecasts produced by the orthogonal GARCH model and the stochastic volatility factor models. For example, the asymptotic tests of Diebold and Mariano (1995) can be applied to allow pairwise comparison of the forecasting models we concerned here and to see if stronger predictive power of a model found is statistical significant. See also Cipollini and Kapetanios (2005) for an application of this test to factor volatility models.

on average.

Interesting results can be seen when comparing only the single local-factor and multi-factor models for both datasets. The statistics show that there are no improvement in forecasting volatility by moving from a single-factor to a multi-factor model. Except for volatility forecasts of the NIK225 stocks we can see from the average MAPE and average Theil coefficient that the two models are comparable. We suspect that the failure in obtaining an improvement in forecasting performance by moving from a single local-factor model to a multi-factor model is due to the fact that the forecasting performance of the model will be worsen when an additional insignificant factor is used for modelling volatility of a stock. We can see from table 2 that not all additional local factors are significant for all stocks in the two datasets need. However, the first local factor always shows significance for all stocks in both datasets. Adding extra insignificant factors to the stocks that needs only one local factor will bring down the goodness of fit of the model and in turn worsen its predictive power. When we use the multi-factor model for modelling, we model volatility of all stocks in a dataset with the same number of local factors, despite the fact that for some stocks but not all, their volatility can be better modelled by one local factor. However, if we take into account of this situation and allow individual stock volatilities from the same dataset to be modelled with different number of local factors according to its need, the overall forecasting performance may be improved.

Individual stocks have their own idiosyncratic characteristics and may expose to different exogenous market factors. Empirically, it can be observed that

although the overall trend of the entire stock market may respond to the same direction when there are changes in market condition, global and local economic environments, individual stocks from the very same market have different degree of reaction to those changes. When modelling and forecasting individual stocks in practice, it is sensible to treat each stock individually by considering their unique characteristics. Stochastic volatility factor models have already allowed the unique characteristics of volatility to be captured by the idiosyncratic part of the model. However, when determining upon the number of factors that should be included in the model, it is more appropriate to consider each stock individually. With this understanding in mind, we advocate the use of a mixed forecasts when using the stochastic volatility factor models to the forecast volatilities of a portfolio of stocks in practice.

In order to get an idea of whether using a mixed forecast can improve forecasting performance of the stochastic volatility factor models in forecasting volatilities of a portfolio of stocks, we carry out an experiment. For each of the stock in the dataset, we look at whether it is better forecasted by the single-factor model or the multi-factor model by comparing the corresponding forecast evaluation statistics, we then pick the smaller value of the two statistics. We do this for every single stock in the dataset, to get an $N \times 1$ vector of forecast evaluation statistics, averaging this vector gives us the average forecast evaluation statistic. These statistics are also reported in table 9. We can see from the table that these average statistics have lower values than those average forecast evaluation statistics of the the single local-factor and of the multi-factor models. This may suggest that overall forecasting performance

can be improved if a mixture of forecasts via single-factor and multi-factor models are used for a portfolio of stocks. However, one should notice that this experiment is not a feasible forecasting procedure. It is because the choice of forecasting model should be made prior to the forecasting exercise being carried out. Therefore, our results here reveal the lower bound of how well our factor models can do in forecasting a bunch of Asian volatilities and thus can somehow provide empirical evidence to show how overall prediction of volatilities in all stocks can be improved by taking into account the number of factors needed for each individual stock in empirical factor modelling.

To conclude, we think the failure of the orthogonal GARCH model to provide better overall performance in our study can be explained by its deficiency in dealing with less correlated system of data. Orthogonal GARCH model works better for highly correlated system of financial time series. This can be seen by comparing the explanatory power of the principal components extracted from our datasets of returns with the findings in the empirical application in the Alexander (2001a) paper. In her study, principal components extracted from highly correlated system such as term structure and crude oil futures and the first few principal components can always explain around 90% of total variation in a dataset. However, our analysis concern stock returns and it can be seen from the analysis in chapter 1 that they are not as highly correlated in general. Large number of principal components are not recommended to use for orthogonal GARCH modelling not only for the ease of analysis and computation and for the aim of reduction in dimensionality, it is also for avoiding extra noise that is embedded in some additional and insignificant principal components to

enter the model. This extra noise can increase the size of the error, causing weak approximation of the variance-covariance matrix, and thus worsen the overall performance of the model. In the case when the large variation is captured by the principal components included in the orthogonal GARCH model, the Ξ term in equation (18) is small. Orthogonal GARCH modelling can be appealing in this situation.

In the case of a less correlated system of data, the first few principal components may not capture enough total variation. If a researcher is to use a large number of the principal components for modelling, he will be worried about having extra noise being included in the model. Using large number of principal components may lead to less stable correlation and less robust volatility estimates. Therefore, in the case of low correlations within a dataset Ξ in equation (18) is larger, approximation in equation (20) is weaker. Consider that in orthogonal GARCH modelling the return residuals are left out for the computation of volatilities and covariance in the dataset. This may result in a weak performance of the model in facing low correlation datasets. This can be proved by our in-sample results that the first principal component used for approximating the comovements in the dataset cannot capture enough dynamics of the system of data. The out-of-sample overall volatility forecast is not better than the stochastic volatility factor model. In contrast to the orthogonal GARCH model, stochastic volatility factor models do not have such deficiency in failing to capture idiosyncratic part of the series in a dataset. As we can see from equations (4) and (22) that the idiosyncratic volatilities also make contributions to volatility modelling and forecasting. This flexibility in modelling

both the common component and idiosyncratic component in the series of a dataset has made the stochastic volatility factor model more desirable than the orthogonal GARCH model for an analysis involving less correlated system of data.

6 Concluding Remarks

In this chapter, we analyse and compare the two factor volatility models that can be applied to large datasets with more than a hundred cross-sectional series. We have discussed an extension of the stochastic volatility factor model from a single local-factor specification to multi-factor specifications. We also look at how common factor and idiosyncratic volatility are modelled in the factor models. Testing of factor significance via a factor augmented regression is performed. Moreover, examination of and comparison between in-sample and out-of-sample performance of the factor models and the orthogonal GARCH model have been carried out. The use of mixed forecasts by the single-factor and multi-factor models for a datasets of stock volatilities is suggested.

Our findings suggest weak significance of the US factors for explaining both Korean and Japanese stock volatilities. There seem to be weak volatility transmission in general between the two Asian markets throughout the sample period. Evidence of misspecification is found in most of the stock returns in the in-sample analysis for both the Korean and Japanese datasets when returns are modelled by an orthogonal GARCH model with only the first principal component. Out-of-sample forecasting shows that the factor models have out-

performed the orthogonal GARCH model in general. Whereas the use of mixed overall forecasts produced by the factor models may be able to further improve forecasting performance. This has lead us to conclude that prediction can be improved by taking into account the idiosyncratic characteristics of individual stocks in empirical modelling.

Weak in-sample and out-of-sample performance of orthogonal GARCH model may imply the model has failed to capture some significance temporal features of Asian stock volatilities. When comparing the orthogonal GARCH model with the stochastic volatility factor models, we can see that the latter are a more flexible specification than the former on dealing with datasets with low correlations as they allow idiosyncratic part of stock volatility to be modelled. Moreover, the common factor is represented by a specification that best describe their temporal features, such as long memory. Further analysis in related topic, may consider a forecasting exercise using an improved version of orthogonal GARCH model. For example, to consider the inclusion of overseas principal components in the model. Moreover, inclusion of long memory in orthogonal GARCH, for example, to allow principal components to have Fractionally Integrated GARCH representation may also make the model more realistic for modelling financial data.

Furthermore, our examination of the idiosyncratic volatilities has shown persistence in their dynamics. This may imply remaining unexplained long memory not captured by the common factors. Further investigation of remaining long memory can be carried out. In particular, if evidence of remaining long memory is found in the idiosyncratic volatility, attention should be paid

to the improvement of the idiosyncratic models that accounts for the remaining long memory.

In addition, the multi-factor models have been applied to Japanese and South Korean volatilities only. Although we fail to provide evidence to show strong volatility transmission during the sampling period, it may be the case that overseas factors are useful in explaining local market volatility in some particular period, for example, in the case of extreme events. Further research work on the multi-factor models can be carried out for some periods when extreme events have happened. Volatility transmission among markets during the period of extreme event can be carried out in the context of factor analysis. Moreover, it may be interesting to look at the relationship between volatility and some other factors that have an effect on it. For example, Kim, Kartsaklas and Karanasos (2006) study the relationship between volume and volatility in the Korean market in relation to the 1997 financial crisis. Similar kind of study may be carried out in the context of factor analysis by modifying the factor models. Furthermore, the study can be extended to cover the entire Asian region to see if the inclusion of more significant overseas factor can improve the practical performance of the multi-factor model.

7 Tables and Figures

Abbreviation used:

SSVF – Single local-factor stochastic volatility factor model

MSVF – Multi-factor stochastic volatility factor model

GARCH – Generalised Autoregressive Conditional Heteroscedasticity model

OGARCH – Orthogonal GARCH model

FAR – Factor-augmented regression

ARFI – Fractionally Integrated Autoregressive process

Table 1: Cumulative R^2 of KOS200 and NIK225 factors		
No. of factors	KOS200	NIK225
1	0.1473	0.1097
2	0.1558	0.1271
3	0.1647	0.1364
4	0.1819	0.1453
5	0.1906	0.1513
6	0.1992	0.1574
7	0.2072	0.1657
8	0.2153	0.1725
9	0.2230	0.1790
10	0.2304	0.1851

Key: Statistics reported in this table is computed via FAR in the form of equation (16) without a set of other observable explanatory variables. They are the average coefficients of determination when K factors are included in the model.

Table 2: No. of times (out of total no. of stocks a factor is proved significant

Parameters in FAR	KOS200	NIK225
constant	0	0
local factor 1	136	211
local factor 2	76	159
local factor 3	69	148
local factor 4	119	147
local factor 5	71	97
regional factor 1	36	62
regional factor 2	40	22
regional factor 3	31	37
regional factor 4	20	54
regional factor 5	14	28
US factor 1	35	50
US factor 2	48	54
US factor 3	33	29
US factor 4	25	34
US factor 5	24	38

Table 3: Average adjusted R^2 of FARs including different factors		
FAR specification	KOS200	NIK225
with only 1 local factor	0.1469	0.1092
with 5 local factors	0.1888	0.1493
with 5 local and regional factors	0.1816	0.1449
with 5 local and US factors	0.1824	0.1450

Table 4: In-sample ARFI(1,d) estimation results for KOS200 factors

	1st factor	3rd factor	4th factor	5th factor
\hat{d}	0.16	0.08	0.18	0.12
AR coefficient	0.61	0.43	0.59	0.36
95% CI for \hat{d}	(0.107, 0.204)	(0.03, 0.13)	(0.127, 0.228)	(0.069, 0.172)

Table 5: In-sample ARFI(1,d) estimation results of NIK225 factors

	1st factor	2nd factor	3rd factor
\hat{d}	0.16	0.14	0.06
AR coefficient	0.67	0.45	0.23
95% CI for \hat{d}	(0.118, 0.211)	(0.094, 0.193)	(0.01, 0.121)

Table 6: Serial correlation test on measurement errors of idiosyncratic volatilities (significance level = 1%)

SSVF		
	No serial correlation in residuals	No serial correlation in squared residuals
KOS200	24	105
NIK225	22	114
MSVF		
	No serial correlation in residuals	No serial correlation in squared residuals
KOS200	21	109
NIK225	45	142
Note: The above figure represents number of stocks.		

Table 7: In-sample GARCH(1,1) estimation of first principal component

		α_0	β_0	β_1	β_2
KOS200	coefficient	-0.012	0.28	0.074	0.92
		(0.09)	(0.10)	(0.01)	(0.01)
	z-statistic	-0.13	2.82	5.11	64.0
		{0.89}	{0.004}	{0.00}	{0.00}
NIK225	coefficient	0.04	3.01	0.11	0.83
		(0.14)	(1.05)	(0.02)	(0.03)
	z-statistic	0.29	2.85	4.96	25.2
		{0.76}	{0.004}	{0.00}	{0.00}
Note: std. errors are reported in parentheses, p-values are reported in curly brackets					

Table 8: Diagnostic test results on return residuals of OGARCH with only the first principal component
(significance level = 1%)

	No serial correlation	No ARCH effect
KOS200	32	6
NIK225	69	16

Note: The above figure represents number of stocks.

Table 9: Forecast performance comparison					
	SSVF	MSVF	SSVF &	GARCH	OGARCH
			MSVF mixed		
KOS200					
Average MAE	2.105	2.113	2.100	2.461	2.157
Average MAPE	0.233	0.234	0.232	0.226	0.234
Average MSE	8.908	9.00	8.892	13.571	9.731
Average RMSE	2.864	2.878	2.862	3.542	2.983
Average U-coef.	0.293	0.295	0.290	0.362	0.305
Average Theil coef.	0.000107	0.000108	0.000107	0.00017	0.000115
NIK225					
Average MAE	1.939	1.944	1.931	2.092	1.902
Average MAPE	0.208	0.208	0.206	0.192	0.194
Average MSE	7.564	7.596	7.541	10.112	7.976
Average RMSE	2.709	2.715	2.705	3.133	2.770
Average U-coef.	0.278	0.279	0.278	0.322	0.284
Average Theil coef.	0.0001	0.0001	0.0001	0.00014	0.000108

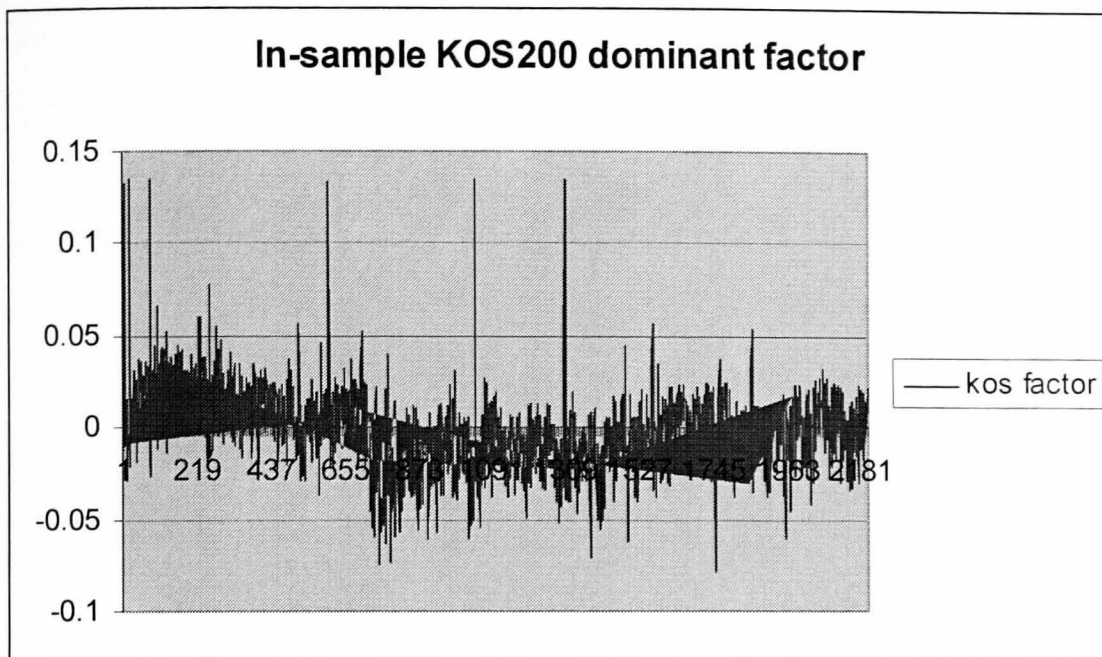


Figure1: In-sample KOS200 dominant factor

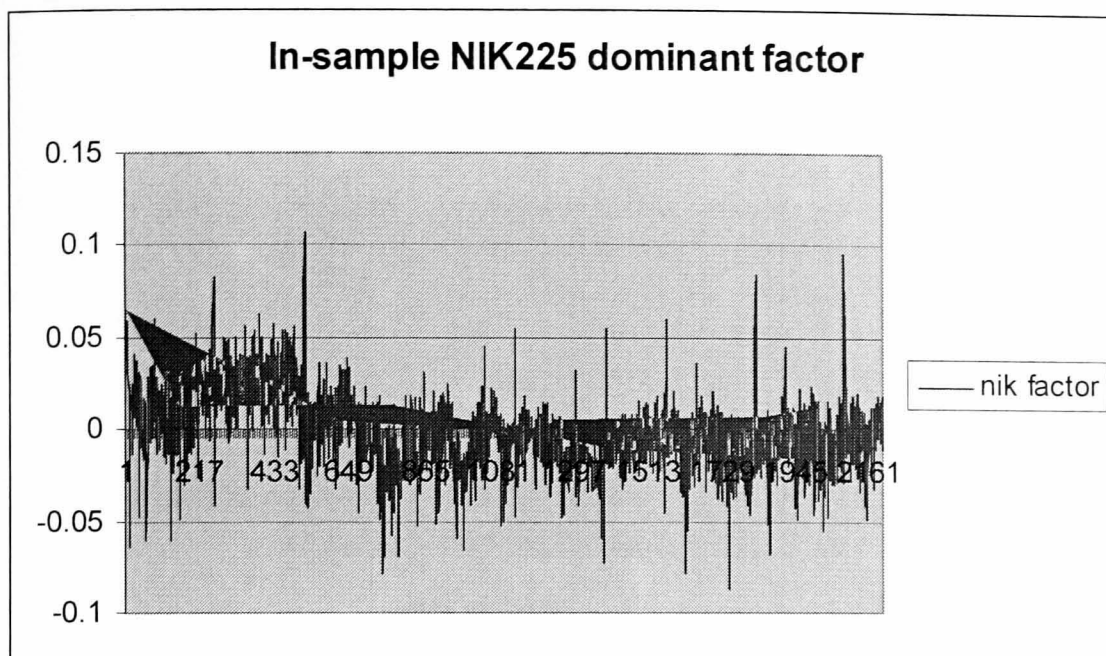


Figure 2: In-sample NIK225 dominant factor

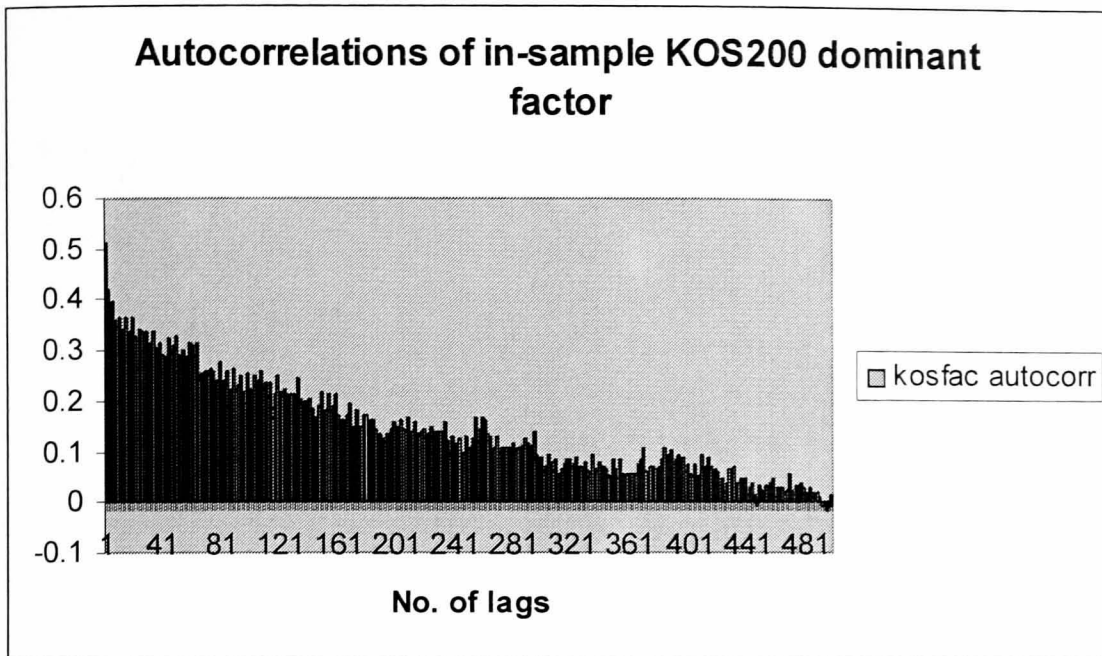


Figure 3: Autocorrelations of in-sample KOS200 dominant factor

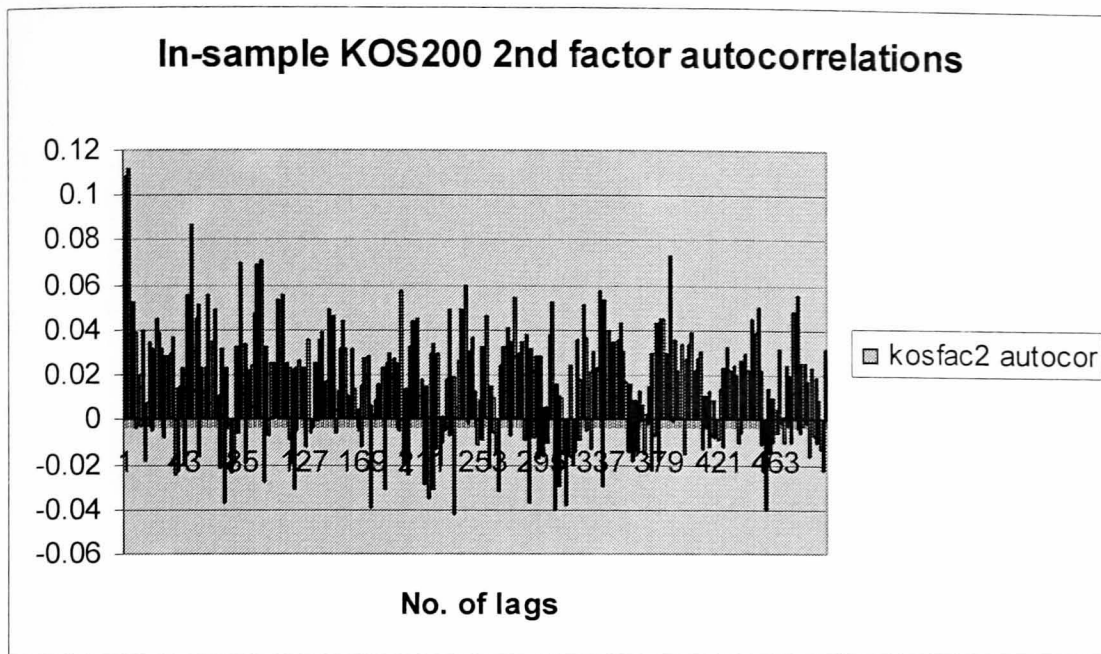


Figure 4: Autocorrelations of in-sample KOS200 second factor

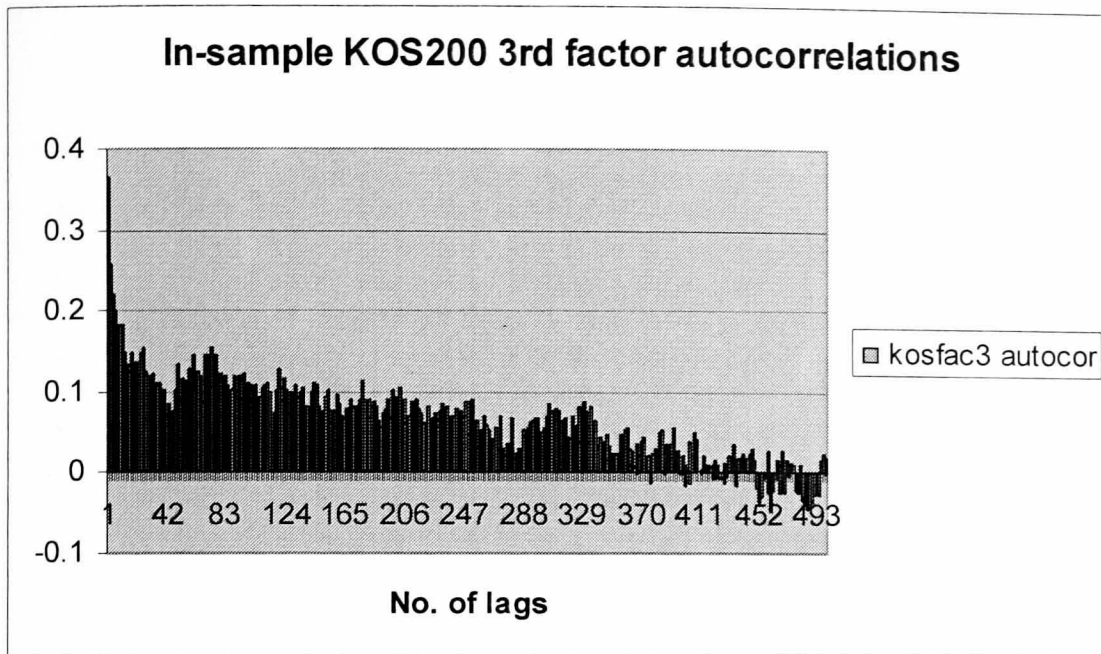


Figure 5: Autocorrelations of in-sample KOS200 third factor

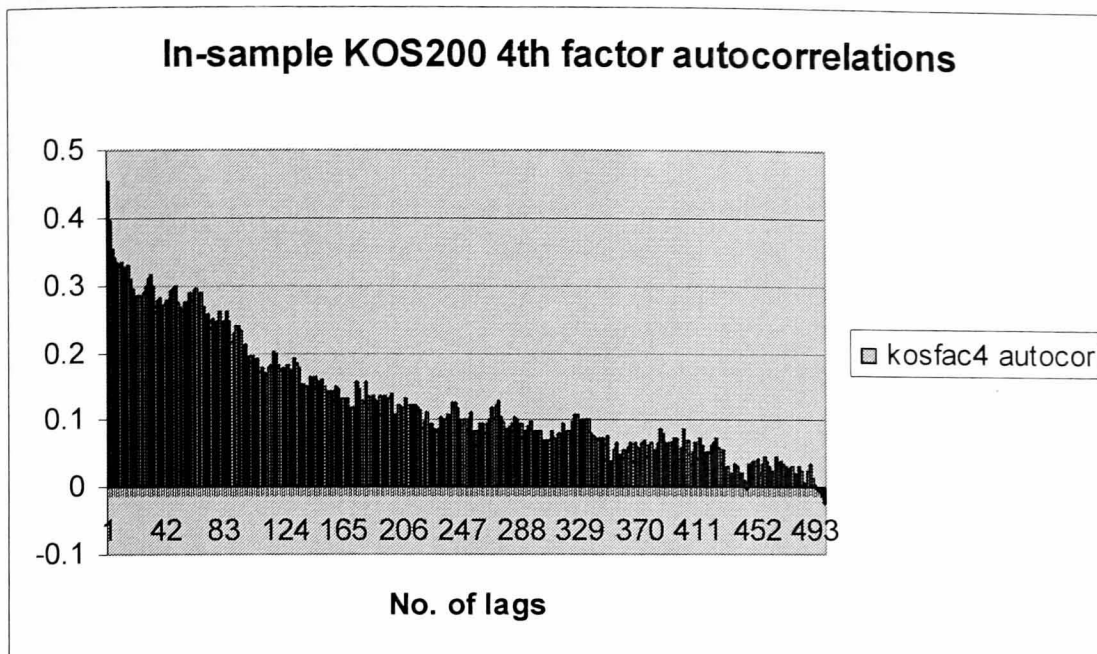


Figure 6: Autocorrelations of in-sample KOS200 fourth factor

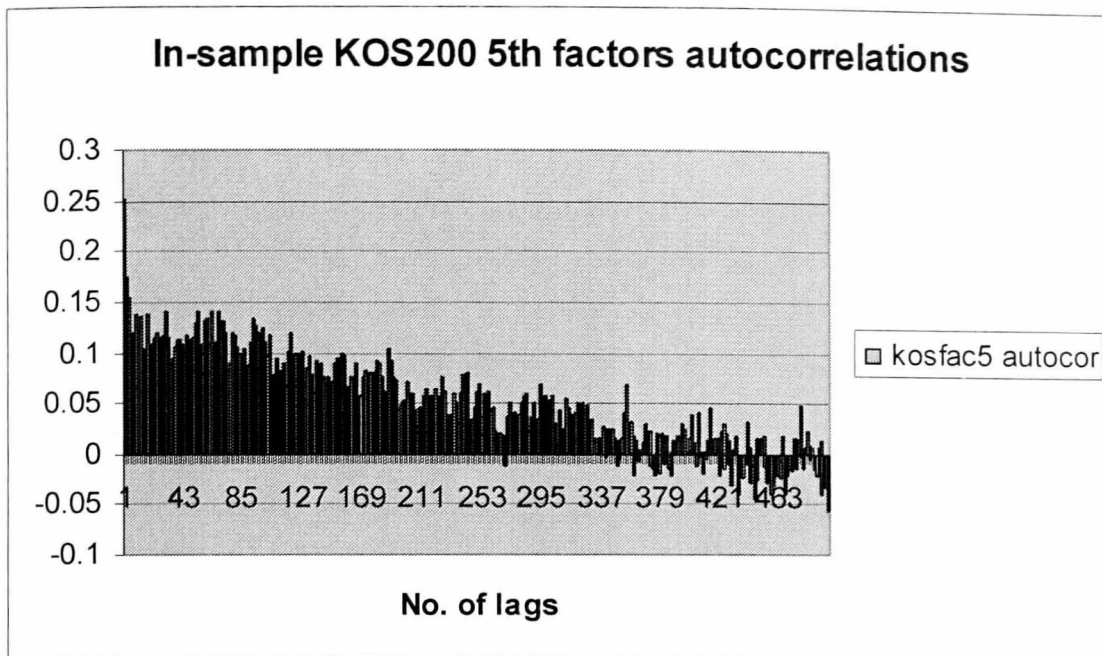


Figure 7: Autocorrelations of in-sample KOS200 fifth factor

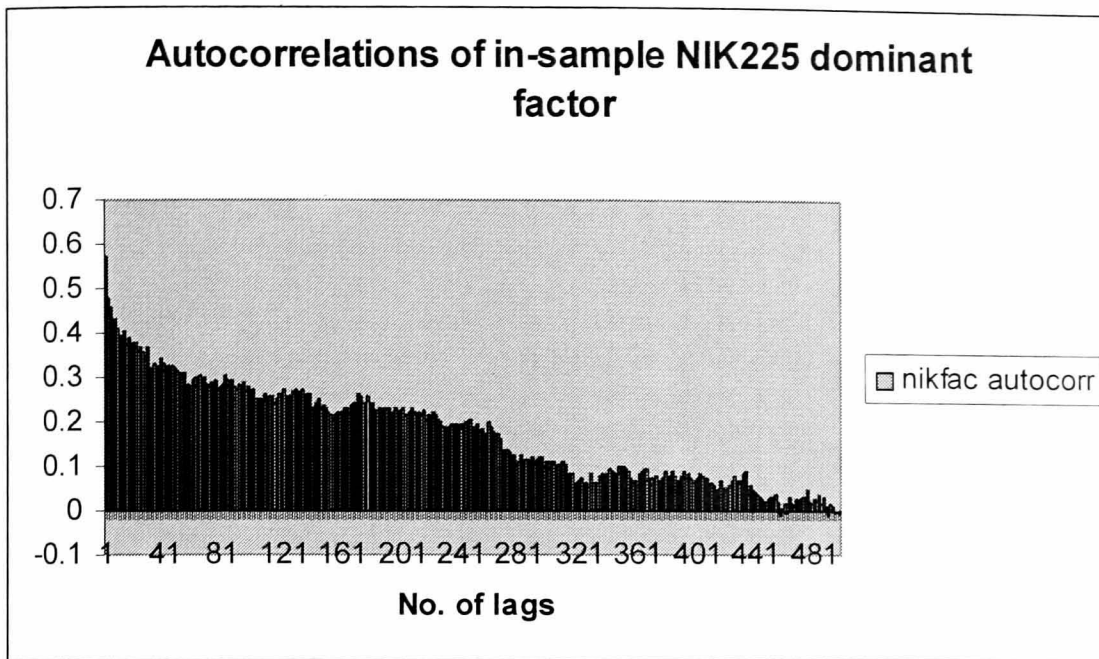


Figure 8: Autocorrelations of in-sample NIK225 dominant factor

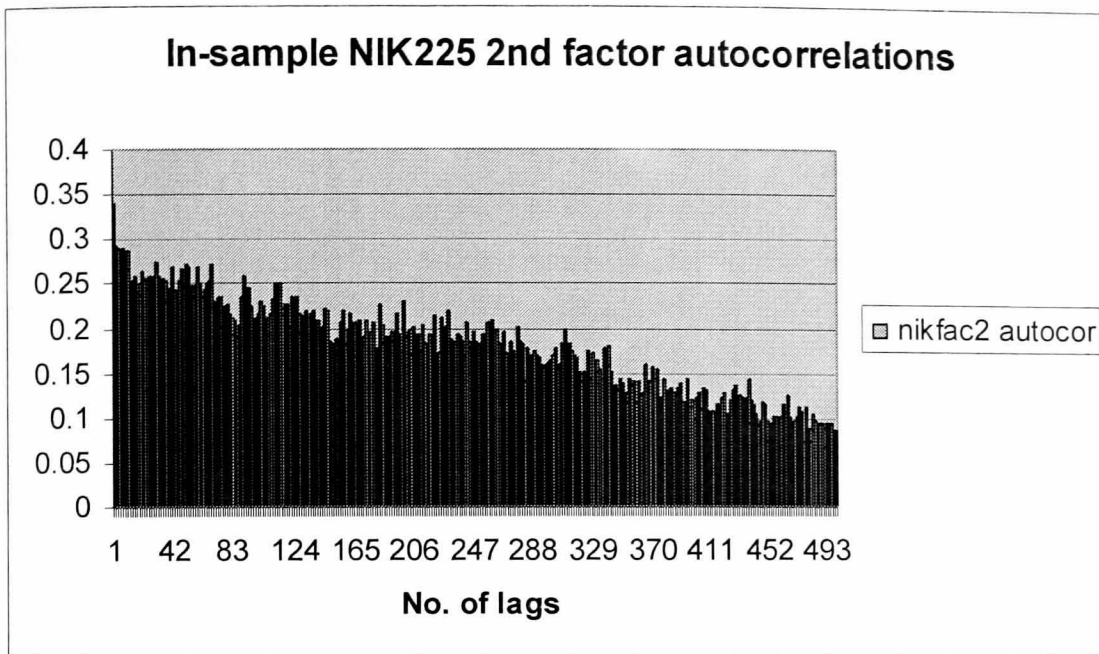


Figure 9: Autocorrelations of in-sample NIK225 second factor

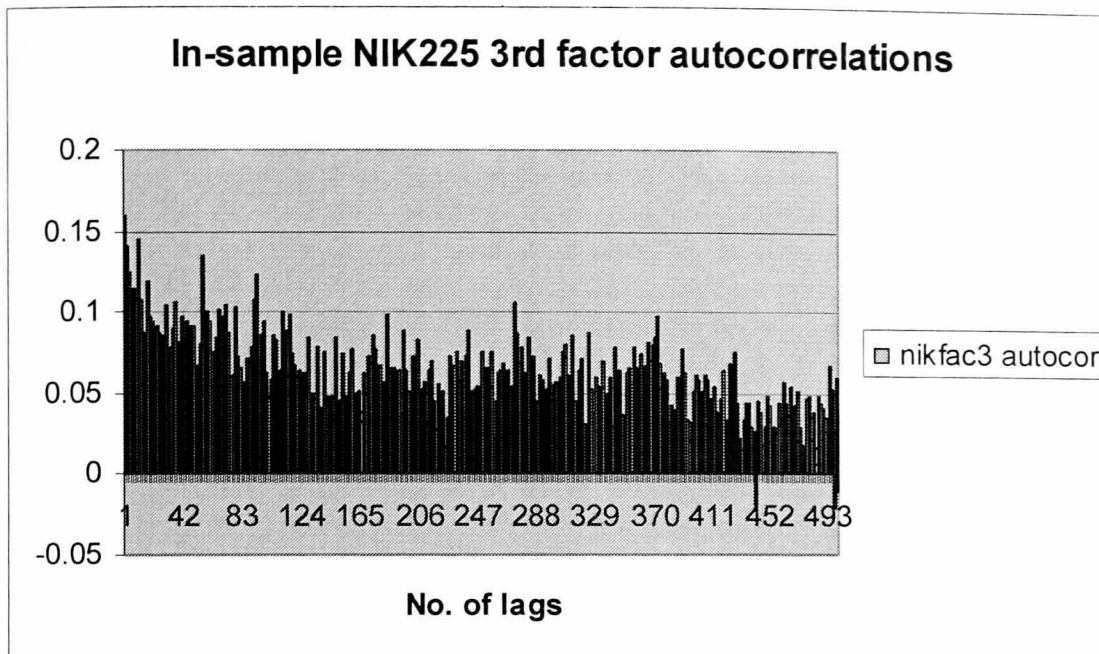


Figure 10: Autocorrelations of in-sample NIK225 third factor

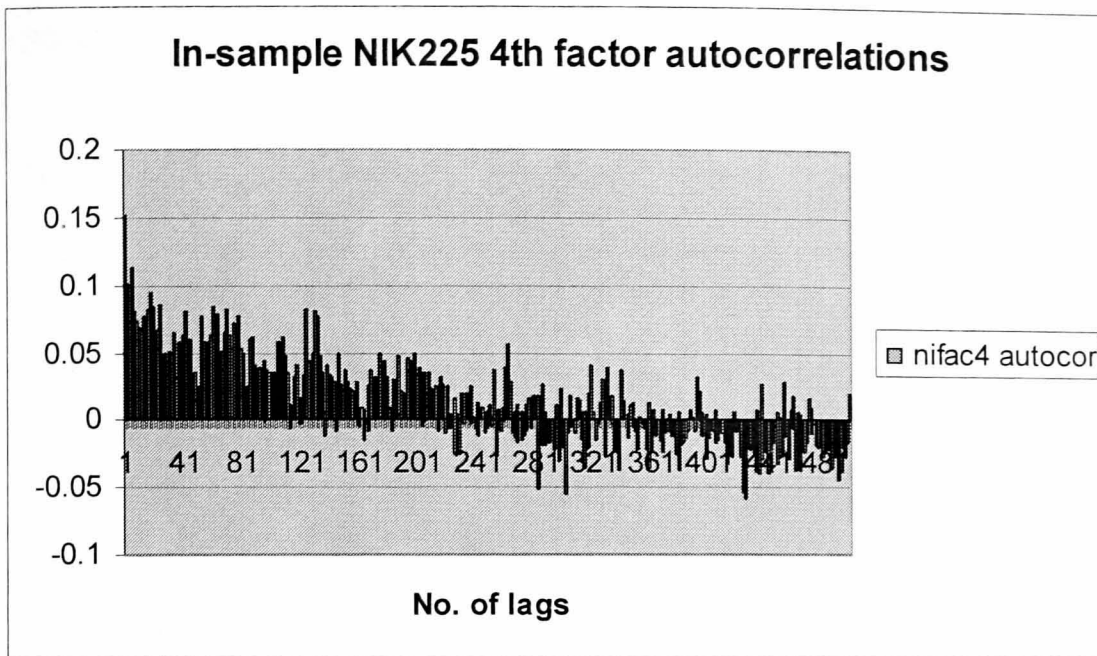


Figure 11: Autocorrelations of in-sample NIK225 fourth factor

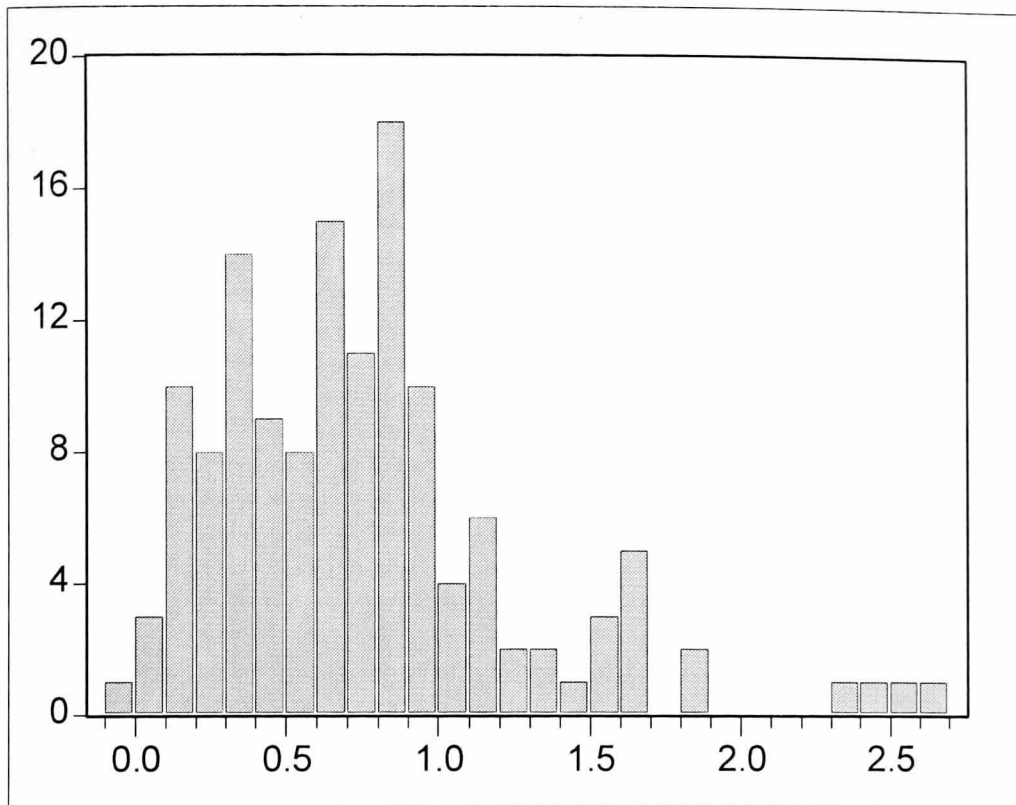


Figure 12: Histogram of hyperparameters in measurement equations of KOS200 idiosyncratic volatilities in SSVF

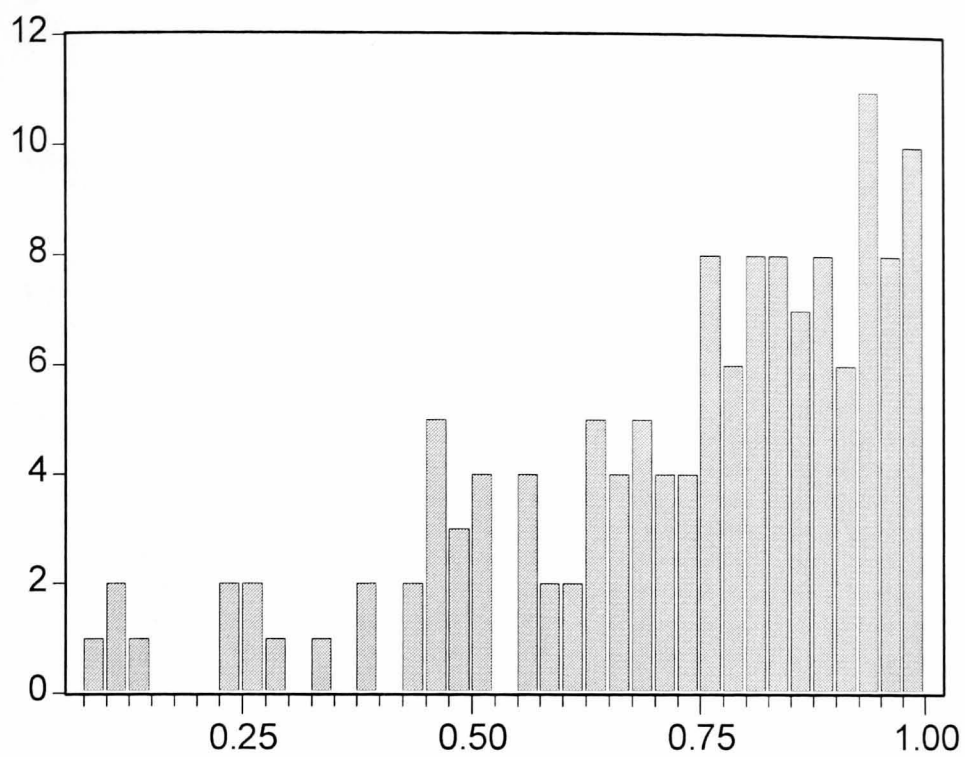


Figure 13: Histogram of hyperparameters in transition equations of KOS200 idiosyncratic volatilities in SSVF

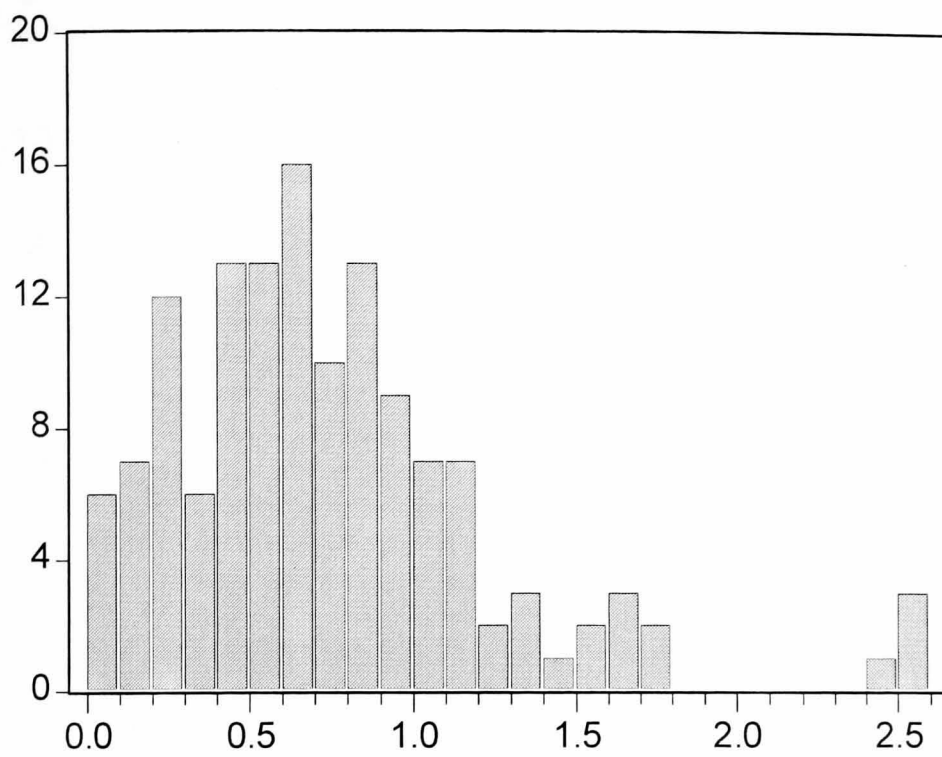


Figure 14: Histogram of hyperparameters in measurement equations of KOS200 idiosyncratic volatilities in MSVF

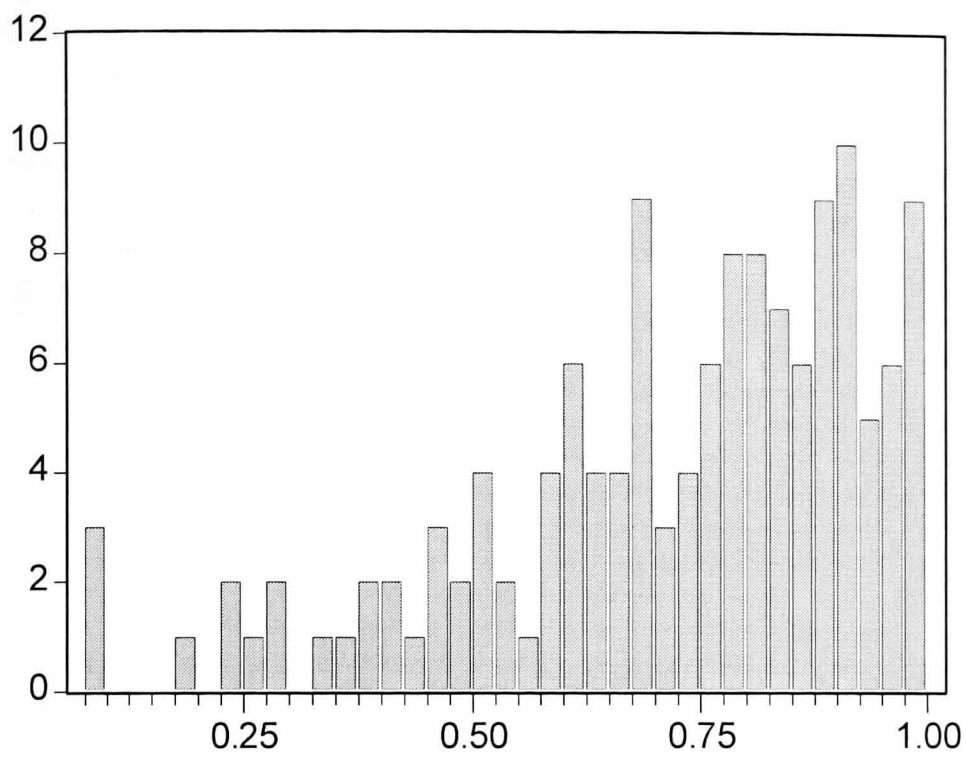


Figure 15: Histogram of hyperparameters in transition equations of idiosyncratic volatilities of KOS200 in MSVF

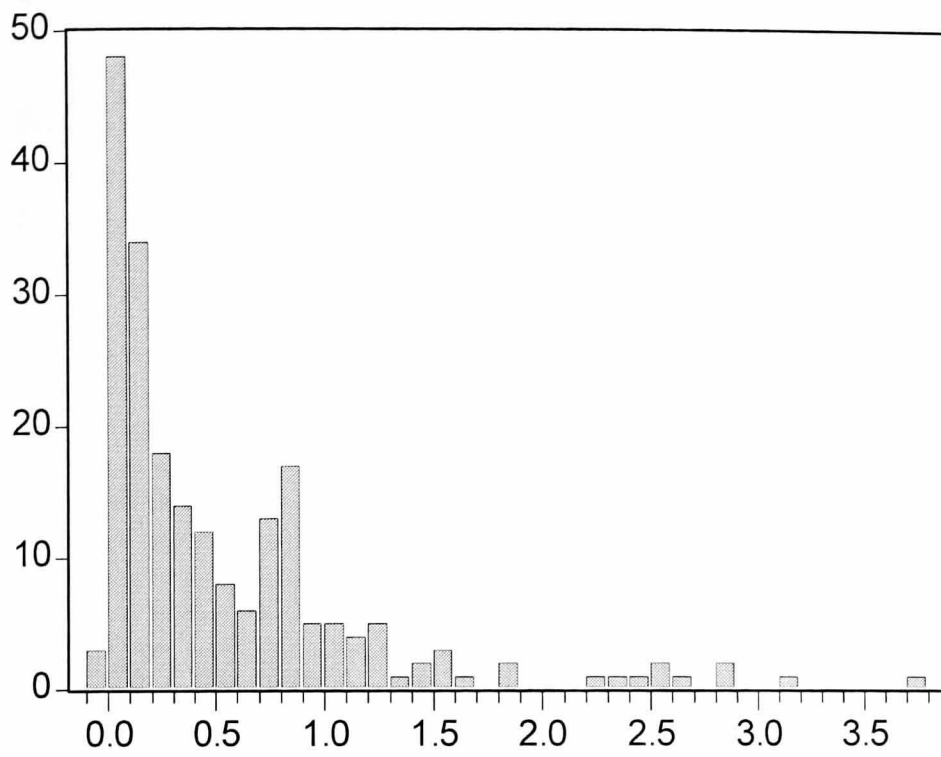


Figure 16: Histogram of hyperparameters in measurement equations of NIK225 idiosyncratic volatilities in SSVF

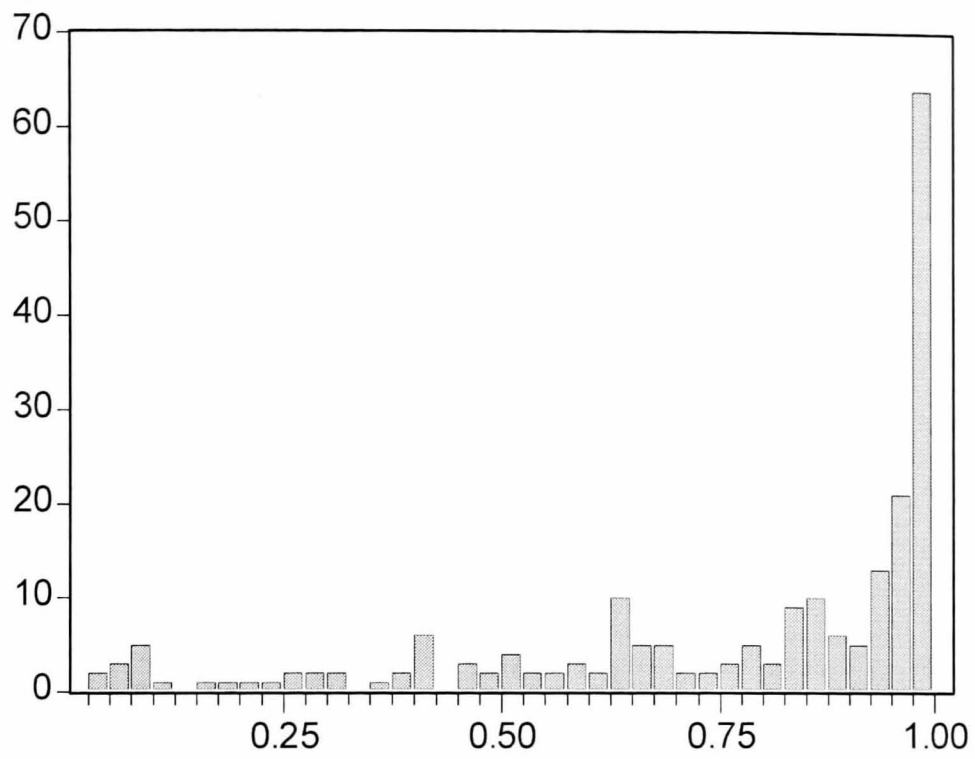


Figure 17: Histogram of hyperparameters in transition equations of NIK225 idiosyncratic volatilities in SSVF

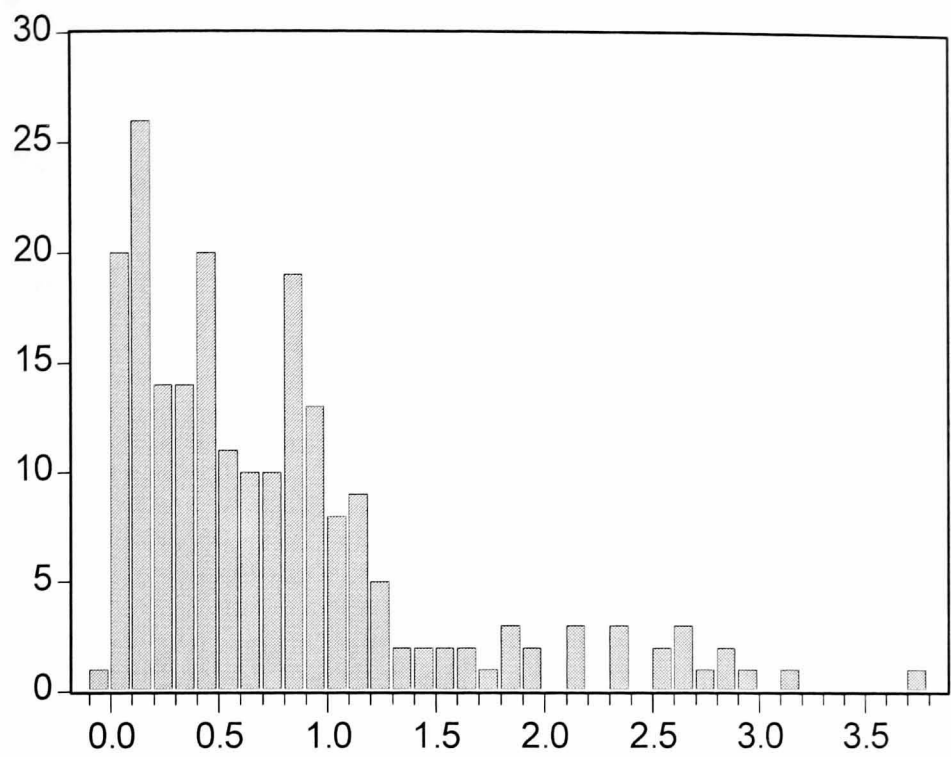


Figure 18: Histogram of hyperparameters in measurement equations of NIK225 idiosyncratic volatilities in MSVF

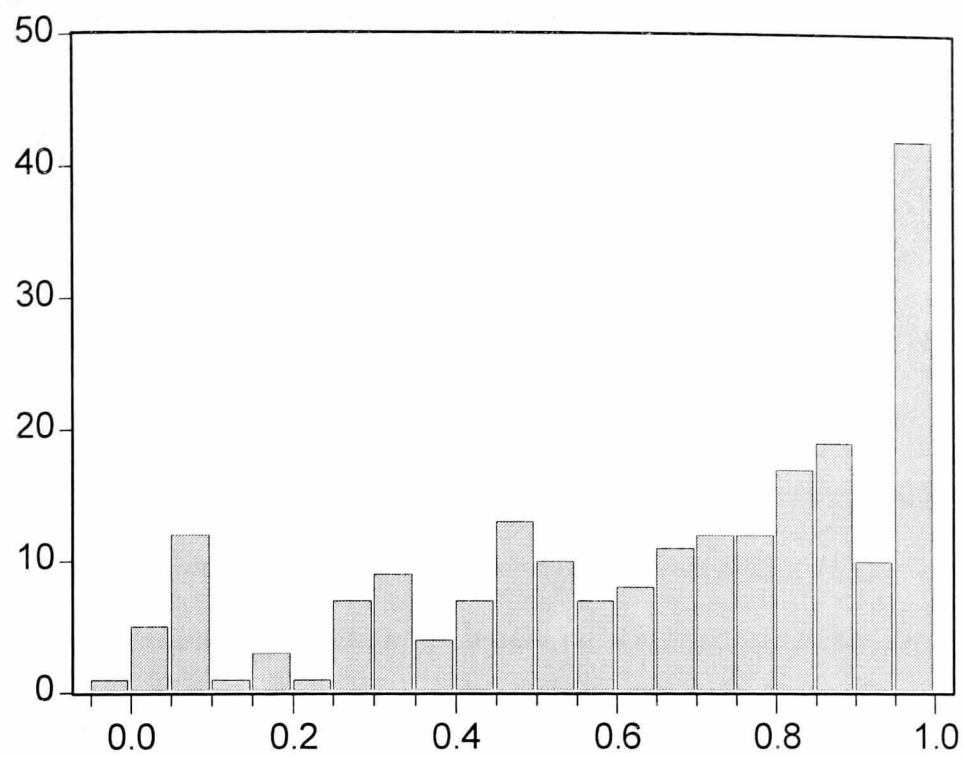


Figure 19: Histogram of hyperparameters in transition equations of NIK225 idiosyncratic volatility in MSVF

Conclusion

This thesis is a study of stock volatility using two comparable factor volatility models for large datasets, the orthogonal GARCH model and the stochastic volatility factor model. We have made an application to the constituent stocks of five Asian indexes. Analysis in the first chapter is carried out using the orthogonal GARCH model. Our empirical results confirm that higher correlation among stock return series contribute to stronger explanatory power of the principal components. We also provide a proof to show that a factor that is capable to represent a large amount of comovements in the dataset of returns can mimic the behaviour of the mean returns. Moreover, a factor extracted from a dataset of returns also exhibit the same distributional characteristics commonly observed in stock returns. We have also given an example to show how GARCH analysis of the entire dataset of returns can be summarised by a univariate GARCH(1,1) analysis of the first principal component. Diagnostic tests on the idiosyncratic returns reveal the fact that more principal components for modelling may not guarantee an improvement in capturing dynamics of the datasets due to extra noise embedded in the relatively less important principal components being introduced into the model.

In the second chapter, our factor analysis is carried out using the stochas-

tic volatility factor model. Factors are extracted from the constituent stock volatilities of the Asian indexes. Our correlation analysis and the examination of explanatory power of factors have shown similar results to that of the first chapter when analysis is carried out on the principal components extracted from dataset of returns. We have found evidence to confirm larger datasets with more data series from the same category may not always be desirable for factor analysis. Investigation of the dynamics of factor estimate has shown that dominant factors from datasets of volatilities that has strong explanatory power exhibit long memory.

An evaluation of the practical performance of the orthogonal GARCH model and the stochastic volatility factor models is implemented in the last chapter of this thesis via a volatility forecasting exercise. We discuss the extension of the original single local-factor model of Cipollini and Kapetanios (2005) into several multi-factor specifications. We test for factor significance using the statistical inference of principal component factor estimate by Bai and Ng (2006). Our results show no stronger volatility transmission in general throughout our sampling period. Our volatility forecasting exercise show that the factor models outperform the orthogonal GARCH models in forecasting the Japanese and South Korean stock volatilities. We argue this finding may be due to the shortcoming of the orthogonal GARCH model in working with system of data series that has low correlations. We also discussed how volatility forecasts for a portfolio of stocks can be improved by using a mixture of single-factor and multi-factor model so that individual characteristics of each stocks can be considered. Our experiment gives indication that improvement may be

achieved by adopting such a procedure.

Throughout our discussion in this thesis, we have made concerns on further improvement of our research work. First of all, in view of empirical studies of the long memory and structural breaks, we believe it will be interested to further examine whether long memory found in the dominant factors of the Asian stock volatilities is caused by some neglected structural breaks. Testing for breaks in factor estimates can be carried out. Secondly, evidence of persistence found in the idiosyncratic volatilities in the stochastic volatility factor models may imply unexplained long memory in this idiosyncratic components. Further examination of remaining long memory and whether idiosyncratic models with long memory can improve the forecasting performance of the factor models may be implemented. Comparison can be made with the orthogonal GARCH model when the principal component is modelled by an univariate long memory GARCH process.

As mentioned in the introductory chapter of this thesis, Asian stock markets have significant roles in international investments and a close linkages with stock markets around the world. Study financial contagion in the case of extreme events in the context of factor volatility models will be an interesting topic to look at. Our study in this thesis does not focus on the impact of extreme events on Asian volatilities and the volatility transmission during period of those events using factor analysis. However, the fact that the sampling periods used by our forecasting exercise is long enough to cover the periods of 1997 financial crisis, the Dot-com bubble and the September 11 terrorists attacks will allow us to further our research into looking at those sub-sampling

periods and allow us to gain insight into the modelling and forecasting of Asian stock volatilities in the context of factor volatility models. We hope follow-on research work for this thesis can contribute to the progression of the factor volatility models for large datasets and improve their power in explaining empirical phenomena.

References

- [1] Alexander, C. (2001a), "A primer on the Orthogonal GARCH model", University of Reading discussion paper
- [2] Alexander, C. (2001b), "Orthogonal GARCH", *Mastering Risk*, Volume 2
- [3] Alexander, C., and Chibuma, A. M. (1997), "Multivariate orthogonal factor GARCH", University of Sussex Discussion Papers in Mathematics
- [4] Andersen, T. G. (1994a), "Bayesian analysis of stochastic volatility models: Comments", *Journal of Business and Economic Statistics*, 12, Issue 4, 389-392
- [5] Andersen, T. G. (1994b), "Stochastic autoregressive volatility: a framework for volatility modelling", *Mathematical Finance*, 4, 75-102
- [6] Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X. (2006), "Volatility Forecast", in G. Elliott, C. W. J. Granger, and A. Timmermann eds, *Handbook of Economic Forecasting*, North Holland
- [7] Andersen, T. G., Bollerslev, T., and Diebold, F. X., and Labys, P. (2001). "The distribution of realised exchange rate volatility", *Journal of the American Statistical Association*, 96, 42-45
- [8] Andersen, T. G., and Sørensen, B. E. (1996), "GMM estimation of a stochastic volatility model: a Monte Carlo study", *Journal of Business and Economic Statistics*, 14, 328-352

- [9] Baba, Y., Engle, R. F., Kraft, D. F., and Kroner, K. F. (1991), "Multivariate simultaneous generalised ARCH", Unpublished manuscript, Department of Economics, University of California,
- [10] Bai, J. (2003), "Inferential theory for factor models of large dimensions", *Econometrica*, Vol. 71, No. 1, 135-171
- [11] Bai, J., and Ng, S. (2002), "Determining the number of factors in approximate factor models", *Econometrica*, 70, 191-221
- [12] Bai, J., and Ng, S. (2006), "Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions", *Econometrica*, 74, 1133-1150
- [13] Baillie, R. T. (1996), "Long memory process and fractional integration in econometrics", *Journal of Econometrics*, 73, 5-59
- [14] Baillie, R. T., and Bollerslev, T. (1989), "The message in daily exchange rates: a conditional variance tale", *Journal of Business and Economic Statistics*, 7, 297-305
- [15] Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996), "Fractionally integrated generalized autoregressive conditional heteroscedasticity", *Journal of Econometrics*, 74, 3-30
- [16] Barndroff-Nielsen, O. E., and Shephard, N. (2002), "Econometric analysis of realised volatility and its use in estimating stochastic volatility models", *Journal of the Royal Statistical Society. Series B* 64, 253-280

- [17] Beran, J. (1994), *Statistics for long-memory processes*, Chapman and Hall
- [18] Beran, J. (1995), "Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models", *Journal of Royal Statistical Society Series B*, 57, No. 4, 689-672
- [19] Beveridge, S., and Oickle, C. (1993), "Estimating fractionally integrated time series models", *Economic Letters*, 43, 137-142
- [20] Bilson, C. M., Brailford, T. J., and Hooper, V. J. (2001), "Selecting macroeconomic variables as explanatory factors of emerging stock market returns", *Pacific-Basin Finance Journal*, 9, 401-426
- [21] Boivin, J., and Ng, S. (2003), "Are more data always better for factor analysis?", National Bureau of Economic Research, Working Paper 9829
- [22] Bollerslev, T. (1986), "Generalized autoregressive conditional heteroscedasticity", *Journal of Econometrics*, 51, 307-327
- [23] Bollerslev, T. (1990), "Modelling the coherence in short-run nominal exchange rates: A multivariate generalised ARCH approach", *Review of Economics and Statistics*, 72, 498-505
- [24] Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988), "A capital asset pricing model with time varying covariances", *Journal of Political Economy*, 96, 116-131

- [25] Bollerslev, T., and Wooldridge, J. M. (1992), "Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances", *Econometric Reviews*, 11, 143-172
- [26] Breidt, F. J., Crato, N., and de Lima, P. (1998), "The detection and estimation of long memory in stochastic volatility", *Journal of Econometrics*, 83, 325-348
- [27] Byström, H. N. E. (2004), "Orthogonal GARCH and covariance matrix forecasting: The Nordic stock markets during the Asian financial crisis 1997-1998", *The European Journal of Finance*, 10
- [28] Chib, S., Nardari, F., and Shephard, N. (2002), "Analysis of high dimensional multivariate stochastic volatility models", Washington University in St. Louis Working Paper
- [29] Chung, C. F., and Baillie, R. T. (1993), "Small sample bias in conditional sum-of-squares estimators of fractionally integrated ARMA models", *Empirical Economics*, 18, 791-806
- [30] Chung, C. F. (1998), "Estimating the fractionally integrated GARCH model", Institute of Economics, Academia Sinica, Working Paper
- [31] Cipollini, A., and Kapetanios, G. (2005), "A stochastic variance factor model for large datasets and an application to S&P data", Queen Mary, University of London, Working Paper 506

- [32] Cox, D. R., and N. Reid (1987), "Parameter orthogonality and approximate conditional inference (with discussion)", *Journal of the Royal Statistical Society Series B*, 49, 1-39
- [33] Danielsson, J. and Richard, J. F. (1993), "Accelerated Gaussian importance sampler with application to dynamic latent factor variable models", *Journal of Applied Econometrics*, 8, S153-S174
- [34] Danielsson, J. (1994), "Stochastic volatility in asset prices: estimation with simulated maximum likelihood", *Journal of Econometrics*, 64, 375-400
- [35] Davidson, J. (2004), "Moment and memory properties of linear conditional heteroscedasticity models, and a new models", *Journal of Business and Economic Statistics*, Vol. 22, No. 1, 16-29
- [36] Diebold, F. X., and Mariano, R. S. (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, 13, 253-263
- [37] Diebold, F. X., and Nerlove, M. H. (1989), "The dynamic of exchange rate volatility: a multivariate latent factor ARCH model", *Journal of Applied Econometrics*, 4, 1-21
- [38] Ding, Z. (1994), "Time series analysis of speculative returns", Ph.D. Thesis, UCSD
- [39] Ding, Z., and Granger, C. W. J. (1996), "Modelling volatility persistence with speculative returns: a new approach", *Journal of Econometrics*, 73, 185-215

- [40] Ding, Z., Granger, C. W. J., and Engle, R. F. (1993), "A long memory property of stock market returns and a new model", *Journal of Empirical Finance*, 1, 83-106
- [41] Duffie, D., and Singleton, K. J. (1993), "Simulated moment of estimation of Markov models of assets prices", *Econometrica*, 61, 929-952
- [42] Dungey, M., Martin, V., and Pagan, A. (2000), "A multivariate latent factor decomposition of international bond yield spreads", *Journal of Applied Econometrics*, 16, 697-715
- [43] Durbin, J., and Koopman, S. J. (2001), *Time series analysis by state space methods*, Oxford University Press
- [44] Enders, W. (2004), "Applied econometric time series", 2nd Edition, John Wiley and Sons, Inc.
- [45] Engle, R. F. (1982), "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", *Econometrica*, 50, 987-1007
- [46] Engle, R. F. (2002), "Dynamic conditional correlation – a simple class of multivariate GARCH models", *Journal of Business and Economic Statistics*, 20, 339-350
- [47] Engle, R. F., and Bollerslev, T. (1986), "Modelling the persistence of conditional variances", *Econometric Review*, 5, 1-5

- [48] Engle, R., and Colcacito, R. (2006), "Testing and valuing dynamic correlations for asset allocation", *Journal of Business and Economic Statistics*, Vol. 24, No. 2, 238-253
- [49] Engle, R. F., Ng, V., and Rothschild, M. (1990), "Asset pricing with a factor ARCH covariance structure: Empirical estimates for treasury bills", *Journal of Econometrics*, 45, 213-238
- [50] Engle, R. F., and Kroner, K. F. (1995), "Multivariate simultaneous generalised ARCH", *Econometric Theory*, 11, 122-150
- [51] Franses, P. H., and van Dijk, D. (2000), *Non-linear time series models in empirical finance*, Cambridge University Press
- [52] Gallant, A. R., Hsieh, D. A., and Tauchen, G. (1997), "Estimation of stochastic volatility models with diagnostics", *Journal of Econometrics*, 81, 159-192
- [53] Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993), "On the relation between expected value and volatility of the nominal excess return on stocks, *Journal of Finance*, 48, 1779-1801
- [54] Granger, C. W. J. (1980), "Long memory relationships and the aggregation of dynamic models", *Journal of Econometrics*, 14, 227-238
- [55] Granger, C. W. J. (1981), "Some properties of time series data and their use in econometric model specification", *Journal of Econometrics*, 16, 121-130

- [56] Granger, C. W. J., and Hyung, N. (2004), "Occasional structural breaks and long memory with an application to the S & P 500 absolute stock returns", *Journal of Empirical Finance*, Vol. 11, Issue 3, 399-421
- [57] Granger, C. W. J., and Joyeux, R. (1980), "An introduction to long memory time series models and fractionally differencing", *Journal of Time Series Analysis*, 1, 15-39
- [58] Greene, W. H. (2003), "Econometric Analysis", 5th Edition, Prentice Hall
- [59] González-Rivera, G. (1998), "Smooth transition GARCH models", *Studies in Nonlinear Dynamics and Econometrics*, 3, 61-78
- [60] Gouriéroux, C., Monfort, A., and Renault, E. (1993), "Indirect inference", *Journal of Applied Econometrics*, 8, S85-S118
- [61] Hamilton, J. D. (1994), *Time series analysis*, Princeton University Press
- [62] Harvey, A. C. (1989), *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press
- [63] Harvey, A. C. (1998), "Long memory in stochastic volatility", in J. Knight and S. Satchell (eds), *Forecasting volatility in the financial markets*, Oxford: Butterworth-Heinemann, pp. 307-320
- [64] Harvey, A., Ruiz, E., and Shephard, N. (1994), "Multivariate stochastic variance models", *Review of Economic Studies*, 61, 247-264
- [65] Hosking, J. R. M. (1981), "Fractional differencing", *Biometrika*, 68, 165-176

- [66] Hull, J., and White, A. (1987), "The pricing of options and assets with stochastic volatilities", *Journal of Finance*, 42, 281-300
- [67] In, F., Kim, S., Yoon, J. H., and Viney, C. (2001), "Dynamic interdependence and volatility transmission of Asian stock markets: Evidence of Asian crisis", *International Review of Financial Analysis*, 10, 87-96
- [68] Jacquier, E., Polson, N. G., and Rossi, P. E. (1994), "Bayesian analysis of stochastic volatility models", *Journal of Business and Economic Statistics*, 12, 371-417
- [69] Jolliffe, I. T. (1986), "Principal Component Analysis", Springer-Verlage New York, Inc.
- [70] Kapetanios, G. (2004), "A new method for determining the number of factors in factor models with large dataset", Queen Mary, University of London, Working Paper 525
- [71] Kim, J., Kartsaklas, A., and Karanasos, M. (2005), "The volume-volatility relationship and the opening of the Korean stock market to foreign investors after the financial turmoil in 1997", *Asian-Pacific Financial Markets*, 12(3), 245-271
- [72] Klalid, A. M., and Kawai, M. (2003), "Was Financial market contagion the source of economic crisis in Asia? Evidence using a multivariate VAR model", *Journal of Asian Economics*, 14, 131-156

- [73] Kim, S., Shephard, N., and Chib, S. (1998), "Stochastic volatility: Likelihood inference and comparison with ARCH models", *Review of Economic Studies*, 65, 361-393
- [74] King, M., Sentana, E., and Wadhvani, S. (1994), "Volatility and links between national stock markets", *Econometrica*, 62, 901-933
- [75] Knight, J., and Satchell, S. (2007), *Forecasting volatility in the financial markets*, Oxford: Butterworth-Heinemann, 3rd edition
- [76] Lobato, I. N., and Robinson, P. M. (1998), "A nonparameteric test for $I(0)$ ", *The Review of Economic Studies*, Vol. 65, No. 3, 475-495
- [77] Li, K. W. (2006), *The Hong Kong Economy: Recovery and Restructuring*, McGraw-Hill Education
- [78] Masih, A., and Masih, R. (1999), "Are Asian Stock Market Fluctuations Due Mainly to Intra-Regional Contagion Effects? Evidence Based on Asian Emerging Stock Markets", *Pacific-Basin Finance Journal*, 7(3,4), 251-282.
- [79] Melino, A., and Turnbull, S. M. (1990), "The pricing of foreign currency options with stochastic volatility", *Journal of Econometrics*, 45, 239-265
- [80] Miyakoshi, T. (2003), "Spillovers of stock return volatility to Asian equity markets from Japan and the US", *Journal of International Financial Markets, Institutions and Money*, 13, 383-399
- [81] Nelson, D. B. (1990), "Stationarity and the persistence in the GARCH(1,1) model", *Econometric Theory*, 6, 318-334

- [82] Nelson, D. B. (1991), "Conditionally heteroscedasticity in asset returns: a new approach", *Econometrica*, 59, 347-370
- [83] Ng, A. (2000), "Volatility spillover effects from Japan and the US to the Pacific-Basin", *Journal of International Money and Finance*, 19, 207-233
- [84] Ooms, M., and Doornik, J. (1999), "Inference and forecasting for fractionally autoregressive integrated moving average models, with an application to US and UK inflation", *Econometric Institute Report 9947/A*, Erasmus University Rotterdam
- [85] Psaradakis, Z., and Sola, M. (1995), "Modelling long memory in stock market volatility: a fractionally integrated generalised ARCH approach", Centre for Economic Forecasting, London Business School, Discussion Paper No, DP 12-95
- [86] Poon, S. H., and Granger, C. W. J. (2003), "Forecasting volatility in financial markets: A review", *Journal of Economic Literature*, Vol. XLI (June 2003), 478-539
- [87] Rabemananjara, R. and Zakoian, J. M. (1993), "Threshold GARCH models and asymmetries in volatility", *Journal of Applied Econometrics*, 8(1), 31-49
- [88] Robinson, P. M. (1995), "Log-periodogram regression of time series with long range dependence", *The Annals of Statistics*, Vol. 23, No. 3, 1048-1072

- [89] Sentana, E. (1995), "Quadratic ARCH models", *Review of Economic Studies*, 62, 639-661
- [90] Shephard, N. (1993), "Fitting non-linear time series models, with applications to stochastic variance models", *Journal of Applied Econometrics*, 8, S135-S152
- [91] Shephard, N. (2005), *Stochastic volatility: selected readings*, Oxford University Press
- [92] So, M. K. P., Lam, K., and Li, W. K. (1997), "An empirical study of volatility in seven Southeast Asian stock markets using ARV models", *Journal of Business and Accounting*, 24(2), 261-275
- [93] Sowell, F. B. (1992), "Maximum likelihood estimation of stationary univariate fractionally integrated time series models", *Journal of Econometrics*, 53, 165-188
- [94] Stock, J. H., and Watson, M. W. (1998), "Diffusion Indexes", National Bureau of Economic Research, Working Paper 6702
- [95] Stock, J. H., and Watson, M. W. (2002a), "Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economic Statistics*, Vol. 22, No. 2, 147-162
- [96] Stock, J. H., and Watson, M. W. (2002b), "Forecasting using principal components from a large number of predictors", *Journal of the American Statistical Association*, Vol. 97, No. 460

- [97] Taylor, S. J. (1982), "Financial return modelled by the product of two stochastic processes – a study of daily sugar prices 1961-1979", In Andersen, O. D. (ed.), *Time series analysis: Theory and Practice*, 1, 203-226. North-Holland
- [98] Taylor, S. J. (1986), *Modelling financial time series*, John Wiley
- [99] van der Weide, R. (2002), "GO-GARCH: a multivariate generalised orthogonal GARCH model", *Journal of Applied Econometrics*, 17, 549-564
- [100] Wright, J. H. (1999), "Long memory in emerging market stock returns", International Finance Discussion Paper No. 650, Board of Governors of the Federal Reserve System
- [101] Whittle, P. (1951), *Hypothesis testing in time series analysis*, Hafner, New York