Department of Economics

An Empirical Study of Asian Stock Volatility Using Stochastic Volatility Factor Model: Factor Analysis and Forecasting

Silvia S.W. Lui

Working Paper No. 581 December 2006 ISSN 1473-0278

Queen Mary
University of London

An Empirical Study of Asian Stock Volatility Using Stochastic Volatility Factor Model: Factor Analysis and Forecasting

Silvia S. W. Lui*
Queen Mary, University of London

4 October 2006
(Manuscript)

Abstract

This paper is an empirical study of Asian stock volatility using stochastic volatility factor (SVF) model of Cipollini and Kapetanios (2005). We adopt their approach to carry out factor analysis and to forecast volatility. Our results show some Asian factors exhibit long memory that is in line with existing empirical findings in financial volatility. However, their local-factor SVF model is not powerful enough in forecasting Asian volatility. This has led us to propose an extension to a multi-factor SVF model. We also discuss how to produce forecast using this multi-factor model.

JEL Codes: C32, C33, C53, G15

Keywords: Stochastic Volatility, Local-factor Model, Multi-factor Model, Principal Components, Forecasting

1 Introduction

Studies of underlying forces that cause variation in stock returns 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 as a proxy of common factors. These variables are believed to have made contribution to

*Department of Economics, Queen Mary, University of London, Mile End Rd., London E1 4NS. Email: S.W.S.Lui@qmul.ac.uk. Special thanks to Professor George Kapetanios for valuable suggestions and comments.

the fluctuation in a particular stock market. For example, Bilson, Brailford and Hooper (2001) employ a multi-factor model to select common explanatory factors for emerging markets from a set of global risk variables and local economic variables. This approach relies on observable significance of the "proxy common factors" in accounting stock return variation. However, this may be quite restrictive in the sense that if stock returns 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 state-space model. Stock and Watson (1998, 2002a) call these common factors "diffusion indexes". These unbserved common factors summarize the information from a large group of driving forces that account for variation in a dataset. These driving forces may be generated by both macroeconomic variables and some other unknown forces. No pre-definition of the common factors is made. Thus, state-space factor analysis is relatively more flexible.

However, there are two types of deficiencies in simple state-space factor analysis. First of all, basic state-space representation may not be sufficient for the analysis of stock returns due to the existence of heteroskedasticity. A simple state-space representation does not capture time varying volatility in financial time series. The famous ARCH and GARCH family models by Engle (1982) and Bollerslev (1986) have been used extensively in financial econometric analysis. However, Harvey, Ruiz and Sheperd (1994) argue that the multivariate version of GARCH family models are inconvenient for estimation and interpretation due to large number of parameters and the need to impose constraints. They therefore, propose a multivariate stochastic volatility (SV) model and most importantly, suggest incorporation of common factor into the model. Although their model is a better alternative than multivariate GARCH models, common factor embedded in the model is assumed to follow a random walk. This will become insufficient if the factor actually have more complex temporal features. The SV model is further generalized into a multivariate stochastic volatility factor (SVF) model and to overcome this constraint by Cipollni and Kapetanios (2005). No prior assumptions to the form of the underlying process of common factor is required in the SVF model. More complex dynamics can thus be captured. Apart from this, their model is very flexible since stochastic volatility consists of a common component and a idiosyncratic component, and these two components are allowed to have different underlying processes. Thorough study of common fluctuaion in a market and fluctuation in return that is unique for individual stocks can then be carried out via analysis of the two components. This model setting is highly applicable to empirical analysis of financial volatility.

Second type of deficiencies in simple state-space factor analysis comes from its conventional estimation method. Maximum likelihood estimation has been widely used to estimate common factors in the state-space factor analysis. However, application of maximum likelihood is not feasible if we are dealing with a large dataset that includes a lot of cross-section series as this involves a large number of parameters being estimated. The suggestion of using principal components estimation for extracting common factors in an approximate dynamic factor model by Stock and Watson (2002) is a breakthrough in factor analysis with large dataset. Both of them, and Bai (2003) show this method provides consistent estimates in large datasets. Bai also study the asymptotic properties of this estimator. In particular, Cipollini and Kapetanios (2004) recommend the use of principal components method to estimate common factors in their SVF model. All these are remarkable developement in dynamic factor analysis.

It can be seen from existing studies of financial volatility forecast that ARCH and GARCH family models and their further developments such as EGARCH of Nelson (1991), IGARCH of Engle and Bollerslev (1986) and FIGARCH of Baillie, Bollerslev, and Mikkelsen (1996); as well as the SV model, have been extensively used. However, due to the limitations of both types of models we have discussed above, they are not preferable in our study. Our forecasting exercise is carried out by using the SVF model and the procedure used in the Cipollini and Kapetanios (2005) paper.

We regard the SVF model they proposed and use in forecasting volatility as a local-factor specification since it only include the dominant factor of US stocks. When using it to forecast US stock volatility, it is not surprising to find out that a local-factor model performs better than other volatility models concerned. However, whether a local-factor model remains sufficient when we are dealing with Asian stock markets and, especially emerging markets, remains unknown. We believe there is a need to further extend their model into a multi-factor SVF model. This belief arises from the fact that volatility transmission from one market to another is actually happening in financial world. A discussion regarding this proposal is included in this paper.

If one has chosen to use a multi-factor model instead of a local-factor model, then the next question one may ask is how to decide on an appropriate specification of a multi-factor model? In particular, how to check out if a common factor is useful in explaining volatility in a local market? In the past few year, an increasing number of literatures on factor analysis has been focusing on analysing the characteristics of common factors estimated by the mean of principal components method, and the empirical applications of factor models. For example, in a study on inferential theory for factor models of large dataset, Bai (2003) derives convergence rate and limiting distribution of factor estimates, factor loadings and common components. Literatures that concern the practical use of factor models centred on producing macroeconomic forecasts and financial forecasts with models that include factor estimates extracted by principal components method. See for example, Stock and Watson (2001, 2002a, 2002b) and Cipollini and Kapetanios (2005).

Although there have been increasing concerns about statistical inferences on parameters in factor models, not a lot of the existing literatures seem to have provided a thorough enough study on this issue. Thus, how to carry out inference on those factor models or factor-augmented models are not well known. Establishing statistical inference of factor models is important because common factors are estimated rather than observed. Therefore, they should be treated with extra care. Until recently, Bai and Ng (2006) have made a significant contribution to the issue by carrying out a detailed study on statistical inferences of prinicipal components factor estimates via the use of factor-augmented regression (FAR). In their study, they derive rate of convergence and limiting distribution of least square parameter estimates in FAR that includes prinicipal components factor estimates as regressors. As a result, confidence intervals of least square estimates, conditional mean and forecast can be constructed. It also allow hypothesis tests to be carried out. Based on their results, we believe significance test on factor estimates can be carried out and the test results will be useful indication for deciding which common factors should go into our multi-factor SVF model.

This paper is an empirical study of Asian stock volatility using the SVF model of Cipollini and Kapetanios (2005). Our empirical analysis consists of two parts. In the first part, we carry out an analysis of the common factors that underlie Asian stock volatility and we study the dynamics of these factor estimates. Whereas the second part concerns a forecasting exercise using the SVF model. Organization of this paper is as follow. Section 2 explains the model and methodology used in this paper. We elaborate how the SV model of Harvey, Ruiz and Sheperd (1994) is extended into the SVF model of Cipollini and Kapetanios (2004). Preliminary statistical analysis is in section 3. Section 4 is a demonstration of our empirical analysis. First of all, we carry out an analysis of the common factors of the constituents of five Asian stock indices extracted by prinicipal component method using data from the entire sample. In particular, we investigate the explanatory power and dynamics of those factor estimates. We then move on to perform a forecasting exercise. We first present some in-sample results, then follow by results of stock volatility forecast using the SVF model. We compare forecasting performance of single local-factor SVF model and univariate State-space model. Having observed the forecasting performance, we believe there is a need to extend the single local-factor SVF model into a multi-factor SVF model. In section 5, we discuss the reason for such an extension and propose several specifications of the multi-factor model. We also explain what procedure we take to determine an appropriate specifications of a multi-factor model and how forecast is produced.

2 Model and methodology

2.1 Stochastic volatility factor model (local-factor specification)

In the paper by Harvey, Ruiz and Sheperd (1994), they propose a multivariate stochastic volatility model (SV). The multivariate SV model is a better alternative to the multivariate version of the GARCH family models for financial time series featured with time-varying volatility and serial correlation, due to the constraints in terms of estimation and interpretation that the latter have. Consider there are N stocks in a dataset. Multivariate SVmodel of daily stock returns takes the following form

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

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

$$\widetilde{y}_{i,t} = a_i + h_{i,t} + \zeta_{i,t} \tag{2}$$

where $\widetilde{y}_{i,t}$ 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. make two suggestions to model the logarithm of 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 SV model, i.e. the $N \times 1$ vector unobserved variance, $h_t = \theta f_t$, where f_t is a K-dimensional vector of common factor, $f_t = (f_{1,t}, \cdots, f_{K,t})'$. Following the latter suggestion, one can model the tranformed squared returns and the common factor using a State-Space representation. 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.

Cipollini and Kapetanios (2005) propose a generalization to the SV model, the Stochastic Volatility Factor (SVF) Model, that improves the shortcomings of the state-space version of SV model mentioned above. In their SVF model, unobserved variance h_t features a common component and a disturbance that is unique to each individual stock. That is,

$$h_{i,t} = \theta_i' f_t + \eta_{i,t} \tag{3}$$

substitute equation (3) into (2) gives us the SVF model

$$\widetilde{y}_{i,t} = a_i + \theta_i' f_t + \omega_{i,t} \tag{4}$$

where $\omega_{i,t} = \eta_{i,t} + \zeta_{i,t}$ is the idiosyncratic volatility of stock i. One thing we can see from equation (4) is that stochastic volatility is determined by a common component $\theta_i' f_t$, and a idiosyncratic component, $\omega_{i,t}$.

Common component constitues the part of stock volatility that is caused by variations in returns on all stocks in the dataset. In other words, it is the contribution to the variation in individual stock return made by market fluctuation. Idiosyncratic component, on the other hand, summarises the part of variation in return that is mainly due to individual stock and so, this is the fluctuation which is unique to the stock. Comparing SV and SVF model, we can see 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 general underlying processes that best describe their empirical temporal properties. Thus, dynamics of financial volatility can be best captured. Moreover, no extra assumptions, in addition to those required by principal component estimation, is required for SVF model to provide consistent factor estimates. All these advantages make the SVF model easy to use, more apply to empirical situation and very tractable.

It is worth mentioning that in their study of S&P 500 constituent stock, Cipollini and Kapetanios (2005) use their multivariate SVF model, which involves only the S&P 500 dominant factor, to forecast US stock volatility. Therefore, their model in the form of equation (4) is actually a local-factor specification (namely local-factor SVF model hereafter). Their finding shows local-factor SVF model outperforms other a selection of volatility models in forecasting volatility. Since the US market is the world leading market, its local fluctuation have a lot more significant impact on other markets in the world than fluctuation of 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 believe its local factor is enough to explain stock volatility of its own market. However, whether local dominant factors of other stock markets in the world, especially those of emerging markets, have contained enough information to produce an accurate volatility forecast for those markets remains a question. From the result of our forecasting exercise, we can see there is a need to further extend the local-factor SVF model into a multi-factor setting. Discussion and proposal of a multi-factor SVF model can be found in section 5. We now move on to the discussion of the estimation method used for SVF model, the principal component method of Stock and Watson (2002), in the next subsection.

2.2 Estimation of common factors using Principal Components

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. It is because 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 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 common 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 assymptotically equivalent to maximum likelihood estimation. In their paper, Stock and Watson (2002b) study the finite sample properties of principal component estimator. They show that 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, under rather general assumptions. Bai (2003) also shows that the necessary conditions for ensuring consistency are asymptotic orthogonality and asymptotic homoskedasticity in idiosyncratic innovations.[1] Consistency in factor estimates can be obtained even in the presence of serial correlations and heteroskedasticity.

It can be seen easily from the SVF model in equation (4) that stock volatility contains two components: the common component that involves the common factor and the idiosyncratic component that summarises variation in returns that is unique to individual stock. For the common factor, Cipollini and Kapetanios (2005) propose that it should be extracted from the dataset of stocks, using principal component estimation. Extending the results of Bai (2003), they point out that principal components estimation can still provide consistent factor estimates if factor process is stationary long memory $ARFIMA(p, d, q)$ with shocks that have finite fourth moment. Their Monte Carlo analysis shows this factor estimation method performs well in estimating the SVF model. We therefore, follow their approach by using the same method for factor estimation. Estimation of factors and factor loadings by this method require minimising the objective function

[1] Bai (2003) calls the restrictions
$N^{-1}\sum_{i=1}^{N}\xi_{i,t}\xi_{i,s} \to 0$, for $t \neq s$
and $N^{-1}\sum_{i=1}^{N}\xi_{i,t} \to \sigma^2$, for all t as N tends to ∞
asymptotic orthogonality and asymptotic homoskedasticity, respectively.

$$V(f,\Theta) = (NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widetilde{y}_{i,t} - \theta_i^{'} f_t)^2 \tag{5}$$

This is analogous to minimising the variance of idiosyncratic volatility $\omega_{i,t}$ in (4). Minimising objective function (5) 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, i.e. $T^{-1}(f'f) = I_k$, where $f = (f_1, ..., f_T)$. Estimating the K 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. $\widetilde{y}_{i,t}$ is the element in the tth row and ith column of matrix $\widetilde{Y}$. The eigenvectors we have extracted are the first K estimated factor series. Several mild assumptions on the factors, factor loadings and innovations are required for consistent estimates being produced. A thorough discussion can be found in Stock and Watson (2002b) (see also Cipollini and Kapetanios (2005)). A summary of these assumptions is shown in appendix 7.1.

2.3 About the common factors

Once common factors are estimated using principal components method, an appropriate specification for factor estimates can be obtained by carefully examining the temporal feature of the factor estimates. This is evident by a lot of existing empirical literatures that financial returns have long memory. If variations in stock returns are underdriven by the common factor, it is logical to think that the common factor should also have long memory nature. This claim is first proved by the analysis of S&P 500 factor in the Cipollini and Kapetanios (2005) paper. They suggest factor estimates follows a $ARFIMA(p, d, q)$ process which takes the following form

$$\Phi(L)(1 - L)^d(f_t - \mu) = \Psi(L)\epsilon_t \tag{6}$$

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

$$(1 - L)^d = \sum_{j=0}^{\infty}\pi_j(d)L^j \tag{7}$$

and

$$\pi_j(d) = (-1)^k \frac{\Gamma(d+1)}{\Gamma(j+1)\Gamma(d-j+1)} \tag{8}$$

where $\Gamma(\cdot)$ denotes the gamma function. The $ARFIMA(p, d, q)$ process is covariance stationary for the fractional differencing operator d lies between -0.5 and 0.5. For $0 < d < 0.5$, f_t exhibit long memory and its autocorrelations show persistence. For simplicity, we drop out the moving average terms in (6) in our empirical analysis, and thus assume the common factor follow only an $ARFI(p, d)$ process. We adopt 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.

2.4 About the idiosyncratic volatility

After removing the common component of the stochastic volatility of a particular stock i, what we are left with in equation (4) is its idiosyncratic stochastic volatility. A univariate State-Space model (SSM) is suggested to be the underlying process of this idiosyncratic component. It means idiosyncratic volatility of stock i is assumed to be driven by some unobservable forces that are summarised by the state vector. That is,

$$\omega_{i,t} = \gamma_i \eta_{i,t} + \zeta_{i,t} \tag{9}$$

$$\eta_{i,t} = \lambda_i \eta_{i,t-1} + \kappa_{i,t} \tag{10}$$

where $\zeta_{i,t} \sim N(0, q_i)$, $\kappa_{i,t} \sim N(0, 1)$, and $\eta_{i,0} \sim N(\alpha_0, p_0)$. $\eta_{i,0}$ is the initial state of stock i and it has a mean of α_0 and variance of p_0.

Equation (9) 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 $q_{i,t}$. $\eta_{i,t}$ is unobservable in general. It is generated by a first-order Markov process as in Equation (10) 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 $p_0 = E(\eta_{i,0}^2) = \frac{1}{1+\lambda_i^2}$. γ_i, λ_i, and q_i are the hyperparameters of the above univariate SSM, which are estimated via prediction error decomposition by Gaussian maximum likeihood using Kalman filter. Harvey, Ruiz and Sheperd (1994) show in their paper that Gaussian maximum likelihood can provide consistent estimates.[2]

2.5 Forecasting volatility using SVF model (local-factor specification)

In this section, we elaborate how volatility forecast is produced using SVF model. We adopt the method prosposed by Cipollini and Kapetanios (2005) to produce our forecasts for Asian stock volatility using their local-factor SVF model. Recursively forecasting scheme is used to produced one-step ahead volatility forecast for each constituent stocks in a dataset. This means when each point forecast is made, the common factor, the factor loadings as well as the idiosyncratic shocks are reestimated. Forecast produced using the SVF model in equation (4) is

$$\widehat{\widetilde{y}_{i,t+1|t}} = a_i + \theta_i' \widehat{f^L_{t+1|t}} + \omega_{\widehat{i,t+1}|t} \tag{11}$$

where $t = T, \cdots, S$ and T is the last period of the estimation sample and S is the last period of the whole sample (i.e. last period of the forecast sample). $\widehat{\widetilde{y}_{i,t+1|t}}$denotes the one-step ahead forecast for the volatility of stock i, $\omega_{\widehat{i,t+1}|t}$ denotes the one-step ahead forecast of idiosyncratic volatility, and $\widehat{f^L_{t+1|t}}$ denotes the one-step ahead forecast for the local dominant factor estimate. Equation (11) is forecast produced by local-factor model since only the dominant factor that is extract from the dataset of a local market is considered. We can see that the forecasting method for SVF model is very different from the conventional forecasting exercise. The final volatility forecast consists of a forecast using estimated data rather than empirical or observed data. Argument on whether this can provide us with an accurate forecast may arise. It is because error occurred due to factor estimation may deteriorate the fitness of SVF model and thus contributes to the disparity between the actual volatility and forecasted volatility. However, the findings by Bai (2003) shows, although there is error in modelling due to the fact that factor is estimated series rather than observed, this error is tiny and can thus be neglected if $\sqrt{T}/N \to 0$. Cipollini and Kapetanios (2005) point

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

out a factor model that involved estimated series as an explanatory varible can provide forecast for the factor estimates and idiosyncratic component.

Equation (16) shows that final forecast of volatility 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 $t = 1, \cdots, T$ be the estimation sample, and $t = T+1, \cdots, S$ be the forecast evaluation sample. So, 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 the constituents of an stock index using principal component method. New estimates of the factor, factor loadings and idiosyncratic shocks are then obtained.

2. Estimate the $ARFI(p, d)$ process of the local dominant factor using AMLE. The estimated parameters of this $ARFI(p, d)$ is then used to form a one-step ahead forecast of the local dominant factor for the next period, i.e.$\widehat{f^L_{T+1|T}}$

3. Then for each stock i in the dataset, we fit a univariate State-Space model into its idiosyncratic shock $\omega_{i,T}$ and estimate it via prediction error decomposition by maximum likelihood using Kalman filter. Next, we form a one-step ahead forecast for the idiosyncratic shock for the next period, i.e. $\hat{\omega}_{i,T+1|T}$.

4. Finally, combine the forecasts form in step 2 and 3, together with the factor loadings estimated for period T, to form the overall forecast of volatility.

$$\widehat{\tilde{y}_{i,T+1|T}} = a_i + \theta_i \widehat{f^L_{T+1|T}} + \omega_{i,\widehat{T+1}|T} \tag{12}$$

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

Our forecasting exercise in section 4.2.2 is carried out for local-factor SVF model. Forecasting volatility using a local factor model clearly ignores the fact that fluctuation in other international financial markets impacts on the fluctuation of a local market. It assumes variation in the local market is explained solely by the local 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 variation 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. Further discussion on multi-factor model and forecasting with this model can be found in section 5.

3 Data and preliminary statistical analysis

Daily data of constituent stocks of five Asian indices are obtained from Datastream. The five indices are NIKKEI 225 (NIK225) and NIKKEI 500 (NIK500) of Japan; Heng Seng Composite Index (HSCI) of Hong Kong; Korean Stock Exchange Composite Index 200 (KOSPI) of South Korea; and Stock Exchange of Singapore All Share Index (SING) of Singapore. The reason for investigating both the NIK225and NIK500 of Japan is for us to look at how size of dataset impacts on the explanatory power of a dominant factor. The sample ranges from 3 January 2000 to 30 July 2004, for a total of 1194 daily observations of stock returns. This is the period after the most difficult time in Asia due to the Japanese banking crisis in the 90's and the 1997 Asian financial turmoil. Daily returns on each constituent stocks i is calculated as

$$y_{i,t} = \ln(p_{i,t}) - \ln(p_{i,t-1})$$

where $y_{i,t}$ denotes the return on constituent stock i at time t, $p_{i,t}$ and $p_{i,t-1}$ denote price 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 KOSPI, 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 KOSPI, 1185 for SING and 1128 for HSCI.

Figure 1 displays the time plot of mean returns of the constituent stocks for the five indices. We can see from these plots that the mean return series appear stationary as what one would expect. Table 1 summarizes descriptive statistics of the mean return series. Mean returns of SING constituents and KOSPI constituents are negative on average during the sample period. Average Mean return on the constituents of HSCI is higher than that on the constituents of the other four indices. Mean return on KOSPI constituents has the highest volatility among the five, whereas mean return on SING constituents has the lowest. 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%.

4 Empirical applications

4.1 Analysis of factor estimates

4.1.1 Explanatory power of factor estimates

First ten factors are extracted from the transformed dataset of constituent stock returns for each of the five indices using principal component method. The first factor estimates for each dataset is plotted in Figure 2. 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 heteroskedasticity and serial correlation in the residuals. Table 2 shows the computed Breush-Godfrey serial correlation LM test statistics and Engle's ARCH LM test statistics, and the probabilities corresponding to these statistics [3]. Clear rejections of no serial correlation and no ARCH effect are found for the residuals of the first factor estimates for all five datasets at 5%. This indicates the estimated factor series are heteroskedastic and serially correlated and their residuals are clearly not *i.i.d.*. It also suggests the model underlying the factor should incorporate these properties.

Table 3 shows cumulative R^2 of these factors. We can see from the table that for KOSPI, NIK225 and NIK500, the first factor explained most of the

[3] 5 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 hypothese of no serial correlation and no ARCH effect, respectively.

variation in the dataset. Additional factors only have marginal contributions towards the explanatory power of the whole set of factors. The first factor of KOSPI is the most powerful one among the five, which explains 10.1% of variation for a total number of 176 stocks. The first factors for both NIK225 and NIK500 explain about 8.1% and 6.1% of the variation in the two datasets of 217 stocks and 481 stocks, respectively. Although they do not seem as powerful as as the first factor of KOSPI, they are quite good in general. We can then conclude the first factors for each of KOSPI, NIK225 and NIK500 are the dominant factors for these datasets.

However, the same story does not apply to the factors of SING and HSCI. The first factor of SING can only explain 5.3% of the variation for a dataset of 227 stocks. This amount is quite low. The first factor of HSCI explains only 1.7% of the total variation. These are not really appealing results in terms of factor analysis using principal components estimation. Existing empirical results show that if a dataset has a factor structure, then the first few factors are expected to have a good performance in accounting variation in the dataset, with the first factor explains most of the variation. (See results and discussion in Stock and Watson (2002), Cipollini and Kapetanios (2005), and Kapetanios (2004)). Our findings suggest that factor model may not be an appropriate specification for both SING and HSCI stock returns. Therefore, follow-on analysis of factor dynamics will be centered on the dominant factors of KOSPI, NIK225 and NIK500.

Moreover, attention should also be paid to the amount of explained variation for the first factors of NIK225 and NIK500 datasets. 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 heteroskedastic and cross-sectionally correlated. If more series 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.[4] In our case, the dominant factors of NIK225 and NIK550 are highly correlated with correlation coefficient

[4]Principal component estimation method by Stock and Watson (2002) shows factors

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 captured a great deal of common components that make large contributions to fluctuation in the Japanese market. NIK500 dataset contains both stocks in NIK225 and other stocks that are relatively less “important” in explaining the fluctuation. These relative less “important” stocks generate noise and therefore, reduce the average size of common components, making the innovations more time and cross-sectionally dependent, and thus lower the explanatory power of the factor estimates. Our results here confirm the suggestion of Ng and Boivin (2003) and consistent with the findings in Cipollini and Kapetanios (2004) paper.

4.1.2 Dynamics of factor estimates

Autoregressive representation of factor estimates In the above subsection, we have already seen two important dynamic properties of the factor estimates for all of our five datasets, i.e. heteroskedasticity and serial correlation. Now, let's take a step further to examine more deeply into their dynamic characteristics. Based on the above findings, we center our analysis on the dominant factors of KOSPI, NIK225 and NIK500 dataset. The first factors of SING and HSCI are also investigated for comparison purpose. $AR(p)$ representation of the following form is fitted into each of these estimated factor series

$$f_{k,t} = \alpha_1 f_{k,t-1} + \alpha_2 f_{k,t-2} + \cdots + \alpha_p f_{k,t-p} \tag{13}$$

where $f_{k,t}$ denotes factor k at time t, and $(\alpha_1, ..., \alpha_p)$ denote the AR coefficients. Table 4 reports the estimated AR coefficients, t-statistics and the corresponding probabilities for the factor estimates of the five datasets. 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

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) for the dominant factors of KOSPI, NIK225 and NIK500; but only 3 lags and 5 lags for SING and HSCI, respectively. It can be seen clearly that 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 explaining a reasonably large amount of variation in stock returns, it should show persistence, and thus sustain a higher order AR process. The first factor of HSCI apparently does not persist. However, the dominant factors of KOSPI, NIK225 and NIK500 seem to have obtained this characteristic. In the next section, we move on to examine the long memory nature of our factor estimates. We focus our analysis on the dominant factors of KOSPI, NIK225 and NIK500 datasets. We concentrate out the first factors of SING and HSCI due to the fact that they have low explanatory power, which evidents factor model being an inappropriate specification that underlies them.

Long memory nature of factor estimates Long memory nature of the dominant factors of these three datasets is further confirmed by the persistence found in their autocorrelations. Figure 3 graphs 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. 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 KOSPI, NIK225 and NIK500 shows persistence even up to 200 lags and they also exhibit hyperbolic decay. This is apparently an evidence of long memory. Autocorrelations of the first factor of HSCI however, does not persist. Figure 4 plots the autocorrelations of the first differenced dominant factors of KOSPI, NIK225 and NIK500 datasets. Their autocorrelations are alternating postively 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 an $ARFI(p, d)$ process to the series of the dominant factors of KOSPI, NIK225 and NIK500 only. We estimate the process by AMLE of the dominant factors of KOSPI, NIK225 and NIK500. Number of lags included in the $ARFI$ process are once again determined by the AIC and SBC, which tell us that $ARFI(1, d)$ should be chosen for all three dominant factors. Table 6 reports the order of autoregressive chosen for the $ARFI$ specification for each of these three dominant factor series, the estimated parameters, the standard errors and the computed 95% confidence interval of $\hat{d}$. We can see from the result 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.

4.2 Forecasting volatility

In this section, we carry out a forecasting exercise on stock volatility using the SVF model. From the findings in section 4.1.1, we can see that only the Korean and Japanese datasets appear to have a factor structure. We therefore carry out our forecasting exercise only on the daily volatility of NIK225 and KOSPI constituent stocks. We split the entire sample period into an estimation sample that covers the period 3 January 2000 to 15 March 2004, and the last 100 days in the entrie sampling period are taken as the forecast evaluation sample and this is corresponding to the period 16 March 2004 to 30 July 2004. Some in-sample results is displayed in the following part of this subsection, then follow by out-of-sample forecasting results using a single-factor SVF model, and univariate SSM as a benchmark forecast.

4.2.1 In-sample analysis

Using data for the period 3 January 2000 to 15 March 2004 of the two indices, we present some in-sample results here. We first apply principal component method to equation (4) to extract common factors from estimation sample. Table 8 shows the cummulated explained variation for the first 10 in-sample factor extracted from each dataset. First in-sample factors still explain most of the variation during the estimation period. This lead us to conclude that the first factor estimates from each of the markets can be regarded as their local dominant factors during the estimation sample. Figures 5 and 6 plots the in-sample first factor estimates from the two datasets. Figure 7 and 8 display the in-sample autocorrelations of the local dominant factors of Korean and the Japanese markets up to 200 lags. Table 6 shows the estimated parameters of the $ARFI(1,d)$ by AMLE. Information criteria suggest $ARFI(1,d)$ is still the most suitable model for the dominant in-sample factors. Results shown in those figures and table for in-sample factors agree with the findings in section 4.1.

From the computed cummulative explained variation of the first 10 in-sample factor estimates, we have seen that the majority of the common fluctuation in the datasets are summarised by their dominant factor and adding further factors can make only tiny contribution. This implies a local dominant factor is enough to explain fluctuation in stock returns of a local market in our case. Following this implication, we move on to examine the idiosyncratic stochastic volatility in the SVF model that contain only one local dominant factor of a market. 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 $\widetilde{y}_{i,t}$.

We fit a univariate SSM as in equation (9) and (10) 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 SSMs. We report our findings using histograms. Figures 9 and 10 are the histograms of estimated γ_i, λ_i of idiosyncratic volatility for the two datasets. We can see from the histograms , the majority of the stocks for both markets have estimated λ_i between 0.8 and 1. This large values of estimated cofficient in the transition equation indicate persistence in the remaining idiosyncratic volatility of the constituents of the two indices.

We also check for remaining serial correlations in measurement errors of idiosyncratic volatility. This is done by first obtaining smoothed estimates of the states, $\eta_{i,t}$ in equation (10), that is denoted as $\widehat{\eta}_{i,t}$. Then we subtract $\widehat{\eta}_{i,t}$ from $\omega_{i,t}$ to obtain measurement errors $\widehat{\zeta}_{i,t}$. We perform Lagrange Multipiler (LM) test to test for serial correlation up to 5 lags in measurement errors at 1 % significance level for every constituent stock in KOSPI and NIK225 indices. Null hypothesis of no serial correlations up to 5 lags is rejected if the probability value corresponding to the computed LM test statistic is less than the size of the test that equals 0.01. Our results show the majority of the stocks have measurement errors that show no remaining serial correlations in our three datasets. For KOSPI, 137 out of 176 stocks are found to have no serial correlations. For NIK225, 169 out of 217 stocks show no signs of serial correlations.

4.2.2 Out-of-sample forecasting: single-factor SVFM v.s. univariate SSM

We carry out our forecasting exercise for last 100 periods of the whole sample to produce 1-step ahead point forecasts using local-factor SVF model. Recursive forecasting procedure is adopted for which our local-factor SVF model is estimated with more data as forecasting move forward in time. Following the five-step approach illustrated in section 2.5, we construct volatility forecast $\widehat{\widetilde{y}_{i,t+1|t}}$ for every period in the forecast evaluation sample. This overall 1-step ahead volatility forecast is formed by combining $\widehat{f^L_{t+1|t}}$ and $\omega_{\widehat{i,t+1}|t}$ as in equation (11). However, attention should be paid to the computation of the final forecasts. Since the data we use for estimation and forecasting are demeaned, we therefore, need to add up the removed mean to the demeaned volatility forecast to obtain the final volatility forecast, $\widehat{\widetilde{y}_{i,t+1|t}}$. So, $\widehat{\widetilde{y}_{i,t+1|t}}$ in equation (11) 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 our datasets.

Since we have large datasets that contains more than a hundred stocks with more than a thousand observations. Thus the number of univariate SSM for idiosyncratic volatility need to be re-estimated along the forecast evaluation period is huge. There are 17600 individual state-space models for KOSPI dataset, and 21700 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.[5]

In order to evaluate the forecasting power of the local-factor SVF model, we use volatility forecast produced by a univariate state-space model that has no factor structure (denote as SSM in Table 7) as a benchmark forecast. We compute two statistics for forecast performance comparison, the average mean absolute prediction error (average MAPE) and the average mean squared prediction error (average MSPE). Our result is reported in Table 7.

[5]It 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 models and they find only 8 models that do not converge.

We can see from the table that forecasting performance of local-factor SVF model seems slightly worse than univariate state-space model. Although in-sample results show the local-factor SVF model is a favourable specification for our data, it is still a little weaker in forecasting Asian stock volatility than univariate state-space model without factor structure. Both forecast evaluation statistics are smaller for SSM forecasts than for local-factor SVF model. However, the difference is quite small. We understand that there are two possibilities that may lower the overall forecasting performance of local-factor SVF model: 1. some of the stocks are badly forecasted by the factor model or 2. the factor model may not be forecasting very well in some subsampling periods. Therefore, we carry out cross-sectional examination and subsampling period investigation to detect for such possibilities. Unfortunately, our examination outcomes show none of these two cases has happened.

Having observed this findings and from what we understood in existing empirical studies of Asian stock volatility, we believe this unfavourable results towards forecasting power of SVF model may actually caused by missing factors (other than the local dominant factor) that are essential in determining stock volatility in a local market . If those missimg factors are proved significant in determining stock volatility of local market, then adding those missing factors to the factor model can improve the in-sample results and forecasting performance. In other words, there is a need for extending a local-factor model into a multi-factor specification. In the following section, we further discuss the need for such an extension and we also propose several specifications of a multi-factor SVF model. How to perform forecasting exercise by using multi-factor model will also be elaborated.

5 Proposal for multi-factor SVF model

A lot of empirical studies provide evidence of volatility tranmission 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)). From these studies and all other relating literatures, we believe if there exists volatility spillovers, then dominant factor in a foreign market that summarises the common market fluctuation in its stock returns should impact on the fluctuation in stock returns of a local market. Therefore, by introducing this dominant factor of the foreign market into the volatility model for stock returns in the local market should increase the goodness of fit and thus, strengthen the forecasting power of the model. This brings about an idea of extending a one-factor (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?

According to existing empirical evidence, we believe apart from the local dominant factor, three other factors should also play essential roles in explaining stock return volatility. These factors are, first of all, a global factor that summarizes the common driving force of stock volatility in US market. Secondly, a regional leading factor that describes the common fluctuation in stock returns of a leading market of the region. Finally, a regional factor that is extracted using data of constituent stocks of all representative indices of the markets in the same region, except the regional leading factor. When these three factors are concerned in the construction of multi-factor model for our study, the global factor is to be extracted from the data of S&P 500 constituents. Regional leading factor is represented by the dominant factor of Japanese market and this factor is extracted from the data of NIK225 constituents. So far, our forecasting exercise in this paper has been carried out for only the Korean and Japanese markets. When our analysis will be further extended to cover the entire Asian region, then the regional factor is to be extracted from constituents of all representative indices of the markets in the Asian region, except the Japanese index constituents. This regional factor describes the common cause of variation in stock returns in Asia except Japan. Japanese stock returns are not included in the dataset for extracting the regional factor because we want to see how much common fluctuation in Asian markets stock returns as a whole can contribute to volatility forecast for a local market, without the impact of variation in Japanese stock returns.[6] Moreover, since we are introducing the Japanese dominant factor as a regional leading factor, we have already singled out and emphasized the importance of Japanese stock returns fluctuation in determining local market volatility.

[6]This can of course be verified to include Japanese 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.

5.1 Multi-factor specifications of SVF model

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

$$\widetilde{y}_t^L = a + \theta f_t^M + \omega_t \tag{14}$$

where $\widetilde{y}_t^L = (\widetilde{y}_{1,t}, \cdots, \widetilde{y}_{N,t})'$ 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 $T \times N$ matrix of idiosyncratic volatility and this is obtained after the common component is removed from the stochastic volatility. 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 SVF model.

Let $\widetilde{y}_{i,t}^L$ denotes volatility of stock i in a local market, f_t^L, f_t^G, f_t^R, and f_t^{RL} denote local dominant factor, global factor, regional factor, and regional leading factor, respectively. θ_i^L, θ^G, θ_i^R and θ_i^{RL} are their corresponding factor loadings. We consider the following 5 specifications of a multi-factor model in equation (14).

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

Stock volatility of local market depends on its dominant factor, and the dominant factor from a leading market of the region. $f_t^M = (f_t^L, f_t^{RL})'$ and $\theta = (\theta_i^L, \theta_i^{RL})'$ are 2-dimensional vector of common factors and factor loadings. i.e.

$$\widetilde{y}_{i,t}^L = a_i + \theta_i^L f_t^L + \theta_i^{RL} f_t^{RL} + \omega_{i,t} \tag{15}$$

Modification 2: Local - Global factor model (L-G)

Dominant factors from local market and US market are used to explain local stock market volatility. $f_t^M = (f_t^L, f_t^G)'$ and $\theta = (\theta_i^L, \theta_i^G)'$ are again 2-dimensional. So,

$$\widetilde{y}_{i,t}^L = a_i + \theta_i^L f_t^L + \theta_i^G f_t^G + \omega_{i,t} \tag{16}$$

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

This model assumes the local dominant factor, together with the dominant regional leading factor and the global factor, are all crucial determinants of the variation in local stock returns. $f_t^M = (f_t^L, f_t^{RL}, f_t^G)'$ and $\theta = (\theta_i^L, \theta_i^{RL}, \theta_i^G)'$ are 3-dimensional in this case. The model becomes

$$\tilde{y}_{i,t}^L = a_i + \theta_i^L f_t^L + \theta_i^{RL} f_t^{RL} + \theta_i^G f_t^G + \omega_{i,t} \tag{17}$$

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

This model concerns local dominant factor, regional factor that summarises the underlying driving force of the the variation of all stock returns in the region, and the global factor being significant in explaining local market volatility. The impact of regional leading market is ruled out. In this case, $f_t^M = (f_t^L, f_t^R, f_t^G)'$ and $\theta = (\theta_i^L, \theta_i^R, \theta_i^G)'$.

$$\tilde{y}_{i,t}^L = a_i + \theta_i^L f_t^L + \theta_i^R f_t^R + \theta_i^G f_t^G + \omega_{i,t} \tag{18}$$

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

In this specification, all common factors are concerned. So, $f_t = (f_t^L, f_t^{RL}, f_t^R f_t^G)'$ and $\theta = (\theta_i^L, \theta_i^{RL}, \theta_i^R, \theta_i^G)'$.

$$\tilde{y}_{i,t}^L = a_i + \theta_i^L f_t^L + \theta_i^{RL} f_t^{RL} + \theta_i^R f_t^R + \theta_i^G f_t^G + \omega_{i,t} \tag{19}$$

There is a reason for not simply to take a model that includes all world, regional, regional leading factors, but to concern different specifications of the multi-factor SVF model. Although some studies of financial contagion have shown intra-regional contagion effects in Asian stock markets fluctuation e.g. Masih and Masih (1999), some other studies show no support for contagion among some markets for some period e.g. Khalid and Kawai (2003). Therefore, we do not want to rule out the possibility that a factor is an important determinant of stock returns volatility for some markets, but not for some others. 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 multi-factor SVF model is straight forward and this is the same as what we do for local-factor SVF model. We estimate each of the common factor included in the multi-factor model using principal components estimation. Once we obatain the factor estimates, we remove them from the stochastic volatility of each stock return series in the dataset and what we are left with is the idiosyncratic stochastic volatility.

Similar to the local-factor SVF model, all common factors in the multi-factor SVF model have a long memory underlying process. We model all the factors in a multi-factor SVF model with an $ARFIMA(p, d, q)$. Idiosyncratic component is the part of stochastic volatility that is unique to each individual stock. Analogous to the local-factor SVF model, this is caused by some driving force with unknown form. And thus, should be modelled by a state-space representation.

5.2 Determining an appropriate specification and testing significance of factor estimates

Once we believe a multi-factor model should be used, the next task is to decide on a correct specification. The questions we need to answer here, are (1) are the chosen common factors significant in explaining local market volatility? (2) whether the model which includes the selected factors is strong enough to explain local market volatility 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 volatility support a local-factor SVF model or a multi-factor SVF model. If it is the latter that is needed, then how many of the local factors, regional leading factors and global factors should be included. If lagged factors are concerned, then how many lags of those factors should be considered. Answer to the first question can be easily drawn by carrying out test for significance on common factors. However, no exisiting 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 inference on factor estimates and the parameters in those 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 (FAR). We apply the inferences developed by Bai and Ng (2006) and use Least Square parameter estimates of FAR for computing test statistics.

Suppose there are N stocks in a dataset, consider the following FAR

$$\widetilde{y}_{i,t} = c + \alpha' f_t + \beta' M_t + e_{i,t} \tag{1}$$

where $i = 1, \cdots, N$ and $t = 1, \cdots, T$. $\widetilde{y}_{i,t}$denotes stock 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, $f_t = (f_{1,t}, \cdots, f_{K,t})$. Common factors contained in matrix f_t is estimated from the $(T \times N)$ dataset of stochastic volatility, $\widetilde{y}_{i,t}$, using prinicipal components. c is a constant. α and β are vectors of least square estimated coefficients of common factors f_t and a set of other observable variables M_t. $e_{i,t}$ is the disturbance. In our case, matrix M_t is not considered.

Bai and Ng (2006) show that under some general assumptions, least square estimates of the above FAR are asymptotically normal and $\sqrt{T}$ consistent if $\frac{\sqrt{T}}{N}$ tends to 0. They also show that in the setting of FAR 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. Morever, 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 heteroskedasticity and cross-section dependence in the idiosyncratic shocks. Given these results, we can carry out significance tests on our factor estimates using FAR.

In the Monte Carlo study by Bai and Ng (2006), confidence intervals of estimated conditional mean is computed using their covariance matrix estimator, CS-HAC (Cross-section and heterokedastic autocorrelation consistent) using FAR as a forecasting model with different chosen combinations of N and T. Their results show that when idiosyncratic errors are heteroskedastic and cross-sectionally correlated , coverage rate for conditional mean is the highest when $N = 100$ and $T = 400$. While coverage rate for forecasting variable is found highest when $N = 100$ and $T = 400$, and when $N = 100$ and $T = 200$. High coverage rate suggests more robust C.I. in general. When computing C.I., error variance in predicting 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 C.I. 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.

When we are to further our empirical analysis, their results will be adopted. 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. And thus, the asymptotic results they develop should be applicable to our case and allow us to perform test for significance on factor estimates. The following procedure will be taken when carry out significance test. We fit a FAR into stochastic volatility of every stocks in each dataset, i.e. there will be N FAR in total. Parameters will then be estimated by least square method. By applying asymptotic theory, we compute t-statistics for testing factors significance in every one of the N regressions. We then move on to check how many times a common factor is found significant out of N model. 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, statistics that provide us with the ideas of goodness of fit will be good indicators. 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^2allows us to pick the specification that is most appropriate to the local market of interest. Another advantage of doing this is less time consumption for a forecasting exercise. Instead of forecasting volatility by everyone of the above specifications for every local market we consider, we use only the one which best describes stocks volatility of that market, and then compare it with the forecasting performance of a local factor model and that of other volatility models. In the case when lagged factors are included in our model, then information criteria is to be used to decide on the number of lags chosen.

5.3 Forecasting with multi-factor SVF model

If additional factors other than local-factor are proved to be essential determinants of stock volatility in local market, volatility forecasts can then be constructed by using an appropriate specification of multi-factor SVF model. Forecasting exercise is done in a similar manner as for the local-factor SVF model. The only difference is that for every period in forecast evaluation sample, we forecast not only the local dominant factorin the first step, but also additional dominant factors included in the chosen specification of multi-factor model using $ARFI(p, d)$. Then, we forecast idiosyncratic stochastic volatility of every stock using univariate state-space model. Overall volatility forecast is formed by combining these forecasts. Forecasts constructed by the five specifications of multi-factor SVF model as in equation (15) to (19) are computed as follow

L-RL factor model:

$$\widehat{\tilde{y}^{L}_{i,t+1|t}} = a_i + \theta_i^L \hat{f}^L{}_{t+1|t} + \theta_i^{RL} \hat{f}^{RL}{}_{t+1|t} + \hat{\omega}_{i,t+1|t} \tag{18}$$

L-G factor model:

$$\widehat{\tilde{y}^{L}_{i,t+1|t}} = a_i + \theta_i^L \hat{f}^L{}_{t+1|t} + \theta_i^{G} \hat{f}^{G}{}_{t+1|t} + \hat{\omega}_{i,t+1|t} \tag{19}$$

L-RL-G factor model:

$$\widehat{\tilde{y}^{L}_{i,t+1|t}} = a_i + \theta_i^L \hat{f}^L{}_{t+1|t} + \theta_i^{RL} \hat{f}^{RL}{}_{t+1|t} + \theta_i^{G} \hat{f}^{G}{}_{t+1|t} + \hat{\omega}_{i,t+1|t} \tag{20}$$

L-R-G factor model:

$$\widehat{\tilde{y}^{L}_{i,t+1|t}} = a_i + \theta_i^L \hat{f}^L{}_{t+1|t} + \theta_i^{R} \hat{f}^{R}{}_{t+1|t} + \theta_i^{G} \hat{f}^{G}{}_{t+1|t} + \hat{\omega}_{i,t+1|t} \tag{21}$$

L-RL-R-G factor model:

$$\widehat{\tilde{y}^{L}_{i,t+1|t}} = a_i + \theta_i^L \hat{f}^L{}_{t+1|t} + \theta_i^{RL} \hat{f}^{RL}{}_{t+1|t} + \theta_i^{R} \hat{f}^{R}{}_{t+1|t} + \theta_i^{G} \hat{f}^{G}{}_{t+1|t} + \hat{\omega}_{i,t+1|t} \tag{22}$$

Empirical applications using multi-factor SVF model is currently being carried out. We expect the multi-factor model will perform better in forecasting volatility than a local-factor model due to the fact that it is more applicable to the real world situation that international stock markets are facing.

6 Conclusion

In this paper, we estimate common factors of Asian stock volatility using principal components method. Analysis of factor estimates and a forecasting exercise using local-factor SVF model is carried out. Results on our factor analysis show approximate factor structure is detected in constituent stock returns of KOSPI 200, NIKKEI 225 and NIKKEI 500; but not in Heng Seng Composite Index and Singapore All Share Index. More importantly, dominant factors of KOSPI 200, NIKKEI 225 and NIKKEI 500 perform quite well in explaining variation in their stock returns. Long memory is also detected in these factors. These findings provide an insight into empirical studies of common factors that contribute to Asian stock volatility with the use of principal components estimation and in the framework other than conventional state-space analysis. However, results on our forecasting exercise show that a local-factor SVF model is slightly weaker than a univariate state-space model in forecasting volatility. We believe this deficiency may be due to missing factor and this has prompted a proposal on a multi-factor SVF model.

Empirical applications using multi-factor SVF model is currently being carried out. We expect the multi-factor model will perform better in forecasting volatility than a local-factor model due to the fact that it is more applicable to the real world situation that international stock markets are facing. Further empirical research concerns boardening the analysis to cover markets in the entire Asian region, especially focusing on emerging markets, in order to get a better understanding on the issue of volatility transmission and financial contagion within the area in the context of factor analysis.

7 Appendix

7.1 Summary of assumptions for principal component estimation

As suggested by Cipollini and Kapetanios (2005), principal components estimation is used for factor estimation of SVF model. This method is a remarkable development in analyzing dynamic approximate factor model with large dataset. It is simple to use and asymptotically equivalent to maximum likelihood estimation. Stock and Watson (2002b) shows that it provides consistent estimates even if idiosyncratic innovations are serially and cross-sectionally correlated under rather general assumptions. Bai (2003) also shows the neccessary conditions for ensuring consistency are asymptotic orthgonality and asymptotic homoskedasticity. We summarize here the use of the estimation method in analyzing dynamic approximate factor model and the assumptions required to ensure consistency.

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

$$x_{i,t} = \lambda_i' f_t + \xi_{i,t}$$

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

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

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 $f = (f_1, ..., f_T)$.

In order words, consistent estimate of f is given by the K largest eigenvectors of matrix XX'. 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.

1. For the factor loadings, $(\Lambda'\Lambda/N) \to I_k$ and $\| \lambda_i \| \leq \bar{\lambda} < \infty$
2. For the factors, $f'f$ has finite unconditional second moment, i.e. $E(f'f)^2 < \infty$. 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 $\rho_{i,i} > \rho_{j,j} > 0$, 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 $\sup_N \sum_{s=-\infty}^{\infty} \left|\gamma_{N,t}(s)\right|$ has finite limit as $N \to \infty$. 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 $N \to \infty$. The size of fourth moment is limited as well, so $\sup_{t,s} N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N} \left|cov(\xi_{i,s}\xi_{i,t}, \xi_{j,s}\xi_{j,t})\right|$ has finite limit as $N \to \infty$.

References

[1] Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X. (Forthcoming), "Volatility Forecast", Handbook of Economic Forecasting (G. Elliott, C. W. J. Granger, and A. Timmermann eds), North Holland

[2] Bai, J. (2003), "Inferential theory for factor models of large dimensions", Econometrica, Vol. 71, No. 1, 135-171

[3] Bai, J., and Ng, S. (2002), "Determing the number of factors in approximate factor models", Econometrica, Vol. 70, No. 4, 191-221

[4] Bai, J., and Ng, S. (2006), "Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions", Econometrica, Vol. 74, No. 4, 1133-1150

[5] Baillie, R. T. (1996), "Long memory process and fractional integration in econometrics", Journal of Econometrics, 73, 5-59

[6] Ballie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996), Fractionally integrated generalised autoregressive conditional heteroskedasticity, Journal of Econometrics, 74, 3-30

[7] Beran, J. (1994), "Statistics for long memory processes", Chapham and Hall

[8] Beran, J. (1995), "Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive intergrated moving average models", Journal of Royal Statistical Society Series B, 57, No. 4, 689-672

[9] 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

[10] Boivin, J., and Ng, S. (2003), "Are more data always better for factor analysis?", Working paper 9829, National Bureau of Economic Research

[11] Bollerslev, T. (1986), Generalised Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327

[12] 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

[13] Chung, C. F. (1998), "Estimating the fractionally integrated GARCH model", Working paper, Institute of Economics, Academia Sinica

[14] 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

[15] Durbin, J., and Koopman, S. J. (2001), "Time series analysis by state space methods", Oxford University Press

[16] Engle, R. F., (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation", Econometrica, 50(1), 987-1007

[17] Engle, R. F., and Bollerslev, T., (1986), "Modelling the persistence of conditional variances", Econometric Review, 5, 1-50

[18] Harvey, A. C. (1990), “Forecasting, structural time series models and the Kalman filter”, Cambridge University Press

[19] Harvey, A. C., Ruiz, E., and Sheperd, N. (1994), “Multivariate stochastic variance models”, Review of Economic Studies, 61, 247-264

[20] Franses, P. H., and van Dijk, D. (2000), “Non-linear time series models in empirical finance”, Cambridge University Press

[21] 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

[22] Kapetanios, G. (2004), “A new method for determining the number of factors in factor models with large datasets”, Queen Mary, University of London, Working paper No. 525

[23] Khalid, 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

[24] Masih, A. M. M., and Masih, R. (1999), “Are Asian stock fluctuations due mainly to intra-regional contagion effect? Evidence based on Asian emerging stock markets”, Pacific-Basin Finance Journal, 7, 251-282

[25] 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

[26] Nelson, D. B. (1991), “Conditionally heteroskedasticity in asset returns: a new approach”, Econometrica, 59, 347-370

[27] Ng, A. (2000), “Volatility spillover effects from Japan and the US to the Pacific-Basin”, Journal of International Money and Finance, 19, 207-233

[28] 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

[29] 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

[30] Stock, J. H., and Watson, M. W. (1998), "Diffusion Indexes", Working paper 6702, National Bureau of Economic Research

[31] Stock, J. H., and Watson, M. W. (2001), "Forecasting output and inflation: the role of asset prices", Journal of Economic Literature, Vol. 42, 1-48

[32] 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

[33] 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

Table1:Descriptive statistics of the mean return series

	SING	**HSCI**	**KOSPI**	**NIK225**	**NIK500**
No. of periods	1185	1120	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
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: Breuch-Godfrey LM test statistics and Engle's ARCH LM test statistics

	SING	**HSCI**	**KOSPI**	**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 3: Cummulative explained variation for the first 10 factor estimates

No. of factors	SING	HSCI	KOSPI	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 4: AR representations of the first factor estimates

AR coeff.	SING	HSCI	KOSPI	NIK225	NIK500
$\boldsymbol{\alpha}_1$	0.334 {11.59}	0.113 {3.786}	0.149 {4.970}	0.076 {2.561}	0.119 {3.991}
$\boldsymbol{\alpha}_2$	0.119 {3.914}	-0.05 {-1.683}	0.235 {7.779}	0.161 {5.369}	0.187 {6.238}
$\boldsymbol{\alpha}_3$	0.126 {4.164}	0.119 {4.001}	0.118 {3.804}	0.150 {4.981}	0.176 {5.798}
$\boldsymbol{\alpha}_4$	0.072 {2.382}		0.061 {1.959}	0.044 {1.452}	0.025 {0.840}
$\boldsymbol{\alpha}_5$	0.151 {5.264}		0.044 {1.446}	0.131 {4.360}	0.106 {3.478}
$\boldsymbol{\alpha}_6$			0.088 {2.926}	0.050 {1.677}	0.039 {1.288}
$\boldsymbol{\alpha}_7$			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 paretheses.					

Table 5: Estimation outputs of ARFI(1,d) for dominant factors of KOSPI, NIK225 and NIK500

	KOSPI	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

Table 6: ARFI(p,d) estimates of in-sample dominant factors

Stock Index	KOSPI	NIKKEI
p	1	1
phi	0.507	0.275469
d	0.24	0.13402188
Std error of d	0.042	0.051410969
95% CI of d	(0.157, 0.322)	(0.033,0.235)

Table 7: Forecasting performance comparison

	KOSPI		NIK225	
	Local-factor SVF	SSM	Local-factor SVF	SSM
Average MAPE	2.02	2.00	1.77	1.76
Average MSPE	5.98	5.90	6.24	6.23

Table 8: Cummulative R^2 of the first 10 factors for the estimation sample

No. of factors	KOSPI	NIK225
1	0.100	0.075
2	0.114	0.092
3	0.122	0.101
4	0.134	0.113
5	0.144	0.123
6	0.156	0.131
7	0.165	0.139
8	0.174	0.147
9	0.183	0.155
10	0.192	0.162

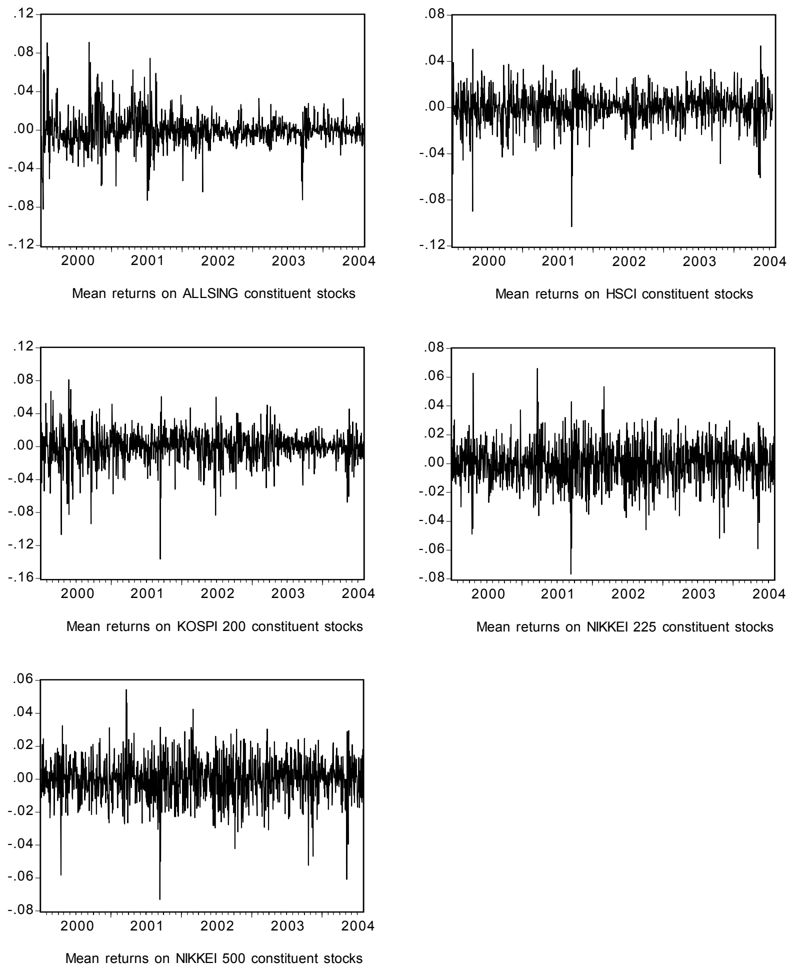


Figure 1: Time plots of mean returns

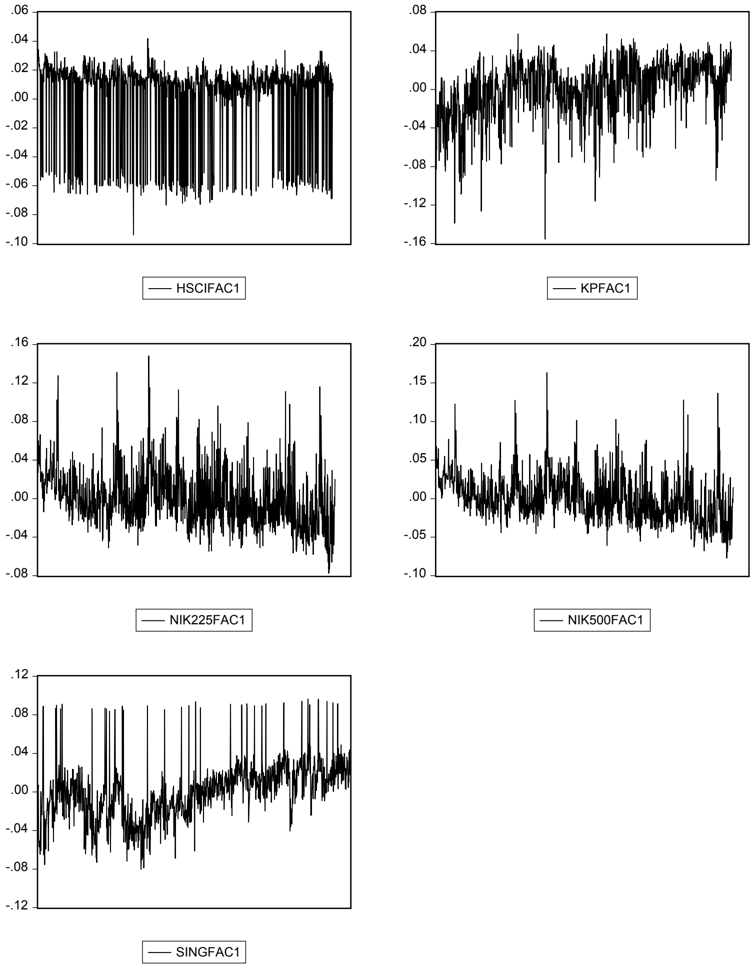
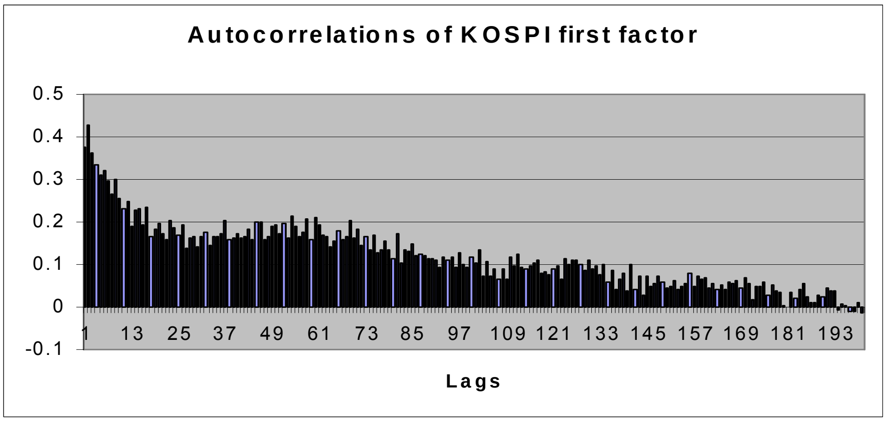
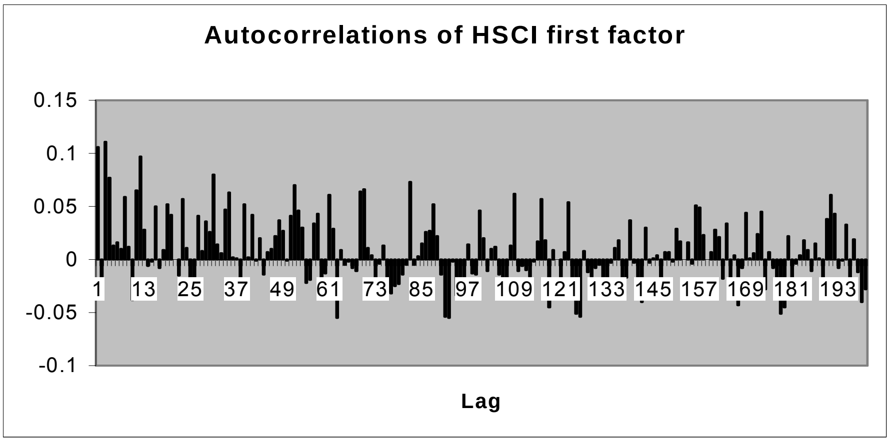


Figure 2: Plots of the first factors





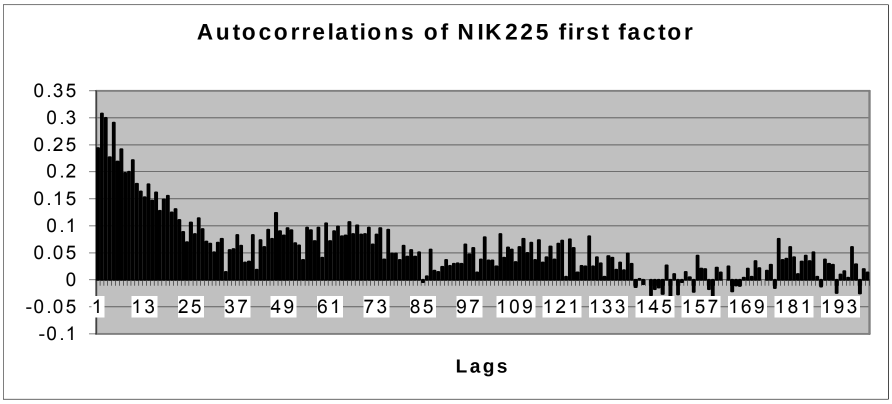
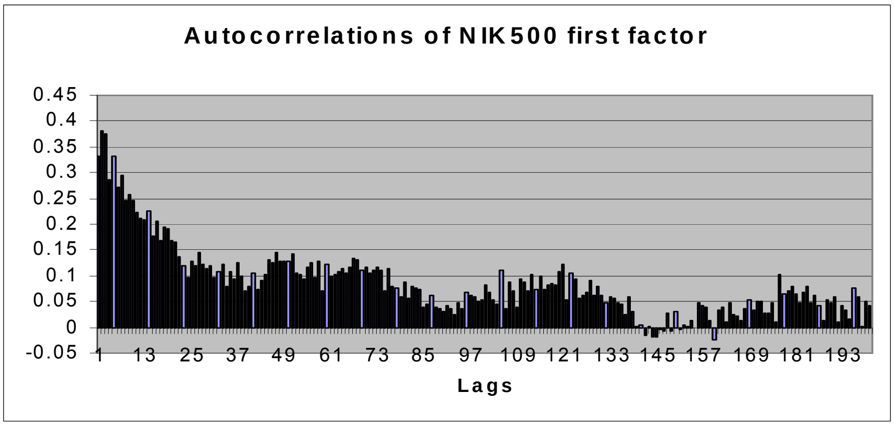


Figure 3: Autocorrelations of the first factors



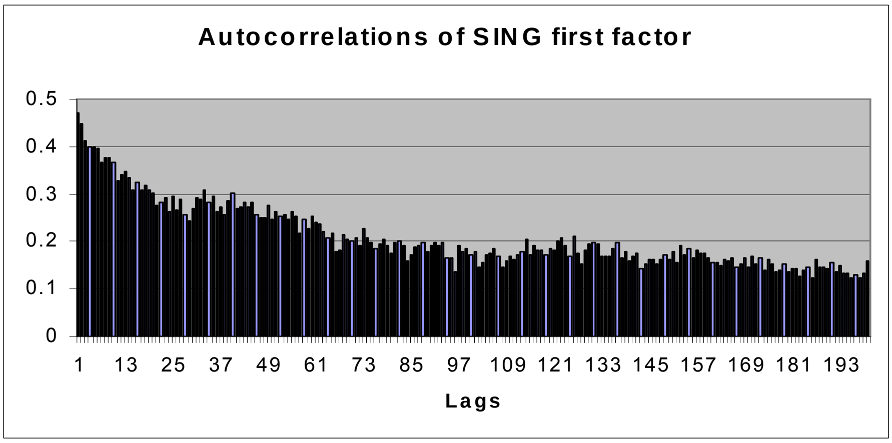
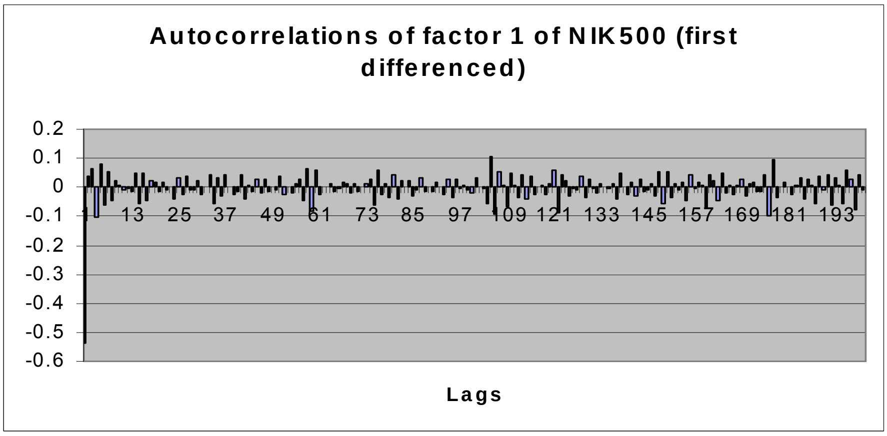


Figure 3 (continued): Autocorrelations of the first factors



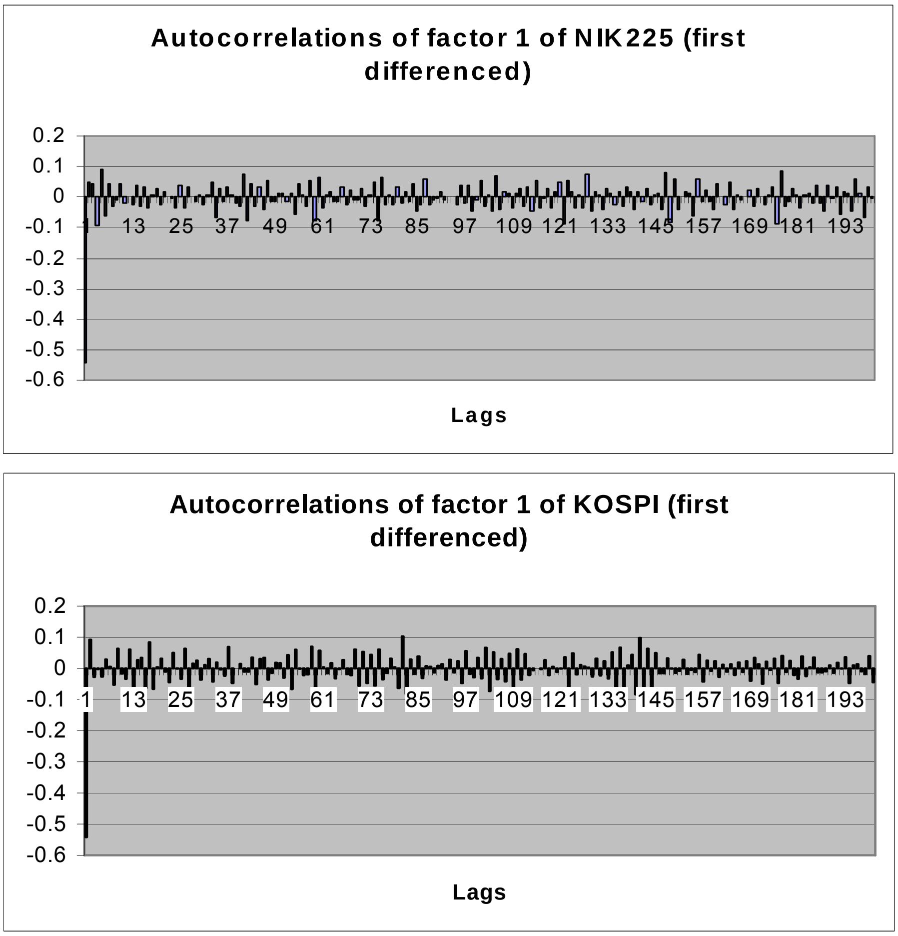


Figure 4: Autocorrelations of the first-differenced dominant factors

Figure 5: In-sample dominant factor of KOSPI
(Period 3 Jan 2000 to 15 March 2004)

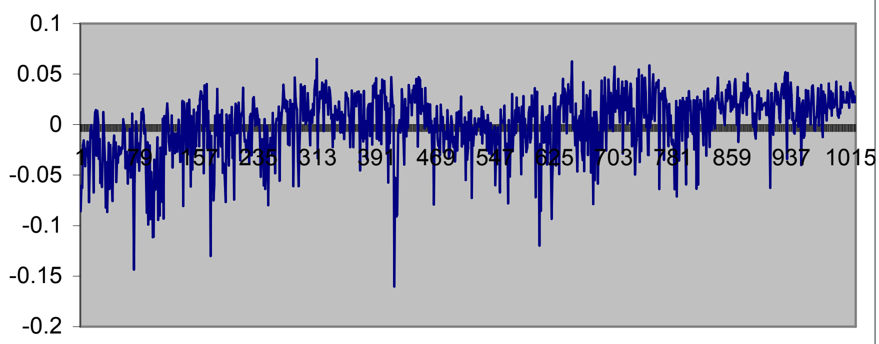


Figure 6: In-sample dominant factor of NIKKEI
(Period: 3 Jan 2000 to 15 Mar 2004)

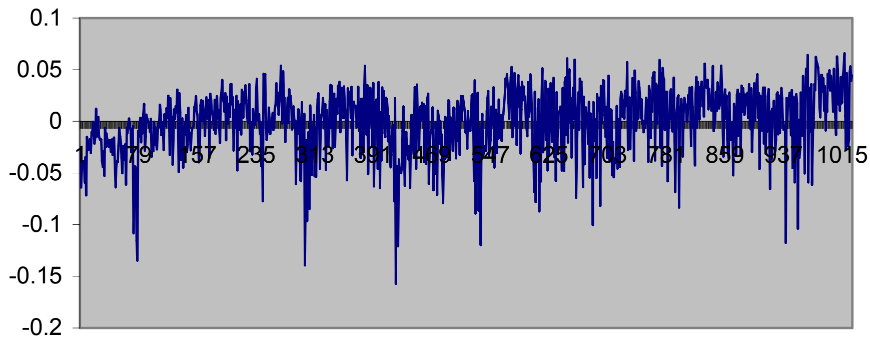


Figure 7: Autocorrelations of KOSPI in-sample first factor (estimation sample: 3 Jan 2000 to 15 March 2004)

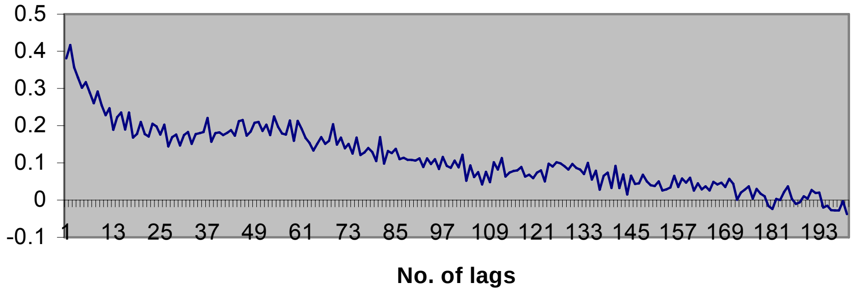


Figure 8: Autocorrelations of NIKKEI in-sample first factor (Period: 3 Jan 2000 to 15 Mar 2004)

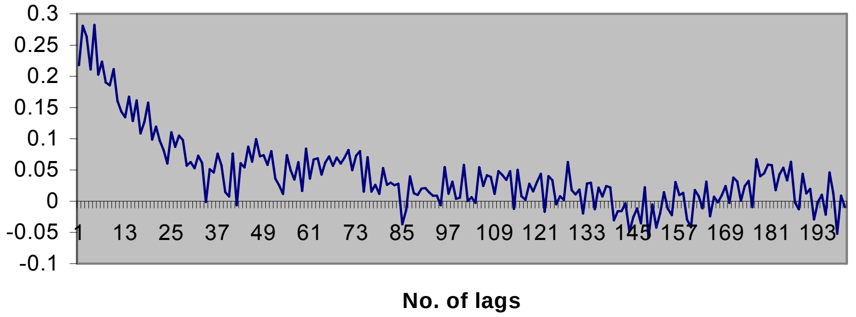


Figure 9: Estimated hyperparameters in SSM of KOSPI idiosyncratic volatility (in-sample data)

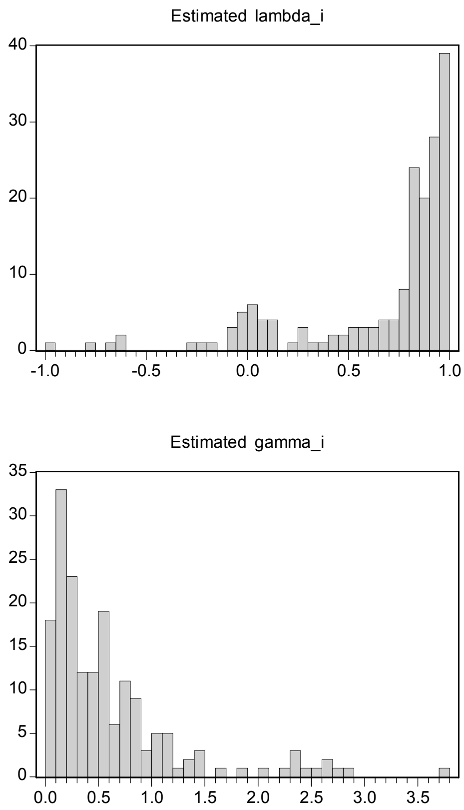
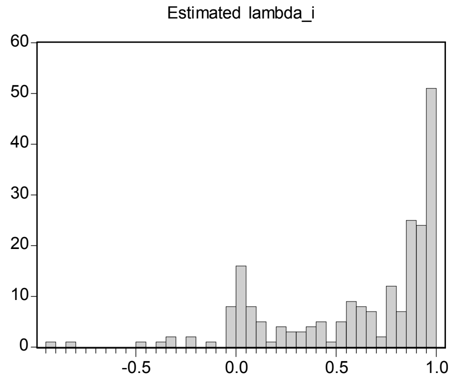
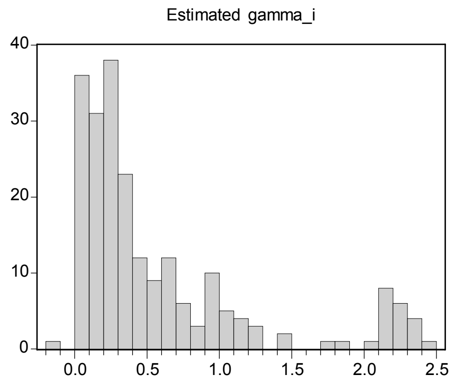


Figure 10: Estimated hyperparameters in SSM of NIK225 idiosyncratic volatility (in-sample data)





Queen Mary
University of London

This working paper has been produced by
the Department of Economics at
Queen Mary, University of London

Copyright © 2006 Silvia S.W. Lui
All rights reserved

Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/papers/wp.htm