UNIVERSITY OF SOUTHAMPTON

International Market Issues in Shanghai Financial Price Behaviour

Leilei Tang

A thesis submitted to the University of Southampton
For the Degree of Doctor of Philosophy

Department of Management
Faculty of Social Science

March 2001

UNIVERISTY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL SCIENCE

DEPARTMENT OF MANAGEMENT

Doctor of Philosophy

INTERNATIONAL MARKET ISSUES IN SHANGHAI FINANCIAL PRICE BEHAVIOUR

By Leilei Tang

This thesis conducts the first-ever detailed examination of the impacts of international stock markets on the Shanghai stock market. The main aim is to investigate whether information originated from other stock markets has adverse effects on the Shanghai stock market, and consequently, to derive some implications for policy agents, investors, and further research.

This study is based on four relatively related research papers, which focus on the impacts of information from different perspectives. It has identified the significant decline of the Shanghai stock price volatility after 1^{st} July 1997. It has also analysed factors from the Hong Kong stock market that impact on Shanghai stock prices. To make this study more robust, the concept of co-persistence in variances is adopted to examine the validity of the methodologies used in the first paper. Moreover, functions of the futures market is also discussed in order to achieve a better understanding of how information coming from the derivative market improves its underlying spot market. This research study supports the main hypothesised notation that we place more emphasis on improving inter-market information linkages. As a by-product the financial portfolio frontier is discussed when the variance-covariance matrix in singular.

Table of Contents

Table of Tables

Table of Tables

ACKNOWLEDGEMENTS

I would like to thank Professor George Mackenzie and Dr. Maozu Lu as my supervisors through this doctoral research. In addition, I would like to acknowledge the financial assistance provided by the Department of Management and Department of Economics, University of Southampton.

My sincere thanks are due to my colleagues Ian Harwood for his encouragement and productive conversations, and Nicola Hilton for her help and advice for obtaining various funds. In particular, I thank all my colleagues in Room 3059, whose company I so much enjoy. Sung Kyu Lee and Yeek Hyun Kim are also acknowledged.

My gratitude is extended to my friends at Southampton Navigators: Mike Spencer, Gez Perry and his wife Alison Perry, Neil Ridegway, Paul Lanbert, Steve McClure and Susan Tailby who always pray for me whenever I have any difficulties in my research.

To Bin He goes particular gratitude for his generous help. My endless gratitude is also dedicated to Anning Ye.

Finally, special thanks are due to my parents and my sister's family for the love and selfless support they have given me all my life.

Chapter 1

Introduction to the research

1.1 Scope and purpose of the thesis

With respect to information flow between financial markets, this thesis stresses two areas which involve critical examinations of how and by what channels the Hong Kong stock market conveys information to the Shanghai stock market, and how the once-lived Shanghai bond futures market, which ceased to exist from October 1995, could provide information to improve its underlying spot market quality. The main purpose of this thesis is to examine the hypothesis that knowledge from current or past price movements in related markets could help to improve the functioning of the Shanghai financial market. The research analysis approach to ascertaining the benefits of allowing information from related markets is to estimate the effects of such information would have on conditional volatility of the Shanghai financial market (i.e. estimating how the conditional volatility varies with augmented information set), assuming that the decline in price volatility would represent the benefits of the inter-market information linkages.

1.2 Motivation for the research

The Shanghai stock market is the earliest and biggest financial market in China. Ever since its inauguration in December 1990, the author had involved in its development. In October 1993, China's first government bond futures market was officially opened in the Shanghai stock market. Soon after the openness of the bond futures market, the futures market was closed and all bond futures products were abolished by the Chinese government in May 1995. Between October 1993 and May 1995, the author had closed connections with the Shanghai International Securities Co., Ltd., which was one of the biggest security companies in China and went bankrupt due to over speculation and adverse political environment. Following these experiences the author became interested in carrying out research into examining the validity of reasons behind the decision by the government to close the Shanghai bond futures market and the claims that the local markets could become more volatile if affected by and integrated with foreign factors.

This thesis is based on the belief that the financial market quality can be improved by disseminating other relevant financial market information. As stated in the following chapters, several studies have been published on the topic of this belief. Why then do we undertake this research? The motivation for selecting the two areas springs from the facts that:

- Ever since the introduction of the study of inter-market information linkages, almost all published studies have referred to developed countries. Insofar there is substantial agreement that the inter-market information improves price discovering efficiency.

- However, many of the literature articles focus on two specific markets without considering the influence of the third market.

- With the exception of papers on several emerging stock markets we are not aware of any journal articles in the English literature that present studies to examine the inter-market information on the Shanghai financial market.

- From the empirical perspective, the Shanghai financial market is of particular interest. Firstly, Shanghai is an interesting emerging market because it is unprecedented that a market could develop so rapidly like Shanghai from being an obscure stock market only a few years ago to having access to one of the major financial centres in the world — the Hong Kong stock market. Consequently, an analysis of information flow between the two stock markets may provide a way to better understand the effects of inter-market information flow. Secondly, the majority of published papers attempt to demonstrate the impacts of the listing derivatives while the Shanghai bond market gives us unique opportunity to understand the impact of the de-listing bond futures in the Shanghai bond market through examining whether the bond futures market had fulfilled a vital role in enhancing the effectiveness of the underlying bond market. Thus such uniqueness provides an additional perspective to better understand a separate but closely related issue for the effects of inter-market information flow.

1.3 Research questions and research approaches

1.3.1 Research questions

This section sets out a number of questions that form the basis of this thesis. So what style of research on the Shanghai stock market is being proposed in this thesis as a way of increasing our understanding of the critical price discovery issues related to the methodological issues? The answer is the econometric analysis of stock prices of research, aimed at producing an understanding of how the information content be incorporated into the stock prices. Thus, in this thesis, we demonstrate what "global" and "local" issues could influence the Shanghai stock prices by regarding the following questions

1. **To what extent has the handover of Hong Kong infiltrated the Shanghai stock market? Are there any "information" issues that impact the Shanghai stock market and implications for policy regulations?**

Traditional segmented stock markets, it could be argued, operates principally on the basis of very limited available local information which leads Bekaert and Harvey (1997) to conclude that the stock prices are more volatile in those markets than developed stock markets. In the past, the Shanghai stock market operated in a similar way. This tradition may be changed by increasing impacts of Hong Kong factors that assist the Shanghai stock market operation and potentially lead to the decline of the Shanghai stock market volatility. If, as seems likely, *capital market liberalizations significantly decrease volatility in emerging markets* (Bekaert and Harvey 1997, p.70), then the Hong Kong information may become an integral part of Shanghai stock prices. The answer to this question leads to answering the next two questions.

2. **Beyond the handover dimension concerning the change of the Shanghai stock market volatility, has the first question left out other potential important dimensions, such as the potential influences of the New York stock market on the Shanghai stock market volatility?**

   Although we expect that the effort to explain the Hong Kong stock market impacts on its counterpart – the Shanghai stock market can benefit from the general literature on volatility spillover between stock markets, it must be also recognized the unique aspects to participation in a third-party involvement no matter how unusual it may seem. Many works have pointed to the significance of information dominance for the US stock market. Then the general expectation is that the relationship between the Hong Kong and Shanghai stock markets is comprised of US components. In this view, the US' potential relevance to explanation of the Shanghai stock prices must be considered also. From this perspective a system is defined including dimensions such as US, Hong Kong, and Shanghai dependent variables.

3. **Are there Hong Kong economic factors that are peculiar to financial markets that have influenced the Shanghai stock price movements?**

Researchers in areas of macroeconomic fluctuations routinely acknowledge interconnections among different economic variables. "Shocks" researchers, for example, have long incorporated different economic shocks such as consumption shocks, credit shocks etc. into their models (for a broad knowledge of this area see Cochrane, 1994). As studied by Lo and MacKinlay (1995), Cheung et al. (1997), research on the relationship between economic activities and financial markets has in recent years addressed issues of stock returns with respect to observable economic variables. The analysis of this question is similar to the above analysis of finding which economic variables that affect stock prices but differs in the methodology.

4. **Finally, had the adoption of financial derivatives provided technical implications for the quality and content of the information in the Shanghai financial market?**

   This question is to assess how valid the criticism by Chinese policy regulators is that the financial derivatives had destabilize the underlying financial market. The relevance of this question to the above three questions is not very obvious. This does not mean, however, that this aspect of analyses is unnecessary. It does serve a purpose of understanding the functions of the information within the Shanghai financial system.

1.3.2 Research questions relating to the research approaches

In the final paragraphs of this subsection, we need some additional contexts for the questions that have been raised here by relating them to the research approaches that we have been involved with.

We have restricted our scope in this chapter to the study of information content in the Shanghai stock prices. Questions (1) to (3) examine the information flow between the Shanghai stock market and other stock markets. The multivariate ARCH family approach, which is a nature measure of the volatility movement, forms the framework for the consideration of the first two research questions. Question (1) explores the changing conditional volatility magnitude for the Shanghai stock market, and the changing role of the Hong Kong stock market for the Shanghai stock market. The second research question looks at the reasons whether or not we should introduce the US factor to the Hong Kong and Shanghai system, and in particular the significance in the debate of the co-persistence in variance. In view of the changing volatility through time is a good representative of new information coming into stock markets, it appears that the analysis of conditional variance through $ARCH$ models might be the appropriate method of analysis in this situation (Of course there are many non- ARCH models that can specify conditional volatility such as Parkinson (1980)'s extreme value method, French et al. (1987)'s $ARMA$ model, Shephard (1996)'s stochastic volatility model).

A different research approach, the structural VAR ($SVAR$) model, is used for the third question which seeks to explore in greater depth the impacts of the Hong Kong stock market on the Shanghai stock market. To date there are very few scholarly research that report on using the structural VAR model to study financial variables. There has been even less research on relationships applying SVAR models for information sources analysis between economic and financial variables. This may due to the data requirements that are needed to support such studies. In recent years, the application of the combination of cointegration and $ARCH$ analysis has become increasingly wide spread. This may be attributed to be numerous successful studies to address the mechanism between futures and its underlying spot markets. To gain a appreciation for how futures market could improve the information efficiency for the spot market, the nonlinear Granger non causality analysis is also applied for the last research question.

1.4 Structure of the thesis and research themes

The thesis is divided into seven chapters, with the first chapter and the final chapter being the introduction and the conclusion. The remaining four chapters are based on four individual papers, each addressing a specific research issue. Each of these chapters contains a specific review of the relevant literature. Having introduced the overall structure of the thesis, a more detailed (chapter by chapter) introduction to the thesis is provided below.

1.4.1 An examination of the volatility transmission mechanism between the Shanghai and Hong Kong stock market:

This is the main theme of the third chapter in which the examination of question (1) revolves around the hypothesis that:

- **When the conditional volatility of the Shanghai stock price is connected to the Hong Kong information, the conditional volatility tends to decrease, reflecting the benefits of the inter-market information flow.**

The basis of the above hypothesis lies in Engle *et al.*'s (1990) meteor shower hypothesis that economic, political, or social events change in one stock market may cause volatility to change in other markets. In the context of their paper, there are two interpretations for the volatility transmission mechanism: One, for the expectation consideration, shocks in a foreign market may alter investors' expectation levels for the domestic market, hence generating the volatility spillover effect.

Two, for the market policy coordination consideration, the volatility spillover effect may be a result of the policy adjustment for the domestic market corresponding to the change of the foreign market policy. These two interpretations can also be considered as the potential channels for the volatility transmission, thus raising the argument that the foreign market can destabilize the domestic market.

It is this argument that forms the motivation for the above hypothesis, investigating the volatility transmission mechanism between the Hong Kong and Shanghai stock markets before and after the handover in July 1997. This empirical analysis is conducted by using the approach proposed by Glosten, Jagannathan and Runkle (1993), which can allow shocks to have asymmetric impacts on conditional volatility processes. An examination and exploration into the volatility transmission mechanism may help investors, by not only in getting a better trading strategy, but also by providing an appropriate assessment for policy concerns.

1.4.2 An overall assessment of the US influence on the relationship between the Shanghai and Hong Kong stock market:

This is the theme of chapter four, in which it extends the results of the third chapter, but the emphasis is somewhat different from the work of chapter two. Theoretical assessments of the effectiveness of a system that is suspected of omitting relevant variables are demonstrated by Lutkepohl (1982), in which he states that "*a low dimensional subprocess contains little information about the structure of a higher dimensional system, ..., the structure of the relation can only be derived by including all important variables in the model. Consequently, since many economic variables are important in the sense that they interact, high-dimensional time series model-building seems to be required*".

In reviewing the journal papers it is apparent that there is a striking measure of agreement: the US stock market has a considerable effect on the rest of world stock markets. Since the dominance of the US stock market has been widely recognized, it can be argued that such a relationship between the Shanghai and Hong Kong stock markets may simply reflect increases in the US's influence rather than the true volatility spillover effects. A further examination of the relationship between the Shanghai and Hong Kong stock markets is therefore needed for their volatility spillover mechanism to be valid. Thus the above quotation from Lutkepohl forms the hypothesis for chapter four that:

- **The dominant US stock market does not influence the results of the volatility spillover between the Shanghai and Hong Kong stock markets, in which the results are generated in a bivariate system.**

Examining the volatility spillover between financial markets whose dimensions are known is a relatively straightforward matter involving nothing more complicated than multivariate ARCH-family models. On the other hand, to consider the US impacts on the volatility spillover between the Shanghai and Hong Kong stock markets with integrated volatility processes seems far more difficult. This problem can be solved with the aid of co-persistent in variance theory proposed by Bollerslev and Engle (1993). The application of the co-persistent in variance theory expands the earlier treatment of the volatility spillover processes in the sense that it brings together the related volatility modelling that has been developed.

Moreover, this theory can be applied by showing how the US factor can be incorporated into the volatility spillover mechanism between the Shanghai and Hong Kong stock markets. In doing so it identifies the valid relationship between the Shanghai and Hong Kong stock markets. It also shows that different treatments the US factor by applying the multivariate ARCH family models and the co-persistent theory should yield the same conclusion on the relationship between the Shanghai and Hong Kong stock markets.

1.4.3 The identification of potential sources that facilitate the linkage between the Shanghai and Hong Kong stock markets:

One of the most distinctive features of the two stock markets occurring since the handover is the information flow that has arisen. So far, we have analyzed the two stock markets as a whole, without focusing on which specific financial and economic variables are responsible for the information flow. Thus the theme of chapter five involves an identification of potential sources that facilitate the linkage between the Shanghai and Hong Kong stock markets. This is necessary because the depth of information flow depends on financial and economic factors, including the level of the stock market development and the extent to which the stock market dominates. It might also hinge on political factors which, however, are beyond the scope of this thesis. The approach adopted for identifying the potential sources is the structural vector auto-regression model ($SVAR$), which provides a convenient framework for taking into account various statistical and econometric issues which arise in identifying a set of independent shocks. In this chapter, it starts with raising the aim of this chapter, and then the basic statistical formulation of the $SVAR$ approach are outlined. This is followed by an identification of the potential shock sources and how they affects the Shanghai stock prices. When discussing the structures VAR analysis for the two stock markets, the innovation accounting examinations are also addressed.

1.4.4 An examination of the impact of the Shanghai bond futures market on the underlying spot bond market:

This is the theme of the sixth chapter, involving an investigation into the problems of the closure of the Shanghai bond futures market by the Chinese policy agency. Assessment of derivative markets has recently become a key area for the international literature since their introductions. Concerns remain, however, about whether derivative markets meet the needs of emerging markets with difficulties in understanding the functions or roles of derivative markets. With its unique history, the examination of the Shanghai bond futures market may help to highlight the needs of such markets. For this reason, examining the relationship between the Shanghai bond futures market and its underlying spot market could provide new insights for understanding the functions of information and become desirable.

A major issue in identifying the needs is whether the bond futures market is necessary, or whether the bond futures market had fulfilled its functions. Several approaches have been applied to examine the functions of the futures market, including the multivariate $ARCH$ model and the nonlinear Granger causality test. It should be noted that this chapter is not inconsistent with the contents discussed in other chapters. In fact, this chapter also draws on the roles of information. Therefore, we view this chapter as complementing other chapters, in that it attempts to bring additional aspects of the way that how information functions between the derivative market and its underlying spot market.

The last chapter draws summary conclusions on the need to recognize the benefits of the information linkages among different markets. It argues that the consequence of the information flow on which the conditional volatility is based can improve price discovery processes. The key issue for policy makers is not to curb such informations linkages. Finally, the contributions and implications of this research are provided.

Chapter 2

Relevant methodology literature review

2.1 Introduction

Many scholars have contributed to a burgeoning literature that sheds new light on how the volatility spillover between stock markets and their financial effects. The purpose of this chapter is to evaluate these past and present literatures, and especially to justify the properties of their methodologies, among which some have been applied in this thesis.

Much of the existing research on the relationship between information arrivals and stock price volatility centres on how major macro-economic statistical data releases affect the underlying stock market volatility (although efforts have also been made to analyze foreign currency markets, capital flow, and other facets of international economic relations). Various studies focus on international stock markets, emphasizing how the rate of information flow among geographically separated stock markets can be measured by the volatility of stock prices. While many other studies have also focused considerable attention on the influence of financial derivatives on the underlying stock market, examining whether financial derivatives are more capable of achieving the required functions. If this ability is crucial to achieving their designed functions such as price discovery, then it may follow that the underlying spot market exhibits less volatile behavior in the financial system. At the same time, however, these research leave various empirical issues unsolved, including which economic factors bear most heavily on information flow and the nature of their effects and whether the openness of the emerging stock markets bring new levels of transparency and liquidity to financial markets.

The resolution of these issues is likely to help clarify whether or not the Shanghai stock market should be opened to the outside world and apply derivatives. This literature analysis is structured on two aspects:

- How is the stock market volatility connected with information, how should stock volatility be defined, and what are the economical and financial policy consequences of stock market volatility spillover?
- Why do most research measure information flows rely solely on the ARCH-family model, the cointegration analysis, and the structural VAR model? This issue is illustrated at greater length later in this chapter.

2.2 The information contents and the econometric analyses of the changing stock prices and conditional volatility through time

Extensive scholarly interest in public information on the stock markets has yet to generate a widely accepted definition of information. Differences over the widely accepted definition of information hinge on the importance of economical and financial proximity. More than twenty years ago, Rozeff and Kinney (1976) find there is abnormal stock returns in January because of the above-average amount of firm-specific information at the turn of the year. Mitchell and Mulherin (1994) employ the number of observable and broad announcements released by daily by Dow Jones & Company to examine the aggregate stock market returns. It, however, cannot be said that they have succeed in finding appropriate proxies for information which would be of much aid in deciding the importance of information for stock markets. While latter research focus on the impacts of macroeconomic, financial, and governmental announcements on stock market prices (On this issue, see French and Roll 1986; Penman 1987; Roll 1988; and Ederington and Lee 1996).

Besides proximity, many papers examine the role of public information arrival as the determinant of financial market volatility (See, for example, Berry and Howe 1994; and Bollerslev and Melvin 1994). Reflecting this position, Stephen A. Ross (1989) notes that " *In an arbitrage-free economy, the volatility of prices is directly related to the rate of flow of information to the market.*" In order to analyses the relationship between price volatility and the flow of information, he defines the form of information process following a stochastic differential equation such that $\frac{ds}{s} = \mu_s dt + \sigma_s dz_s$, where the variance of information flow σ_s^2 equals the variance of price change σ_p^2 in the absence of arbitrage. While there is no ideal definition, Ross has suggested a good idea that information can be measured. Exactly how to apply the measurement of the effects of information on prices and price volatility remains different.

The above papers, for example, consider the impact of the available information on the behavior of prices and price volatility within a single financial market. While others consider the impacts of information flow between two different regional markets, and still others consider it more than two regions. Furthermore, considerable interests have been focused on how the derivatives market, such as futures, stock index futures, and options, can bring new information and subsequently influence the underlying market price and price volatility. Setting aside the issue of how the information should be defined, this chapter focuses primarily on the information influence on the measurement for the volatility spillover among bivariate, multivariate regions and the relationship between the derivatives market and the underlying spot markets.

2.2.1 Information measurement in the ARCH model perspective

Some evidence considerations

The development of measures for information has important financial policy and investment strategy implications. Such measures should necessarily entail the interaction of dynamic of price volatility and the information involved, and of the transmission of information among stock markets. In an early study on price volatility and its international market transmission, Roll (1989) provides a survey paper on how and why the crash of October 1987 propagates internationally. Even though most of Roll's review papers suffer from selection bias problem (Daniel B. Nelson, 1989), Roll provides a striking consensus on the existence of international volatility influences. The existence of volatility spillover effect is equivalent to a international information flow. Consistent with this proposition, a contagion model of international volatility spillover developed by King and Washwani (1989) reveal that, in their words, "*In a non fully revealing equilibrium, price changes in one market will...in a real sense depend on price changes in other countries through structural contagion coefficients. Mistakes or idiosyncratic changes in one market may be transmitted to other markets, thus increasing volatility*". We have seen that much of the literature have noticed that information releases and transmission are the potential sources of much of market volatility. Many papers are also aware of the issues of heteroskedasticity of the return variates involved.

In Diebold and Lopez (1994)'s modelling volatility dynamics literature review (pp.435), the random price change error per unit of calendar time (ε_t) is a function of information arrival of calendar time (I_t), $\varepsilon_t = \sum_{i=1}^{I_t} \delta_i$, with $\delta_i \overset{i.i.d}{\backsim} N(0, \eta)$. And after a transformation, $\varepsilon_t = (\eta I_t)^{1/2} z_t$, with $z_t \overset{i.i.d}{\backsim} N(0, 1)$. Thus, ε_t characterized by conditional heteroskedasticity is connected with information that is incorporated into stock prices. Ederington and Lee (1993) examine the null hypothesis that the variance returns are constant by applying the Brown-Forsythe-modified Levene test, showing the hypothesis of homoskedasticity is overwhelmingly rejected with the macroeconomic news arrival. So the measures must have been carefully designed, firstly, to establish an effective connection with the time varying stock price return and volatility; secondly, to target volatility process which is compatible with information generating process. In other words, these measures should take into account the non-linear dynamic properties of information generating process that whether information is reflected in stock prices immediately and disappears or its effects on volatility persist over a longer period.

Much of the literature on the sources of information movement among the stock markets draws upon on modeling conditional variances as the function of past conditional variances and past squared innovations such as the autoregressive conditional heteroskedasticity ($ARCH$) and its generalized autoregressive conditional heteroskedasticity ($GARCH$) processes. Developed primarily by Engle (1982) and Bollerslev (1986) respectively, their works on conditional volatility process serve as a point of departure for the following analysis. Tsay (1987) extends Engle's $ARCH$ model to his Conditional Heteroscedastic Autoregressive moving Average ($CHARMA(p,q,r,s)$) model, which is unfortunately unpopular among econometricans perhaps due to its complexity in model building. Preferential the general autoregressive conditional heteroscedastic ($ARCH$) models have several quality properties:

- liberalizing conditional variances to capture volatility clustering (i.e. conditional variance heteroskedasticity) while keeping unconditional variances constant for statistical convenience or even allowing them inexistent. As Engle (1982) explained, the conditional variance of the current random stock return errors ε_t (i.e., $R_t = E(R_t|I_{t-1})+\varepsilon_t$, where I_{t-1} is the σ-field generated by all the available information at time $t-1$) is a function of its own serially correlated squared lagged errors induced by the dependent arrival information. That is assuming that $\varepsilon_t|I_{t-1} \backsim N(0,h_t)$ where $h_t|I_{t-1} = \alpha_0+\alpha(L)\,\varepsilon_t^2$ with $\alpha_0 > 0, \alpha(L) = \sum_{i=1}^{q} \alpha_i L^i$, $\alpha_i \geq 0$ and $L^i\varepsilon_t^2 = \varepsilon_{t-i}^2$ for $i = 1, ..., q$. All the roots of polynomial $1-\alpha_1 z - \alpha_2 z^2 - \cdots - \alpha_p z^p = 0$ lie outside the unit circle (i.e. $\|z\|_\infty > 1$) to guarantee the unconditional variance being homoscedastic. and conditional variance being finite. Just how long a shock persists in conditional variance process hinges largely on the order of the lag q which can be determined by some information criteria;

- not only distinguishing the normality distribution assumption for the conditional random errors ε_t from the unconditional nonmoral distribution for random errors ε_t, but also keeping the flexibility to admit other parametric conditional distribution assumptions for ε_t. The unconditional distribution of stock returns or ε_t has heavy tails compared with normal distribution with its fourth moments being 3. Engle (1982) again shows that the conditional fourth moment property generated by $ARCH(1)$ process for ε_t conditional on past information is larger than that of the standard normal distribution. That is $\frac{E\left(\varepsilon_t^4 | I_{t-1}\right)}{\sigma_\varepsilon^4} = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3$, provided $3\alpha_1^2 < 1$ for the existence of the kurtosis; consequently seizing some parts of the leptokurtosis for the stock returns. However, the conditional normal distribution assumption can not apprehend sufficiently the whole excess unconditional fourth moments. As a result, the conditional t distribution and generalized error distribution (GED) assumptions are utilized to describe the conditional leptokurtic distribution for ε_t by Engle and Bollerslev (1986) and Nelson (1991) respectively, while some other non-normal conditional distribution are applied by researchers as well;

- evolving abundant source of $ARCH$ processes to pursue the empirical regularities of stock returns while keeping the ability to illustrate some unique characteristics of information flow processes for both from within a single stock market *vs.* international markets and the derivative market *vs.* its underlying spot market. Generally, negative pieces of information provoke stock markets more volatile than positive pieces of information (i.e. the asymmetric effects for stock returns) and this particular feature enriches the ARCH model family from symmetric to asymmetric $ARCH$ models. Hentschel (1995) even works out, in his words, "*a family of models of generalized autoregressive heteroskedasticity* ($GARCH$) *that encompasses all* *the popular existing GARCH models"* at the expense of sacrificing the simplicity property of the $GARCH$ models. Even if these asymmetric conditional variance functions may become cumbersome in order to accommodate the asymmetric effect, they can nonetheless enhance our understanding of the two information flow hypotheses provided by Engle, Ito and Lin (1990): heat waves and meteor showers.

The key structure of the family of $ARCH$ models characterize the idea that there is a interaction between conditional variances expressing various properties of stock returns and information transmissions. The incorporation of different pieces of information set from other foreign markets into local markets, which is of importance for many decision strategies, is similar to the information formation within a single market. Considerable interests have been emerged from the investigation of such interactions in terms of different markets. We focus primarily on this area as well. A number of phenomena attributed to information flow are referred to collectively as "volatility spillover" effects (e.g., Hamao, Masulis, and Ng, 1990; Hamao, 1991). In their univariate $ARCH$ model, which is in contrast to an autoregressive model (McQueen, 1992), the foreign information represented by the most recent squared residuals estimated from foreign stock returns are directly added into the expression for the local conditional variance functions. That is, $h_t^{local}|I_{t-1} = ARCH\ Model + \sum_{i=1}^{n} \delta_i X_{it}^{foreign}$, where X_{it} is the most recent foreign squared stock return residuals in foreign market i, and n denotes the number of foreign markets. The parameter δ_i measures the volatility spillover effects (i.e., meteor shower) from foreign market i to the local stock market. The evidence indicates significant volatility spillover from the New York stock market to the London and Tokyo stock markets, even though this evidence may suffer the incorrect inference test as argued by Susmel and Engle (1994).

One particular research field that directly addresses the volatility spillover effects concerns the way that multivariate conditional variance functions for different markets are estimated simultaneously to help using information more efficiently (Bollerslev, Engle, and Nelson, 1993). However, the multivariate ARCH models are troubled by the overwhelming large number of parameters that need to be estimated (see the $VECH$ model of Bollerslev, Engle, and Wooldridge, 1988). Soon thereafter, it became apparent that many approaches that could reduce the dimensionality of the parameter space while maintain the conditional covariance matrix positive definite by imposing plausible restrictions have been developed (for reviews on this issue, see Engle, Ng, and Rothschild, 1990; Engle and Kroner, 1995). Improvement from univariate to multivariate models and factor-ARCH models can also help us to examine how economic, political, and market environment cause changes in volatility more directly (see Bollerslev, 1990; Karolyi, 1995). Finally, a relatively large proportion of the research on hypothesis volatility spillover is concerned with the impact of the introduction of derivative markets on underlying cash markets or the impact of large markets on small markets. Multivariate conditional functions are capable of reconstructing the time varying inter-markets dependence generating process (e.g., Chan, Chan, and Karolyi, 1991; Shastri, Sultan, and Tandon, 1996; Conrad, Gultekin, and Kaul, 1991).

Conditional variance persistent and co-persistence considerations

As have already mentioned, the Autoregressive Conditional Heteroskedasticity ($ARCH$) models have rich 'semantic' structures and it is likely that we could use them to measure information more appropriate than simply the approximate measurement of unconditional approaches. Despite of this distinguishing feature, economtricans have devoted relatively less attention to the issue that, for example in the $GARCH(p, q)$ model, $h_t|I_{t-1} = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}$, there may exist an approximate unit root in the autoregressive polynomial for the model; i.e. the polynomial $1 - \left(\alpha_1 + \cdots + \alpha_q + \beta_1 + \cdots + \beta_p\right) = 0$ has a root exactly on the unit circle. An issue is referred as integrated in variance or IGARCH model by Engle and Bollerslev (1986), in which the second and the fourth unconditional moments are infinite for the integrated ARCH models. For the integrated GARCH model, shocks to the conditional variance s steps in the future h_{t+s} do not die out because the weights on past shocks equal to those of the most recent shocks. As a result, the volatility has persistent impact on the stock prices that is, as remarks by Engle and Bollerslev, *"the current information remains important for the forecasts of the conditional variances for all horizons"*.

This property for the conditional variance is called persistence in variance. It worth noting that this persistent effect does not necessarily mean that the IGARCH model depends on the initial value over time. Nelson (1990) points out in his analyzing the properties of $GARCH\,(1,1)$ that the persistence property path for the $IGARCH\,(1,1)$ depends entirely on the signs of $E\left[\ln\left(\beta+\alpha z_{t-i}^2\right)\right]$, where $z_t = \varepsilon_t/\sqrt{h_t}$, with $E\left[z_t\right] = 0$,and $E\left[z_t^2\right] = 1$. One of the important results that Nelson obtains for the IGARCH$(1,1)$ process is that even though the unconditional variances may not exist, their process is still strictly stationary but not covariance stationary. "*This illustrates the important point that the behavior of a martingale can differ very sharply from the behavior of a random walk*" as he points out. However, early efforts to analyze the integrated variances were heavily criticized by the possible presence of structural change in the unconditional variances (see Diebold (1986), Lamoureux and Lastrapes (1990), Cai (1994)). Lately, research on integrated variances have been revived, especially in research on co-persistence variances and long-memory variances. Much such analyses suggest that increased information flows among stock market have impacts on the applications in financial time series data, leading researchers to develop new directions that can promote deeper understanding the role of information.

A related line of research suggests that when the variance persistence in a univariate framework is extended to a multivariate framework, the information influence of groups of conditional variances can affect the portfolio selection strategies and the related econometric consequences. Bollerslev and Engle (1993) present a theoretical consideration on cointegration for conditional variances. They view the common persistence in conditional variances as implying the existence of one or more long-run relationship(s) among the individual persistent conditional variances. For ease of reference, it is useful to repeat some Bollerslev and Engle's definitions and theorems. Let the stochastic variance process is defined to be

$$\limsup_{t\to\infty} \left| E_s\left(vech\left(\mathbf{H}_t\right)\right) - E_0\left(vech\left(\mathbf{H}_t\right)\right)\right| \equiv \limsup_{t\to\infty} \left|\left(\mathbf{H}_t^*\left(s\right)\right)\right|$$

Where E_i represents the expectations operator conditional on current and past values of the process of $\mathbf{H}_t$ at time i; $vech\left(\bullet\right)$ is an abbreviation for "vector-half" operator and $\mathbf{H}_t$ is the $n \times n$ conditional variance-covariance matrix. This definition implies that whether the stochastic variance process $\mathbf{H}_t$ almost surely uniformly converge to random variables depends on the least upper boundeness of $\mathbf{H}_t^*\left(s\right)$.

The non-singular conditional variances $\mathbf{H}_t^*\left(s\right)$ converge if and only if $\left\|\lambda\right\|_\infty = Max\left|\lambda_i\right| < 1$, where λ_i is the characteristic roots of $\mathbf{H}_t^*\left(s\right)$ with $i = 1, ..., n\left(n+1\right)/2$. That is, $\limsup_{t\to\infty} \left|\left(\mathbf{H}_t^*\left(s\right)\right)\right|$ almost surely uniformly converges to zero by the strong law of large numbers for all $s > 0$.

If, however, $\left|\lambda_i\right| = 1$ for some i, then $\limsup_{t\to\infty} \left|\left(\mathbf{H}_t^*\left(s\right)\right)\right| \neq 0$; Bollerslev and Engle define this situation as *persistent in variance.* However, in a similar spirit to the definition of persistence in mean (see Engle and Granger 1987), if a non-zero $n \times 1$ vector γ, $\gamma \in \acute{R}^n$, can make linear combinations of each persistent variances, i.e. $\limsup_{t\to\infty} \left|\left(\mathbf{H}_t^*\left(s\right)\right)\right| \neq 0$, to be covariance stationary such that

$$\limsup_{t\to\infty} \left|E_s\left(\gamma'\left(\mathbf{H}_t\right)\gamma\right) - E_0\left(\gamma'\left(\mathbf{H}_t\right)\gamma\right)\right| = \limsup_{t\to\infty} \left|vec2\left(\gamma\right)'\left(\mathbf{H}_t^*\left(s\right)\right)\right| = 0$$

almost sure for all $s > 0$, where the $vec2\left(\cdot\right)$ is an operator for γ, then the process of the conditional variances $\mathbf{H}_t^*\left(s\right)$ is called *co-persistence in variances.* Bollerslev and Engle (1993) successfully apply their theoretical work to examine the information source for exchange rates. They find that the common persistence in the volatility in several foreign exchange market can be attributed to dollar-related news.

2.2.2 Long run relationships in systems of equations

Cointegration and Unit Root Analysis

The needs to form the concept of cointegration rest on the interests of econometricians to distinguish long term solutions and short term solutions among groups of financial or economic variables. In the preceding subsection, we have reviewed the ways that these concerns might operate separately or in combination to analysis information behavior in variances. But extensive work has been done in attempting to define and measure the distinction between long term and short term relationships among those financial and economic variables on level. Johansen (1996) suggests that it is easier, more accurate, more consensual, and more beneficial for econometricians to identify long run relation with combination of simultaneous values and short term relation with dynamic values. A long run equilibrium relationship is viewed as having achieved a correct specification while it is free from characteristics of spurious regression, in which two or more independent non-stationary variables display strong correlation when regressing one variable on the other.

From the cointegration perspective, the long run relationship between non stationary variables, $\mathbf{Y}_t$, is not only stable, that is actual deviations from this long run relationship being a median-zero stationary process (Banerjee *et al.* 1993, pp. 4), but also the long term relationship remains perceptually constant unless the structure of the underlying system changes. That is, within broad limits, the long run equilibrium state associated with non-stationary variables remains invariant as information arrives—the relationship of long solutions keeps the same whether the information shocks persist forever in the underlying variables or not. This implies that cointegration 'filters' the persistent shocks for the combination of the underlying variables. This is so because if the underlying non-stationary variables share a common stochastic trend(s) (Johansen, 1996), the cointegration can extract this common trend by taking suitable linear combinations for the integrated of order one process $\boldsymbol{\beta}'\mathbf{Y}_t$. Because, based on the Beveridge-Nelson decomposition, the $I(1)$ process $\mathbf{Y}_t$ can be represented by a reduced number of common random walks plus a stationary component as $\mathbf{Y}_t = \mathbf{Y}_0 + \mathbf{C}(1)\sum_{i=1}^{p}\boldsymbol{\varepsilon}_i + \mathbf{C}^*(L)\boldsymbol{\varepsilon}_t$, the suitable linear combination $\boldsymbol{\beta}'\mathbf{Y}_t$ can remove the random walks elements $\sum_{i=1}^{p}\boldsymbol{\varepsilon}_i$ if $\boldsymbol{\beta}'\mathbf{C}(1) = \mathbf{0}$.

The first strategy adopted for testing cointegration relations was Engle and Granger's (1987) two-step OLS procedure with a strong assumption that there exists only one unknown long run relationship among the cointegrated non-stationary variables. Engle and Yoo (1987) provide the correct critical values for testing the integrated order of the estimated OLS residuals. Heavily built on the multivariate Beveridge-Nelson (1981) decomposition of the non-stationary variables, Stock and Watson (1988) directly test the cointegrated model through testing whether the number of common stochastic trends will reduce from k to m, then there will be $k-m$ cointegrating vectors. The levels of variables can then be expressed by the m common trends plus stable components. The common trend representation can provide a convenient approach for the long run forecasts in the cointegrated systems. King et al. (1987) defined the common trend as the technology shocks in their paper. It is obvious that these tests are non-likelihood-based approaches.

A second strategy stemmed from the requirements that a majority of research on economic and financial analysis have been based on studies of multivariate variables analysis. With some important exceptions, relatively little has been done testing directly two variables. Accordingly, likelihood based inference tests for cointegration analysis have been adopted. These techniques could enrich our knowledge of cointegrating relationships among the interested variables. However, the initial difficulty for likelihood-based test is that test statistics are no longer standard distributed due to the presence of unit root processes. Given that the cointegration relations can be a priori known, or unknown, or some known and some unknown, different test statistics have to be applied. For the reason that most of the times we have no prior information for knowing exact cointegrating relations, Johansen (1988) proposes the most important full information maximum likelihood ratio statistics estimating the number of cointegration relationships, cointegrating parameter vectors and inference for the cointegration space through the error-correction model (ECM) [1]. He first defines the process $\mathbf{Y}_t\,(n \times 1)$ following a $VAR\,(p)$ in levels, $\mathbf{Y}_t = \mathbf{C} + \mathbf{\Pi}_1\mathbf{Y}_{t-1} + \cdots + \mathbf{\Pi}_p\mathbf{Y}_{t-p} + \boldsymbol{\epsilon}_t$, and reparametrises it into the error correction form under certain restrictions as

$$\Delta\mathbf{Y}_t = \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{\mathbf{k-1}}\Delta\mathbf{Y}_{\mathbf{t-p+1}} - \boldsymbol{\alpha\beta}'\Delta\mathbf{Y}_{\mathbf{t-p}} + \boldsymbol{\epsilon}_t \quad t = 1, ..., T$$

where the size of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ matrices is $r \times n$ and their rank is r with $r \leqslant n$, and the white noise process $\boldsymbol{\epsilon}_t$ is normally distributed as $\boldsymbol{\epsilon}_t \sim N\,(\mathbf{0}, \mathbf{\Omega})$. The conditional distribution of $\Delta\mathbf{Y}_t$ based on past information is also normally distributed as

$$\Delta\mathbf{Y}_t \mid \mathbf{Y}_{t-1,}\mathbf{Y}_{t-2}, \cdots \sim N\left(\Delta\mathbf{Y}_t - \textstyle\sum_{i=1}^{p-1}\mathbf{\Gamma}_i\Delta\mathbf{Y}_{\mathbf{t-i}} + \boldsymbol{\alpha\beta}'\Delta\mathbf{Y}_{\mathbf{t-p}}, \mathbf{\Omega}\right)$$

The variance-covariance matrix $\mathbf{\Omega}\,(n \times n)$ is positive definite. Apart from the constant elements, the log likelihood function of $\Delta\mathbf{Y}_t$ subject to the restriction $\boldsymbol{\alpha\beta}'$ and the unit matrix Jacobian factor is given by

$$L\left(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{p-1}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{\Omega}\right) = -0.5T\log|\mathbf{\Omega}| - 0.5\textstyle\sum_{t=1}^{T}$$
$$\left\{\left(\Delta\mathbf{Y}_t - \textstyle\sum_{i=1}^{p-1}\mathbf{\Gamma}_i\Delta\mathbf{Y}_{\mathbf{t-i}} + \boldsymbol{\alpha\beta}'\Delta\mathbf{Y}_{\mathbf{t-p}}\right)'\mathbf{\Omega}^{-1}\left(\Delta\mathbf{Y}_t - \textstyle\sum_{i=1}^{p-1}\mathbf{\Gamma}_i\Delta\mathbf{Y}_{\mathbf{t-i}} + \boldsymbol{\alpha\beta}'\Delta\mathbf{Y}_{\mathbf{t-p}}\right)\right\}$$

The only interested matrix is $\boldsymbol{\beta}$ whose rank determines the number of cointegrating relationships. So in order to maximize the parameter matrix $\boldsymbol{\beta}$ without considering the rest matrices $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{p-1}, \boldsymbol{\alpha}$,and $\mathbf{\Omega}$, we could concentrate the likelihood functions $L\left(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{p-1}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{\Omega}\right)$ three times[2].

Firstly, maximize the likelihood function $L\left(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{p-1}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{\Omega}\right)$ by assuming $\mathbf{\Omega}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as given so as to concentrate out $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{p-1},$. This is equivalent to projecting

[1] Without considering any economic or financial theories, Johansen assumes that the number of cointegrating relationships are totally unknown to econometricans before testing for cointegrating vectors. Watson (1994) presents the Wald tests statistics if exact cointegrating relationships are of a priori knowledge under the alternative hypothesis.

[2] For details see Bierens (1994), Hamilton (1994) and Johansen (1995). It should be noted that Ahn and Reinsel (1990) propose the iterative least square approach to estimate cointegrating vectors.

$\Delta\mathbf{Y}_t$ and $\boldsymbol{\alpha\beta}'\Delta\mathbf{Y_{t-p}}$ onto the space spanned by $\sum_{i=1}^{p-1}\boldsymbol{\Gamma}_i\Delta\mathbf{Y_{t-i}}$. In doing so, two auxiliary OLS regressions are used to generated first difference stationary residuals $\mathbf{R}_{0t}$ and level non-stationary residuals $\mathbf{R}_{pt}$ by regressing $\Delta\mathbf{Y}_t$ and $\mathbf{Y_{t-p}}$ on $\Delta\mathbf{Y}_{t-1},\ldots,\Delta\mathbf{Y}_{t-p+1}$, respectively as

$$\begin{aligned}\Delta\mathbf{Y}_t &= \boldsymbol{\zeta}_1\Delta\mathbf{Y}_{t-1}+\cdots+\boldsymbol{\zeta}_{k-1}\Delta\mathbf{Y}_{t-p+1}+\mathbf{R}_{0t}\\ \mathbf{Y}_{t-k} &= \boldsymbol{\xi}_1\Delta\mathbf{Y}_{t-1}+\cdots+\boldsymbol{\xi}_{k-1}\Delta\mathbf{Y}_{t-p+1}+\mathbf{R}_{pt}\end{aligned}$$

The concentrated log likelihood functions is given by

$$\begin{aligned}L^c(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Omega}) &\equiv L\left(\boldsymbol{\Omega},\boldsymbol{\Gamma}_{1(\boldsymbol{\alpha},\boldsymbol{\beta})},\ldots,\boldsymbol{\Gamma}_{k-1(\boldsymbol{\alpha},\boldsymbol{\beta})},\boldsymbol{\alpha},\boldsymbol{\beta}\right)\\ &= -\tfrac{1}{2}T\log|\boldsymbol{\Omega}| - \tfrac{1}{2}\sum_{t=1}^{T}\left[\left(\mathbf{R}_{0t}+\boldsymbol{\alpha\beta}'\mathbf{R}_{pt}\right)'\boldsymbol{\Omega}^{-1}\left(\mathbf{R}_{0t}+\boldsymbol{\alpha\beta}'\mathbf{R}_{pt}\right)\right]\end{aligned}$$

Secondly, seek to isolate $\boldsymbol{\alpha}_{(\beta)}$ regressing $\mathbf{R}_{0t}$ on $\boldsymbol{\beta}'\mathbf{R}_{pt}$ based on the assumption that both $\boldsymbol{\beta}$ and $\boldsymbol{\Omega}$ are fixed. We obtain the OLS expression $\boldsymbol{\alpha}_{(\beta)}$ for $\boldsymbol{\beta}'$ such that $\mathbf{R}_{0t} = \left(\boldsymbol{\beta}'\mathbf{R}_{pt}\right)\boldsymbol{\alpha}_{(\beta)}+\epsilon_t$

$$\widehat{\boldsymbol{\alpha}}'_{(\beta)} = -\left(\boldsymbol{\beta}'\mathbf{R}_{pt}\mathbf{R}'_{pt}\boldsymbol{\beta}\right)^{-1}\boldsymbol{\beta}'\mathbf{R}_{pt}\mathbf{R}'_{0t} = -\left(\boldsymbol{\beta}'\mathbf{S}_{pp}\boldsymbol{\beta}\right)^{-1}\boldsymbol{\beta}'\mathbf{S}_{po}$$

With $\mathbf{S}_{pp}=\mathbf{R}_{pt}\mathbf{R}'_{pt}$ and $\mathbf{S}_{p0}=\mathbf{R}_{pt}\mathbf{R}'_{0t}$. Finally, fixed $\boldsymbol{\beta}$ to concentrate $\boldsymbol{\Omega}$ out, we get

$$\begin{aligned}\widehat{\boldsymbol{\Omega}}_{(\beta)} &= \left(\mathbf{R}_{0t}+\widehat{\boldsymbol{\alpha}}_{(\beta)}\boldsymbol{\beta}'\mathbf{R}_{pt}\right)\left(\mathbf{R}_{0t}+\widehat{\boldsymbol{\alpha}}_{(\beta)}\boldsymbol{\beta}'\mathbf{R}_{pt}\right)'\\ &= \mathbf{S}_{00}-\mathbf{S}_{0p}\boldsymbol{\beta}\left(\boldsymbol{\beta}'\mathbf{S}_{pp}\boldsymbol{\beta}\right)^{-1}\boldsymbol{\beta}'\mathbf{S}_{p0}\end{aligned}$$

With $\mathbf{S}_{00} = \mathbf{R}_{0t}\mathbf{R}'_{0t}$ and $\mathbf{S}_{0p} = \mathbf{R}_{0t}\mathbf{R}'_{pt}$. Let us manipulate the last expression of $L^c(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Omega})$, we get

$$\begin{aligned}&\sum_{t=1}^{T}\left[\left(\mathbf{R}_{0t}+\boldsymbol{\alpha\beta}'\mathbf{R}_{pt}\right)'\boldsymbol{\Omega}^{-1}\left(\mathbf{R}_{0t}+\boldsymbol{\alpha\beta}'\mathbf{R}_{pt}\right)\right]\\ &= \sum_{t=1}^{T} tr\left[\boldsymbol{\Omega}^{-1}\left(\mathbf{R}_{0t}+\boldsymbol{\alpha\beta}'\mathbf{R}_{pt}\right)'\left(\mathbf{R}_{0t}+\boldsymbol{\alpha\beta}'\mathbf{R}_{pt}\right)\right]\\ &= tr\left[\boldsymbol{\Omega}^{-1}\sum_{t=1}^{T}\left(\mathbf{R}_{0t}+\boldsymbol{\alpha\beta}'\mathbf{R}_{pt}\right)'\left(\mathbf{R}_{0t}+\boldsymbol{\alpha\beta}'\mathbf{R}_{pt}\right)\right] = \text{constant}\end{aligned}$$

Then the final concentrated log likelihood function becomes

$$L^c(\boldsymbol{\beta}) = -\tfrac{1}{2}T\log\left|\widehat{\boldsymbol{\Omega}}_{(\boldsymbol{\beta})}\right| = -\tfrac{1}{2}T\log\left|\mathbf{S}_{00}-\mathbf{S}_{0p}\boldsymbol{\beta}\left(\boldsymbol{\beta}'\mathbf{S}_{pp}\boldsymbol{\beta}\right)^{-1}\boldsymbol{\beta}'\mathbf{S}_{p0}\right|$$

Now the problem of maximization of $L^c(\boldsymbol{\beta})$becomes the minimization of the function $\left|\mathbf{S}_{00}-\mathbf{S}_{0p}\boldsymbol{\beta}\left(\boldsymbol{\beta}'\mathbf{S}_{pp}\boldsymbol{\beta}\right)^{-1}\boldsymbol{\beta}'\mathbf{S}_{p0}\right|$. Where $\mathbf{S}_{00}$ and $\mathbf{S}_{pp}$ denote the variances for the difference residuals $\mathbf{R}_{0t}$, and level residuals $\mathbf{R}_{pt}$, respectively; and $\mathbf{S}_{0p}$ and $\mathbf{S}_{p0}$ denote the covariance for the two generated residual series. By minimizing the function $\left|\widehat{\boldsymbol{\Omega}}_{(\boldsymbol{\beta})}\right|$, we can solve the eigenvalues of the matrix $\begin{bmatrix} \mathbf{S}_{00} & \mathbf{S}_{0p} \\ \mathbf{S}_{p0} & \mathbf{S}_{pp} \end{bmatrix}$, i.e., $\left|\boldsymbol{\lambda}\mathbf{S}_{pp} - \mathbf{S}_{p0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0p}\right|$, in which the number of the largest non-zero eigenvalues r, which turn out to be the rank of the symmetric matrix $\begin{bmatrix} \mathbf{S}_{00} & \mathbf{S}_{0p} \\ \mathbf{S}_{p0} & \mathbf{S}_{pp} \end{bmatrix}$, are derived by the computation of the r largest squared canonical correlations between $\mathbf{R}_{0t}$ and $\mathbf{R}_{pt}$ and tested by two likelihood ratio test statistics. Consequently, the matrix $\boldsymbol{\beta}\,(n \times r)$ has to be the first r columns of the corresponding matrix of eigenvectors, while the rest $n - r$ eigenvalues must be zero. This is because if not all variables are independent in $\mathbf{Y}_t$ then the matrix $\begin{bmatrix} \mathbf{S}_{00} & \mathbf{S}_{0p} \\ \mathbf{S}_{p0} & \mathbf{S}_{pp} \end{bmatrix}$ must be singular, consequently there is at least one eigenvalues must be zero. Perhaps it is worth noting that the eigenvectors needs to be normalized to make them unique. To obtain canonical correlations simply demonstrates that the concentrated maximum likelihood and canonical correlation methods are aimed at measuring the extent of the correlation between the stationary $\boldsymbol{\Delta}\mathbf{Y}_t$ and the linear combination of non-stationary $\mathbf{Y}_t$, $\boldsymbol{\beta}'\mathbf{Y}_t$. Thus the rank of matrix $\boldsymbol{\beta}$ is r, determining the number of cointegration vectors which in turn determine the equilibrium relationships among the variables in $\mathbf{Y}_t$.

Detail motivations by applying maximum likelihood estimation of parameters and canonical correlation analysis are proposed by Hamilton (1994, chapter 20). Johansen (1991) again provides further results on maximum likelihood estimation and likelihood ratio tests for cointegration allowing for constant term and seasonal dummy variables.

There are still other methods of estimating cointegration vectors available such as Boswijk's (1994) conditional cointegration analysis, in which the cointegration parameters of interest can be estimated conditional on weakly exogenous variables and the Wald test can be applied to examine the significance of the error correction variable, and Park and Phillips (1988, 1989)'s triangular system ECM representation, in which no dynamics are need to be estimated, simplifying the estimation and asymptotic inference procedures. The triangular representation can be treated as a special case for the ECM representation, while the Stock and Waston's common trend representation can also be derived from the triangular representation (see Hamilton 1994, p.578). Although the above measures differ in some details, their broad approaches and ideas are quite similar. Their aim is to find long term equilibrium solutions among a set of non-stationary variables.

While there has been much research on the structural properties of cointegration in economics, research has generated relatively little progress and still remains controversial in detailing international financial markets by using the concept of cointegration. Most papers on financial market efficiency and economic variables impacts on stock prices apply E. Fama's classical methodologies.

Granger (1986) perhaps is the first author to address the links between cointegration and financial market efficiency. While Granger's assertion that the presence of cointegrated financial prices is against the market efficient theory is certainly true in some situations, the problem he identifies is neither unique to all financial markets, (for example, financial futures market and its underlying spot market) nor necessarily, the most difficult the researcher faces. Applying Johansen's (1988) methodology, Kasa (1992) presents evidence concerning the number of common stochastic trends in the stock markets of five major industrial countries and he finds the levels of total return indices are strongly cointegrated around a single common stochastic trend. Richards (1995) argues that Kasa's findings *"is due to a failure to adjust asymptotic critical values to take account of the small number of degrees of freedom that remain in the Johansen's MLE procedure"*. Obviously, this is a strong argument because cointegration tests indeed suffer potential size distortions in high dimensional systems (Ho and Sorensen 1996) and bias towards the lag length specification and non-normal errors in systems (Cheung and Lai 1993).

However, this is not to say that nothing at all is known about the significance for cointegrating relations among financial variables. These studies on "pitfalls" of Johansen's ML test remind us to pay attention to the robustness of likelihood ratio tests. As a case in point, applying Johansen's maximum likelihood procedure and using the same sample data, Diebold et al. (1994) have reexamined Balllie and Bollerslev (1989)'s earlier tests on nominal dollar spot exchange rate and found no evidence of cointegration for the system of spot exchange rates. Diebold et al. then conclude that their results support to Granger's (1986) claim that cointegration should not present in efficient markets. Baillie and Bollerslev (1994) reply Diebold et al.'s argument and claim that "*However, a more detailed examination of the deviations from the estimated cointegrating relationship reveals that the exchange rates may well be tied together through a long memory* $I(d)$ *type process, rather than an* $I(0)$ *process*".

Central to the links between cointegration relations and financial markets and much more consensus are the effects of futures market on its underlying spot market. There should exist an equilibrium state between domestic futures market and its own spot market or foreign futures market and domestic spot market (i.e. this is defined as composite arbitrage) due to arbitrage opportunities. The arbitrage process represent the limitless opportunities of wealth through riskless profit. Duffie (1996 p.22) defines "*there is no arbitrage if and only if there is a strictly increasing linear function* $F : L \longrightarrow \Re$ *such that* $F(\bullet) = 0$ *for any trading strategy at any time.*" For financial markets, this linear function for futures and spots is exactly the idea of cointegration. The concept of cointegration may even apply in options market such as Put-Call-Parity theorem etc. Returning to the claim, no cointegration in efficient markets may appear bias in this situation. The links between the identification of cointegration and conditional variances remain important in financial econometrics area, although they have not been studied in sufficient depth. The existing literature does not furnish an adequate understanding of how to identify a set of cointegrated variances and how the information has been affected among stock markets.

One of the most distinctive features of the econometrics literature occurring since mid 80's is the autoregressive time series with a unit root framework in which they arose. The presence unit roots in time series has led to the formation of cointegration tests. On the one hand, the identification of cointegration with equilibrium has been established under the support of the unit root asymptotic distribution analysis, which has attempted to derive the limiting distribution properties for integrated variables under *"a general class of functional limit theory on metric spaces"* (Phillips 1987 p.278). On the other hand, *"Johansen's procedure provides an alternative approach to testing for unit roots in univariate series"* (Hamilton 1994, p.646).

The available evidence suggests that, from the international stock markets standpoint, most financial prices are integrated of order *one*. The classical asymptotic results for stationary time series variables can not be applied to test the presence of unit roots because these conventional inference of the asymptotic distributions of parameter estimators are divergent. These conventional results are actually special cases of the weak convergence Weiner process (Phillips (1988) gives a direction flag for unit roots literature). Except for few unit root tests, most unit root tests assume the null hypothesis of unit root process against the alternative of stationary process. The reason of such considering is that, for the researcher, a more important consideration maybe avoiding an early decision of accepting the variables being stationary. Moreover, the consequence of accepting the wrong null hypothesis is less serious than accepting the wrong alternative hypothesis. Testing unit root in time series and the asymptotic critical values are mainly proposed by Fuller (1976), Dickey and Fuller (1979), Phillips (1987), and Phillips and Perron (1988). To determine statistical inference for unit root process, we need assume continuous function $Y_T(r)$ to be the sum of the first r fraction of a sequence of independent normal variables U_t with mean zero and finite variance $\sigma^2 > 0$ as

$$Y_T(r) = T^{-0.5} \sum_{t=1}^{[Tr]} U_t$$

where $Var(Y_T(r)) = \frac{[Tr]}{T} \to r$ as $T \to \infty$ and $r \in [0,1]$. Then it follows the functional central limit theorem ($FCLT$) that $Y_T(\cdot)/\sigma$ converge to the continuous stochastic Wiener process $W(\cdot)$. This peculiar feature of this result is the most important steps tone for Fuller (1976) and Dickey and Fuller (1979)'s results while the rest of deriving statistical inference are just algebra manipulations. Phillips (1987), and Phillips and Perron (1988) relax the restrictive assumptions of U_t by allowing U_t to be serially correlated and heteroskedastic such that $\lim_{T\to\infty} E\left[\left(T^{-0.5}\sum_{t=1}^{T} U_t\right)^2\right] \neq \lim_{T\to\infty} \left[T^{-1}\sum_{t=1}^{T} E(U_t)^2\right]$. Thus the former two unit root tests might be viewed as special cases for the latter two tests.

Linear and nonlinear Granger Non-Causality

Granger causality is, at its roots, a theory of information knowledge. Using F-statistic test, Granger (1969) refers to it as an attempt to well predict one variable X based on all the information including past information of X itself and past information of other variable Y. Sims (1972) further defines the concept as whether the past, present, and future information of Y contribute significantly to the explanation of X. Hendry (1995, pp.176) expresses concerns that one major problem with above definition needs "*the universe of information to ascertain the effect of deleting the history of any one variable,*" and the other problem with the definition is the potential loss of information "*if the conditional distribution contains information about the parameters of the marginal distribution...which determines the parameters of interest.*" In fact, the concept of Granger non-causality is widely used in limited information sets. Starting with Engle, Hendry, and Richard (1983), the concepts of Granger non-causality, weak (and strong) exogeneity and (later) cointegration have been closely connected. Urbain (1992) and Johansen (1992) propose the cointegration estimation in the conditional single equation with less parameters, which is equivalent to the full system estimation, based on exogeneity.

Engle and Granger (1987) concern whether a set of non-stationary vector variables are cointegrated of having an equilibrium relationship. A necessary condition for the presence of cointegration is the existence of at least one way Granger causality (Granger 1988). As Caporale and Pittis (1997) put it in the $VAR(1)$ case, the presence of cointegration implies the eigenvalues of the parameter matrix for the VAR system being less than 1. If, therefore, the off-diagonal elements in the parameter matrix must not all be zero, then one way Granger causality must exists. Relying on testing Granger non-causality in unrestricted VAR system can lead to misinterpretation of what is tested because the test statistics may easily suffer from diverge limiting distribution and nuisance parameter problems if some or all variables in the VAR have unit root processes.

Phillips (1995) proposes a fully modified VAR estimation ($FM - VAR$) approach to solve these problems, in which the Wald test of causality has a standard χ^2 distribution without worrying non-stationary variables by taking into account of serial correlations in the residuals and the effects of second-order bias. This does not, as is seemed however, lead to specific methodological prescriptions that the FM-VAR approach should be the only approach for Granger causality research. However, as Toda and Phillips (1993) put it, '*Causality tests are valid asymptotically as χ^2 criteria only when there is sufficient cointegration.... The precise condition for sufficiency involves a rank condition on a sub-matrix of the cointegrating matrix. Since the estimates of such matrices in levels VAR's suffer from simultaneous equations bias there is no valid statistical basis for determining whether the required sufficient condition applies.*' Hence, as they argue if there exist cointegrations in the system, "*testing for causality in ECM's is more promising than in levels VAR's.*" Testing causality within the ECM framework can enable researchers to gain direct access to testing and distinguishing short run Granger non-causality and long run causality.

To complicate things further, sometimes it is desirable for researchers to test Granger non-causality in the nonlinear framework, in which the aim and methodology are totally different from those in the linear framework shown above. Although identified by Baek and Brock as comprising only little of Granger non-causality research, interest in and examples of uncovering non-linear causality in financial market research have grown (Hiemstra and Jones 1994, Brooks 1998). Hiemstra and Jones (1994) set out a non-parametric statistical causality testing procedure, a valuable approach to the study of relationships among financial variables in the domain of information clustering flow and '*the richer types of asset behavior that nonlinear models provide researchers.*'

2.2.3 Structural vector autoregressive (SVAR) analysis

Some evidence considerations

In thinking about the interactions among sets of financial and/or economic variables, we draw upon the literature on Structural Vector Autoregressive ($SVAR$) analysis in general, as well as work on Impulse Response Functions ($IRFs$) technique, first introduced by Sims (1980). The structural VAR literature is relevant to this thesis, because we expect the method of identifying sources within sets of financial and/or economic variables to have much in common with that of the interactions among sets of financial and/or economic variables.

The reduced-form Vector Autoregressive theory (VAR) has been more popular than the simultaneous equation model. The simultaneous equation model was an important functional form in econometric model building and it "*arises from economic theory in terms of the operations of markets and the simultaneous determination of economic variables through an equilibrium model (*Hausman 1983, pp.392).*"* Most of the VAR literature begins with Sim's (1980) puzzle about the suspicious treatment some variables being exogenous and the 'incredible identifying restrictions on short run coefficients' made in the traditional simultaneous equation model. Since the VAR model treats all observed variables as endogenous such that the *a priori* assumptions are weak (which is important indeed), every variable is explained by predetermined values of its and other endogenous variables. However, the consequence is the difficulty in identifying how the unobserved exogenous shocks affect each observed variable (Cooley and LeRoy (1985) is a good critique of the VAR model literature).

The SVAR literature indeed recognizes the importance of VAR's attributes: the unrestricted VAR's ability of inverting from VAR to the Moving Average (MA) representation, providing further insights into the dynamics of model; the ability of characterizing Granger-Non-Causality, giving further insights into the policy analysis; and more important the flexibility of incorporating integrated variables, allowing researchers to understand long run relationships within the system have a well established place in explanations of large scale variables interactions. These resources enable researchers to bear all the costs of applying the unrestricted VAR model even if it is of little of inherent interest.

However, the information of interest is locked in the variance-covariance matrix of the unrestricted reduced VAR residuals, so how do we unlock information that embedded in the unrestrained reduced VAR model to have economic interpretations? The empirical $SVAR$ literature sheds surprisingly light on how to transform the reduced from VAR relation into structural systems and even to estimate structural cointegrating relationships within the system. For example, in a study of traditional interpretation of the Keynesian economy model in the United States, Blanchard (1989), assuming the existence of five economical structural disturbances, constructs a structural model by applying a set of just-identifying restrictions based on the Keynesian model to start from the reduced form innovations to a set of uncorrelated structural innovations. His model is defined as C-model for the structural VAR model (see Amisano and Giannini 1997).

Blanchard and Quah (1989) similarly show how innovation variances estimated from reduced form VAR impose restrictions on recovering structural disturbances and structural parameters. Moreover, what makes their paper influential is that they recognize the presence of unit root processes for the variables, the potential permanent effects of disturbance shocks on the level of one of the variables. Then the long run economic theory restrictions based on the sum of the structural MA coefficient matrix can provide additional restrictions for the recovering identification.

Other evidence favors such long-run structural VAR models. In the study of measuring the long run effects of a common stochastic productive trend on real business cycle and economic fluctuation in the United States, King et al. (1991), another very influential and much-cited work often abbreviated as $KPSW$, further expand Blanchard and Quah's methodology by imposing cointegration restrictions on the sum of the structural MA coefficient matrix $\mathbf{\Gamma}(1)$ (such that $\mathbf{\Gamma}(1) = [\mathbf{A}, \mathbf{0}]$ and $\mathbf{A} = \boldsymbol{\beta}_{\perp}$, where $\boldsymbol{\beta}$ is the cointegrating vectors) to identify the permanent components which are uncorrelated with the remaining transitory components. King et al.'s main idea is to partition the structural disturbances comfortably into mutually uncorrelated permanent innovations (which can be determined from the reduced MA form) and temporary innovations, further more, each permanent innovation can be separated if there are more than one permanent innovation.

In addition to the imposition of restrictions to provide the VAR framework with structural content, the recent work of Amisano and Giannini (1997) points to the significance of "structural cointegrating VAR approach" for econometric modelling. The purpose of this approach is not only to bring together those properties of the structural VAR and cointegration analysis but also to derive the devotions of the long run structural shocks that are thought to be of interest to researchers. Because the "*the long-run reduced form shocks are embedded within an otherwise unrestricted log-linear VAR model of a given order in the variables of interest to obtain a cointegrating VAR model which incorporates the structural long-run relationships as its steady state solution* (Garrat et al. 1999)." The treatment of structural cointegrating VAR model is fairly exhaustive.

Theoretical & statistical analysis for SVAR and impulse response functions (IRFs)

A major body of work in the SVAR literature is constructed by Amisano and Giannini at the university of Brescia and the university of Pavia in Italy. Their work will be elaborated on here and has contributed significantly to the methodological basis of the work utilized in chapter 5. They coined the term 'AB-models' with it two special models namely 'K-models' and 'C-models' to reflect the nature of the strategy process to estimate the structural model:

"*In this kind of structural model, it is possible to model explicitly the instantaneous links among the endogenous variables, and the impact effect of the orthonormal random shocks hitting the system ...* through the $\mathbf{A}$ matrix, *induce a transformation on the* $\boldsymbol{\epsilon}_t$ *disturbance vector, generating a new vector* $(\mathbf{A}\boldsymbol{\epsilon}_t)$ *that can be conceived as being generated by linear combinations through* $\mathbf{B}$ *matrix of* n *independent orthonormal disturbances, which we will refer as* $\mathbf{e}_t$." (p.19)

In establishing the SVAR, it is necessary to start with the unrestricted VAR model given by

$$\mathbf{A}(L)\,\mathbf{y}_t = \boldsymbol{\epsilon}_t, \quad \text{where } \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Omega})$$

with the loglikelihood function for the reduced model $\mathbf{y}_t$ apart from the constants being

$$L_n = -\tfrac{T}{2}\log|\boldsymbol{\Omega}| - \tfrac{T}{2}tr\left(\boldsymbol{\Omega}^{-1}\widehat{\boldsymbol{\Omega}}\right), \quad \text{where } \widehat{\boldsymbol{\Omega}} = \textstyle\sum_{t=1}^{T}\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'$$

And if $\mathbf{A}$, $\mathbf{B}$ are invertible $n \times n$ matrices, then define the structural AB-model as

$$\mathbf{AA}(L)\,\mathbf{y}_t = \mathbf{Be}_t, \quad \text{where } \mathbf{A}\boldsymbol{\varepsilon}_t = \mathbf{Be}_t \text{ and } \mathbf{e}_t \sim N(\mathbf{0}, \mathbf{I})$$

with the loglikelihood function for the structural model being

$$L_n^s = \frac{T}{2}\log|\mathbf{A}|^2 - \frac{T}{2}\log\left(|\mathbf{B}|^2\right) - \frac{T}{2}tr\left(\mathbf{A}'\mathbf{B}'^{-1}\mathbf{B}^{-1}\mathbf{A}\widehat{\boldsymbol{\Omega}}\right)$$

In fact, the $\mathbf{A}$, $\mathbf{B}$ matrices with $2n^2$ parameters altogether cannot be estimated based on just $n(n+1)/2$ distinct elements of $\boldsymbol{\Omega}$ which is generated from the unrestricted VAR form. In other words, the information matrix for $\begin{bmatrix} vec\mathbf{A} \\ vec\mathbf{B} \end{bmatrix}$ is singular. Consequently, further restrictions on **A**, **B** matrices are needed by feeding information to them. The process of $SVAR$ model estimation is similar in analytical approach to that of VAR, making full information maximum likelihood estimation but also monitoring condition for identification, and further analysis from the score algorithm perspective is given in Amisano and Giannini (1997).

It is interesting to note that the work discussed above is focused on two of the analytical dimensions in structural analysis, namely structures of identification and structures of estimation. The third dimension is structures of cointegration, and this is the main focus of the paper by Garrat et al. (1999). It provides a more radical comprehensive analysis than the work on SVAR reviewed so far, in that it views the structural cointegrating VAR as an economic consistent modelling strategy that ' *it provides an explicit link between the estimated model residuals and the structural shocks of the underlying economic model.'* However, the work of Garrat et. al. is mainly based on ECM in emphasizing the long run structure of the macro-economy..

The theoretical summary above can be made more concrete by noting some of the power effects from the analysis on Impulse Response Functions (IRFs) technique which provides the basis for identifying the systems's responses to innovations in the variables[3]. We mainly focus on the asymptotic distributions of IRFs from the paper by Lűtkepohl (1990) and, as usual, the book by Amisano and Giannini , which are the final piece of work reviewed in this subsection. The two works draw heavily on Baillie (1987)'s work in emphasizing the limiting distribution of estimated moving average (MA) representation coefficients matrices, and the way in which MA representation is always intricately bound up with vector autoregressive (VAR) representation.

The computation and the relevant asymptotic distribution for the accounting innovations may start from the alternative companion form of the VAR (p)representation

[3]The detail defination for the impulse response functions (IRFs) can be found in Watson (1994).

$\mathbf{A}(L)\,\mathbf{y}_t = \epsilon_t$ with $E(\epsilon_t \epsilon_s') = \mathbf{\Omega}$ if $t = s$ and zero otherwise. The companion form has exactly the same information as the VAR representation. Let

$$\mathbf{Y}_t = \begin{bmatrix} \mathbf{y}_t \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p+1} \end{bmatrix} \quad \text{and} \quad \mathbf{\Upsilon} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{bmatrix}$$

such that

$$\underset{(np\times 1)}{\begin{bmatrix} \mathbf{y}_t \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p+1} \end{bmatrix}} = \underset{(np\times np)}{\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{bmatrix}} \underset{(np\times 1)}{\begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{y}_{t-2} \\ \vdots \\ \mathbf{y}_{t-p} \end{bmatrix}} + \underset{(np\times 1)}{\begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}}$$

In matrix notation this representation can be written as

$$\mathbf{Y}_t = \mathbf{\Upsilon}\mathbf{Y}_{t-1} + \eta_t$$

This companion form exploits the fact that it is easy to convert to the Wold representation because of its first order form, which is given by

$$\mathbf{y}_t = \sum_{j=0}^{\infty} \mathbf{J}\mathbf{\Upsilon}^j \mathbf{J}' \varepsilon_{t-j}$$

where $\mathbf{J} = [\mathbf{I} : \mathbf{0}]$ is the $(n \times np)$ dimensional matrix.[4] An explicit form of the structural

[4]The $\mathbf{Y}_t = \mathbf{\Upsilon}\mathbf{Y}_{t-1} + \eta_t$ process can generally be written in MA representation as: $\mathbf{Y}_t = \sum_{j=0}^{\infty} \mathbf{\Upsilon}^j \eta_{t-j}$. So that $\mathbf{y}_t = \mathbf{J}\mathbf{Y}_t = \sum_{j=0}^{\infty} \mathbf{J}\mathbf{\Upsilon}^j \eta_{t-j} = \sum_{j=0}^{\infty} \mathbf{J}\mathbf{\Upsilon}^j \mathbf{J}' \varepsilon_{t-j}$. Where $\mathbf{J}\mathbf{\Upsilon}^j\mathbf{J}'$ is the MA representaion coefficient matrices.(see Baillie 1987, p.106-111; Amisano and Giannini 1997, chapter 5)

MA representation for the AB-model is given by

$$\mathbf{y}_t = \sum_{j=0}^{\infty} \mathbf{J}\boldsymbol{\Upsilon}^j\mathbf{J}'\left(\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{e}_{t-j}$$

It then follows that the coefficient matrix $\mathbf{C}_j$, $\mathbf{C}_j = \mathbf{J}\boldsymbol{\Upsilon}^j\mathbf{J}'\left(\mathbf{B}^{-1}\mathbf{A}\right)^{-1}$, is the impulse responses of the system. Next we need to derive the asymptotic distributions for the impulse responses matrix. In order to avoid the difficulty in defining matrix differentiation, the algebra of vectors and $0-1$ matrices notations can be employed. It can be recognized that the structural MA representation $\mathbf{C}_j$ is based on $\mathbf{A}, \mathbf{B}$, and $\mathbf{A}_i$ with $i = 1, \ldots, p$ in VAR representation. Then we begin with taking *vecs* to get the asymptotic normal distributions for $vec\left(\mathbf{A}_i\right)$ and $vec\left(\mathbf{B}^{-1}\mathbf{A}\right)$ which is[5]

$$\sqrt{T}\left[\left(vec\left(\widehat{\mathbf{A}_i}\right) - vec\left(\mathbf{A}_i\right)\right) \;\vdots\; \left(vec(\widehat{\mathbf{B}^{-1}\mathbf{A}})^{-1} - vec\left(\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\right)\right] \sim N\begin{bmatrix} \boldsymbol{\Sigma}\otimes\mathbf{W}^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Theta} \end{bmatrix}$$

[5]Suppose the consistent estimator of $vec\left(\mathbf{A}_i\right)$ is asymptotically normally distributed as

$$\sqrt{T}\left(vec\left(\widehat{\mathbf{A}_i}\right) - vec\left(\mathbf{A}_i\right)\right) \sim N\left(\mathbf{0}, \boldsymbol{\Omega}\otimes\mathbf{W}^{-1}\right)$$

where $\mathbf{W} = p\lim \frac{1}{T}\mathbf{Y}_{t-1}\mathbf{Y}_{t-1}'$. And the asymptotically normally distribution for the consistent estimator of $vec\left(\mathbf{B}^{-1}\mathbf{A}\right)^{-1}$ is

$$\sqrt{T}\left(vec(\widehat{\mathbf{B}^{-1}\mathbf{A}})^{-1} - vec\left(\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\right) \sim N\left(\mathbf{0}, \boldsymbol{\Theta}\right)$$

where

$$\boldsymbol{\Theta} = (\grave{\mathbf{a}})\left\{\left[\; \mathbf{I}\otimes\mathbf{B}^{-1} \quad -\left(\mathbf{A}'\mathbf{B}'^{-1}\right)\otimes\mathbf{B}^{-1} \;\right]\boldsymbol{\Sigma}_{ab}\begin{bmatrix} \mathbf{I}\otimes\mathbf{B}'^{-1} \\ -\left(\mathbf{B}^{-1}\mathbf{A}\right)\otimes\mathbf{B}'^{-1} \end{bmatrix}\right\}(\acute{\mathbf{a}})$$

with $\grave{\mathbf{a}} = \left(\left(\mathbf{B}^{-1}\mathbf{A}\right)'^{-1}\right)\otimes\left(\mathbf{B}^{-1}\mathbf{A}\right)^{-1}$ and $\acute{\mathbf{a}} = \left(\left(\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\right)\otimes\left(\mathbf{B}^{-1}\mathbf{A}\right)'^{-1}$. The asympotic normal distribution for the $\mathbf{A}$ and $\mathbf{B}$ parameters with certain restritions is

$$\sqrt{T}\left(\begin{bmatrix} vec\widehat{\mathbf{A}} \\ vec\widehat{\mathbf{B}} \end{bmatrix} - \begin{bmatrix} vec\mathbf{A} \\ vec\mathbf{B} \end{bmatrix}\right) \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_{ab}\right)$$

On applying Lutkepohl (1989)'s proposition 2 based on the Taylor Expansion Theorem, the expression for the asymptotic distribution with respect to structural impulse responses coefficient matrices $vec\left(\mathbf{C}_0, \ldots, \mathbf{C}_h\right)$ can be derived as

$$\sqrt{T}\left[\left(\widehat{\mathbf{C}}_0, \ldots, \widehat{\mathbf{C}}_h\right) - \left(\mathbf{C}_0, \ldots, \mathbf{C}_h\right)\right] \sim N\left(\mathbf{0}, \mathbf{\Sigma}\left(h\right)\right)$$

Where $\mathbf{\Sigma}\left(h\right)$ is some combinations of $\mathbf{\Sigma} \otimes \mathbf{W}^{-1}$ and $\mathbf{\Theta}$ such that

$$\mathbf{\Sigma}\left(h\right)_{ij} = \mathbf{G}_i\left(\mathbf{\Sigma} \otimes \mathbf{W}^{-1}\right)\mathbf{G}_j' + \left(\mathbf{I}_n \otimes \mathbf{J}\mathbf{\Upsilon}^j\mathbf{J}'\right)\mathbf{\Theta}\left(\mathbf{I}_n \otimes \mathbf{J}\mathbf{\Upsilon}^j\mathbf{J}'\right)'$$

The purpose of compiling the above technical results involved is to synthesize a perspective on what the impact of one set of economic or financial variables would have as a result of changes in another set of variables and as a way of gaining a deeper understanding of the nature and process of IRFs themselves. From the above process, it is clear that IRFs only describe the dynamic properties of sets of variables rather than an aid to the model evaluation. Pesaran and Shin (1998) note another point, in the context of the comments of IRFs, that impulse response function analysis are subject to a model's estimated parameters, for example $\mathbf{\Upsilon}$ and $\mathbf{\Omega}$, no matter how inconsistently and inefficiently they are estimated.

Chapter 3

Volatility transmission between the Hong Kong and Shanghai stock markets

3.1 Introduction

This chapter investigates the volatility transmission mechanism between the Hong Kong and Shanghai stock markets before and after the transfer of Hong Kong sovereignty from Great Britain to China in July 1997. Our analysis is based upon an asymmetric ARCH model which allows us to decompose innovations into positive and negative components. Three indices are examined: the Hang Seng, Shanghai and the China-Affiliated (red-chip) indexes. Our results support the following propositions:

1. The squared stock index return residual series in both markets exhibit significant positive autocorrelations and the existence of long memory processes;

2. There are significant time variation in the estimated conditional variance of the three indexes in both markets;

3. There are statistically signiffcant volatility spillover effects from the Hong Kong to Shanghai stock markets both before and after July 1997;

4. The impact of this relationship has been to reduce the volatility of the Shanghai market, especially after July 1997.

Despite the huge difference between them, the Hong Kong and China's economies have been increasingly intertwined and inter-dependent since the early 1990s. The Chinese based Hong Kong companies raised US$ 14bn in the Hong Kong stock market in 1997, compared with a relatively meagre US$ 1.3bn in 1996. A plan of listing 250 top-quality Chinese companies in Hong Kong and establishing a developed set of derivative and debt products in the next five years has been launched by the Chinese government. At present, two of those Chinese-based or China-related corporations are included in the calculation of the Hang Seng Index. Furthermore, many Hong Kong local property companies, which dominate the Hang Seng Index of blue-chip companies, are heading across the border to invest in Shanghai infrastructure and housing markets. On the other hand, the Shanghai stock market has undergone major structural changes in the last few years and has become more accessible to foreign investors. And indeed Hong Kong companies are among the major foreign investors in Shanghai. The complexity of the relationship between the Hong Kong and the Shanghai markets — the political, economic and geographical closeness on one hand and the huge diversity in systems between the two markets on the other hand — is often considered as one of the major sources of volatility which affects both markets. After the handover in 1997, with increasing number of the China-affiliated companies (*red-chip*)[1] listed in Hong Kong, and along with them came some practice that was previously unfamiliar to many Hong Kong investors, there has been concern that the Hong Kong market may become more volatile as the result of more direct involvement from China. One the other hand, anxiety was also developed in the Shanghai market before

[1]The *red-chip* companies were first listed at the Hong Kong stock exchange on Jan. 1, 1993. These companies are registered in Hong Kong, but controlled by the ministries of central government, local and provincial governments, or state-owned enterprise groups of China.

and after the handover when financial link and transaction with Hong Kong reached historical record level. With a record number of major Shanghai companies cross-listed in Hong Kong, there has been a seemingly lead-lag relationship between the Hong Kong and Shanghai indices, caused by different time schedule in responding to the events that affect the both markets. The worries that the volatility will become excessive and uncontrollable as the result of more foreign involvement has resulted in stricter control on investment in the Shanghai share market.

This chapter seeks to discuss these concerns by analyzing the conditional second moment of the Hong Kong and Shanghai data using the Autoregressive Conditional Heteroskedasticity ($ARCH$) model and, hence, to establish empirically the nature and direction of the volatility contagion and spillover between the two markets. In the early studies, the contagion model has often been used to identify the effect of volatility spillover (*e.g.* King and Wadhwani (1989)). But it was found that the contagion model was inconsistent with fundamental news and was not capable of identifying the original shocks (Roll (1989)). The ARCH family models have since become the main framework in which various hypotheses associated volatility transmission have been tested and analyzed (see, among others, Hamao, Maulis, and Ng (1990), Engle, Ito and Lin (1990), Susmel and Engle (1994), Karolyi (1995), Koutmos and Booth (1995) and De Santis and Imrohoroglu (1997)). De Santis and Imrohoroglu (1997) show that there is evidence that the existence of volatility spillover from the developed markets does not necessarily increase the volatility in the emerging markets after liberalization.

In this paper, we apply the approach adopted by Glosten, Jagannathan and Runkle (1993) to examine the evidence from daily closing prices of the Shanghai, Hang Seng and *red-chip* indices from January 1, 1993 to August 25, 1998. The conditional volatility processes of the two markets are compared, with attention focused on testing the hypothesis that the shocks in stock return (the so-called meteor shower effects) originated in Hong Kong (or Shanghai) may have destabilizing effects on the other market. We are particularly interested in the period after the handover to China the Hong Kong sovereignty, and the impact of this historical event on the underlying conditional volatility process. Considerable time variation is found in the estimated conditional variances of the Shanghai stock market, but the same finding is not established for the Hong Kong stock market. This asymmetry in the conditional volatility spillover, which was also found in previous research (*e.g.* Bae and Karolyi (1994)), plays a major role in this paper and bears important implications. One of the important features of our analysis is the decomposition of the innovations of stock market into positive and negative components, which enables us to compare with previous research and assess the change of market volatility of both markets under varying market conditions.

The properties of the three series of daily return indices are discussed in the next two sections. The results in section 3.2 show that the squared stock return residual series have significant positive autocorrelations, their large magnitudes and their order of lags can be considered as indications that the stock returns are long memory processes. It is shown that as far as the relationship of the indices are concerned Hong Kong has indicated a leading role while Shanghai has been lagging. In section 3.3, the conditional volatility processes for the three series are analyzed using GJR model — an asymmetric ARCH type model. The results also show that there are significant volatility spillover effects from the Hong Kong to the Shanghai stock market before and after the handover of Hong Kong, but not other way around. This, however, does not imply that the innovations of Hong Kong de-stabilize the Shanghai stock market. On the contrary, our results suggest that the Shanghai stock market can benefit from more integration with the Hong Kong stock market.

3.2 Statistical properties of the Hong Kong and Shanghai Indices

Three sets of daily data are analyzed in this chapter. They are the daily closing prices of the Shanghai index, the Hang Seng index, and the China-Affiliated Corporate index (the *red-chip*) from January 1, 1993 to August 25, 1998. All the three indices are calculated by the market-value-weighted method. As Shanghai and Hong Kong are in the same time zone, the problem relating to non-synchronous trading hours does not occur in our analysis. The stock index returns are defined as,

$$R_{it} = (\ln P_{it} - \ln P_{it-1}) * 100 \tag{3.1}$$

where P_{it} is the daily closing index price for day t with i representing the three indexes individually. To facilitate our analysis, the capital gains are continuously compounded and dividend yields are not excluded from the returns as their values are so insignificant and their inclusion will not affect the final results.

3.2.1 Data description

A summary of the descriptive statistics for the three index return series R_{it} are presented in Table 3.1. These figures show that both the Hang Seng and Shanghai indices have positive and stable drifts around the mean of 0.025%. Among the three series, the Shanghai index has the highest sample unconditional variance which is almost twice the value of that of the Hang Seng index. While the Hang Seng index has the highest return and the lowest unconditional variance. The return series of the China-Affiliated Corporations (*red-chip)* index has negative drift of −0.03%, indicating the risk premium for investment in the *red-chip* shares is negative for this sample period. The sample variance of the *red-chip* index is also higher than that of the Hang Seng index return series, which seems consistent with the phenomenon that lower than average risk premium induces more speculative activity and therefore increases its volatility.

The Shanghai index return series is also seen having the largest outlier and consequently having larger skewness and excess kurtosis than those of the Hang Seng index. Both the Shanghai and the Hang Seng index return series skewed to the right, while the skewness of the *red-chip* is negative, −0.282, indicating that probability of negative returns is greater than that predicted by a symmetric distribution. The Shanghai index also has considerably higher excess kurtosis than other series, implying that big shocks of either sign are more likely to be observed in Shanghai market. The null hypothesis of normality for each stock index return series is rejected by the overwhelmingly significant test statistic and it is worth noting that it may be the excess kurtosis contributing the rejection more that the skewness.

Table 3.1 Summary statistics for stock index return series

Statistics	Mean	St.d	Skew.	Ex.Kurtosis	Min	Max	Normality Test
Hang Seng	0.025	1.836	0.1478*	12.417*	-14.73	17.25	1955.6*
Shanghai	0.025	3.014	1.3221*	14.867*	-17.90	28.86	1291.3*
Red-chips	-0.03	2.663	-0.282*	12.014*	-17.65	16.74	1840.9*

Notes: St.d = standard deviation; Skew = skewness; Ex.kurtosis = excess kurtosis. The skewness and kurtosis are normally distributed with means 0 and 3 and variances 6/T and 24/T respectively in large samples; The normality test statistic is a function of the skewness and excess kurtosis, examining whether the distributions of the two statistics are normally distributed; The null hypothesis of the normal distribution is be rejected at the 5 percent level, if a test statistic of more than 5.91 is observed; * denotes significance at the 5 percent.

3.2.2 The correlations estimators for the stock index return and residual series

Table 3.2 provides the estimates of the autocorrelations from lag 1 through lag 6 for the three raw return series. In addition, the estimated autocorrelations of the raw residuals and squared residuals are also summarized in this table. The estimated raw residual series, $\{\widehat{\epsilon}_{it}\}$, are generated by using the following auto-regressive regression:

$$R_{it} = \sum_{l=0}^{L} \phi_l R_{i,t-l} + \epsilon_{it} \qquad (3.2)$$

Where R_{it} are daily returns for the Hang Seng, *Red-chip*, and Shanghai indexes individually, and the appropriate minimum lag length L is determined by testing the estimated serial correlations for ϵ_{it} proposed by Richardson and Smith (1994).[2] The optimal lag length L is 3 for the Hang Seng and Red-chip indexes and 1 for the Shanghai index.

The estimated autocorrelations of the three indexes are separately analyzed over the overall period (January 15, 1993 to August 25, 1998); over the pre-handover period (January 15, 1993 to June 30, 1997); and over post-handover period (July 1, 1997 to August 25, 1998). For the whole three sample periods, the Ljung-Box (LB) statistics, testing the hypothesis of serial correlations for time series, indicate that all the three raw returns have significant estimated autocorrelations. The autocorrelations for each series of daily returns vary across sub-periods. The absolute value of the average autocorrelations for the post-handover sub-period exceeds that of the pre-handover sub-period for the three indexes. The estimated magnitude of autocorrelation for the *Red-chip* index is the highest among the three indexes and its first order autocorrelation are significantly from zero exceeding the value of 0.14 for the whole periods.

---

[2]Fuller (1996) demonstrates the sample correlation coefficients are asymptotically normally distributed as

$$\sqrt{T}\widehat{\rho}\,(j) \sim N\,(0,1)$$

After taking into account of heteroskedasticity, Richardson and Smith (1994) propose the serial correlation test statistic for testing the hypothesis of no autocorrelations in the stock returns, such that

$$\sqrt{T}\widehat{\rho}\,(j) \sim N\,(0,1+c_j)$$

where $c_j = cov\left(R_{it}^2, R_{it-j}^2\right) \diagup \left[var\left(R_{it}\right)\right]^2$. Then the serial correlation test statistic is asymptotically distributed as χ^2 with l degrees of freedom as

$$T\sum_{j=1}^{l}\widehat{\rho}^2 \diagup (1+c_j) \sim \chi_l^2$$

The test of the oppropriate minimum lag length L falls into the repeating procedures that first test $l = 0$, if the hypothesis of no autocorrleation in the estiamted residual series $\widehat{\epsilon}$ is rejected, then test $l = 1, 2, ...$ until the no hypothesis is accepted.

Table 3.2 Autocorrelation estimators for daily return and residual series

	order of autocorrelation							
Variables	$\widehat{\rho}_1$	$\widehat{\rho}_2$	$\widehat{\rho}_3$	$\widehat{\rho}_4$	$\widehat{\rho}_5$	$\widehat{\rho}_6$	$LB\,(12)$	P-value
				Overall period (01/01/93—28/08/98)				
r_{hk}	0.014	-0.040	0.128	-0.040	0.006	-0.039	47.525	[0.000]
r_{red}	0.141	-0.003	0.079	0.019	0.063	-0.079	81.217	[0.000]
r_{sh}	-0.028	0.028	0.055	0.031	0.007	0.010	31.795	[0.001]
$\widehat{\epsilon}_{hk}$	0.000	0.003	0.006	0.005	0.004	-0.005	9.026	[0.701]
$\widehat{\epsilon}_{red}$	0.000	0.002	0.007	0.010	0.006	-0.005	13.185	[0.356]
$\widehat{\epsilon}_{sh}$	0.002	-0.001	-0.005	0.004	-0.007	-0.001	11.153	[0.516]
$(\widehat{\epsilon}_{hk})^2$	0.398	0.218	0.274	0.246	0.166	0.114	587.98	[0.000]
$(\widehat{\epsilon}_{red})^2$	0.289	0.239	0.312	0.209	0.286	0.138	682.27	[0.000]
$(\widehat{\epsilon}_{sh})^2$	0.078	0.245	0.156	0.086	0.034	0.120	178.04	[0.000]
				Pre-handover period (01/01/93—30/06/97)				
r_{hk}	0.064	0.031	0.022	-0.008	-0.018	-0.054	13.326	[0.346]
r_{red}	0.143	0.047	0.011	-0.002	0.045	-0.055	43.176	[0.000]
r_{sh}	-0.027	0.034	0.056	0.029	0.005	0.012	27.457	[0.006]
$\widehat{\epsilon}_{hk}$	0.003	0.012	-0.052	0.024	-0.014	-0.010	18.89	[0.267]
$\widehat{\epsilon}_{red}$	0.001	0.003	-0.001	0.010	-0.005	-0.019	8.265	[0.752]
$\widehat{\epsilon}_{sh}$	-0.001	0.007	-0.003	0.002	-0.011	-0.001	11.042	[0.525]
$(\widehat{\epsilon}_{hk})^2$	0.194	0.184	0.110	0.082	0.198	0.101	232.52	[0.000]
$(\widehat{\epsilon}_{red})^2$	0.266	0.149	0.072	0.071	0.039	0.017	149.61	[0.000]
$(\widehat{\epsilon}_{sh})^2$	0.070	0.239	0.150	0.080	0.027	0.114	128.26	[0.000]
				Post-handover (01/07/97—28/25/98)				
r_{hk}	-0.037	-0.110	0.213	-0.083	0.019	-0.027	29.365	[0.003]
r_{red}	0.128	-0.048	0.098	0.015	0.066	-0.098	30.072	[0.003]
r_{sh}	-0.037	-0.082	0.049	0.054	0.017	-0.013	20.342	[0.061]

Continued...

Table 3.2 Autocorrelation estimators for daily return and residual series

Variables	$\widehat{\rho}_1$	$\widehat{\rho}_2$	$\widehat{\rho}_3$	$\widehat{\rho}_4$	$\widehat{\rho}_5$	$\widehat{\rho}_6$	$LB\,(12)$	P-value
	order of autocorrelation							
				Post-handover (01/07/97—28/25/98)				
$\widehat{\epsilon}_{hk}$	-0.045	-0.059	0.099	-0.019	0.011	-0.009	10.106	[0.607]
$\widehat{\epsilon}_{red}$	-0.019	-0.059	0.051	0.017	0.007	-0.029	10.601	[0.563]
$\widehat{\epsilon}_{sh}$	0.051	-0.013	-0.042	0.022	0.061	-0.018	14.903	[0.246]
$(\widehat{\epsilon}_{hk})^2$	0.389	0.173	0.251	0.221	0.108	0.059	94.16	[0.000]
$(\widehat{\epsilon}_{red})^2$	0.218	0.168	0.261	0.139	0.234	0.061	70.15	[0.000]
$(\widehat{\epsilon}_{sh})^2$	0.270	0.258	0.019	0.036	0.045	0.027	44.19	[0.001]

The identification of autocorrelation coefficients of return residuals and their transformation residuals generated from equation (3.2), $\{\widehat{\epsilon}_{it}\}$ and $\{\widehat{\epsilon}_{it}\}^2$ respectively, could explain why each return series depends on its own past values for long lag lengths. The Ljung-Box statistics of the three estimated raw residuals $\{\widehat{\epsilon}_{it}\}$, in contrast, are insignificant for both the overall period and the two sub-periods, indicating the autocorrelations of raw return series can be filtered by the autoregressive regression. However, all the three estimated squared residual series $\{\widehat{\epsilon}_{it}\}^2$ show high positive and significant autocorrelations at various lags for any reasonable significant levels. For all indexes at all periods, the absolute values of the estimated autocorrelations for the squared residuals are larger than those of residuals at most lags. Such substantial autocorrelations in the squared return residual series suggest the second moments of the three indexes are dependent and vary through time.

Table 3.3 Cross-correlations for each pair index return and residual series

	R_{ij}			$\widehat{\epsilon}_{ij}$			$\widehat{\epsilon}_{ij}^2$		
Cross-correlations	Overall	Pre-	Post-	Overall	Pre-	Post-	Overall	Pre-	Post-
corr$[hk_{t-2}, red_t]$	0.002	0.068	-0.057	0.002	0.086	-0.072	0.189	0.129	0.131
corr$[hk_{t-1}, red_t]$	0.112	0.088	0.119	0.014	-0.009	0.019	0.306	0.163	0.268
corr$[hk_t, red_t]$	0.742	0.649	0.813	0.736	0.643	0.808	0.710	0.493	0.715
corr$[hk_t, red_{t-1}]$	-0.010	0.036	-0.055	-0.017	0.023	-0.059	0.308	0.101	0.268
corr$[hk_t, red_{t-2}]$	-0.037	0.022	-0.089	-0.010	0.053	-0.071	0.240	0.081	0.197
corr$[hk_{t-2}, sh_t]$	-0.025	-0.027	-0.039	-0.025	-0.033	-0.019	-0.026	-0.025	0.035
corr$[hk_{t-1}, sh_t]$	-0.022	0.022	-0.223	-0.006	0.036	-0.175	-0.022	-0.030	0.141
corr$[hk_t, sh_t]$	0.025	0.016	0.087	0.018	0.010	0.063	-0.014	0.006	0.058
corr$[hk_t, sh_{t-1}]$	0.020	0.000	0.116	0.019	0.007	0.082	-0.017	-0.023	0.118
corr$[hk_t, sh_{t-2}]$	-0.025	-0.048	0.041	-0.033	-0.050	0.001	-0.017	-0.015	0.102
corr$[sh_{t-2}, red_t]$	-0.004	-0.000	-0.023	-0.025	-0.004	-0.035	-0.019	0.004	0.068
corr$[sh_{t-1}, red_t]$	0.014	0.064	-0.129	-0.006	0.063	-0.093	-0.024	-0.020	0.090
corr$[sh_t, red_t]$	0.084	0.094	0.157	0.018	0.054	0.147	-0.003	0.020	0.076
corr$[sh_t, red_{t-1}]$	0.032	0.006	0.152	0.019	0.011	0.151	-0.025	-0.036	0.140
corr$[sh_t, red_{t-2}]$	-0.014	-0.007	-0.052	-0.033	-0.034	-0.060	-0.020	0.008	0.049

Notes: Overall stands for the whole sample period (01/01/93—28/08/98) including the pre-handover period and post-handover period , Pre- stands for the pre-handover period (01/07/97—28/25/98), and the post-stands for the post-handover period (01/07/97—28/25/98).

As a preliminary examination, the nature of cross-relationship among the three indexes is shown in Table 3.3, applying the cross-correlation analysis on both the contemporaneous and lead-lag correlation estimations. These estimated correlation coefficients are also analyzed separately over the overall period, over the pre-handover period, and over the post-handover period. The analysis of cross-correlation properties may reveal the tendency of the cross market information aggregation or the direction of information flow to some extents even though the cross correlation analysis can be influenced by the stale price effects and transaction costs etc..

For the results of the estimated correlations between stock index returns (the second column to the fourth column), the Hang Seng index and the *Red-chip* index move closely over time and they have the highest average correlation values for the entire sample periods, ranging from 0.649 (pre-handover) to 0.813 (post-handover). There is an apparent lead at the return level for the Hang Seng index over the *Red-chip* index for 1-day (about 0.112 for the overall period, 0.088 and 0.119 for the pre-handover, post-handover period respectively). While, on average, the smallest correlation magnitudes are associated with the relationship between the daily returns in the Hang Seng index and the daily returns in the Shanghai index. It is interesting to note that both the contemporaneous correlation (being 0.016) and first order lead-lag correlation (being 0.000 and 0.022 respectively) between the Hang Seng and the Shanghai index returns are statistically insignificant, suggesting there is almost no cross correlation between them before the handover. The correlations for this pair indexes, however, increases significantly from 0.016 to 0.087 after the handover. The magnitudes of the lead-lag correlations increase significantly after the handover too.

The changed pre- and post-handover correlation patterns should not simply be attributed to Lo and MacKinlay (1990)'s hypothesis, in which large capitalization stocks lead small ones. Because the capitalization of the Hang Seng index is consistently larger than that of the Shanghai index for the whole our sample period. We can also find the inter-temporal stability of the correlations between the *Red-chip* index and the Shanghai index is weak over the two sub-periods. The three index return residuals (the fifth column to the seventh column) also display the same pattern of correlations as the index return levels, except that the values of correlations for the former residual series are much smaller than those of the latter return series.

In addition to the estimated correlations for the raw returns R_{ij} and its residuals $\widehat{\epsilon}_{ij}$, the estimated correlations for the squared return residuals $(\widehat{\epsilon}_{ij})^2$ are also presented in the last three columns in Table 3.3. The examination of the correlations for the squared residual series could reveal the relationship between each index in the second-moments level and provide a foundation for modelling volatility (Chan and Karolyi 1991, and Susmel and Engle 1994). The pairwise contemporaneous correlations in the squared residuals over the overall period are HK/RED, 0.710, HK/SH, −0.014, and SH/RED, −0.003. And the pairwise correlations for all series also vary over the two sub-periods that on average the post-handover coefficients of cross correlation are much higher than the pre-handover coefficients of cross correlation. It shows that the correlation increases from 0.006 to 0.058 between the Hang Seng index and the Shanghai index, from 0.493 to 0.715 between the Hang Seng and the *Red-chip* index, and from 0.020 to 0.076 between Shanghai and *Red-chip* indexes, respectively. These significantly big differences seem to suggest that the information connection between each pair indexes increases after the handover.

Particularly, it needs to be noted that the strong lead and lag correlation relationships exist among each pair indexes after the handover period. Furthermore, such cross correlations are non-symmetric for both $\widehat{\epsilon}$ and $\widehat{\epsilon}^2$ series after the handover. The first order lead-lagged cross-correlation between the Shanghai and *Red-chip* index residual series, $corr\left[\widehat{\epsilon}_{sh,t}, \widehat{\epsilon}_{red,t-1}\right]$, is 0.151, where as the first order lead-lagged cross correlation for the two series $corr\left[\widehat{\epsilon}_{sh,t-1}, \widehat{\epsilon}_{red,t}\right]$ is −0.09. The lead-lag cross correlation for squared residuals $corr\left[\widehat{\epsilon}^2_{sh,t}, \widehat{\epsilon}^2_{red,t-1}\right]$ being 0.140 is higher than $corr\left[\widehat{\epsilon}^2_{sh,t}, \widehat{\epsilon}^2_{red,t-1}\right]$ being 0.090. While the lead-lagged cross correlations between the Hang Seng index and the Shanghai index have the similar non-symmetric pattern.

Note that when comparing the cross correlation magnitudes between the Hang Seng and Shanghai squared return residuals for the pre-handover period with the post-handover period, the cross correlations of the squared return residuals for the two stock markets rise significantly. This seems to indicate that the variances for both markets are closely connected in some ways after the handover, the issue which will be analyzed in the next chapter by the co-persistence in variance analyses. The deep patterns of the Hong Kong stock market leading the Shanghai stock market can be investigated by examining the structure of conditional volatility transmission between the two markets. Because it is known that the information flow is more pervasive in terms of the second conditional moments (Ross (1989)). This is the subject of the next section 3.3.

3.3 Modelling the movement for the conditional volatility

There are very good agreements among stock investors, policy regulators and academic researchers of the need for objective measurements that would serve to measure both the linkages between international asset prices and the comovement between international markets. In an effort to achieve this goal, a number of investigations have been made by examining cross-market or assets time-varying variances and covariances based on the current available information set. The statistically significant autocorrelations and cross correlations in the squared residuals of the three indexes in section 3.2 suggest that it will be fruitful to characterize the movement in the conditional volatility by the $ARCH$ family models, which will be able to provide more insight into the relations among the three indexes. The $ARCH$ (*Autoregressive Conditional Heteroskedasticity*) models and their variants have provide a powerful framework to analyses the movement in the conditional volatility. Therefore we apply the bivariate GJR model proposed by Glosten, Jagannathan, and Runkle (1993) (henceforth denoted as GJR) to examine whether there exist any information flow between each pair indexes and how the properties of the information flow change after the handover.[3]

It is obvious that this approach, as well as other variants of the multivariate approaches, is much more sophisticated than the univarite models.[4] But precisely because of this complexity that it estimates all parameters simultaneously, the bivariate approach can exploit all information to capture the interdependence between variance and covariance for each pair of the indexes. The bivariate GJR approach could not only examine the magnitudes of the conditional volatility spillover effects across indexes but also capture the response of volatility spillover effects to different signs of shocks, the so called 'leverage effects'.

The specifications of the bivariate GJR process system for $\boldsymbol{\sigma}_t^2$ vector are expressed as:

[3]The main reason that we choose the asymmetric GJR model over the symmetric $GARCH$ model is based on Wooldridge's (1990, 1991) robust regression-based Lagrangian Multiplier (LM) specification tests. Following his procedure, we set up the null hypothesis H_0 that the conditional variance processes for the three indexes are generated by the usual $GARCH\,(1,1)$ and the alternative hypothesis H_1 is that of the $GJR\,(1,1)$ model, such that

$$H_0 : \sigma_{0t}^2 = GARCH$$
$$H_1 : \sigma_{1t}^2 = GJR$$

Next we test the validity of H_0 against that of H_1. In this case, the Lagrangian Multiplier test statistic for the hypothesis is

$$0.5\left[\sum_{t=1}^{T}\frac{1}{2\widehat{\sigma}_{0t}^2}\frac{\partial\sigma_{1t}^2}{\partial\theta}\left(\frac{\widehat{\epsilon}_t^2}{\widehat{\sigma}_{0t}^2}-1\right)\right]'\left[\sum_{t=1}^{T}\left(\frac{1}{\widehat{\sigma}_{0t}^2}\frac{\partial\sigma_{1t}^2}{\partial\theta}\right)\left(\frac{1}{\widehat{\sigma}_{0t}^2}\frac{\partial\sigma_{1t}^2}{\partial\theta}\right)'\right]^{-1}\left[\sum_{t=1}^{T}\frac{1}{2\widehat{\sigma}_{0t}^2}\frac{\partial\sigma_{1t}^2}{\partial\theta}\left(\frac{\widehat{\epsilon}_t^2}{\widehat{\sigma}_{0t}^2}-1\right)\right]$$

Where θ is the vector of conditional volatility parameters for the alternative GJR model. To construct the robust LM tests for whether the alternative is to be preferred, let restriction indicator $\widetilde{\lambda}_t = \frac{\partial\sigma_{1t}^2/\partial g}{\sqrt{\widehat{\sigma}_{0t}^2}}$ denotes the standardized first derivative of σ_{1t}^2 with respect to the asymmetric parameter(s) g and let $\nabla_\theta\widetilde{v}_t = \frac{\partial\sigma_{1t}^2/\partial\theta}{\sqrt{\widehat{\sigma}_{0t}^2}}$ denotes the standardized first derivative of σ_{1t}^2 with respect to the rest of the conditional volatility parameters for the alternative GJR model. Then we run the regression $\widetilde{\lambda}_t$ on $\nabla_\theta\widetilde{v}_t$ to generated the estimated residuals $\widetilde{r}_t$. Finally we run the regression 1 on $\left(\frac{\widehat{\epsilon}_t}{\widehat{\sigma}_{0t}^2}-1\right)*\widetilde{r}_t$, examining whether the specification statistics TR^2 is larger than the χ_1^2 significant values, in which R^2 is the coefficient of determination for the last regression.

[4]Baillie and Bollerslev (1990), Chan et al. (1991), and Karolyi (1995) apply bivariate GARCH conditional variance processes with constant conditional correlation to study international transmissions of stock returns and volatilities.

$$\mathbf{R}_t = \boldsymbol{\phi}_0 + \sum_{l=1}^{L} \boldsymbol{\phi}_l \mathbf{R}_{t-l} + \boldsymbol{\epsilon}_t$$

$$\boldsymbol{\epsilon}_t | \mathbf{I}_{t-1} \sim N\left(\mathbf{0}, \boldsymbol{\sigma}_t^2\right)$$

$$\begin{aligned} \begin{bmatrix} \sigma_{ii,t}^2 \\ \sigma_{jj,t}^2 \end{bmatrix} &= \boldsymbol{\varpi} + \boldsymbol{\beta} \begin{bmatrix} \sigma_{ii,t-1}^2 \\ \sigma_{jj,t-1}^2 \end{bmatrix} + \boldsymbol{\alpha} \begin{bmatrix} \epsilon_{ii,t-1}^2 \\ \epsilon_{jjh,t-1}^2 \end{bmatrix} \\ &+ \mathbf{G} \begin{bmatrix} \epsilon_{ii,t-1}^2 \\ \epsilon_{jj,t-1}^2 \end{bmatrix} S_{t-1}^- + \mathbf{g} \begin{bmatrix} D_t \epsilon_{ii,t-1}^2 \\ D_t \epsilon_{jj,t-1}^2 \end{bmatrix} S_{t-1}^- + \boldsymbol{\theta} D_t \end{aligned}$$

$$\sigma_{ij,t} = \left(\rho_{ij} + \vartheta_{ij} D_t\right) \sqrt{\sigma_{ii,t}^2 \sigma_{jj,t}^2} \tag{3.3}$$

Where the return residual vectors given by $\boldsymbol{\epsilon}_t = \left[\epsilon_{ii,t}, \epsilon_{jj,t}\right]'$ (for i, j represent the Hang Seng, Shanghai, and *Red-chip* index respectively and $i \neq j$) are assumed that their conditional joint distributions are normal using the available information set $\mathbf{I}$ as of time $t-1$. The Berndt, Hall, Hall, and Hausman (1974) numerical derivative algorithm is used to estimate the following parameters for the bivariate GJR model. $\boldsymbol{\varpi} = \begin{bmatrix} \omega_{ii} \\ \omega_{jj} \end{bmatrix}$, which is the constant vector; $\boldsymbol{\beta} = \begin{bmatrix} \beta_{ii} & \beta_{ij} \\ \beta_{ji} & \beta_{jj} \end{bmatrix}$, which is the coefficient matrix on the lagged variance; $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_{ii} & \alpha_{ij} \\ \alpha_{ji} & \alpha_{jj} \end{bmatrix}$, which is the lagged squared error term coefficient matrix; $\mathbf{G} = \begin{bmatrix} G_{ii} & G_{ij} \\ G_{ji} & G_{jj} \end{bmatrix}$, which is the coefficient matrix on the asymmetric effect, one of the stylized facts for the stock index returns; and $\mathbf{g} = \begin{bmatrix} 0 & g_{ij} \\ g_{ji} & 0 \end{bmatrix}$, which is designed to examine how the asymmetric effect of one index affects the conditional variance for the other index after the handover. Note if the $\mathbf{G}$ and $\mathbf{g}$ are null matrix, the GJR model reduce to the standard GARCH model. S_{t-1}^{-} is the sign dummy variable that is set to unity if the innovations, $\epsilon_{ii,t}$ or $\epsilon_{jj,t}$, are negative and zero otherwise, thus allowing the different impacts of different sign of innovations on conditional variances. The most single important factor in the recent event that contributed to the structural change for each index is the return of sovereignty of Hong Kong to China. Undoubtedly, this unusual event may have profound long term effects on the economy, the political system and social structure for both stock markets as a whole. To capture the short term dynamic effect of the handover on the volatility of each stock index, an extra dummy variable D_t is included. D_t is the handover dummy that is set to unity after the handover and zero otherwise. $\boldsymbol{\theta} = [\theta_{ii}, \theta_{jj}]'$,which is the coefficient vector on the handover dummy. We assume the conditional correlations ρ_{ij} are constant so that the covariances are the proportional to the product of the corresponding conditional variance. Finally, we follow Longin and Solnik (1995) extending the conditional correlations by adding another term $\vartheta_{ij} D_t$ to test the change of the correlations after the handover.

3.4 Results and evidence

The maximum log-likelihood joint estimation for each pairwise indexes are reported in Table 3.4. Panel A documents the results for conditional variance-covariance of each pairwise indexes. Values in parentheses are the robust t-statistics proposed by Bollerslev and Wooldridge (1988), and i, j denote target and source index series respectively. Strong ARCH effects have been found in all the three pairwise series, as the estimates for the main diagonal of the coefficients of the lagged conditional variance matrix, β_{ii} and β_{jj}, are all significantly positive for very individual index in each pair indexes, and the outcomes for the main diagonal of the coefficients of the lagged squared residual matrix, α_{ii} and α_{jj}, are all significantly different from zero. The asymmetric behavior of the conditional variance has been confirmed existing in all indexes. The positive and significant values of the main diagonal of the matrix, G_{ii} and G_{jj}, are consistent with the leverage hypothesis that the conditional volatility is higher when stock index price declines.[5]

The above ARCH effects analyses indicate that the conditional variance in each index can be explained on the basis of its own past uncorrelated random innovations. The properties of the cross-index volatility spillover process are determined by the off-diagonal coefficients of the matrices which are $\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{G}$, and $\mathbf{g}$. In the column of 'Hang Seng vs. *Red-chip'* in Panel A of Table 3.4, the volatility spillovers, measured by the significance of the coefficients β_{ij} vs. β_{ji} and α_{ij} vs. α_{ji}, exhibit asymmetric direction. The estimated values clearly suggest that the shocks originated from the Hang Seng index have spillover effects on the *Red-chip* index, as $\beta_{red,hk} = -0.1448\,(t = 5.82)$ and $\alpha_{red,hk} = 0.0692\,(t = 3.61)$. However, the reverse effects from *Red-chip* to Hang Seng index are not statistically significant at 5% level, as $\beta_{hk,red} = -0.008\,(t = 1.29)$ and $\alpha_{hk,red} = 0.0118\,(t = 1.19)$. The volatility spillover process is further investigated with the asymmetric effects caused by the negative shocks for the whole sample period and after the handover period. The coefficients $G_{red,hk}$ and $g_{red,hk}$ (*i.e.* $G_{ji} = 0.0633\,(t = 2.68)$ and $g_{ji} = 0.5453\,(t = 4.63)$) are statistically significant from zero, suggesting that negative innovations from the Hang Seng index influence volatility to the *Red-chip* index more than positive innovations for both the overall and post-handover periods. But the asymmetric response can not be found from the *Red-chip* index to the Hang Seng index, as the coefficients $G_{hk,red}$ and $g_{hk,red}$ (*ie.* $G_{ij} = 0.0062\,(t = 0.65)\,and\ g_{ij} = 0.1022\,(t = 0.66)$) are insignificant accepting the null hypothesis that the true population value of coefficients $G_{hk,red}$ and $g_{hk,red}$ are zero.

[5] For full description of the leverage hypothesis see Black (1976), Christie (1982), and Nelson (1991).

Table 3.4 Maximum Likelihood Estimations of Bivariate GJR Models

	Hang Seng vs.Red-chip	*Hang Seng vs.Shanghai*	*Shanghai vs.Red-chip*
	(i *vs.* **j)**	**(i** *vs.* **j)**	**(i** *vs.* **j)**
	Panel A: Estimates of the coefficients of the variance-covariance equation		
ω_{ii}	0.0651	0.0556	0.7024
	(5.86)	(3.55)	(8.95)
ω_{jj}	0.1134	1.1690	0.2744
	(6.37)	(10.2)	(7.84)
β_{ii}	0.8789	0.8745	0.7880
	(62.2)	(80.4)	(46.8)
β_{ij}	-0.008	0.0019	0.0735
	(-1.29)	(0.74)	(22.5)
β_{ji}	-0.1448	-0.0135	0.0061
	(-5.82)	(-3.17)	(1.46)
β_{jj}	0.9251	0.7422	0.6761
	(73.3)	(34.6)	(167.4)
α_{ii}	0.0622	0.0408	0.1160
	(5.90)	(4.29)	(5.58)
α_{ij}	0.0118	0.0010	-0.0419
	(1.19)	(0.91)	(-7.24)
α_{ji}	0.0692	0.0160	0.0019
	(3.61)	(2.17)	(0.83)
α_{jj}	0.0852	0.1012	0.1971
	(7.59)	(5.44)	(19.6)
G_{ii}	0.0333	0.1130	0.0612
	(3.16)	(6.64)	(3.16)

Continued...

	Hang Seng vs.Red-chip (i *vs.* j)	*Hang Seng vs.Shanghai* (i *vs.* j)	*Shanghai vs.Red-chip* (i *vs.* j)
G_{ij}	0.0062	-0.0029	0.0126
	(0.65)	(-1.30)	(2.31)
G_{ji}	0.0633	-0.0073	0.0045
	(2.68)	(-0.72)	(0.98)
G_{jj}	0.1666	0.0969	0.0228
	(5.96)	(4.94)	(2.83)
g_{ij}	0.1022	-0.3893	0.6044
	(0.66)	(-1.25)	(2.45)
g_{ji}	0.5453	0.5138	0.2008
	(4.63)	(3.84)	(6.56)
θ_{ii}	0.1035	0.2884	0.8798
	(1.21)	(1.66)	(5.59)
θ_{jj}	-1.1016	-1.1906	-0.5127
	(-4.31)	(-9.53)	(-13.9)
ρ_{ij}	0.6820	0.0353	0.0896
	(49.1)	(1.14)	(3.12)
ϑ_{ij}	0.1041	0.0874	0.0228
	(4.20)	(2.65)	(2.83)
LogL	-2511	-3363	-3964
Panel B: Residual diagnostic tests of the variance-covariance equation			
$Q\,(12)$	8.390	7.908	11.18
$Q^2\,(12)$	3.046	1.360	4.352
sign bias test	0.436	0.044	-0.804
negative size bias test	-1.644	0.353	0.975
positive size bias test	-1.127	-0.446	-0.355
joint test	1.348	0.142	0.768

Notes: The values in parentheses are statistic t-values. The regression for the volatility specification tests are as follows:

Sign bias test:	$z_t^2 = a + bS_{t-1}^- + \delta' \left(\frac{1}{\sigma_{0t}^2}\frac{\partial\sigma_{1t}^2}{\partial\theta}\right) + e_t$
Negative size bias test:	$z_t^2 = a + bS_{t-1}^-\varepsilon_{t-1} + \delta' \left(\frac{1}{\sigma_{0t}^2}\frac{\partial\sigma_{1t}^2}{\partial\theta}\right) + e_t$
Positive size bias test:	$z_t^2 = a + b(1 - S_{t-1}^-)\varepsilon_{t-1} + \delta' \left(\frac{1}{\sigma_{0t}^2}\frac{\partial\sigma_{1t}^2}{\partial\theta}\right) + e_t$
Joint test:	$z_t^2 = a + b_1 S_{t-1}^- + b_2 S_{t-1}^-\varepsilon_{t-1} + b_3(1 - S_{t-1}^-)\varepsilon_{t-1} + \delta' \left(\frac{1}{\sigma_{0t}^2}\frac{\partial\sigma_{1t}^2}{\partial\theta}\right) + e_t$

where z_t^2 is the squared standardized return residuals as $\left(\epsilon_t/\sqrt{\sigma_{0t}^2}\right)^2$, and S_{t-1}^- is a dummy variable that takes one if $\varepsilon_{i,t} < 0$ and zero otherwise. Individual tests are t-tests for coefficients b for the first three regressions. The joint test is an LM test which is asymptotically distributed as a chi-squared distribution with three degrees of freedom for the last regression.

In the column of 'Hang Seng vs. Shanghai', as usual, the second moment cross relationships between the two indexes can also be determined by examining the significance of the off-diagonal coefficients of the $\boldsymbol{\beta}, \boldsymbol{\alpha}$, $\mathbf{G}$ and $\mathbf{g}$ matrices. The results show there are no significant volatility spillover from the Shanghai index to the Hang Seng index, which is indicated by $\beta_{hk,sh}$ and $\alpha_{hk,sh}$ $\left(ie.\beta_{ij} = 0.0019\,(t = 0.74)\right)$ and $\left(\alpha_{ij} = 0.0010\,(t = 0.91)\right)$. However, the volatility spillover coefficients from the Hang Seng index to the Shanghai index, presented by the estimated value of $\beta_{sh,hk}$ and the estimated value of $\alpha_{sh,hk}$ $\left(ie.\beta_{ji} = -0.0135\,(t = 3.17) \text{ and } \alpha_{ji} = 0.016\,(t = 2.17)\right)$, are statistically significant from zeros. So the conditional variance of the Shanghai index not only depends on its own past innovations but also depends on past shocks induced by the Hang Seng index. It should be noted that even though neither the off-diagonal coefficients of the $\mathbf{G}$ matrix is significant, (represented by G_{ij} and G_{ji} respectively), the after handover asymmetric effect coefficient from the Hang Seng index to the Shanghai index $g_{sh,hk}$ $(ie.g_{ji} = 0.5138\,(t = 3.84))$ is also significant from zero.

This seems to suggest the information flow from the Hong Kong stock market to the Shanghai stock market enhanced after the handover and the negative shocks from Hong Kong may have larger impacts on the Shanghai market than the corresponding positive shocks. The results confirm that the bad news in Hong Kong do have stronger impact on the Shanghai market that the good news. The asymmetric mechanism of volatility transmission has profound implications to the relationship between the two markets, and in particular to the further open up of the Shanghai Stock market. In the last column of 'Shanghai vs. *Red-chip*', the source of information flow is still the *Red-chip* index being traded in the Hong Kong stock market, but the Shanghai index can also provide information to the *Red-chip* index after the handover (coefficient $g_{ji} = 0.2008\,(t = 6.56)$). Particularly, the negative shocks originated in the Shanghai index can affect the volatility of the *Red-chip* index.

The evidence of the handover effect on the conditional variance process for each index is obtained by examining the handover dummy coefficient vector $\boldsymbol{\theta}$. In the columns of 'Hang Seng vs. *Red-chip*' and 'Hang Seng vs. Shanghai', the coefficients θ_{ii} for the Hang Seng index are insignificant being from zero ($ie.\theta_{hk} = 0.1035\,(t = 1.21)$ and $\theta_{hk} = 0.2884\,(t = 1.66)$) for the two columns respectively. These results indicate that the handover has no effect on conditional volatility of the Hang Seng index. However, the coefficients θ_{jj} ($ie.\theta_{red} = -1.1016\,(t = 4.31)$ and $\theta_{sh} = -1.1906\,(t = 9.53)$) are significant and negative, indicating the handover has significant influences on the conditional variance of both the *Red-chip* index and the Shanghai index. Further, the magnitudes of their conditional variances decline after the handover, which is consistent with the handover coefficient vector $\boldsymbol{\theta}$ in the third column of the pair 'Shanghai vs. *Red-chip*'.

In the correlation between the Hang Seng index and the *Red-chip* index, the estimated correlation coefficient, $\rho_{hk,red}$ ($ie.\rho_{ij} = 0.6820\,(t = 49.1)$), are significant for the whole period. And the correlation of the two indexes increases after the handover, as the estimated correlation coefficient, $\vartheta_{hk,red}$ ($ie.\vartheta_{ij} = 0.1041\,(t = 4.20)$), are positive and significant. The correlation relationship between the Shanghai index and *Red-chip* index almost show the same pattern as the correlation relationship between the Hang Seng index and the *Red-chip* index. It is interesting to note that although the correlation coefficient between Hang Seng index and Shanghai index is not significant (coefficient

$\rho_{ij} = 0.0353\,(t = 1.14)$), the tendency of linkage seems to occur between the two indexes after the handover (coefficient $\vartheta_{ij} = 0.0874\,(t = 2.65)$), further confirming that the Shanghai stock market is indeed subject to the influences coming from the Hong Kong stock market after the handover.

The standardized residual diagnostic statistics of the bivariate GJR model are provided in Panel B. The values of LB (12) and $LB^2(12)$ statistics for the standardized residuals and squared standardized residuals indicate that most of the dependences have been captured in the model specification. In particular, the values of $LB^2(12)$ statistic have been reduced significantly when compared with those in Table 3.2, suggesting that most of the dependency has been captured in the proposed model specifications and no remaining nonlinear structure exists in the normalized residuals.

Four other diagnostic tests for asymmetric effects are also reported. We first consider the effect of the sign of the past shocks, the test for which was first introduced by Engle and Ng (1993) who proposed the use of dummy variable S_{t-1}^- which takes value one if ϵ_{t-1} is negative and zero otherwise. The effects of the size of the negative and positive shocks are also measured by the coefficients of variables $S_{t-1}^-\epsilon_{t-1}$ and $\left(1 - S_{t-1}^-\right)\epsilon_{t-1}$, and finally, the joint test is carried out for the joint hypothesis of the above three effects. The test results appear to be insignificant for the three pairs models, hence we could conclude that the asymmetric effects have been adequately accommodated by the specification of the conditional volatility in equation 3.3.

The effects on volatility spillover effects from one index to the other are also plotted from figure 3-1 to figure 3-6 because of their greater ease of interpretation. These figures are called *The News Impact Surface.*[6] In the three-dimensional graphs to follow, for a given X, Y, Z axis combination, sometimes different rotational views are used to better illustrate the conditional variance features for each index. Each figure displays the impact of the conditional volatility of one index when the shocks are varied by the introduction of

[6]Bae and Karolyi (1994) extend Engle and Ng(1993)'s News Impact Curve into the three-dimensional figure— the News Impact Surface.

shocks from another index. The shocks designate the squared residuals from each index. The scale of the residuals runs from negative minimum values to positive maximum values but reported as (1,1500), centered at 751. (Here we follow Bae and Karoly (1994)'s notation).

A three-dimensional graph is first shown in figure $(3-1)$. The conditional volatility of the Hang Seng index is a matrix for which the elements define the height of a surface over the underlying grid constructed by innovations from the Hang Seng index and the Shanghai index. This perspective figure shows how the shocks from the Hang Seng and Shanghai index can influence the conditional variances of the Hang Seng index. When Hang Seng shocks are far from centre, increasing both signs of Hang Seng shocks has the effect of increasing conditional variances of the Hang Seng index. Further, it can be easily observed that negative shocks on the variances are larger than positive shocks, showing the asymmetric effects. However, varying the magnitudes of the Shanghai shocks has little effect on Hang Seng variances. These types of behavior prove the Hang Seng index has larger influence on Shanghai variances than does the Shanghai index.

Figure $(3-2)$ is a graph which can be compared with figure $(3-1)$. In figure $(3-2)$, the Shanghai index variance is a matrix for which the elements define the height of a surface over the underlying grid constructed by innovations from the Hang Seng index and the Shanghai index. The general shapes of the corresponding curves in each figure are broadly similar. The conditional variance of the Shanghai index responds asymmetrically to the different signs of its own shocks. However, in the curve in figure $(3-2)$, the shocks of the Hang Seng index have impacts on the conditional volatility of the Shanghai index as the shape becomes curved subject to Hang Seng shocks. (While the shape is flat corresponding to Shanghai shocks in figure $(3-1)$).

Figure $(3-3)$and figure $(3-4)$ illustrate how shocks from the Hang Seng index and the *Red-chip* index influence the variances of each index individually. The behaviors of the two graphs are very similar to those displayed in figures $(3-1)$ and $(3-2)$ respectively and is not surprising since the Hang Seng index dominates the information flow. The last pair three-dimensional graph are shown in figure $(3-5)$ and figure $(3-6)$, in which we have studied how the Shanghai and *Red-chip* shocks have impacts on each index individually. Figure $(3-5)$ shows the international news impact curve is bowl-shaped, indicating that the conditional variances of the Shanghai index depend not only on its own shocks but also the *Red-chip* shocks. Figure $(3-6)$ shows the dependence of the *Red-chip* variances on its own shocks and Shanghai shocks. Two peaks can be observed for the *Red-chip* index variances around the left corners.

In summary, the Hong Kong stock market has played a dominant role in the volatility transmission over the sample period straddling the handover of the handover of Hong Kong. As a mature and established international financial centre, the extra information that Hong Kong brings to Shanghai in the process of volatility spillover benefits the latter. In fact, this provides a plausible explanation to the significant reduction of volatility in the Shanghai stock market after the handover. As a long-term result, the Shanghai market may gradually improve the liquidity and speed up the stock price revealing process, which would eventually lead to more transparency of the market.

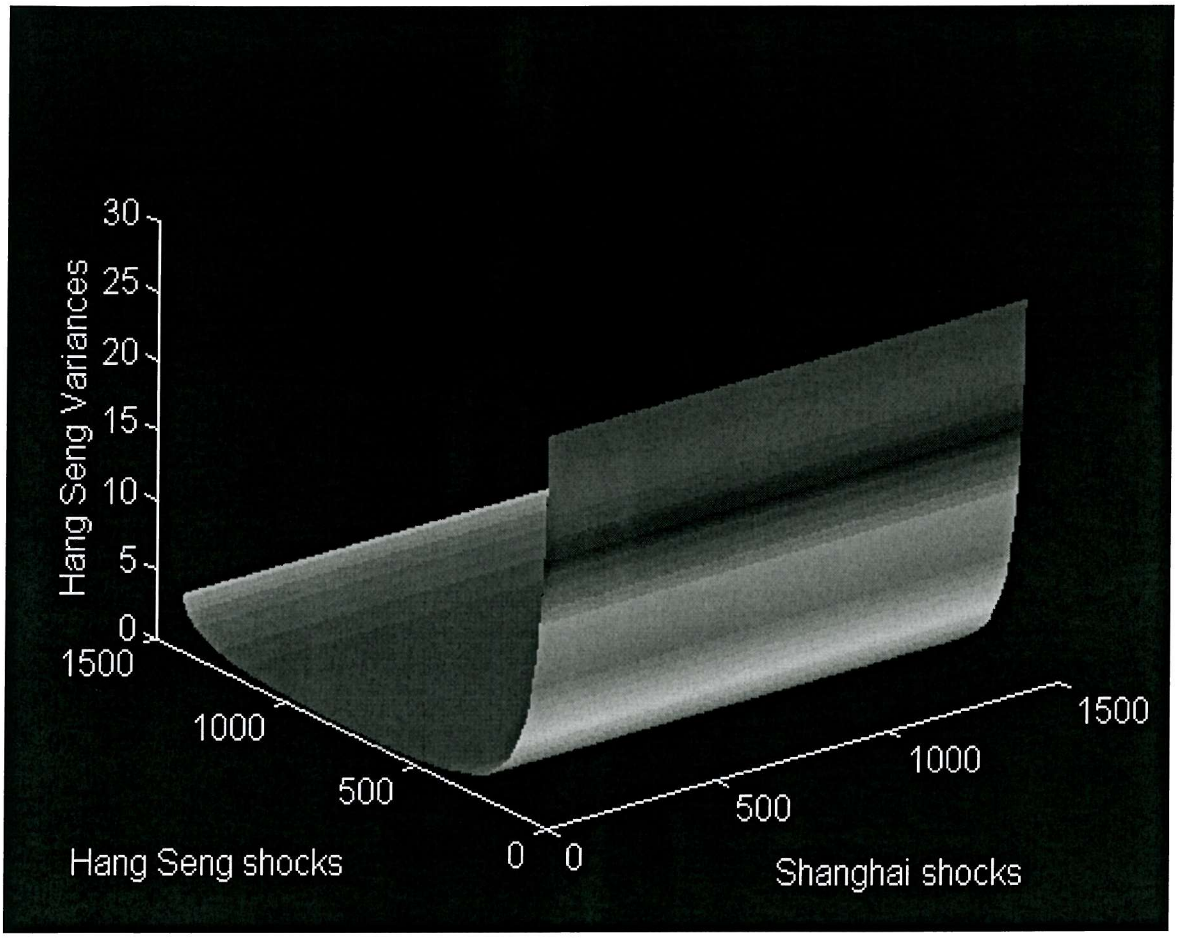

Figure 3-1: The three-dimensional plot showing the Hang Seng index variance as the function of its own shocks versus the Shanghai index shocks. The behavior is very similar to that seen in figure 3-3.

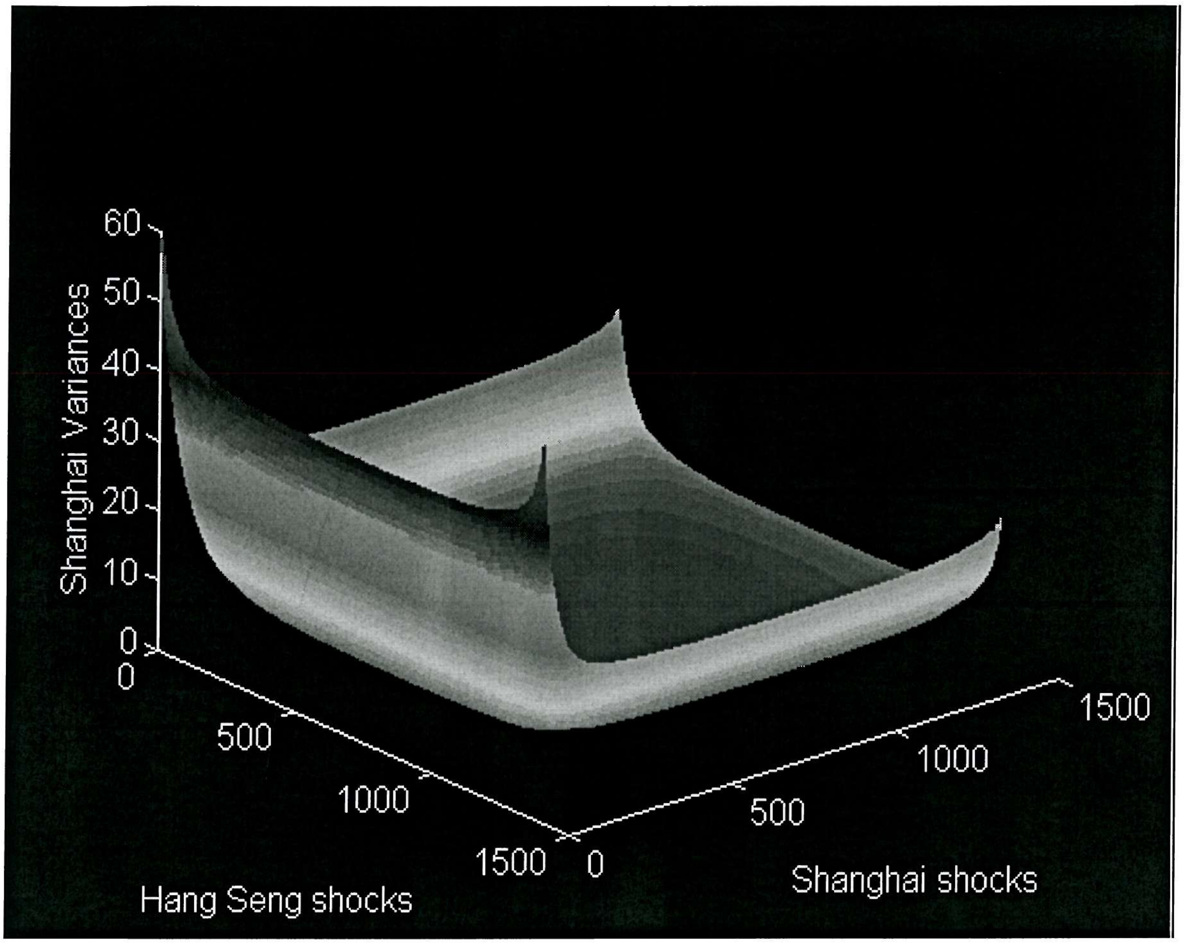

Figure 3-2: The three-dimensional plot showing the Shanghai index variance as the function of its own shocks versus the Hang Seng index shocks. The behavior is very similar to that seen in figure 3-4.

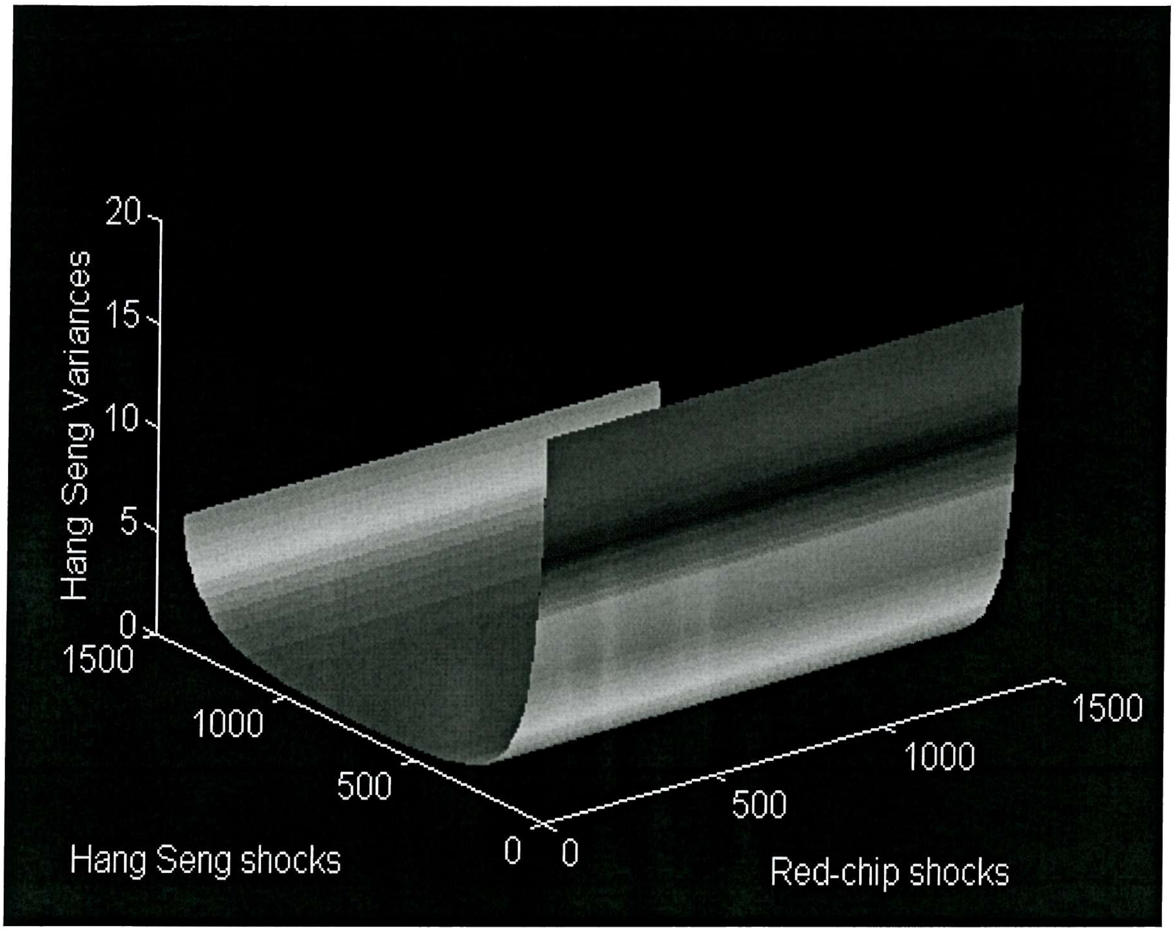

Figure 3-3: The three-dimensional plot showing the Hang Seng variance as the function of its own shocks and the Red-chip index shocks. In this graphic we note that the Hang Seng index variance responds asymmetrically only to its own shocks.

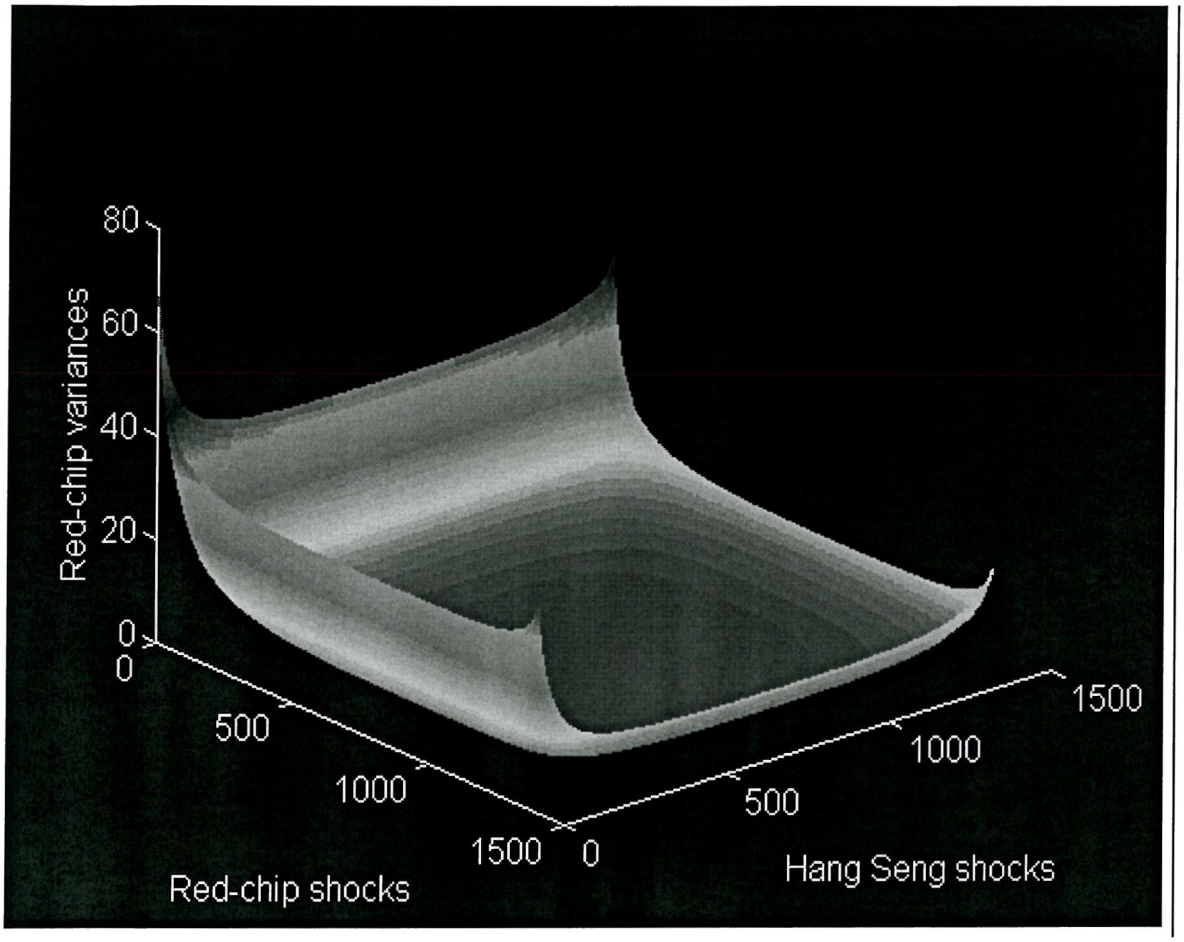

Figure 3-4: The three-dimensional plot showing the Red-chip index variance as the function of its own shocks and the shocks corresponding to the Hang Seng index. In this graphic we see the influences of the Hang Seng index over the Red-chip index.

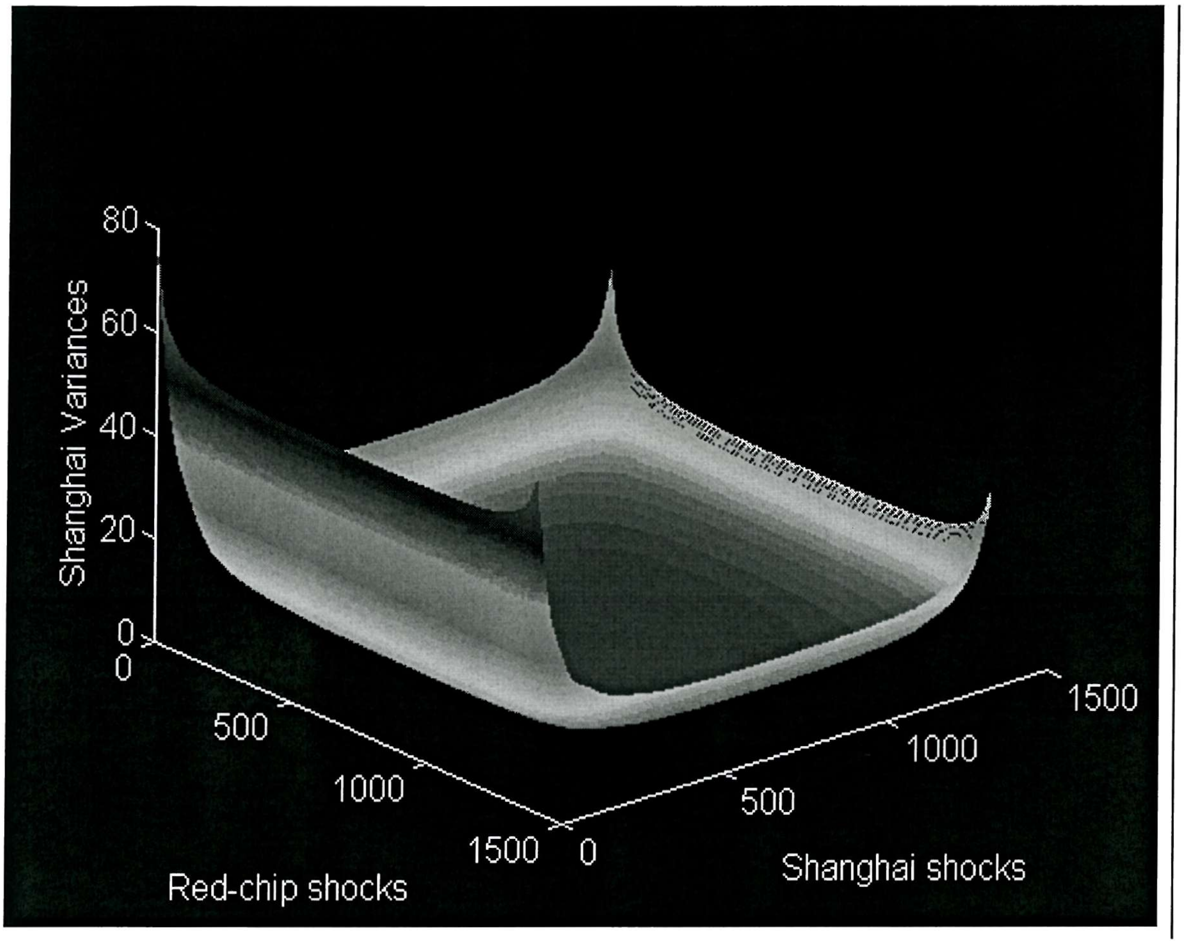

Figure 3-5: The three-dimensional plot showing the Shanghai index variance as the function of its own shocks and the Red-chip index shocks. In this graph, the variance of the Shanghai index depends on both index shocks.

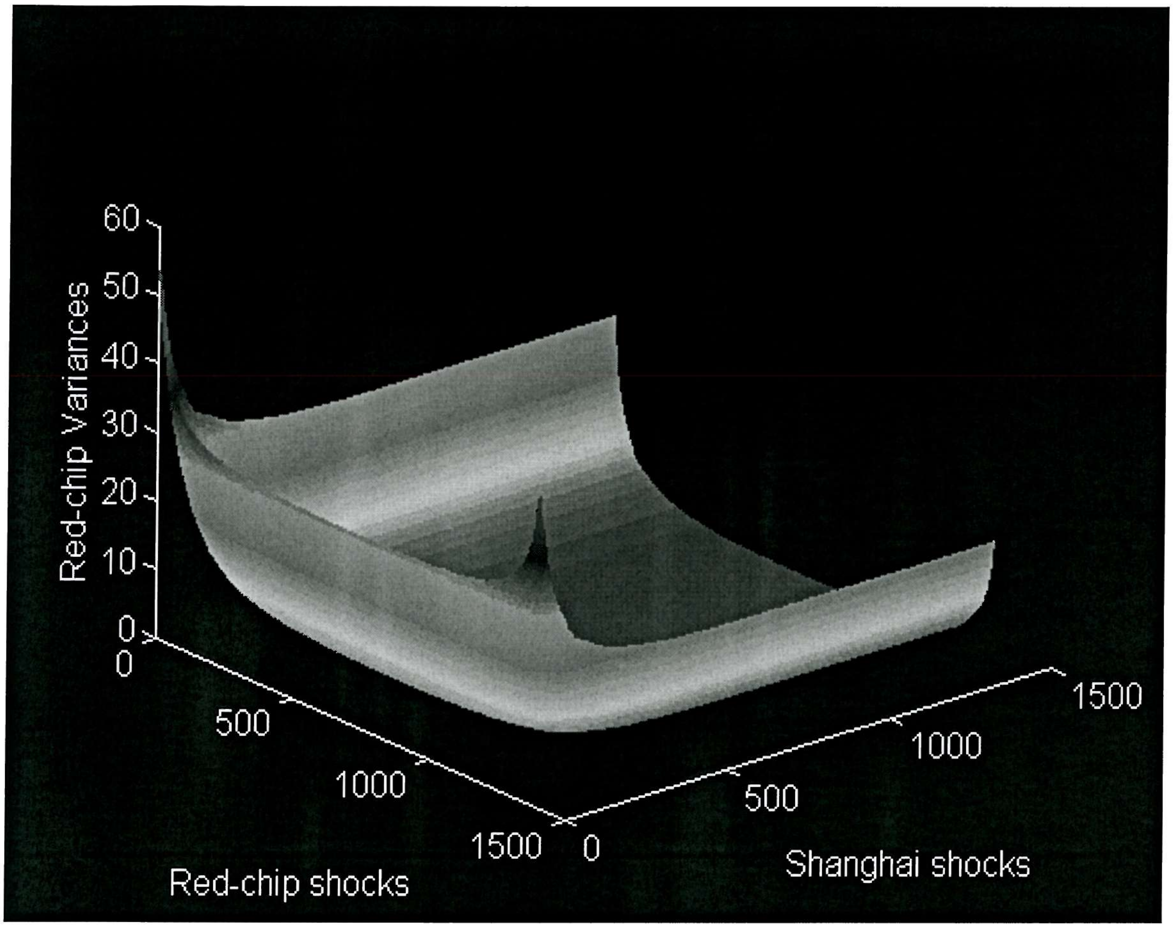

Figure 3-6: The three-dimensional plot showing the Red-chip index variance as the function of its own shocks and the Shanghai index shocks. In this graph, the negative shocks from the Shanghai index can affect the variance of the Red-chip index after the handover.

3.5 Conclusion

The empirical estimation results and the graphical presentation of news impact curves included above are intended to investigate the volatility transmission mechanism between the Hong Kong and Shanghai stock markets. The study is based on the analysis of the stock return series of the Shanghai, Hang Seng and *red-chip* indices using three bivariate asymmetric GJR models. The examination of these results reveals several implications for the interpretation and understanding the effects of inter-market information linkages.

Firstly, the empirical evidence suggests that the dynamics of returns and volatility spillover effects are asymmetric: the volatility transmission from the Hong Kong stock market to the Shanghai stock market is significant, and the post-handover Shanghai stock market becomes less volatile than the pre-handover. While the Hong Kong stock market is not affected by the shocks originated in the Shanghai stock market both before and even after the handover.

The findings can be attributed to the existence of market structure differences between the two markets. The Hong Kong stock market is approaching more developed, while the Shanghai stock market is still in the early stage of development thus lacks of standards in many areas. For instance, in the Hong Kong stock market the information disclose requirements, auditing standards, accounting standards, and credit rating criteria of listing stocks are much more strident than those in the Shanghai stock market. In this case, the Hong Kong stock market is the information driver with respect to the Shanghai stock market. As the fundamental information for the shares flowing from Hong Kong to Shanghai market proceeded, it improves the equity price efficiency and reduces the sentiments of the noise traders (i.e. most Shanghai stock investors) oscillating between optimum and pessimism, which affect stock returns and cause stock market over volatile (Lee et al. (1991)). Moreover, the more efficient equity prices and more extra information may restrict over-speculation, inside trading, and even manipulation in the Shanghai stock market,[7] thus, avoiding the distortion from the stock price to its underlying fundamentals.

Secondly, the Hong Kong and Shanghai stock markets are informatively connected through the China-affiliated Corporation index (*Red-chips).* The above empirical results show that the effects of red-chip equities trading on the Hong Kong stock market can also improve the stock price discovery processes for the Shanghai stock market. It needs to be noted that the significant volatility spillover from the Hang Seng index to the Shanghai stock market suggests the Shanghai index is linked with the red-chip index and this linkage between the Hang Seng and Shanghai indexes significantly increases after the handover period. This finding is consistent with the theoretical analysis by Domowitz et al. (1997) that the cross-listing could make more information incorporated into equity prices. As a result, the improving information linkage could increase the liquidity and price efficiency for the Shanghai stock market. The more liquid stock market would reduce the liquidity compensation (Stulz (1997)), following by infrequent trading induced volatility, bid-ask error induced volatility (Jorion (1995)), and noise induced volatility, for which the Chinese regulators is anxious to avoid. The conditional variances play a key role in the information flow process rather than the " large stocks leading small stocks" phenomenon.[8]

Thirdly, our results show that there is a variance causal link in the sense of Granger between the Hong Kong and Shanghai markets, in which the *red-chip* index plays a vital role. This finding is further strengthened by the discoveries of the asymmetric volatility spillover mechanism from Hong Kong to Shanghai. The shocks in the Hang Seng and *red-chip* indices are shown having significant contemporaneous impacts on the volatility of the Shanghai index, but the reverse effects are not established. While there in no *red-chip* stocks listed on the Shanghai stock exchange. Contrary to the common belief, the

---

[7]Kupiec and Sharpe (1991) propose an equilibrium model and find that irrational speculators may create excess stock price volatility.

[8]Madavan (1996) argues the information flow can always reduce volatility and improves market quality in large markets. However, the information flow may have adverse effects in thin markets.

handover of Hong Kong to China does not affect the continuity of the Hong Kong market, and the Shanghai stock market may benefit from it.[9] These findings imply that further opening up the Shanghai financial market and forming closer link with Hong Kong may benefit the Shanghai stock market greatly. As the extra information brought to Shanghai through the volatility spillover from Hong Kong reduces its uncertainties and conditional volatilities. Finally, further rigorous tests are applied to examine the validity or to add more clarity to this study's findings in the next chapter.

[9]The effects of the currency exchange between Chinese yuan and Hong Kong dollar are also taken into account in the volatility transmission patterns. We find that the currency effects can be ingored. This is probably because of both currencies being strictly pegged with US dollars.

Chapter 4

Modelling New York's influences on the Hong Kong & Shanghai stock markets: The common persistence and error correction analysis in variances

4.1 Introduction

A closely related issue to the international stock markets information flow effect concerns the role of information within a broader system. Although the research in the last chapter has been conducted on the existence of volatility volatility spillover effects directing from the Hong Kong stock market to the Shanghai stock market, the argument still remains that the majority of studies on the direction of volatility spillover have found that the New York stock market is the stem of the 'meteor shower' (Engle et al. 1990). Thus, in spite of last chapter's results, the suspicion of the existence of the volatility spillover between the Hong Kong stock market and the Shanghai stock market arises because the volatility spillover phenomenon may just reflect the increasing centralized dominance and control of New York's common source of influences. The next step in developing the understanding of relationships among stock markets is to extend the analysis to consider the aspects of information interactions among stock markets in a broad system. This chapter is mainly concentrated on this aspect of suspicion, that of examining common persistence in conditional variances between the New York, Hong Kong and Shanghai stock markets.

In a study of cointegration and long run relationships, Granger (1991) raises the issue of cointegration among a set of unit root process in variances and he defines it as "cointegrated in variances". However, it is Bollerslev and Engle (1993) who uncover necessary and sufficient conditions for the existence of linear combinations of persistent variances displaying no persistence in variances. They define it as co-persistence in variance of a multivariate $GARCH\,(p,q)$ process. As discussed more extensively in Bollerslev and Engle (1993), co-persistence in variance occurs provided all the $I\,(1)$ variance processes have the same common long-run components. A similar concept to that adopted by Hansen (1995) may help to understand the concept of common persistence in variance. As Hansen explains, the generalized least squares (GLS) estimator of martingale difference sequences (MDSs) with non-stationary volatility processes are not asymptotically normal or mixture normal distributed. But when there is a "cointegration relation in the variance processes" for the MDSs, the GLS estimator has a standard normal distribution or variance mixture of normal distribution.

The use of cointegration as a fundamental concept in variance experiments is a reasonable extension (and interesting parallels) of some lines of thought. We review this theory to show the viability and how compelling the use of co-persistence in variances can be, because it may suggest that the volatility movement between persistence variance processes reflects the presence of common components. Then, we examine the long run relationships between the three indexes in two procedures. First, the integration variances for the three indexes are analyzed in which the unit roots have been found to be an important determinant of volatility behavior in information memory. Second, the system of multivariate generalized ARCH models with the concept of co-persistent in variances are applied to verify the unique volatility spillover from the Hong Kong stock market to the Shanghai stock market and this unique volatility spillover is not affected by the New York stock market.

4.2 Process in the modeling of co-persistence in conditional variances

Bollerslev and Engle (1993) draw a theoretical extension from the persistence conditional mean to the persistence conditional variance. Persistence in conditional mean refers to whether information is stored in memory of the underlying stochastic process permanently and persistence in conditional variance refers to the case with which '*current information to stochastic variance process remains for the forecasts of the conditional variances for all horizons*' (Engle and Bollerslev 1986, pp.27). Then a logically equivalent way of extending the theorem of cointegration concerning persistence in mean is to say that a suitable linear combination of several persistent variances could result in co-persistent in variance. It follows directly from the impact of co-persistence in variance on stock prices that stock prices could become stable even if the level of individual stock prices are very sensitive to the extent of persistence in individual variance[1].

The fact is that much of the contemporary literature on cointegration focus on cointegration in mean, so there is widespread statistical consensus that there might exist cointegrating relationships among $I(1)$ variables; however, there is a lack of consensus on persistent variance properties (Most research on persistence in variance focus on $GARCH$ $(1,1)$ models). Perhaps the reason for the lack of consensus on this issue is the difficulty associated with the features for persistence in variance, in which a shock to

[1] Poterba and Sumers (1986) give an numerical example about the effect of persistence in volatiltiy on the level of share prices.

the conditional variances can have persistent effects such that the conditional variance process looks like a random walk, but the unconditional variance is strictly stationary and ergodic (Nelson, 1990).

In order to characterize the persistence in variance and co-persistence in variance properties, suppose that the conditional variance of ε_t based on I_{t-1}, the set of information available at time $t-1$, is described by a $GARCH\,(1,1)$ process,

$$\varepsilon_t | I_{t-1} \sim N\,(0, h_t)$$

$$h_t = c + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$$

By recursive substitution and the law of iterated expectations

$$E_t\left(\varepsilon_{t+s}^2\right) = E\left(E\left(\varepsilon_{t+s}^2 | I_t\right)\right) = E\left(h_{t+s}\right) = c + (\alpha+\beta)^s E\left(\varepsilon_t^2\right); \quad \text{for } t > s > 0$$

and if $\alpha + \beta < 1$, the unconditional variance converge to the constant

$$_u\sigma_{t+s}^2 = E_t\left(\varepsilon_{t+s}^2\right) = \frac{c}{1-(\alpha+\beta)} \tag{4.1}$$

Thus, the sum of $(\alpha+\beta)$ in the $GARCH\,(1,1)$ specification is the sensitive indicator of how information is stored in the conditional variance memory process. If, on the other hand, as $\alpha + \beta \longrightarrow 1$ that is the roots of the characteristic polynomial $\det\left[1 - \alpha\left(\lambda^{-1}\right) - \beta\left(\lambda^{-1}\right)\right] = 0$ lie on the unit circle, the speed of information decay measured by the sum of $(\alpha+\beta)$ will not diminish such that the multi-step advances of the conditional variance can not reach the unconditional variance. However, the IGARCH volatility process is not affected by the very past initial observations. Engle and Bollerslev (1986) then defined this property of the conditional variance process as '*Integrated-GARCH*'.

Bollerslev and Engle (1993) present a formal definition of common persistence in variance based on evidence from the integrated *GARCH* model. Linear combinations of integrated processes such as $I(1)$ are usually integrated processes. However, if the $(N \times N)$ multivariate integrated variance process $\{\mathbf{H}_t\}$ share some common stochastic components, and if there exists a non-trivial $(N \times 1)$ vector γ (which is a necessary condition for co-persistence in variance), some particular suitable linear combinations of $\{\mathbf{H}_t\}$ may become stationary. Then we say that the vector γ is called the co-persistent vector provided that

$$\limsup_{t \to \infty} \left| \gamma' \mathbf{H}_t^* (s) \gamma \right| = 0$$

holds almost surely for all $t > s > 0$ and $\mathbf{H}_t^*(s)$ is given by such that $\mathbf{H}_t^*(s) \equiv E_s(vec(\mathbf{H}_t)) - E_0(vec(\mathbf{H}_t))$, where $E_t(\bullet)$ denotes that we take expectation on $(\bullet)$ at time t. This follows from the fact that even if the limit of the least upper largest amount by which $E_s(vec(\mathbf{H}_t))$ differs from $E_0(vec(\mathbf{H}_t))$ for any s is not zero uniformly, i.e. $\limsup_{t \to \infty} |\mathbf{H}_t^*(s)| \neq 0$ almost surely for some $s > 0$, but the form $\limsup_{t \to \infty} \left|\gamma' \mathbf{H}_t^*(s) \gamma\right|$ will tend to zero uniformly, suggesting that equilibrium relationships among the conditional variances can be achieved across individual persistent variance systems by pre- and post-multiplying the weight vector γ. It is apparent that the expression for $\gamma' \mathbf{H}_t^*(s) \gamma$ has the same structure as the expression for the cointegration in mean.

Suppose that the multivariate *GARCH* conditional variance-covariance process $\mathbf{H}_t$ for the $N \times 1$ vector ε_t process is given by Engle and Kroner (1995)'s *BEKK* parameterization with the summation limits k being *one*.[2]

[2] As Engle and Kroner (1995) point out that the assumption of the summation limits k being one is equivalent to imposing some unnecessary restrictions. So there need at least 9 distinct elements in $\mathbf{A}_{1k}$ matrices for a general enough model in our case (There are also the same requirements for $\mathbf{B}_{1k}$ matrices). However, there are difficulties in achieving convergence for the parameter estimations after allow $k = 4$ (i.e. $k = n^2$) with the sufficient identifying restrictions proposed by the proposition 2.3 in their paper.

$$\varepsilon_t | I_{t-1} \sim N(\mathbf{0}, \mathbf{H}_t)$$

$$\mathbf{H}_t = \mathbf{C}_0^{'}\mathbf{C}_0 + \mathbf{A}_1^{'}\varepsilon_{t-1}\varepsilon_{t-1}^{'}\mathbf{A}_1 + \mathbf{B}_1^{'}\mathbf{H}_{t-1}\mathbf{B}_1 \tag{4.2}$$

where $\mathbf{C}_0$, $\mathbf{A}_1$, and $\mathbf{B}_1$ are $(N \times N)$ square parameter matrices with $\mathbf{C}_0$ being triangular. The $BEKK$ parameterization ensures the conditional covariance matrix $\mathbf{H}_t$ to be positive definite because $\mathbf{H}_t$ is expressed in quadratic forms. In order to analyses persistence properties in the variance-covariance $\mathbf{H}_t$ more easily, we can vectorize both sides of equation (4.2) to obtain

$$\begin{aligned} vec(\mathbf{H}_t) = (\mathbf{C}_0 \otimes \mathbf{C}_0)' \, vec(\mathbf{I}_N) + (\mathbf{A}_1 \otimes \mathbf{A}_1)' \, vec\left(\varepsilon_{t-1}\varepsilon_{t-1}^{'}\right) \\ + (\mathbf{B}_1 \otimes \mathbf{B}_1)' \, vec(\mathbf{H}_{t-1}) \end{aligned} \tag{4.3}$$

Where the notation $\otimes$ is the Kronecker tensor product creating large regular block matrices.[3]

Similar to the univariate integrated GARCH $(1,1)$ model, the unit root process for $\mathbf{H}_t$ requires the roots of the characteristic matrix polynomial

$$\det\left[\mathbf{I} - (\mathbf{A}_1 \otimes \mathbf{A}_1)' \mathbf{z} - (\mathbf{B}_1 \otimes \mathbf{B}_1)' \mathbf{z}\right] = 0 \tag{4.4}$$

to lie on the unit circle, i.e. $z_i = \lambda_i^{-1} = 1$. The values of λ_i can be calculated from the Jordan form of the square matrix $\left[(\mathbf{A}_1 \otimes \mathbf{A}_1)' + (\mathbf{B}_1 \otimes \mathbf{B}_1)'\right]$, that is, there exists a non-singular matrix $\mathbf{S}$ such that $\mathbf{S}^{-1}\left[(\mathbf{A}_1 \otimes \mathbf{A}_1)' + (\mathbf{B}_1 \otimes \mathbf{B}_1)'\right]\mathbf{S} \ = \mathbf{J}$, where $\mathbf{J}$ is a block diagonal matrix with eigenvalues λ_i being its diagonal elements. Thus, the process $\mathbf{H}_t$ is non-stationary in variance if some $|\lambda_i| \geqslant 1$, indicating shocks to the process $\mathbf{H}_t$ remain persistent over time.

[3]For more details of the properties of the Kronecker tensor product and the *vec* operator see Magnus and Neudecker (1986).

To determine the sufficient condition for the identification of the co-persistence in variance, we first need to determine the eigenvectors $\mathbf{x}$ of the polynomial (4.4). Given that we have some $|z_i| \leqslant 1$ corresponding to$|\lambda_i| \geqslant 1$ ($z_i = \lambda_i^{-1}$), it then follows that the sufficient condition for the multivariate integrated variance process $\{\mathbf{H}_t\}$ being co-persistence in variance is

$$(\gamma_i \otimes \gamma_i')\, \mathbf{x}_i = 0; \tag{4.5}$$

for some nontrivial vector $\mathbf{x}_i$.[4]

It is virtually certain that the existence of co-persistent in variance for $\{\mathbf{H}_t\}$ implies the presence of a univariate GARCH (1,1), $h_t^* = c^* + \alpha_1^* (\boldsymbol{\gamma}' \boldsymbol{\varepsilon}_t)^2 + \beta_1^* h_{t-1}^*$, representation for $\{\boldsymbol{\gamma}'\boldsymbol{\varepsilon}_t\}$. As a result, the sum of the scalar parameters of this univariate GARCH (1,1) $\alpha_1^* + \beta_1^*$ must be less than one.

4.3 Co-persistent in variance analysis

4.3.1 Results of GARCH and integrated-GARCH estimation

The data used are daily closing returns of the NYSE S&P 500, Hang Seng, and Shanghai indexes from the DATASTREAM. The sample periods are exactly the same as those of the data from last chapter, starting from January 1, 1993 to August 25, 1998.

Using the uncorrelated stock index return residual series, we examine non-stationary variance within the GARCH framework by estimating the standard $GARCH\ (1,1)$ and $IGARCH\ (1,1)$ with the imposed restriction $\alpha + \beta = 1$. The restriction of integrated variances are performed by the classical misspecification tests including the likelihood ratio (LR) and robust Lagrangian multiplier (LM) statistics, each with one degree of freedom. Table 4.1 reports the results from the quasi-maximum likelihood estimation of the $GARCH$ and $IGARCH$ for the three indexes respectively. The parameter esti-

[4]For comprehensive proofs of the issues addressed in this subsection, see Bollerslev and Engle (1993).

mates are obtained by the simplex algorithm to generate initial values, and then by the numerical derivative BFGS algorithm. The adequacy of the models are then checked by Ljung-Box Q-Test for analyzing the squared normalized residual serial correlations, where these diagnostic results for each index are also reported in panel A, Table 4.1.

The Ljung-Box Q statistics on the squared standardized residuals support the validity of the maintained specification models. The parameter estimates for each index conditional variance imply high persistence shocks to the conditional variance process. To examine the extent of variance shocks persistence, the likelihood ratio (LR) and robust Lagrangian multiplier (LM) tests reported with the null hypothesis that the conditional variance process for each index is specified by $IGARCH\,(1,1)$. It is should be noted that even though the conditional variance processes for the three indexes do not meet the fourth moment conditions, the loglikelihood ratio test could still be applied for testing the limiting distribution of the estimators. This is a big distinction from unit root in the conditional mean process (Lumsdaine 1996). However, Lumsdaine (1995) points out that the loglikelihood ratio test has no robust test. So the robust LM tests that he proposed are also applied.

The results of likelihood ratio (LR) and robust Lagrange multiplier (LM) statistics are statistically insignificant based on the $p-values$ in brackets, indicating the existence of long persistence in each conditional variance process. The conditional distribution for integrated $GARCH$ processes can still be defined even though their second and fourth unconditional moments do not exist. Next, we investigate whether a long-term solution would be yield between the three indexes that linear combinations of variances removes unit roots even if the variance process for each index individually are not stationary. This is the objective of this chapter.

Table 4.1. Quasi-Maximum Likelihood GARCH & IGARCH Estiamtion

(i) S&P 500

$$GARCH \qquad h_t = \underset{(3.19)}{0.0000} + \underset{(7.27)}{0.0471}\varepsilon_{t-1}^2 + \underset{(128.9)}{0.9482}\, h_{t-1}$$

$$\log L_u = 5018.26 \qquad LB^2\,(12) = 6.98\,[0.801]$$

$$IGARCH \qquad h_t = \underset{(3.51)}{0.0000} + \underset{(7.42)}{0.0486}\varepsilon_{t-1}^2 + \underset{(--)}{0.9514}\, h_{t-1}$$

$$\log L_R = 5017.60 \qquad LB^2\,(12) = 7.08\,[0.793]$$

$LR = -2 * (\log L_R - \log L_u) = 1.32\,[0.251]$; *Robust* $LM = 2.30\,[0.129]$

(ii) Hang Seng

$$GARCH \qquad h_t = \underset{(5.20)}{0.0000} + \underset{(9.50)}{0.1222}\varepsilon_{t-1}^2 + \underset{(64.78)}{0.8789}\, h_{t-1}$$

$$\log L_u = 3953.17 \qquad LB^2\,(12) = 5.42\,[0.909]$$

$$IGARCH \qquad h_t = \underset{(6.40)}{0.0000} + \underset{(10.87)}{0.1306}\varepsilon_{t-1}^2 + \underset{(--)}{0.8694}\, h_{t-1}$$

$$\log L_R = 3951.90 \qquad LB^2\,(12) = 5.24\,[0.919]$$

$LR = -2 * (\log L_R - \log L_u) = 2.54\,[0.111]$; *Robust* $LM = 3.11\,[0.078]$

(iii) Shangahi

$$GARCH \qquad h_t = \underset{(10.41)}{0.0000} + \underset{(5.22)}{0.2240}\, \varepsilon_{t-1}^2 + \underset{(20.27)}{0.7380}\, h_{t-1}$$

$$\log L_u = 3306.67 \qquad LB^2\,(12) = 0.734\,[0.999]$$

$$IGARCH \qquad h_t = \underset{(10.48)}{0.0001} + \underset{(7.50)}{0.2615}\varepsilon_{t-1}^2 + \underset{(--)}{0.7384}\, h_{t-1}$$

$$\log L_R = 3305.84 \qquad LB^2\,(12) = 0.783\,[0.999]$$

$LR = -2 * (\log L_R - \log L_u) = 1.66\,[0.198]$; *Robust* $LM = 3.02\,[0.082]$

Notes: Numbers in parentheses are the robust t-values which are robust to departures from the conditional normality assumptions and the numbers in brackets are p-values. $LB^2(12)$ is asymptotically distributed as $\chi^2(12)$. The 5% χ^2 critical values with 12 degrees of freedom is 21.0. Both the LR and robust LM tests are distributed as $\chi^2(1)$ with 5% critical value being 3.84. Both White (1982, theorem3.5) and Lumsdaine (1995) proposes the robust LM statistics as

$$LM^r = T\left[\left(\frac{\partial l_t}{\partial \theta}\right)' J(\theta)^{-1}\left(\frac{\partial g}{\partial \theta}\right)\right]\left[\left(\frac{\partial g}{\partial \theta}\right)' J(\theta)^{-1} I(\theta) J(\theta)^{-1}\left(\frac{\partial g}{\partial \theta}\right)\right]^{-1} \left[\left(\frac{\partial g}{\partial \theta}\right)' J(\theta)^{-1}\left(\frac{\partial l_t}{\partial \theta}\right)\right]$$

where θ is the parameter vector for the conditional variance process;

$$L_T(:,\theta) = -(2T)^{-1}\left(\sum_{t=1}^{T}\ln h_t(\theta) - \sum_{t=1}^{T}\frac{\varepsilon_t(\theta)}{h_t(\theta)}\right);\ \frac{\partial L_t}{\partial \theta} = \begin{bmatrix} E_0\left(\frac{1}{2h_t(\theta)}\left[\frac{\varepsilon_t^2(\theta)}{h_t(\theta)} - 1\right]\frac{\partial h_t(\theta)}{\partial \alpha}\right) \\ E_0\left(\frac{1}{2h_t(\theta)}\left[\frac{\varepsilon_t^2(\theta)}{h_t(\theta)} - 1\right]\frac{\partial h_t(\theta)}{\partial \beta}\right)\end{bmatrix};$$

$$I(\theta) = E_0\left[\frac{\partial l_t}{\partial \theta}\cdot\frac{\partial l_t}{\partial \theta'}\right] = E_0\left[\frac{1}{4h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial \theta}\frac{\partial h_t(\theta)}{\partial \theta'}\left(\frac{\varepsilon_t^4(\theta)}{h_t^2(\theta)} - 1\right)\right];$$

$$J(\theta) = E_0\left[\frac{1}{2h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial \theta}\frac{\partial h_t(\theta)}{\partial \theta'}\right] = E_0\begin{bmatrix} \frac{\varepsilon_{t-1}^4(\theta)}{h_t^2(\theta)} & \frac{\varepsilon_{t-1}^2(\theta)h_{t-1}(\theta)}{h_t^2(\theta)} \\ \frac{\varepsilon_{t-1}^2(\theta)h_{t-1}(\theta)}{h_t^2(\theta)} & \frac{h_{t-1}^2(\theta)}{2h_t^2(\theta)}\end{bmatrix};$$

and $\frac{\partial g}{\partial \theta} = [1, 1]$

4.3.2 Results of the co-persistence analysis

The results of running QML estimation for equation (4.2) to determine the roots of the characteristic matrix polynomial are shown in Table 4.2.

Table 4.2 QML Estimation for Bi-variate GARCH Models

(*i*) S&P 500 *vs.* Hang Seng

$$H_t = \begin{bmatrix} \underset{(1.36)}{0.000} & \underset{(0.41)}{-0.001} \\ \underset{(0.41)}{-0.001} & \underset{(1.72)}{0.002} \end{bmatrix} \begin{bmatrix} \underset{(1.36)}{0.000} & \underset{(0.41)}{-0.001} \\ \underset{(0.41)}{-0.001} & \underset{(1.72)}{0.002} \end{bmatrix}'$$

$$+ \begin{bmatrix} \underset{(4.84)}{0.160} & \underset{(3.04)}{-0.264} \\ \underset{(0.99)}{0.018} & \underset{(7.09)}{0.382} \end{bmatrix} \varepsilon_{t-1}\varepsilon'_{t-1} \begin{bmatrix} \underset{(4.84)}{0.160} & \underset{(3.04)}{-0.264} \\ \underset{(0.99)}{0.018} & \underset{(7.09)}{0.382} \end{bmatrix}'$$

$$+ \begin{bmatrix} \underset{(125.5)}{0.986} & \underset{(3.62)}{0.059} \\ \underset{(0.88)}{-0.007} & \underset{(46.9)}{0.934} \end{bmatrix} H_{t-1} \begin{bmatrix} \underset{(125.5)}{0.986} & \underset{(3.62)}{0.059} \\ \underset{(0.88)}{-0.007} & \underset{(46.9)}{0.934} \end{bmatrix}'$$

$$\widehat{\gamma} = \begin{bmatrix} \underset{(--)}{1.000} & \underset{(5.10)}{-0.058} \end{bmatrix}' \quad \widehat{\alpha}_1^* = \underset{(3.87)}{0.766} \quad \widehat{\beta}_1^* = \underset{(2.72)}{0.151}$$

(*ii*) S&P 500 *vs.* Shanghai

$$H_t = \begin{bmatrix} \underset{(2.12)}{0.000} & \underset{(0.32)}{-0.001} \\ \underset{(0.32)}{-0.001} & \underset{(3.54)}{0.009} \end{bmatrix} \begin{bmatrix} \underset{(2.12)}{0.000} & \underset{(0.32)}{-0.001} \\ \underset{(0.32)}{-0.001} & \underset{(3.54)}{0.009} \end{bmatrix}'$$

$$+ \begin{bmatrix} \underset{(5.44)}{0.176} & \underset{(1.09)}{0.059} \\ \underset{(0.69)}{0.004} & \underset{(6.43)}{0.442} \end{bmatrix} \varepsilon_{t-1}\varepsilon'_{t-1} \begin{bmatrix} \underset{(5.44)}{0.176} & \underset{(1.09)}{0.059} \\ \underset{(0.69)}{0.004} & \underset{(6.43)}{0.442} \end{bmatrix}'$$

$$+ \begin{bmatrix} \underset{(143.3)}{0.983} & \underset{(0.69)}{-0.014} \\ \underset{(0.54)}{-0.002} & \underset{(21.5)}{0.862} \end{bmatrix} H_{t-1} \begin{bmatrix} \underset{(143.3)}{0.983} & \underset{(0.69)}{-0.014} \\ \underset{(0.54)}{-0.002} & \underset{(21.5)}{0.862} \end{bmatrix}'$$

$$\widehat{\gamma} = \begin{bmatrix} \underset{(--)}{1.000} & \underset{(0.79)}{0.0049} \end{bmatrix}' \quad \widehat{\alpha}_1^* = \underset{(59.4)}{0.955} \quad \widehat{\beta}_1^* = \underset{(2.87)}{0.042}$$

Continue...

Table 4.2 QML Estimation for Bi-variate GARCH Models

(*iii*) Hang Seng *vs.* Shanghai

$$H_t = \begin{bmatrix} \underset{(2.67)}{0.002} & \underset{(0.59)}{-0.001} \\ \underset{(0.59)}{-0.001} & \underset{(4.10)}{0.009} \end{bmatrix} \begin{bmatrix} \underset{(2.67)}{0.002} & \underset{(0.59)}{-0.001} \\ \underset{(0.59)}{-0.001} & \underset{(4.10)}{0.009} \end{bmatrix}'$$

$$+ \begin{bmatrix} \underset{(6.47)}{0.291} & \underset{(2.49)}{-0.014} \\ \underset{(0.61)}{-0.007} & \underset{(6.10)}{0.416} \end{bmatrix} \varepsilon_{t-1}\varepsilon_{t-1}' \begin{bmatrix} \underset{(6.47)}{0.291} & \underset{(2.49)}{-0.014} \\ \underset{(0.61)}{-0.007} & \underset{(6.10)}{0.416} \end{bmatrix}'$$

$$+ \begin{bmatrix} \underset{(53.9)}{0.951} & \underset{(0.04)}{0.0003} \\ \underset{(1.16)}{0.0069} & \underset{(23.6)}{0.872} \end{bmatrix} H_{t-1} \begin{bmatrix} \underset{(53.9)}{0.951} & \underset{(0.04)}{0.0003} \\ \underset{(1.16)}{0.0069} & \underset{(23.6)}{0.872} \end{bmatrix}'$$

$$\widehat{\gamma} = \begin{bmatrix} \underset{(--)}{1.000} & \underset{(1.95)}{-0.021} \end{bmatrix}' \quad \widehat{\alpha}_1^* = \underset{(21.7)}{0.864} \quad \widehat{\beta}_1^* = \underset{(3.77)}{0.117}$$

Notes: Numbers in parentheses are the robust t-values which are robust to departures from the conditional normality assumptions. The first element in γ vector is normalized to unity. The parameters $\widehat{\alpha}_1^*$ and $\widehat{\beta}_1^*$are for univariate GARCH(1,1) process given by $\widehat{\gamma}\varepsilon_t$.

By comparing the estimation results of S&P 500 stock index *vs.* Hang Seng stock index with those of Hang Seng *vs.* Shanghai stock indexes, the two cases would seem to illustrate several similarities. Both cases emerge as being alike in the following ways:

- In both cases, eigenvalues (not necessarily distinct) and eigenvectors of the Kronecker products $(\mathbf{A}_1' \otimes \mathbf{A}_1' + \mathbf{B}_1' \otimes \mathbf{B}_1')$ can be determined with either Jordan decomposition of the Kronecker products or the characteristic polynomial

$$\left|\mathbf{I} - (\mathbf{A}_1 \otimes \mathbf{A}_1)' \boldsymbol{\lambda}^{-1} - (\mathbf{B}_1 \otimes \mathbf{B}_1)' \boldsymbol{\lambda}^{-1}\right| = 0$$

as the characteristic equation, where the matrix $\mathbf{I}-(\mathbf{A}_1 \otimes \mathbf{A}_1)' \boldsymbol{\lambda}^{-1}-(\mathbf{B}_1 \otimes \mathbf{B}_1)' \boldsymbol{\lambda}^{-1}$ is singular. One can see a similar pattern that the variances are highly persistent from the eigenvalues $\widehat{\boldsymbol{\lambda}}$. For example, in the case of S&P 500 stock index *vs.* Hang Seng stock index, the ordered eigenvalues are

$$\widehat{\boldsymbol{\lambda}}_{sp \longleftrightarrow hs} = [0.9943, 0.9828, 0.9788, 0.9702]'$$

while in the case of Hang Seng *vs.* Shanghai stock indexes, the ordered eigenvalues are

$$\widehat{\boldsymbol{\lambda}}_{hs \longleftrightarrow sh} = [0.9877, 0.9504, 0.9491, 0.9363]'$$

- In both cases, there exist nontrivial vector $\widehat{\gamma}$ though the vector $\widehat{\gamma}$ for Hang Seng *vs.* Shanghai is marginally significant with robust $t = 1.95$, a finding could enable us to examine co-persistence in each pair conditional variances. For S&P 500 vs. Hang Seng, the estimated $\widehat{\gamma}$ can make some eigenvectors, $\widehat{x}_i$, with respect to each corresponding eigenvalue to be zero that is, $(\gamma_2 \otimes \gamma_2')\,\widehat{x}_2 = (\gamma_3 \otimes \gamma_3')\,\widehat{x}_3 = 0$; and $(\gamma_2 \otimes \gamma_2')\,\widehat{x}_2 = (\widehat{\gamma}_4 \otimes \widehat{\gamma}_4)\,\widehat{x}_4 = 0$ for the pair of Hang Seng *vs.* Shanghai. Thus, in two cases, the nontrivial vector $\widehat{\gamma}$ are indeed co-persistent vector. In particular, the sums of $(\alpha_1^* + \beta_1^*)$ are low especially for S&P vs. Hang Seng system, suggesting long term solutions exist in S&P *vs.* Hang Seng index system and in Hang Seng *vs.* Shanghai index system.

However, it emerges from the system of S&P *vs.* Shanghai that the estimated co-persistent vector $\widehat{\gamma}$ that linearly combines the conditional variances of S&P and Shanghai stock indexes is highly insignificant, that is, $\widehat{\gamma} = [1, 0.0049]$ with $t = 0.79$. Further, it is noticeable that all elements in the block diagonal matrix $\mathbf{J}$ are larger than one, suggesting their conditional variance process are not co-persistent. It is also noticeable that there seems to be no volatility spillover between the New York and Shanghai stock markets as the volatility spillover coefficients from S&P 500 to the Shanghai index, presented by the estimated value of β_{12} and the estimated value of α_{12}

$(ie.\beta_{12} = 0.059\,(t = 1.09)$ and $\alpha_{12} = -0.014\,(t = 0.69))$, are insignificant. It should be emphasized that the co-persistent analysis only provide evidence of the long run equilibrium relationships but not directions of the volatility spillover.

4.4 The error correction model (ECM) analysis

Although the volatility spillover from the Hong Kong stock market to the Shanghai stock market is explicitly analyzed, we need an additional justification for which the three bivariate systems applied in the previous section as the way of analysis should not be considered to be invalid. With respect to the literature on bivariate systems, the inference of bivariate system approach may easily suffer from not including a potential relevant causing variable (Lütkepohl, 1982), which is a highly relevant argument and a necessary consideration to this study. Before drawing any firm conclusions, it is thus worth examining the results by other measures[5].

So the purposes of this section are to justify the results estimated above from a different perspective and to demonstrate the validity of these bivariate systems with an experiment, in which we apply the error correction model (ECM) for the three variance

[5]We also estimate the three conditional variance processes in a larger three-dimensional framework, still applying Engle and Kroner's multivariate $GARCH - BEKK\,(1,1)$ model. The estimation results for the trivariate simulataneous GARCH (1,1) compared to those for the bivariate system indicate that there is no major change in the results. The estimation results for the three-dimensional framework are as follows:

$$H_t = \begin{bmatrix} \underset{(12.74)}{0.1535} & \underset{(8.47)}{-0.1797} & \underset{(0.774)}{0.0691} \\ \underset{(1.44)}{0.0092} & \underset{(18.64)}{0.2679} & \underset{(1.09)}{-0.030} \\ \underset{(0.03)}{-0.000} & \underset{(0.16)}{-0.001} & \underset{(21.79)}{0.4284} \end{bmatrix}' \varepsilon_{t-1}\varepsilon'_{t-1} \begin{bmatrix} \underset{(12.74)}{0.1535} & \underset{(8.47)}{-0.1797} & \underset{(0.774)}{0.0691} \\ \underset{(1.44)}{0.0092} & \underset{(18.64)}{0.2679} & \underset{(1.09)}{-0.030} \\ \underset{(0.03)}{-0.000} & \underset{(0.16)}{-0.001} & \underset{(21.79)}{0.4284} \end{bmatrix}$$

$$+ \begin{bmatrix} \underset{(426)}{0.9870} & \underset{(4.62)}{0.0376} & \underset{(0.47)}{-0.0154} \\ \underset{(1.44)}{-0.003} & \underset{(238)}{0.9579} & \underset{(0.13)}{0.019} \\ \underset{(0.38)}{0.001} & \underset{(0.56)}{0.002} & \underset{(78.3)}{0.8613} \end{bmatrix}' H_{t-1} \begin{bmatrix} \underset{(426)}{0.9870} & \underset{(4.62)}{0.0376} & \underset{(0.47)}{-0.0154} \\ \underset{(1.44)}{-0.003} & \underset{(238)}{0.9579} & \underset{(0.13)}{0.019} \\ \underset{(0.38)}{0.001} & \underset{(0.56)}{0.002} & \underset{(78.3)}{0.8613} \end{bmatrix} + C_0'C_0$$

where $H_t = [h_{us,t}, h_{hk,t}, h_{sh,t}]'$

processes in a multivariate framework. A major strength of the ECM dynamic specification in understanding relationships among the three index conditional volatilities is its clear emphasis on long run equilibrium relationships which is time invariant. Another strength is its emphasis on identifications of weak exogeneity variables while taking into account of cointegration restrictions. The ECM dynamic specification can also focus on restrictions on the loading matrix deciding which cointegrating vector enter which particular variable (see Urbain (1992) and Johansen (1992) on the use of error correction models (ECM) within a single equation conditional framework).

The long-run equilibrium relationships among the three variance precesses based on maximum likelihood estimation of the error correction model (ECM) is given by

$$\Delta\boldsymbol{\varepsilon}_t^2 = \mathbf{C} + \boldsymbol{\xi}_1\Delta\boldsymbol{\varepsilon}_{t-1}^2 + \boldsymbol{\xi}_2\Delta\boldsymbol{\varepsilon}_{t-2}^2 + \cdots + \boldsymbol{\xi}_k\Delta\boldsymbol{\varepsilon}_{t-k+1}^2 + \boldsymbol{\Pi}\boldsymbol{\varepsilon}_{t-1}^2 + \boldsymbol{\eta}_t$$

where $\Delta\boldsymbol{\varepsilon}_t^2 = [\Delta\varepsilon_{us}^2, \Delta\varepsilon_{hk}^2, \Delta\varepsilon_{sh}^2]'$, the cointegration restriction parameters $\boldsymbol{\Pi}$ having reduced ranks as $\boldsymbol{\Pi} = -\boldsymbol{\alpha\beta}'$, and the rest of parameters are defined as the usual ECM[6]. Three hypotheses are specifically considered:

[6]We wish to specify cointegrating relationships for the conditional volatility process H_t, however, there is no direct error correction represenation for H_t since the conditional H_t process is the determined process rather than the random process based on past information. To establish such relationships, let $v_t \equiv \varepsilon_t^2 - H_t$, in which H_t is the GARCH (p,q)representation as

$$H_t = C + A(L)\,\varepsilon_t^2 + B(L)\,H_{t-1}$$

where L denotes the lag operator such that $L^i y_i \equiv y_{t-i}$.

On substituting H_t into v_t , the unobserable ε_t^2 process can be expressed as an $VAR(k)$ process

$$\varepsilon_t^2 = C + (A(L) + B(L))\,\varepsilon_t^2 + \eta_t$$

where $\eta_t = (I - B(L))\,v_t$.

In practice, it may be instructive to reparametrize the VAR specification of ε_t^2 into its equivalent representation; the error correction model (ECM) in order to solve co-persistence in variance, (For the detail procedures of the reparametrization see Hamilton, 1994, pp.580)

$$\Delta\varepsilon_t^2 = C + \xi_1\Delta\varepsilon_{t-1}^2 + \xi_2\Delta\varepsilon_{t-2}^2 + \cdots + \xi_k\Delta\varepsilon_{t-k+1}^2 + \Pi\varepsilon_{t-1}^2 + \eta_t$$

where:

$$\xi_s = -\left[(A_1 + B_1) + (A_2 + B_2) + \cdots + (A_k + B_k)\right]$$

for $s = 1, 2, ..., k-1$, and

- Hypothesis (1): no variance cointegrating relationship exists in the system.
- Hypothesis (2): the New York's variable and the Shanghai's variable would not enter each other's long run equilibrium relationships by imposing restrictions on the matrix $\boldsymbol{\beta}'$ $(r \times 3)$ of cointegrating vector(s).
- Hypothesis (3): restrictions on the loading matrix $\boldsymbol{\alpha}$ $(3 \times r)$ are testing one certain cointegrating vectors would appear one and only one particular equation.

On the basis of Akaike information criterion (AIC) the optimal lag length is of 2. The statistical results of cointegration tests for the rank of matrix $\mathbf{\Pi}$ with constant unrestricted are tabulated in Panel A, Table 4.3. The maximum eigenvalue and trace statistic tests illustrate that null hypothesis (1) of no cointegrating relationship is strongly rejected and there are two long run equilibrium relationships among the three variables (i.e. $rank\,(\mathbf{\Pi}) = 2$), which are expected.

$$\Pi = -(I_n - (A_1 + B_1) - (A_2 + B_2) - \cdots - (A_k + B_k)) = -\alpha\beta'$$

Table 4.3 Cointegration and Likelihood Ratio Hypothesis Tests

Panal A: Likelihood ratio test for the rank of matrix $\mathbf{\Pi}$

Null Hypothesis	Eigenvaule	Maximum Eigenvalue	Trace Statistic
$r = 0$	0.0254	34.78** (21.0)	65.75** (29.7)
$r \leqslant 1$	0.0210	29.08** (14.1)	30.97** (15.4)
$r \leqslant 2$	0.0014	1.887 (3.8)	1.887 (3.8)

Panel B: Restrictions hypothesis test for matriices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ *in which* $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$

Restricitons	LR statistic
$\alpha_{31} = \alpha_{22} = \beta_{31} = \beta_{12} = 0$	$\chi^2(4) = 6.4302\,[0.1692]$
$\alpha_{11} = \alpha_{12} = 0$	$\chi^2(2) = 0.2423\,[0.8727]$

Notes: (1) The error correction model (ECM) is given by

$$\begin{bmatrix} \Delta\varepsilon^2_{us,t} \\ \Delta\varepsilon^2_{hk,t} \\ \Delta\varepsilon^2_{sh,t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{bmatrix} \begin{bmatrix} \varepsilon^2_{us,t-1} \\ \varepsilon^2_{hk,t-1} \\ \varepsilon^2_{sh,t-1} \end{bmatrix} + \mathbf{C} + \sum_{i=1}^{k-1} \xi_i \Delta\varepsilon^2_{t-i} + \eta_t$$

where the subscripts 1, 2, 3 denote the New York, Hong Kong, and Shanghai stock markets respectively. (2) r is the number of cointegrating vectors and the critical values for the maximum eigenvalue and trace statistics at the 95% significance level are reported in parentheses. (3) The values in brackets are p-values. (4) ** denotes significance at the 5% level.

Now the illustration of the relevance of the New York stock market to the volatility spillover process between the Hong Kong and Shanghai stock markets is subject to the statistical tests for hypothesis testing on matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$ on the basis of the defined rank of $\boldsymbol{\Pi}$. The likelihood ratio test on matrices $\boldsymbol{\alpha}$ (3×2) and $\boldsymbol{\beta}'$ (2×3) can also provide 'information' on whether cointegration test results are consistent with respect to the bivariate examination results given in the previous section. Panel B provides the results for such restriction tests based on hypotheses (2) and (3) .

We first test the restrictions on the coefficients of the two long-run equilibrium relationships among the three variables entering the cointegrating space spanned by the two rows of $\boldsymbol{\beta}'$. Two identifying restrictions are imposed on β_{31} and β_{12} (i.e. $\beta_{31} = 0$ and $\beta_{12} = 0$). The restriction $\beta_{31} = 0$ implies that Shanghai does not appear in the long run relationship between New York and Hong Kong and $\beta_{12} = 0$ suggests that New York does not enter the long run relationship between Hong Kong and Shanghai. These two restrictions are imposed based on hypothesis (2) of no relationship between New York and Shanghai. One can inquire that even if the restrictions on $\boldsymbol{\beta}'$ such that $\beta_{31} = 0$ and $\beta_{12} = 0$ are accepted, the sum of the two cointegrating vectors can also be formed a long run stationary equilibrium relationship which could enter the Shanghai variable $\Delta h_{sh,t}$. Subject to this situation, two additional restrictions are imposed on some elements of the loading matrix $\boldsymbol{\alpha}$, α_{31} and α_{22} (*i.e.* $\alpha_{31} = 0$, $\alpha_{22} = 0$), which are equivalent to requiring no New York volatility in the Shanghai volatility equation. Therefore, the null hypotheses (2) and (3) that New York volatility do not enter the Hong Kong and Shanghai cointegrating relationship and Shanghai volatility equation are verified if and only if these four restrictions are accepted.

These restrictions are examined by the means of likelihood ratio tests and the results are presented in Panel B, Table 4.3. The null hypotheses are not rejected, thus may implying that there has no a stable long run link between the New York and Shanghai stock markets. Next, we examine whether the New York stock market is exogenous for the Shanghai and Hong Kong system by testing whether α_{11} and α_{12} are of zero. The last two restrictions are accepted, implying all the relevant information concerning Shanghai and Hong Kong volatilities is contained in the Shanghai *vs.* Hong Kong subsystem. This investigation is in agreement with the findings in the last section. Particularly, the results of volatility spillover effects from the Hong Kong to the Shanghai stock markets are of adequacy under the specification of the three bivariate systems.

4.5 Conclusion

This chapter has aimed to highlight the issue concerning the potential US' influences on the volatility spillover between the Hong Kong stock market and Shanghai stock market. And, in particular, this chapter has taken into account the view of New York as volatility spillover domination, the issue which has been left to be addressed from the previous chapter. However, with respect to volatility spillover, we have seen that the dominance is not an inevitable consequence for such volatility spillovers.

Instead of just directly seeking whether there exists any volatility spillover from US to Shanghai, the main idea of this chapter is to provide evidence whether there exist long run solutions between the three stock markets by using the theory of co-persistent in conditional variances, where a suitable nonzero vector that makes linear combination of integrated variances being stationary can be found. Such an approach seems to be appropriate when the three index variances are integrated of order one by the preliminary $IGARCH$ analysis.

A deep co-persistence in the variance analysis illustrates the presence common components both in the pair of US and Hong Kong system and in the pair of Hong Kong and Shanghai system but not in the system of US and Shanghai. The volatilities for the combination of S&P 500 index and Hang Seng index and the combination of Hang Seng index and Shanghai index are stationary. When taking into account the results of previous chapter, the common components for the Hong Kong and Shanghai stock markets may be considered as the *red-chip* shares. In other words, the *red-chip* shares may be the potential causes for the volatility spillover from the Hong Kong stock market to the Shanghai stock market. The reason why there is no direct link between S&P index and Shanghai index may be that there is no proper channels allowing the volatility spillover between the two markets.

Therefore, the main contribution of this further study with respect to the last chapter is to provide some evidences in recognizing the volatility spillover from the Hong Kong stock market to the Shanghai stock market being valid. The findings of co-persistent in variances between these two stock markets could provide us additional insight on why the Shanghai stock market becomes less volatile after the handover. The reason might rest on the fact that stock market volatility is closely associated with various information. The long relationship between the two stock markets in conditional variances can reduce the impacts of current shocks on the Shanghai stock market. Numerous measurements yield consistent results. It is important to note that, in line with last chapter, the results of this chapter should not lead the view that stock market control is a necessary part of protective action but that this needs to be analyzed in a proper way in order to gain the benefits from more information being incorporated in stock price discovery process. A further issue needs to mention that New York's total dominance is not possible in the relatively isolated Shanghai stock market may be due to lack of proper channels between New York and Shanghai. That whether there exist any other channels for volatility spillovers from the Hong Kong stock market to the Shanghai stock market is explored further in the next chapter.

Chapter 5

What Factors Within The Hong Kong Stock Market Move The Shanghai Stock Prices?

5.1 Introduction

The main theme of this chapter is to examine the potential sources of volatility spillovers from the Hong Kong capital market to the Shanghai stock market. It seeks to identify and examine the relative importance of financial and economic variables originating in the Hong Kong capital market that might affect the Shanghai stock price movements. The empirical methodology that is applied is based on the Structural Vector Autoregressive model ($SVAR$) and its innovation accounting exercises, suggested by Amisano and Giannini (1997).

As have mentioned in the previous chapters, scholars of financial economics have focused considerable attention on the influence of volatility transmission between different stock markets. This is because volatility examination is the point of most direct interaction between investors and many asset pricing models. The early research efforts have focused mainly on the directions of the volatility spillovers which represent the direction of the information flow among these stock markets. This is because sound understanding of such information flow is of paramount importance to understanding many fascinating questions regarding portfolio selections, hedging, and risk management problems. However, relatively few researchers have attempted to identify what the ingredients of the information flow are, (except for King et al. (1994) in which they decompose the source of changes between different stock market volatility into observable factors by applying factor analysis from ten macroeconomic variables and unobservable factors. They even apply the Kalman filter to take into account the difference of information set available to econometrican), and how long these ingredients from one market that may have effects on another stock market may last. There are, however, a number of studies have examined the relationships among economic state variables, such as, interest rates and spreads, the term structure, and Gross National Product (GNP) etc., and the expected stock returns (See Harvey 1988, Fama and French 1988, Chen 1991, and Patelis 1997).

This brings us to the aim of this study, which is to link the volatility spillovers and the economic state variables, and comes to conclusions that offer partial support for the contentions in the literature. This research suggests that integration of both market is warranted; the Hong Kong economic and financial factors do have influences on the Shanghai stock market. At the same time, however, the research suggests that drawing policies attempting to curb foreign capital influences on domestic market is inappropriate. This is because international stock markets volatility spillover is an outcome of information interaction among stock markets responding to both domestic and international factors.

This chapter proceeds in three steps. First, (section 5.2), the structural vector autoregressive ($SVAR$) model is described to show that the specification of the structural vector autoregressive is more appropriate than that of the reduced form VAR. Researchers have already discussed the inadequacy of the reduced-form VAR, for example, its inability to uncover the exact exogenous shocks to each individual variable in the vector. The SVAR model overcomes this problem. Identification of the sources is achieved by imposing restrictions on the reduced-form VAR. Next, section 5.3 covers the characteristic properties of each variable utilized in this paper. It also covers the model specification and diagnostics. Finally, section 5.4 deals with the relationship between the Hong Kong market and the Shanghai stock market.

Empirical results show that the Hang Seng interest rate shocks appear to be the most significant factor that influences the Shanghai stock prices, although other financial shocks represented by the Hang Seng index trading volumes also play important roles in explaining the fluctuations in the Shanghai stock prices.

5.2 The structural VAR methodology

The focus of this chapter is on locating the potential economic and financial factor sources within the Hong Kong capital market that may influence the Shanghai stock prices movements for the post-handover period. Structural VAR model has been used in constructing and estimating the findings. The theoretical appeal that forms the basis for the application of the structural VAR approach rather than the reduced form or unrestricted VAR approach is explained as follows.

The reduced-form or unrestricted VAR model has become one of the most important approach of econometricans because it is a convenient way to identify the relationships among economic variables without the need to link it to any particular economic theories. All the variables in the vector are equally treated as endogenous in the VAR system.[1] The Wold theorem is the major theorem for the identification of the unrestricted VAR. As usual, the resulting reduced form VAR for a vector of n-dimensioned autoregressive endogenous variables $\mathbf{Y}_t$ is shown as

$$\mathbf{Y}_t = \mathbf{\Pi}(L)\,\mathbf{Y}_{t-1} + \boldsymbol{\epsilon}_t, \quad t = 1, ..., T \tag{5.1}$$

[1]There are no observable exogenous variables in the vector because the exogeneity of those variables are not crediable (Sims 1980).

where $\mathbf{Y}_t$ is the $(n \times 1)$ vector of endogenous variables, $\mathbf{\Pi}(L)$ is a p^{th} degree matrix polynomial in the lag operator L such that $L^i\mathbf{Y}_t = \mathbf{Y}_{t-i}$, that is, $\mathbf{\Pi}(L) = \sum_{i=1}^{p} \mathbf{\Pi}_i\mathbf{L}^i = \mathbf{\Pi}_1\mathbf{L} + \mathbf{\Pi}_2\mathbf{L}^2 + ... + \mathbf{\Pi}_p\mathbf{L}^p$, where all of the $\mathbf{\Pi}$ matrices are square matrices. p is an nonnegative integer and $\boldsymbol{\epsilon}_t$ is a multivariate Gaussian distributed white noise vector errors with $\mathbf{0}$ mean, $\mathbf{\Sigma}$ variance-covariance matrix. Notice that the $\boldsymbol{\epsilon}_t$'s are serially uncorrelated but $\boldsymbol{\epsilon}_t$ may be contemporaneously correlated across different elements within the $\boldsymbol{\epsilon}_t$ vector.

However, there have been endless criticisms and arguments on the interpretation and atheoretical approach of the unrestricted VAR ever since its introduction by Sims (1980).[2] The unrestricted VAR decomposes the residuals $\boldsymbol{\epsilon}_t$ into orthogonal shocks by applying the Choleski decomposition of the variance-covariance matrix $\mathbf{\Sigma}$. As a result the matrix $\mathbf{\Sigma}$ is decomposed into the unique product of lower triangular matrix $\mathbf{U}$ and its transposed of the lower triangular matrix $\mathbf{U}^\intercal$(i.e., $\mathbf{\Sigma} = \mathbf{U}\mathbf{U}^\intercal$), through which each shocks is uncorrelated with the other shocks. The top element of lower triangular matrix $\mathbf{U}$ responds to its own exogenous shock while the bottom element responds to not only its own shock but also all the elements above it. One problem with this decomposition is that the results are not robust to changing variable orderings. This identification makes the interpretation difficult because the same element responds to different other elements each time with the change of the ordering of the variables in $\mathbf{Y}_t$. Another problem, (See Cooley and LeRoy (1985)), is that there is no identification of the instantaneous relationships among the estimated residuals $\boldsymbol{\epsilon}_t$ because they are linear combinations of the exogenous shocks. As a result, the correlation structure of the reduced form $\mathbf{\Sigma}$ matrix conceal the information of the instantaneous relationships. It should be noted that these two problems are closely related through the Choleski factorization of the variance-covariance matrix for the residuals $\boldsymbol{\epsilon}_t$.[3]

[2]Sims (1980) imposes a Wold causal chain on the contemporaneous relationships between shocks and endogenous variables to achieve identification. The absence of a constant term in equation 5.1 is without loss of generality, since any one of the elements of $\mathbf{Y}_t$ can be set to unity.

[3]Bernanke and Mihov (1995) give an example showing how the reduced form VAR residuals matrix ϵ_t are linear combinations of the mutually uncorrelated "primitive" error terms from the simultaneous marcoeconomic system, thus each element in ϵ_t including several such "primitive" error terms.

Responding to the no distinct response patterns from the reduced form VAR residuals $\boldsymbol{\epsilon}_t$ and the difficulties in identifying unobserved structural shocks from the observable VAR-based residuals $\boldsymbol{\epsilon}_t$, the Structural VAR approach has been developed by introducing two invertible matrices and imposing restrictions on them to structuralize the unrestricted VAR.[4] Equation 5.1 can be equivalently written as

$$(\mathbf{I} - \mathbf{\Pi}(L))\,\mathbf{Y}_t = \boldsymbol{\epsilon}_t$$

Where $\mathbf{I}_n$ is the unit matrix. Setting $\mathbf{\Lambda}(L) = \mathbf{I}_n - \mathbf{\Pi}(L) = \mathbf{I}_n - \mathbf{\Pi}_1 L - \mathbf{\Pi}_2 L^2 - ... - \mathbf{\Pi}_p L^p$, we get

$$\mathbf{\Lambda}(L)\,\mathbf{Y}_t = \boldsymbol{\epsilon}_t, \quad t = 1, ..., T \tag{5.2}$$

Then multiplying two $(n \times n)$ invertible matrices $\mathbf{A}$ and $\mathbf{B}$, which could give distinct interpretations to the model, on equation 5.2 yields

$$\begin{aligned} \mathbf{A}\mathbf{\Lambda}(L)\,\mathbf{Y}_t &= \mathbf{A}\boldsymbol{\epsilon}_t \\ \mathbf{A}\boldsymbol{\epsilon}_t &= \mathbf{B}\mathbf{e}_t \end{aligned} \tag{5.3}$$

Where $\mathbf{e}_t$ is a the orthogonal disturbances matrix and is distributed as $(\mathbf{0}, \mathbf{I}_n)$. The two invertible matrices $\mathbf{A}$ and $\mathbf{B}$ could make the endogenous $\mathbf{Y}_t$ $(n \times 1)$ variable movements as a function of the structural or orthonormal shocks whose variance-covariance matrix is $\mathbf{B}\mathbf{B}^{\intercal}$. Thus the structural shocks have certain effects on the levels of the endogenous variables and identify distinct economic interpretations. The $\mathbf{A}$ matrix impose simultaneous relationships on the endogenous variables and the $\mathbf{B}$ matrix is closely related to the long run restriction for structural shocks. The product of $\mathbf{A}\mathbf{\Lambda}(L)$ may (should) not be unit matrix. Amisano and Giannini (1997) and Giannini *et al.* (1995) define this

[4]Generally, the main idea of structuring reduced form VAR is quite similar to Sims' (1980) Choleski factorization. Sims just multiplies the reduced form VAR by using the inverse of the Choleski factor of $\mathbf{\Sigma}$ rather than another one or two invertible square matrices. Granger-causality tests are not good alternatives for unrestricted VAR models because the right-hand variables may not be orthogonal. Canova (1995) provides a brillant explanation of the philosophy for the structural VAR.

structural VAR model as AB-model .

One special case for the AB-model is that when **B** is a unit matrix, then the AB-structural model becomes contemporaneous structural VAR model. The contemporaneous parameter matrix **A** can provide the structural intepretations (Blanchard and Watson (1986)). The other special case is that when **A** matrix is the unit matrix, then the AB-structural model becomes long-run structural model whose structural interpretations are captured by the long run structural parameter matrix **B**. The long run structural VAR model is well applied by Blanchard and Quah (1989). Keating (1992) gives detail comparisons between the two special cases structural VAR models.[5] The AB-model applies both contemporaneous and long-run parameter matrices to identify the structural VAR model (Gali (1992), Amisano *et al.* (1995), Amisano and Giannini (1997) and Giannini *et al.* (1995)).

The improvement from the unrestricted VAR to structural VAR models is that the structural VAR model can provide insights into the unique shocks on each endogenous variables in $\mathbf{Y}_t$ by introducing sufficient identifying restrictions.[6] Furthermore, the behavioral variables of the structural models can be presumed to be policy invariant. In general, the structural VAR model represents a simple structural model in which we could identify whether the Shanghai stock prices are driven by the both current and past shocks from the Hong Kong economic and financial variables. Further, the structural responses of the Shanghai stock prices to the Hong Kong shocks can further be measured by the associated impulse response functions and forecast variance error decomposition functions.

---

[5]Sims' (1980) triangular representation can also be treated as a special case of the structural AB model.

[6]In the AB-model, there are $2n^2$ unkown parameters need to be estimated in the **A** and **B** matrices. There are, however, only $n(n+1)/2$ known conditions - the variance-covariance matrix $\widehat{\boldsymbol{\Sigma}}$ - which are estimated from the unrestricted VAR. So addition restrictions have to be imposed on the **A**, **B** matrices.

5.3 The unrestricted VAR specification and misspecification diagnostic tests

Prior to analyzing the structural identifications and specifications, the specification and several misspecification diagnostic tests for the unrestricted VAR must be performed by using PcFiml and the RATS procedures MALCOLM. This is because that the well defined misspecification tests for the unrestricted VAR are crucial for the solutions of the structural parameters of interests. Three issues need to be addressed for the specification of the unrestricted VAR (for detail see Canova 1995).

The first issue is the decision of which the potential financial or economical variables within the Hong Kong stock market that may influence the Shanghai stock prices after the handover should be put into the vector autoregression (VAR) system. There are many potential candidate variables such as dividend yields, interest spreads, inflation rates, corporate cash flows, and other possible information variables (Campbell and Shiller (1988, 1991), Campbell (1993), and Fama and French (1988)). In principle, all these variables might affect the Shanghai stock prices. Unfortunately, not all the data that are likely to have influences on the Shanghai stock prices can be collected. The directly observable weekly data are employed from September 1997 to October 1999, which is a time series of 131 observations.

The data used in the unrestricted VAR to examine the relationships between the two markets are defined and justified as follows: The three-month Hong Kong Interbank Offered Rate (HIBOR) futures contracts are used as a representative Hong Kong economic variable. The Hong Kong Interbank Offered Rate (HIBOR) is the benchmark short-term interest rate quoted by twenty banks designated by Hong Kong Association of Banks. Because the fluctuations of the Hong Kong dollar interest rate have increased since the Asian financial crisis in 1997, the Hong Kong Futures Exchange (HKFE) introduce the three-month HIBOR futures contracts on 26 September 1997 allowing investors to manage their interest rate exposures.

While other potential economic and financial data were limited by data availability, the Hang Seng index stock weekly trading volumes are utilized as the proxy for the number of unobservable public information flowing from the Hong Kong stock market to the Shanghai stock market. Research by Gallant *et al.*(1992), Campbell *et al.* (1993), and Hiemstra and Jones (1993) have found volume data contain useful information for stock prices, especially stock market volumes have predictive power for stock index volatility. Here, however, the use of the stock volumes is not without questions, because the volumes may also include the interest rate information. The measure of the stock index volume is the total number of shares traded each week. The Shanghai stock price index is a value-weighted, arithmetic index of all traded stocks.

The second issue is the selection of the type of deterministic components and the choice of the optimal lag length of the vector autoregression. The deterministic components may include restricted, unrestricted constants, linear trends, or other non-stochastic regressors. Different deterministic can induce various asymptotic distributions for unit root and cointegration tests. Thus, the prerequisite for valid inference on the potential stochastic features of the variables and the determination for estimation procedure are dependent on choosing proper deterministic components.

Only a unrestricted constant is chosen for the VAR including the above three variables. This unrestricted constant is especially important for cointegration system because it will make the error correction terms have a common mean. In order to eliminate serial autocorrelation from the estimated residuals $\widehat{\varepsilon}_t$ in equation (5.1) and keep $\widehat{\varepsilon}_t$ series as the martingale difference process, a relatively long lag length with allowance degree-of-freedom is required so that residuals do not include any informations from omitting conditioning variables. However, the "curse of dimensionality" and the loss of estimation power will occur with too many lags i.e. too many lags can cause near collinearity, which is insufficient data information, or even exact collinearity (Canova 1995). Generally, the appropriate lag length for the correct specification of the model are determined by Akaike's AIC, HQ, Schwarz's SIC, and LR tests.

The selection of the lag length of the VAR is closely related with the choice of variables.(see Johansen 1995, pp.21). The optimal lag length test results for the three variables are presented in panel A of Table 5.1. The selection criteria do not show consistent results. The maximum lag length equal to 8 suggested by the AIC, HQ, and SIC criteria. While the LR tests indicates that five lags are appropriate for the whole system. Thus, we choose the lag length at five. Re-running the systems with longer length of lags does not change the results too much. As shown in panel B, there is no ARCH effects for variables at all lags except for the Shanghai stock prices at lag 1, then the variances of each estimated residuals cannot be predicted by the their corresponding past squared residuals.

The ARCH tests are quite convincing evidence that the VAR is not misspecified by omission of a relevant regressor. The Bera-Jarque normality tests for the residuals for each variable and the whole system are also reported in panel C. The residuals for the Hang Seng index stock volumes and the Shanghai stock prices are considered to be normal while the residuals for the Hong Kong interest rates are non-normal. It is clearly shown the rejection of normality for the whole system is because of excess kurtosis (the fourth moments) of the Hong Kong interest rates. However, skewness (the third moments) has a stronger effect on the cointegration estimation than kurtosis so the rejection of normality do not influence the estimation results. Overall, the diagnostic tests generally support the statistical validity of the VAR specification over the sample period.

Table 5.1 The Dianostic Tests for Reduced-Form VAR Model

1997 : 09 − −1999 : 10

A: Lag Length Determination

$lag\,(h)$	$LR\,(L\,(h)\,/L\,(h-1))$	$[p-value]$	AIC	HQ	SIC
1	N/A	N/A	-16.288	-16.162	-15.977
2	12.560	0.184	-16.234	-16.014	-15.690
3	15.007	0.091	-16.204	-15.890	-15.427
4	8.014	0.533	-16.105	-15.696	-15.095
5	16.842	0.051	-16.094	-15.591	-14.851
6	26.095	0.002*	-16.174	-15.576	-14.698
7	3.034	0.963	-16.026	-15.334	-14.317
8	6.799	0.658	-15.915	-15.129	-13.973

B: ARCH Effect Tests

ARCH order	HKINT	HKVOL	BSHA
1	1.4325	0.066	4.534*
$[p-value]$	[0.235]	[0.798]	[0.036]
3	1.106	0.398	1.639
$[p-value]$	[0.352]	[0.755]	[0.187]
5	0.676	0.669	1.263
$[p-value]$	[0.643]	[0.648]	[0.288]

C: Normality Tests

	IIKINT	HKVOL	BSHA
SK	17.97*	0.491	0.001
$[p-value]$	[0.00]	[0.484]	[0.969]
KUR	23.10*	0.513	0.226
$[p-value]$	[0.00]	[0.474]	[0.635]
SK&KUR	41.71*	1.004	0.227
$[p-value]$	[0.00]	[0.605]	[0.892]

Notes: [a]$LR\left(L\left(h\right)/L\left(h-1\right)\right)$ test is the likelihood ratio test for the maximum lag analysis in VAR. $L\left(h\right)$ is the maximized likelihood corresponding to a VAR(h). The selected maximum lag is $\left(h-1\right)$, where h is the first lag which the null hypothesis is not rejected. (Mosconi, 1999). An asterisk denotes rejection of the null at the 5% significance level. AIC is Akaike's information criterion , HQ is Hannan-Quinn's criterion and SC is Schwarz'ss criterion. The maximum values of AIC and SC determine the appropriate lag length for the VAR model. [b]ARCH(h) test is the joint significance test of the squared residuals in the regression on a constant and h-lagged squared residuals. The values in the brackets are p-values.

The final issue concerns the statistical properties of the variables employed in the VAR. It requires that the variables should be stationary or transformed into stationary variables so that the conventional asymptotic limiting theory is valid. If, however, the processes of the variables are integrated, the functional central limit theorem (FCLT) has to be applied to take into account that the integrating processes have different distributions from stationary processes. Usually cointegration is the optimal way for the transformation if the variables are integrated of order larger than zero. Only through cointegration analysis are the static regressions among integrated variables meaningful. Thus testing for the order of integration in variables is of vital importance for the following estimations.

The unit root test results employing the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests for the three variables in levels are presented in Table 5.2. The ADF and PP unit root tests are essentially the same and they have the same asymptotic distributions. The major difference is the treatment for the higher-order autocorrelation terms of the first difference variables. The main function of the higher-order autocorrelation is to ensure no remaining information left in the estimated residuals. The ADF unit root test just put the first difference higher terms into the unit root test regression models. While the PP unit root test estimated by a non-parametric correction to the standard statistics assumes that the unit root regression residuals is a mixing process rather than white noise. So the PP test accounts for the autocrrelations in the unit root regression residuals. As a result the Phillips-Perron unit root test is robust to serial correlation and time-dependent heteroskedasticity by incorporating autocorrlated product residual variance estimates with specified truncation lags. When the unit root regression residuals are independently, identically distributed, The ADF parametric test is a special case for the non-parametric test due to autocorrlated products being zero.

As shown in the panel A of table 5.2, all ADF and PP test statistics for the levels for the Hong Kong interest rates and the Shanghai stock prices are well below the 5% critical values, then both tests fail to reject the null hypothesis of unit root in the levels for both series. While the ADF and PP unit root tests reject the null hypothesis of non-stationary in the levels for the share trading volumes of the Hang Seng index. In summary, these unit root tests lead to the conclusion that the Hong Kong interests and the Shanghai stock prices are $I(1)$ processes, and the share trading volumes of Hang Seng index is $I(0)$ process.

Table 5.2 Unit-root & Cointegration Rank Statistics for Reduced-form VAR

1997 : 09 − −1999 : 10

A: Unit-Root Statistics

	ADF		Phillips-Perron					
Variables	τ_μ	τ_τ	$Z(t_\alpha)$	$Z(\alpha^*)$	$Z(\Phi_1)$	$Z(t_{\alpha^*})$	$Z(\Phi_2)$	$Z(\Phi_3)$
HKINT	-1.85	-2.94	-3.83	-10.14	2.65	-2.30	4.93*	7.39*
HKVOL	-2.36	-3.24*	-5.34*	-41.12*	12.48*	-5.00*	9.64*	14.46*
BSHA	-2.20	-1.72	-1.66	-5.55	2.65	-2.15	1.80	2.36
Significance level 5%	-3.45	-2.89	-3.45	-13.70	4.71	-2.89	4.88	6.49

B: Cointegration Rank Statistics (with unrestricted constants)

H_0	$r = 0$	$r \leqslant 1$	$r \leqslant 2$
Trace test	35.04*	16.22*	3.34
95% critical values	29.68	15.41	3.76

C: Estimated Cointegration Vectors

Vector	HKINT	HKVOL	BSHA
Restricted coefficients	1.00	0.00	0.074
Standard errors	(N/A)	(N/A)	(0.014)

χ^2 test of restrictions $\beta 1 = (1\ 0\ *)$

$\chi^2(1) = 0.242$ with $p - value = [0.622]$

Notes: * denotes rejection at the 95% critical values.

Testing cointegration usually requires non-stationary variables of the same order of integration, however, the cointegration relationships could exist when the system has a mix of $I(0)$ and $I(1)$ variables. It is noted that the cointegration ranks increase correspondingly with adding stationary $I(0)$ variables into multivariate model.[7] The examination of whether the three variables are cointegrated is conducted by Johansen's full information maximum likelihood testing procedure. As stated in Granger representation theorem, equation (5.1) has an equivalent vector error correction form which includes both short- and long-run information for $\Delta\mathbf{Y}_t$, $(\Delta = (1 - L)$, the first difference operator), through estimating short term and error-correction term matrices of parameters. Johansen (1988) geniusly estimates the error correction form by conducting two auxiliary regressions. The two auxiliary regressions are estimated by OLS regression, thus generating 'level' residual series and 'difference' residual series respectively. The number of the cointegrating relationships can be obtained by calculating canonical correlations between those two residual series. The main motivation for calculating canonical correlation is to decide how many stationary cointegrating vectors can achieve the highest correlations with the stationary $\Delta\mathbf{Y}_t$ processes.[8]

The Johansen's rank test results of cointegrating relationships among the three variable series are also reported in panel B of Table 5.2. The null hypothesis of non-stationary with no cointegration $(H_0 : r = 0)$ and at most one cointegrating vector $(H_0 : r \leqslant 1)$ are rejected at the 5 percent significant level on the basis of the trace statistic. While the null hypothesis of at most two cointegrating two cointegrating vectors $(H_0 : r \leqslant 2)$ are

[7]That the variables should be integrated of the same order (more than zero) is not a neccessary nor sufficient condition for cointegration. Suppose we have three variabes x_{1t}, x_{2t}, and x_{3t}. The order of integration of the first two variable series is $I(1)$ and cointegrated with a relationship, say, $z_t = x_{1t} + \beta * x_{2t}$. Whereas x_{3t} is stationary. Thus we would have cointegrating ranks equal to 2, then proceed by imposing and testing overidentifying restrictions for the cointegrating structure as $\begin{bmatrix} * & * & 0 \\ 0 & 0 & * \end{bmatrix}^{\mathsf{T}}$ where $*$ denotes free elements.

[8]Banerjee *et al.* (1993) and Hamilton (1994) give detail canonical correlation analysis and its application on testing reduced rank of cointegrating systems.

easily accepted. This leads to the conclusion that there are at most two cointegrating vectors, which is consistent with the footnote 7 analysis. The cointegrating relationship between Hong Kong interest rate and Shanghai stock price is only statistical regularity rather than any well defined economic equilibrium relationships. However, deviations of the Hong Kong interest rate and Shanghai stock price series from the statistical regularity are transitory in nature and are appropriately modeled in the structure of error correction models (Engle and Granger 1987).

5.4 The structural VAR identification and dynamic simulation

The correct specification and identification of the cointegrating relationships and the estimation of the VAR parameters $\widehat{\mathbf{\Pi}}$, $\widehat{\mathbf{\Sigma}}$ (equation 5.1) can be of central importance as the basis for a successful identification of the Structural VAR analysis. Amisano and Giannini (1997) present full detail analysis on the asymptotic distributions of the parameters for Structural VAR with cointegrated variable series. Thus Structural VAR combines the long run equilibrium conditions through the error correction term and the instantaneous correlations estimated by accounting innovation analysis over a long time horizon. (i.e. impulse response analysis and forecast error variance decomposition functions). This allows us to examine directly whether the Shanghai stock prices are affected by the nature of the Hong Kong interest rates and the Hang Seng index trading volumes.

Having estimated the rank of the cointegrating space above, the identification of the cointegrating relationships could be given by imposing zero restrictions on the matrix of long run coefficients $\boldsymbol{\beta}$.[9] Since rank $(\boldsymbol{\beta}) = r$, $\boldsymbol{\beta}$ is exactly identified if we normalize $\boldsymbol{\beta}$ by imposing r restrictions for each row in different columns and additional $r-1$ restrictions

[9]The reduced form VAR $\mathbf{A}(L)\,\mathbf{Y}_t = \epsilon_t$ has an isomorphic error correction form

$$\mathbf{\Psi}(L)\,\mathbf{Y}_t = -\mathbf{\Lambda}(1)\,\mathbf{Y}_{t-k} + \epsilon_t$$

for each row of $\boldsymbol{\beta}$. However, the data generating process of the Hong Kong stock trading volume series is stationary. Then over-identifying restrictions are imposed on the matrix of long run coefficients $\boldsymbol{\beta}$ of

$$\boldsymbol{\beta} = [\mathbf{H}_1\boldsymbol{\varphi}_1, \mathbf{H}_2\boldsymbol{\varphi}_2]$$

$$\mathbf{H}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{H}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{5.4}$$

Where matrices $\mathbf{H}_{1,..,r}$ are restriction hypotheses that need to be tested on each of the r cointegration relations, and $\boldsymbol{\varphi}_{1...r}$ is the vector of parameters to be estimated in each cointegration relation. Such tests of long-run exclusion are testing that the stock index trading volumes have no long-run relationship with the Hong Kong interest rates and Shanghai stock prices, while the two latter variables do not enter the trading volumes cointegration space (i.e. the trading volumes are cointegrated with themselves due to their stationary properties). The estimated cointegrating vector between Hong Kong interest rate and Shanghai stock price is $HKINT + 0.0744_{\ (t=5.20)} \times BSHA = 0$.[10] A likelihood-ratio test against the unrestricted value of $\boldsymbol{\beta}$ is performed. The test results in an asymptotic $\chi^2(1)$ distribution and the over-identification restriction is accepted with $\chi^2(1) = 0.24269$ with $p - value = 0.62227$ significance level.

In an attempt to examine whether and how the Hong Kong interest rates and Hong Kong stock trading volumes affect the Shanghai stock price movements, we have to approach the structural VAR and take into account of the existence of the cointegrating

---

Where $\boldsymbol{\Psi}(L) = \mathbf{I}_n - \sum_{j=1}^{k-1} \boldsymbol{\Psi}_j L^j$ and k is the optimal lag optimal length (i.e. $k = 5$ here), $\boldsymbol{\Psi}_j = \sum_{i=j}^{k} \boldsymbol{\Lambda}_i$. $\boldsymbol{\Lambda}(1)$, containing information on the long run relationships, is an n-dimensional square matrix with rank r. The rank r equal the number of cointegration relationships. It is common to decompose $\boldsymbol{\Lambda}(1)$ into two $n \times r$ matrices with each of rank r, namely, $\boldsymbol{\Lambda}(1) = \alpha\beta^{\intercal}$. α denotes the speed of adjustment to desequilibrium and β is the long run coefficients. Tests of restrictions on this isomophic form can be conducted by imposing restrictions on matrices α and β.

[10] The other cointegrating vector is just the trading volume series itself due to its stationary property.

relationships among the three variables. Nevertheless, the issue of how to identify restrictions on $\mathbf{A}$ and $\mathbf{B}$ matrices in equation (5.3) rises. Since most papers apply the structural VAR models to examine joint behavior of macroeconomic variables such as output, unemployment, prices, and wages etc., the set of identification restrictions that are imposed on matrices $\mathbf{A}$ and $\mathbf{B}$ in those papers can be based on economic theory such as the Keynesian model (e.g., Blanchard (1989), Gali (1992), Giannini *et al.* (1995), and Amisano *et al.* (1995)). There is, however, no specific economic or financial theory concerning the relationship among these three variables, so the identifying restriction is not motivated by any meaningful economic theories. This section tries to consider the following set of restrictions on $\mathbf{A}$ and $\mathbf{B}$ matrices with respect to the relation between reduced-form and structural innovations, equation (5.3).

$$\mathbf{A}\boldsymbol{\epsilon}_t = \mathbf{B}\mathbf{e}_t$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ b_{31} & 0 & b_{33} \end{bmatrix} \tag{5.5}$$

where $\boldsymbol{\epsilon}_t = [\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t}]^\intercal$ and $\mathbf{e}_t = [e_{1t}, e_{2t}, e_{3t}]^\intercal$. The subscripts 1, 2, and 3 for the elements in matrices $\mathbf{A}$, $\mathbf{B}$ and vectors $\boldsymbol{\epsilon}_t$, $\mathbf{e}_t$ represent the Hong Kong short-term interest rate, Hong Kong stock trading volumes and the Shanghai stock prices respectively. The zero identification restrictions criterion on $\mathbf{A}$ and $\mathbf{B}$ matrices are based on assumptions that the Shanghai stock prices do not affect the Hong Kong short-term interest rate and stock trading volumes, because we have spelled out clearly in last chapter that the volatilities of the Shanghai stock market have no effect on those of the Hong Kong stock market after the handover. Alternatively, the expression (5.5) corresponds to the linear system which can be written in the explicit form:

$$\begin{bmatrix} \epsilon_{1t} = b_{11}e_{1t} \\ \\ \epsilon_{2t} + a_{21}\epsilon_{1t} = b_{22}e_{2t} \\ \\ \epsilon_{3t} + a_{31}\epsilon_{1t} + a_{32}\epsilon_{2t} = b_{31}e_{1t} + b_{33}e_{3t} \end{bmatrix}$$

The first row of the linear system states that the reduced form shocks in Hong Kong short term interest rate entirely depend on its own structural innovation e_{1t}. While the reduced form shocks in Hong Kong stock trading volumes, given the interest rate shocks,depend on structural innovation e_{2t}. The last row of the system describes that the reduced form shocks in Shanghai stock prices are allowed to respond to innovations in all variables. With respect to comparison with the restriction perspectives based on Kynesian economic model, it can be argued that these identification restrictions are not immune from arbitrariness. However, they are at least robust to the issue of the ordering identification assumptions which the reduced form VAR suffers. It is, of course, desirable for the restrictions to be consistent with empirical evidence.

Conditional on these identification restrictions, there are seven free elements in **A** and **B** matrices that need to be estimated while there are only six known parameters contained in the reduced form residuals $\widehat{\mathbf{\Sigma}}$ estimated by the reduced form (5.2). Thus, at least one restrictions are need for exactly recovering **A** and **B** matrices. The reduced form VAR with a cointegrated system of variables can be represented in its isomorphic error-correction model whose structural form enables us to generate additional restrictions on **A** matrix.[11] Two restrictions are imposed on the 'structural loadings' $\mathbf{A}\boldsymbol{\alpha}$ as

[11]Suppose that the system can be written as a finite order vector autoregrssion form as equation 1:

$$\mathbf{\Lambda}(L)\mathbf{Y}_t = \epsilon_t, t = 1, ..., T,$$

This equation can then be rewritten in error correction form as

$$-\alpha\beta^{\mathsf{T}}\mathbf{Y}_{t-p} + \Xi(L)\Delta\mathbf{Y}_t = \epsilon_t$$

$$\mathbf{A}\boldsymbol{\alpha} = \boldsymbol{\gamma} = \begin{bmatrix} * & * \\ 0 & * \\ * & 0 \end{bmatrix} \tag{5.6}$$

Where $*$ denotes the free elements. Imposing $\gamma_{21} = \gamma_{32} = 0$ illustrates that the stationary Hong Kong stock trading volumes does not affect the cointegrated relationship between the Hong Kong interest rates and Shanghai stock prices, thus the equilibrium error of the stock trading volumes does not enter the change of Hong Kong short term interest rates and that of Shanghai stock prices. The over-identified structural model so obtained for the **A** and **B** matrices is summarized in Table 5.3.

The estimating results show that all element values in the **A** and **B** matrices are significant from zero except that a_{21} is not significantly different from zero, whose $t-value$ equals to 0.01. The over-identification is then accepted by the LR test (i.e., the over-identification LR test is distributed as chi-squared(1) = 0.95621 with significance level of 0.32814. For the Shanghai stock prices structural equation, the Hong Kong short-term interest rates affect the Shanghai stock prices negatively in both current and long run terms (i.e., the significant coefficient $a_{31} = -5.938$ with $t - value = -4.94$ and $b_{31} = -0.039$ with $t - value = -3.31$). And the Hong Kong stock trading volumes also show a significant effect on the Shanghai stock prices contemporaneously (i.e., the significant coefficient $a_{32} = 0.051$ with $t - value = 5.52$).

Where $\Xi(L) = (1 - L)^{-1} (\mathbf{\Lambda}(L) - \mathbf{\Lambda}(1) L^p)$. Then the AB structual form of the error correction form is

$$-\mathbf{A}\alpha\beta^{\intercal}\mathbf{Y}_{t-p} + \mathbf{A}\mathbf{\Phi}(L)\Delta\mathbf{Y}_t = \mathbf{B}\mathbf{e}_t$$

Let $\gamma = \mathbf{A}\alpha$. In Giannini *et al.* (1995) and Amisano and Giannini (1997) the matrix γ is defined as 'structural loadings' and the elements γ_{ij} represents the effects of the j^{th} equilibrium error on the i^{th} component of $\Delta\mathbf{Y}_t$.

Table 5.3 The Estimation of Structural VAR Parameters

$1997:09--1999:10$

(i) $\mathbf{A}\boldsymbol{\epsilon}_t = \mathbf{B}\mathbf{e}_t$ (The AB-model is over identified)

$$\epsilon_{1t} = \underset{(14.42)}{0.010}\, e_{1t}$$

$$\epsilon_{2t} + \underset{(0.01)}{5.710}\, \epsilon_{1t} = \underset{(14.42)}{0.386}\, e_{2t}$$

$$\epsilon_{3t} - \underset{(-4.94)}{5.938}\, \epsilon_{1t} + \underset{(5.520)}{0.051}\, \epsilon_{2t} = \underset{(-3.310)}{-0.039}\, e_{1t} + \underset{(14.42)}{0.061}\, e_{3t}$$

LR test for one over-identification restrictions:

$\chi^2(1) = 0.95621$ with significance level of 0.32814

The innovation accounting exercises such as impulse response function and the forecast error variance decomposition further can offer insights into the dynamic effects of the structural innovations on the three endogenous variables, which have by far been the most applied ways for researchers. In this analysis, the innovation accounting exercises are utilized to evaluate the extent and nature of how the Shanghai stock prices being influenced by the Hong Kong financial and economical variables. All the information about such dynamic effects can be derived from the structural vector moving average representation (SVMA), which is an equivalent form of the structural VAR.[12]

[12]The MA representation, which is the Wold representation by successive substituting on the lagged variables of $\mathbf{Y}_t$ form the VAR representation (i.e., equation 5.2) when $\mathbf{Y}_t$ is covariance stationary, can offer further insights into the dynamic of the model and is shown as

$$\mathbf{Y}_t = \mathbf{C}(L)\,\boldsymbol{\epsilon}_t$$

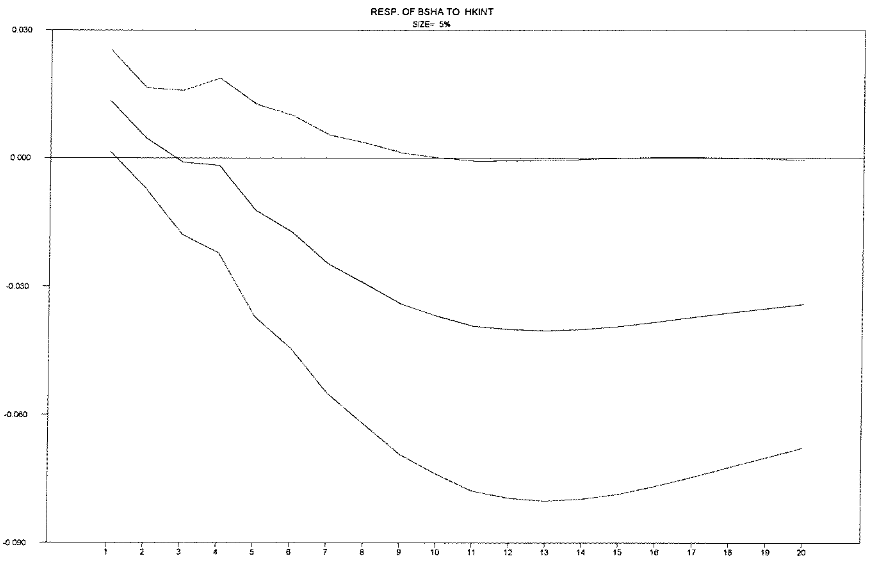

Figure 5-1: Impulse responses of Shangahi share index to Hong Kong interbank rate

Figure $(5-1)$ and figure $(5-2)$ display the estimated structural 20-week impulse responses for the levels of the variables to a standard deviation in each of the structural shocks with two standard error confidence intervals.[13] The dark line in the figures report the point estimates of the response of the levels of each of the variables to a one percentage point perturbation to each of the three shocks. The lighter lines on either side of these point estimates represent the two standard deviation error intervals to justify the precision of the impulse response functions estimates.[14]

Figure $(5-1)$ plots the Hong Kong short term interest rate shocks produce dynamic

Where $\mathbf{C}(L) = \Gamma^{-1}(L)$ and $\mathbf{C}(L) = \mathbf{I} + \mathbf{C}_1 L + \mathbf{C}_1 L^2 + ...$ The variance-covariance matrix of ϵ_t is symmetric and positive definite.

[13]Here only presents the estimated structural impulse response functions for each of the structural shocks on the Shanghai stock prices. Impluse responses for the interest rates and stock volumes in the sample are not illustrated but are available.

[14]Here, the confidence intervals for impulse responses is automatically ploted by the Monte Carlo integration procedure in the computer package RATS. The confidence intervals can also be evaluted by means of standard bootstrap techniques.

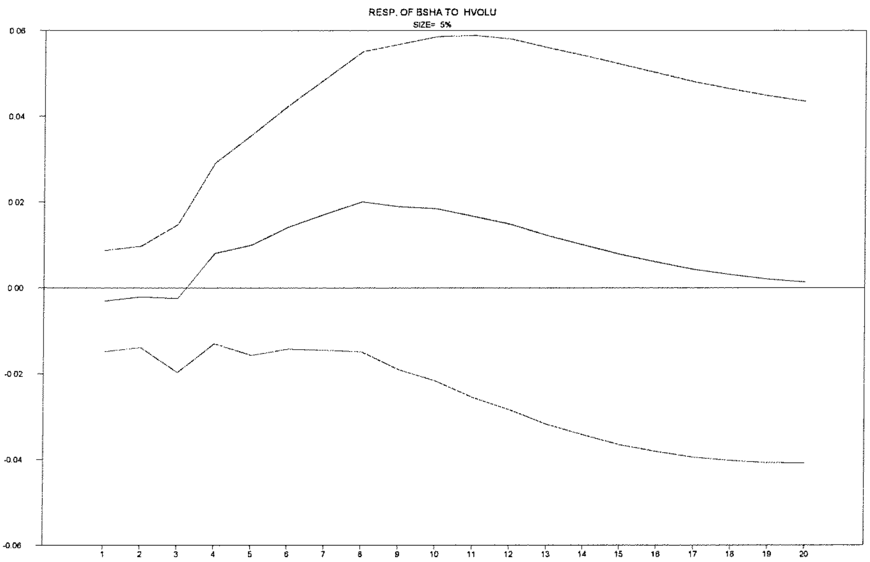

Figure 5-2: Impulse responses of Shanghai stock index to Hong Kong stock volume shock

variables. Table 4 and Figure $(5-3)$ provide the variance decomposition of 5-week, 10-week, and 15-week ahead forecasts of the average amount of the variance of the variance of each variable attributable to each shock. Standard $t - values$ for these estimators are in parentheses. The table indicates that the Hong Kong short term interest rates fully accounts for its own variance, suggesting the interest rates are relatively exogenous for the system. While the interest rates exert substantial influence on the Hong Kong stock trading volumes. At the horizon of 15 weeks, for example, the percentage of forecast error variance of the interest rates attributable to the latter is 4.82 percent.

Table 5.4 FEVD at Selected Horizons

1997 : 09 − −1999 : 10

Forcast error variance decomposition of BSHA			
	BSHA	HKINT	HKVOLU
1 week	0.9516 (56.6)	0.0460 (28.80)	0.0024 (27.41)
3 week	0.9811 (26.5)	0.0172 (5.68)	0.0017 (24.29)
5 week	0.9764 (21.7)	0.0155 (10.76)	0.0080 (33.51)
10 week	0.8890 (7.73)	0.0803 (12.56)	0.0307 (9.46)
15 week	0.8280 (3.49)	0.1427 (8.81)	0.0294 (7.12)

However, our most concerned is how the forecast error variance decomposition of the Shanghai stock prices can be allocated to the Hong Kong market financial variable sources. At 15 weeks more than 14.27 percent of the fluctuations in the Shanghai stock prices are explained by innovations in the Hong Kong interest rates. As the time horiresponses of the Shanghai stock prices, whose signs of responses are consistent with the long-run equilibrium relationships between the variables. All the responses are statistically significant. It shows that positive interest rate shocks usually generate an immediate response in the Shanghai stock prices of about 1.5 percent per week temporarily. Then the interest rate shocks lower the Shanghai stock prices, with the decline bottoming at the thirteenth weeks at about 9 percent. The economic theory suggests that the stock prices equal the expected present value of future net dividend flows. When Hong Kong tightens monetary policy, the Chinese companies listed in the Hong Kong stock market may have difficulty in credit money market, which in turn declines expected future cash flow and raise the discount factor. The decline stock prices of those companies lower the Shanghai stock prices. However, such an effect does not necessarily volatile the Shanghai stock prices. The effects of the Hong Kong interest rates shocks on the Shanghai stock prices last much longer than the effects of other shocks represented from the Hong Kong stock trading volumes. The main reason for the interest rate shocks not dying out in the long run is because the Hong Kong interest rates and the Shanghai stock prices are cointegrated, so the shocks to the system will move the system to a new equilibrium.

The impulse response function for the effect of the Hong Kong stock trading volumes on the Shanghai stock prices is shown in Figure $(5-2)$. The response of the Shanghai stock prices to a 1 percent impulse in the Hong Kong stock trading volumes is positive with a lag of one or two periods. Such responses reach peak at the eighth period and gradually declined to be ceased in the long run, consistent with the stationary properties of the stock volumes. This is because the persistence of responses to shocks reflects the unit root proprieties of the variables. In all, it can be noted that the Hong Kong interest rates have a permanent effect on the Shanghai stock prices, while the response of Shanghai stock prices to other Hong Kong information shocks generally die off after several weeks.

Further evidence on the importance of interest shocks and volume shocks in Hong Kong can be examined by the decomposition of the variance of forecast errors in the

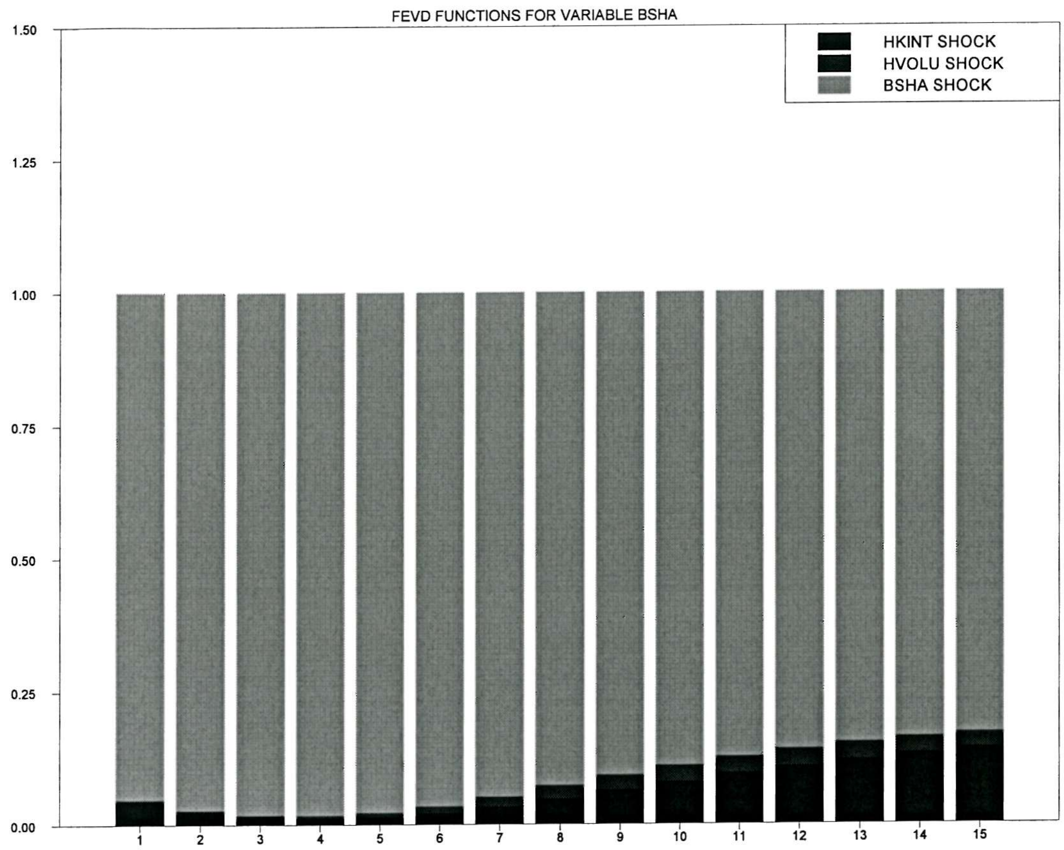

Figure 5-3: Variance decomposition of the Shangahi index error forecast

zons getting longer, the higher proportion of the Shanghai stock price variances can be attributed by the Hong Kong interest rates. The Shanghai stock prices decomposition results are consistent with the earlier results provided by the with-in sample cointegration analysis. Other shocks from the Hong Kong stock market, represented by the Hang Seng index volumes, also contribute significantly in explaining almost 2.94 percent of shocks to the Shanghai stock prices at 15 weeks.

5.5 Conclusions

Based on the empirical results in chapter three and four that there is significant volatility spillover from the Hong Kong stock market to the Shanghai stock market, this chapter attempts to diagnose the potential sources of variables in Hong Kong that may influence the Shanghai stock price movements after the handover. The system is specified and estimated by using the structural vector autoregressive (SVAR) model. The results of the SVAR tests suggest that the Hong Kong short term interest rate is a significant factor that can affect the Shanghai stock price movements. The evidence reported in this chapter have also revealed that the Shanghai stock market is also influenced by other economic and financial variables represented by the Hang Seng trading stock volumes.

It should be noted that the limited sample size, coupled with an inability to get enough the other various financial and economical data, mean that this study might suffer from not including many relevant factors that may influence the formation and effects of the Shanghai stock prices. A higher dimension samples might help to provide a formal general framework in a more structural sense, in which we can identify exactly which factors contribute, and how much they contribute to the Shanghai stock prices.

However, these findings together with the literature in the area of information flow can still provide us a better understanding the role of information in the stock price discovery process. One field we have not touched but merits greater attention is how the derivative market could bring the information to the underlying spot market, thus providing additional aspects of the way how information is represented, how information functions. So in the course of next chapter, we focus on the relationships between the futures market and spots market, ways of representing information, and ways of relating several empirical methodologies.

Chapter 6

An analysis of the relationship between the spot and futures markets: Evidence from the Shanghai bond futures market

6.1 Introduction

The creation of derivative markets is a vital theoretical and empirical issue, and raises many questions on regulation and speculation issues. The derivative market is the point of most direct interaction between the futures market and the underlying spot market, and between the financial market and the economy. There have existed several excellent papers regarding the relationships between the futures market and the spot market. Most of the existing papers focus on the, important and conflicting, issue of whether or not the futures market would destabilize the underlying spot market following the introduction of the futures market. The emphasis of this chapter addresses the question of whether or not the Shanghai bond futures market had fulfilled its functions. It aims at assessing whether it was appropriate for the Chinese government to order the short lived bond futures market to be closed down just because one of China's biggest investment companies and several other small ones went bankrupt as the result of failure speculating on bond futures. The government justified the closure by claiming that the Shanghai bond futures market had destabilized the underlying bond spot market and the stable social economy as well. Although the closure has provoked widespread controversy, the validity of the claims are not, however, well documented. In fact, no research studying the issue of whether the closure has reduced investor's exposure to uncertainties has been presented since the closure. The lack of research in this futures closure is surprising, because very few scenarios like this provide us such an unique opportunity to study the relationship between futures and spots from the closure perspective. Consequently, this chapter tries to provide an empirical examination of this issue and to fill the existing gap in lack of research. Therefore, we view this chapter as complementing to other chapters, in that this chapter attempts to bring additional aspects of the role of information in the financial price discovery processes.

Although previous research can provide some guidance for this issue, it is not an appropriate guide for the short lived Shanghai bond futures market because it lasted only one and half years. Furthermore, little information is known regarding the role the bond futures market played in the Shanghai capital market until now. Information regarding the financial market can not claim to be well established if there has been no study about the dramatic changes that have had great influences on it. The motivation for this chapter is twofold: first, the fact that very little is known about the closure, second, the importance of understanding the relationship between bond futures and spot markets, pertaining directly to decisions concerning the implication of regulation policies, risk management decision and the appropriate design of the futures.

The most important functions of futures market are risk hedging and price discovering (Driskill & MaCafferty(1982), Figlewski (1984), Silber(1985), and Turnovlkey & Campell (1985)). The effectiveness of futures markets in hedging risk is dependent on the their own efficiency (Antoniou & Foster (1994)). In efficient markets, the futures prices should be unbiased predictions of the future spot prices. The theoretical equilibrium relationship between the price of futures and the price of the underlying spot is determined by the cost of carry through no-arbitrage process (Stein(1961), Kawaller et al. (1987), and Serletis & Scowcroft (1991)). A full exposition of the theory and some of its uses is given by Stoll & Whaley (1990), whose reference contains detailed references. In what follows, however, we will examine the nature of the relationships between the bond futures and spots prices on the return levels and volatility levels.

Questions that guided this study included: What is the impact of introducing bond futures on the underlying bond market? Does the short-lived bond futures fulfil its designed functions and improve the efficiency of the spot market? What are mechanisms of volatility linkages in the two markets? Finally, does the outcome of the closure result in a reduction in spot market risks? Undertaking these questions is an emphasis on the identification of the bond futures with spots. It argues that, despite the short lived history of Shanghai bond futures, bond futures can improve the information discovery process and reduce the underlying bond market risks. It further argues that eliminating the derivative instrument may not result in the stability of the Shanghai bond market.

6.2 The empirical examination methodologies

The no-arbitrage condition for the futures prices and the underlying spot prices rests upon the simultaneous equilibrium mechanics of the two prices. However, it is likely that relative movements of both prices will indeed drift away from such mechanics in the short run, and perhaps it is because of this that the futures market is often be blamed for destabilizing the financial market by the regulators. It is, therefore, useful to consider the short and long run relationship for the two price series, together with their implications for the behaviors of the bond futures that had been closed by the policy makers.

The concepts of stationary and non-stationary are the core issues for the study of financial time series. The state of stationary or non-stationary determines parameter estimations and statistical inferences, which is important to analyzing the underlying relationship between bond futures and spot markets. The conventional statistical techniques are inappropriate if the futures prices and bond prices are non-stationary processes. Estimating models in first-differences, however, may not be a satisfactory solution. Because it will only focus on short-run dynamics and remove the important long-run equilibrium information between the two variables if one variable could be predicted by the other. The theory of cointegration, first introduced by Granger (1981) and Engle and Granger (1987), can provide us a superb theoretical basis for estimating the presence of long-run equilibrium and the short-run dynamics for the bond futures and spots price series, even though they may be non-stationary in levels. Thus, the starting point for the analysis is the examination of whether there exists cointegration relationships between the two price series. A necessary pre-condition to estimate whether or not the two variables are cointegrated is to identify the integration order of the underlying variables. The augmented Dickey and Fuller (1979) tests, Phillips and Perron (1988), and Kwiatkowski et al.(1992)-KPSS tests are applied to test the stochastic properties of the data.

6.2.1 The bivariate EGARCH analyses

In addition to understanding the long run relationship between the two prices on level, the understanding of the movements between conditional variance processes for the two price series can provide further insights for the derivative markets. Because the second moments dominate the information flow over the first moments and higher-order variations (Karatzas and Shreve, 1991). Understanding the second moments transmission between bond futures market and spot market is not only important for evaluating direct investment and asset allocation decisions (Bekaert and Harvey (1997)), but also the fundamental factor to find appropriate strategies such as hedging to reduce various market risks, for which the regulators are of most concerned.

The bivariate constant correlation exponential generalized autoregressive conditioning heteroskedasticity (EGARCH) is applied to take into account of the asymmetric volatility transmission. The conditional variance-covariance processes are specified in the following system:

$$\Delta S_t = \mu_s + \alpha_s \widehat{z}_{t-1} + \sum_{i=1}^{n} \theta_{ss} \Delta S_{t-i} + \sum_{i=1}^{n} \theta_{sf} \Delta F_{t-i} + \epsilon_{st}$$

$$\Delta F_t = \mu_f + \alpha_f \widehat{z}_{t-1} + \sum_{i=1}^{n} \theta_{fs} \Delta S_{t-i} + \sum_{i=1}^{n} \theta_{ff} \Delta F_{t-i} + \epsilon_{ft} \tag{6.1}$$

$$\begin{bmatrix} \epsilon_{st} \\ \epsilon_{ft} \end{bmatrix} \mid \mathbf{I}_{t-1} \sim N(\mathbf{0}, \mathbf{H}_t)$$

$$\mathbf{H}_t \equiv \begin{bmatrix} \sigma_{st}^2 & \sigma_{sf,t} \\ \sigma_{sf,t} & \sigma_{ft}^2 \end{bmatrix} \tag{6.2}$$

$$\begin{aligned} \ln \sigma_{st}^2 &= \omega_s + \gamma_s \ln \sigma_{st-1}^2 + \beta_s \left(\frac{\epsilon_{st-1}}{\sqrt{\sigma_{st-1}^2}} \right) + \beta_f \left(\frac{\epsilon_{ft-1}}{\sqrt{\sigma_{ft-1}^2}} \right) \\ &+ \eta_{ss} \left[\left| \frac{\epsilon_{st-1}}{\sqrt{\sigma_{st-1}^2}} \right| - \sqrt{\frac{2}{\pi}} \right] + \eta_{sf} \left[\left| \frac{\epsilon_{ft-1}}{\sqrt{\sigma_{ft-1}^2}} \right| - \sqrt{\frac{2}{\pi}} \right] \end{aligned} \tag{6.3}$$

$$\begin{aligned} \ln \sigma_{ft}^2 &= \omega_f + \gamma_f \ln \sigma_{ft-1}^2 + \beta_f \left(\frac{\epsilon_{ft-1}}{\sqrt{\sigma_{ft-1}^2}} \right) + \beta_s \left(\frac{\epsilon_{st-1}}{\sqrt{\sigma_{st-1}^2}} \right) \\ &+ \eta_{ff} \left[\left| \frac{\epsilon_{ft-1}}{\sqrt{\sigma_{ft-1}^2}} \right| - \sqrt{\frac{2}{\pi}} \right] + \eta_{fs} \left[\left| \frac{\epsilon_{st-1}}{\sqrt{\sigma_{st-1}^2}} \right| - \sqrt{\frac{2}{\pi}} \right] \end{aligned} \tag{6.4}$$

$$\sigma_{sf,t} = \sigma_{fs,t} = \rho \sqrt{\sigma_{st}^2 \sigma_{ft}^2} \tag{6.5}$$

Where $\mathbf{I}_{t-1}$ is the past information set available at time $t-1$, $\mathbf{H}_t$ denotes the 2×2

variance-covariance matrix conditional on past information set at time $t-1$, and ϵ_{st} and ϵ_{ft} represents the vector of disturbances from each equation in (6.4) respectively, so as to avoid the model misspecification for not taking into account the steady-state equilibrium between the futures and spot prices. The distribution of vector disturbances is assume to be normal distributed with zero mean and H_t variance.

The EGARCH model is superior to the symmetric GARCH model.[1] The coefficients restrictions for EGARCH is far less than those of GARCH. The EGARCH model also successfully capture the leverage or asymmetric effects uncovered by Black (1976). The terms $(\frac{\epsilon_{st-1}}{\sqrt{\sigma^2_{st-1}}})$ and $(\frac{\epsilon_{ft-1}}{\sqrt{\sigma^2_{ft-1}}})$ capture the sign effects, in which the positive and negative news have asymmetric influences on the conditional volatility. And the terms $\left[\left|\frac{\epsilon_{st-1}}{\sqrt{\sigma^2_{st-1}}}\right| - \sqrt{\frac{2}{\pi}}\right]$ and $\left[\left|\frac{\epsilon_{ft-1}}{\sqrt{\sigma^2_{ft-1}}}\right| - \sqrt{\frac{2}{\pi}}\right]$ measure the size effects. The advantage of bivariate specification compared to univariate specification is that the bivariate specification can capture the temporal dependencies in the conditional variances and covariances without losing any efficiency for the analysis of the linkage of the bond futures and spot markets. As a result, it can improve the power of testing cross market volatility spillover. The cross market volatility spillover effect between the futures and spot markets are specified by coefficients η_{sf} and η_{fs}. In order to simplify the estimation process, we assume the conditional correlation between spot and futures conditional variances is constant first proposed by Bollerslv (1986).

6.2.2 The non-linear Granger causality analyses

Although we are capable of implementing the direction of causation through the error correction representation, in which the linearity is assumed, such an assumption is often too optimistic and has lower power than nonlinear causality assumption (Brooks, 1998). We ought to investigate the effect that nonlinear Granger causality between the bond futures and the underlying spots is likely to have upon the relationship between the

[1]Nelson (1991) gives full detail descriptions of the advantages of EGARCH over GARCH model.

two prices. And we also ought to compare whether the linear and nonlinear Granger causality generate consistent results. It is therefore appropriate for us first to discuss why the Granger causality may be nonlinear for the two price series and what value is added for the non-linear Granger causality test over the linear Granger causality test.

It is natural to assume that the volatility clustering effect and variations in the rate of information arrival are responsible for the nonlinear structure in financial prices, in which Hiemstra and Jones (1994) give a more detail explanation. The non-linearity perhaps also due to the leverage effect that negative shocks have larger impacts on financial price movements than positive shocks. Furthermore, the potential volatility spillover effect between the futures and spot markets may contribute the nonlinear Granger causality process for the two price series because the volatility linkage is the indicator for the information flow between markets. Then the motivation for this analysis comes from the fact that the non-linear Granger causality approach has higher power than the linear Granger causality approach in deciding whether the Shanghai bond futures can provide more information to its underlying spot market. The non-parametric non-linear Granger causality approach is approached by Hiemstra and Jones (1994) and Brooks (1998).

To describe their procedures, we shall consider two stationary $EGARCH$ filtered futures return residuals, $\epsilon_{ft}/\sqrt{\widehat{\sigma}^2_{ft}}$, and spots return residuals, $\epsilon_{st}/\sqrt{\widehat{\sigma}^2_{st}}$, where $\widehat{\sigma}^2_{ft}$ and $\widehat{\sigma}^2_{st}$ are the conditional variances estimated by the bivariate $EGARCH$ model described in the previous subsection. Define the m length lead series of $\epsilon_{st}/\sqrt{\widehat{\sigma}^2_{st}}$ as $\epsilon_{st}/\sqrt{\widehat{\sigma}^2_{st}}^{\,m}$, and the Lf, Ls lag series of $\epsilon_{ft}/\sqrt{\widehat{\sigma}^2_{ft}}$ and $\epsilon_{st}/\sqrt{\widehat{\sigma}^2_{st}}$ respectively as $\epsilon_{f,t-Lf}/\sqrt{\widehat{\sigma}^2_{f,t-Lf}}^{\,Lf}$ and $\epsilon_{s,t-Ls}/\sqrt{\widehat{\sigma}^2_{s,t-Ls}}^{\,Ls}$. Two different performance measures are used to compare the results for the conditional probability ratios based on different conditions. The first performance measure is the conditional probability ratio, $C1$, such that the probability of the largest change of the all lead and lag spots series and lag futures series being less than any value e within time $t-\tau$ over the probability of the largest change of all past spots and futures series being less than any value e within time $t-\tau$, i.e.

$$C1 = \frac{\Pr\left(A < e, \quad B < e \right)}{\Pr\left(C < e, \quad D < e \right)} \tag{6.6}$$

Where $A = \left\| \epsilon_{s,t-Ls} / \sqrt{\widehat{\sigma}^2_{s,t-Ls}}^{m+Ls} - \epsilon_{s,\tau-Ls} / \sqrt{\widehat{\sigma}^2_{s,\tau-Ls}}^{m+Ls} \right\|$,

$B = \left\| \epsilon_{f,t-Lf} / \sqrt{\widehat{\sigma}^2_{f,t-Lf}}^{Lf} - \epsilon_{f,\tau-Lf} / \sqrt{\widehat{\sigma}^2_{f,\tau-Lf}}^{Ls} \right\|$,

$C = \left\| \epsilon_{s,t-Ls} / \sqrt{\widehat{\sigma}^2_{s,t-Ls}}^{Ls} - \epsilon_{s,\tau-Ls} / \sqrt{\widehat{\sigma}^2_{s,\tau-Ls}}^{Ls} \right\|$,

$D = \left\| \epsilon_{f,t-Lf} / \sqrt{\widehat{\sigma}^2_{f,t-Lf}}^{Lf} - \epsilon_{f,\tau-Lf} / \sqrt{\widehat{\sigma}^2_{f,\tau-Lf}}^{Ls} \right\|$.

The second performance measure is the conditional probability ratio, $C2$, such that the probability of the largest change of the all lead and lag spots series being less than any value e within time $t - \tau$ over the probability of the largest change of all past spots being less than any value e within time $t - \tau$, i.e.

$$C2 = \frac{\Pr\left(\left\| \epsilon_{s,t-Ls} / \sqrt{\widehat{\sigma}^2_{s,t-Ls}}^{m+Ls} - \epsilon_{s,\tau-Ls} / \sqrt{\widehat{\sigma}^2_{s,\tau-Ls}}^{m+Ls} \right\| < e \right)}{\Pr\left(\left\| \epsilon_{s,t-Ls} / \sqrt{\widehat{\sigma}^2_{s,t-Ls}}^{Ls} - \epsilon_{s,\tau-Ls} / \sqrt{\widehat{\sigma}^2_{s,\tau-Ls}}^{Ls} \right\| < e \right)} \tag{6.7}$$

It follows that when testing the adequacy of the null hypothesis that the futures do not Granger cause spots, the difference of the two ratios, often denoted by CS, is asymptotically normally distributed as

$$\sqrt{n} * CS \overset{a}{\sim} N\left(0, \sigma^2\left(m, L_s, L_f, e\right)\right)$$

Where $CS = C1 - C2$. The parameter m is known as the embedding dimensions such that $\{\sigma_t^2\}^m = \{\sigma_{t-m+1}^2, ..., \sigma_t^2\}$.

6.3 The data and preliminary statistics description

This section examines the impacts of futures listings and delistings on the underlying spot government bonds. The empirical tests are based on daily closing prices. The observations are from October 1993 to May 1995, which was the life time for the bond futures market. The delivery months for the futures contracts are March, June, September, and December. In order to avoid the problems of thin trading, the actively traded nearest contracts are used until the first trading day of the delivery month.

Table 6.1 provides the preliminary statistical properties for the futures and the underlying spot return series, including the mean, variance, measures for skewness and kurtosis, and the normality tests. The skewness and kurtosis measurements indicate both return series are positively skewed and highly leptokurtic compared to normal distribution. The normality test rejects the hypotheses that the daily returns for both series are normally distributed. The values of Ljung-Box (LB) Q statistics for 6 lags of the autocorrelation function are significant at any reasonable confidence level, indicating the presence of serial correlation in futures and spots returns. The $ARCH\,(4)$ test statistics reveal the presence of volatility heteroskedasticity for the two return series.

Table 6.1 Preliminary statistics for bond futures & spot returns

Series	Mean	Std	Skew	Kurtosis	Normality test	$Q(6)$	$ARCH(4)$
ln(S_t/S_{t-1})	0.000	0.003	1.418	8.460	106	77.2	54.6
ln(F_t/F_{t-1})	0.001	0.003	2.025	7.469	178	34.9	27.0

Notes: The normality test is a chi-squared distribution with two degrees of freedom with critical value at the 95 percent confidence is 5.99. The Q(6) is a test for serial correlation with six lags. The critical value at the 95 percent confidence is 12.6. ARCH(4) test is the ARCH test with four lags for the own squared returns. The critical value at the 95 percent confidence is 9.49.

Table 6.2 Unit-root test statistics for bond futures and spot prices

	Levels		First differences		Critical values	
Test statistics	Futures	Spots	Futures	Spots	1%	5%
KPSS tests						
	0.466**	0.852**	0.044	0.049	0.216	0.146
ADF tests						
τ (*without trend*)	-0.972	-0.071	-7.33**	-8.21**	-3.47	-2.87
τ_μ (*with trend*)	-2.81	-2.26	-7.35**	-8.20**	-3.98	-3.42
PP tests						
$Z(t_{\widehat{\alpha}})$	2.539	4.601	-15.9**	-13.9**	-2.58	-1.95
$Z(t_{\alpha^*})$	-0.715	-0.061	-16.2**	-14.9**	-3.44	-2.87
$Z(t_{\widetilde{\alpha}})$	-2.447	-2.045	-16.2**	-14.9**	-3.98	-3.42
$Z(\Phi_1)$	3.675	10.63**	131**	112**	6.43	4.59
$Z(\Phi_2)$	4.227	8.482**	87.3**	74.3**	6.09	4.68
$Z(\Phi_3)$	3.002	2.186	131**	111**	8.27	6.25

Notes: Full descriptions for the six different test statistics: $Z(t_{\widehat{\alpha}})$, $Z(t_{\alpha^*})$, $Z(t_{\widetilde{\alpha}})$, $Z(\Phi_1)$, $Z(\Phi_2)$, and $Z(\Phi_3)$ of Phillip-Perron tests can be found in Baillie and Bollerslev (1989). The truncation lag parameter l used for the Phillip-Perron tests is used a window choice of w(s,l)=1-[s/(1+l)] where the order is the highest significant lag from either the autocorrelation or partial autocorrelation function of the first-difference series. We use a truncation of lag of four. The KPSS tests, contrary to the former two tests, examine the null hypothesis of stationary against the alternative of non-stationary, which is rejected if the value of KPSS statistics is greater than the critical values.. The asterisk ** indicates rejection of null hypothesis of 99 percent level.

Table 6.2 reports the results of unit root properties of the bond futures and spots prices. By examining the conventional Augmented Dickey-Fuller and Phillips-Perron test statistics, the null hypothesis of a unit root for the bond futures and spot prices cannot be rejected. While the trend stationary null can be rejected by the $KPSS$ tests, confirming both series are non-stationary processes on levels. The $Z\left(\Phi_1\right)$ statistics reject the random walk without a drift for the spots series. We also conduct the same testing procedures in the first differences of the two series. The null hypothesis for both ADF tests and PP tests are rejected, while the null for KPSS is not rejected when the unit root tests were applied to the first differences of each variable, indicating that each variable is in fact integrated of the same order one, $I(1)$. While the level of each price series can be arbitrarily large or small and the tendency for them to revert to their means is low, each price series are difference stationary. Therefore, we will test whether there exist a stationary linear combination of the two price series being integrated of order zero.

6.4 Empirical results

6.4.1 Cointegration and ECM analysis

After the determination of the unit root properties for the bond futures prices and spot prices, the Johansen and Juselius (1990)'s multivariate cointegration test is applied to uncover whether the existence stochastic trend in both price series have a long-run equilibrium relationship. Next, we apply the error correction model (ECM) to examine whether changes in futures prices could affect those of spot prices, and vice versa. The Johansen and Juseliue's multivariate cointegration procedure is sensitive to the specification of the lag length of VAR[2]. To determine the appropriate lag length, we apply the likelihood ratios on VAR lengths from 8 to 1 lags. The chosen lag length for each series is two lags.

[2]The issue of setting the appropriate lag length is important. While large value for lag length may result in overparameterisition which affects the estimation of cointegration rank, small values may distort the size of the tests.(Cheung and Lai (1993)).

The results of cointegration estimation based on the trace test and the maximum eigenvalue test are reported in Table 6.3. Two null hypotheses are tested from no cointegration (i.e., $r = 0$) to one cointegrating vector (i.e., $r = 1$), which are presented in the first column. The second column is the eigenvalues ordered from high value to low value. The maximum eigenvalue and trace test statistics are in the third and fourth column respectively. The 5 percent critical values are also presented for each test individually. The null hypothesis is that the two variables are not cointegrated, which is equivalent to testing that the number of cointegrating vector is zero. This hypothesis is rejected by sufficiently large values of both maximum eigen value (λ-Max) and trace test statistics. While the next null hypothesis of $r \leqslant 1$ can not be rejected at the 5% significant level, indicating the presence of one cointegrating vector for both series.

Table 6.3 Multivariate cointegration tests for bond futures and spot prices

Null hypothesis	$\widehat{\lambda}_i$	Test statistics		Test statistics	
of rank r		$\lambda - Max$	95%	Trace	95%
$r = 0$	0.062	22.5*	11.4	23.4*	12.5
$r \leqslant 1$	0.002	0.85	3.8	0.85	3.8
		Cointegrating vectors: $(1, -\beta)'$			
$\widehat{\beta}$		1.00			
LR-test		Chi^2(1)=1.20			
p-value		[0.2729]			

Notes: λ-Max denotes the maximum eigenvalue test (r vs. r+1). Trace denotes the trace test (r vs. no restrictions). The Likelihood Ratio test statistics for the restriction on the cointegrating vector of (1,-1), which is chi-square distributed with one degree of freedom. The asterisk * indicates rejection at the 5% significant level.

The identification of one cointegration relationship suggests that even though both series contain stochastic trends (i.e. non-stationary) individually, and may drift apart in the short-run, their linear combination is stationary and will drive them converge to their long-run equilibrium relationship[3]. The existence of cointegration for the bond futures and spot series illustrates that the one variable series contain information to forecast the other variable. Thus the once existed bond futures in the Shanghai bond market could protect investors against the risks of spot prices just as any other futures are designed to do so. Having determined the existence of cointegration between them, however, we need to uncover more about their long term equilibrium relationship.

With the bond futures being an unbiased predictor of the future its underlying bond prices, there may exist an unique cointegrating vector between the two prices, say **(1, -1)**. If so, their relative price movements could provide investors hedging instruments[4]. Therefore, Table 6.3 also provides the results of cointegration restriction tests that the estimated $\widehat{\beta}$ in cointegrating vector should be **-1.** The restriction that is conducted by the Likelihood Ratio tests which are distributed as chi-square with one degree of freedom[5]. The imposing restriction can be accepted since the likelihood ratio test statistics is estimated at 1.20 ($\chi^2(1)$ distribution) and its p-value is 0.27. So the result suggests that the imposing cointegrating relationship **(1, -1)** support the necessary existence of the futures market.

Conditional on the long-run cointegration relationship between the two time series, it is now possible to estimate the VECM with the error-correction term and lagged values of the first differences of each series explicitly included. The OLS estimation for the error correction parameters and the parameters on the lagged variables are presented in Table 6.4.

[3]Brenner and Kroner (1995) present full theoretical explainations how some of the spot and futures prices should be cointegrated with cointegrating vector (1, -1).

[4]Hedging is the process by which one attempts to elimnate the risks associated with the postitions in the spot bond contracts.

[5]We follow Johansen and Juselius (1990) to test restriction on β, with null hypothesis being $\beta = H\phi$. This restriction test is tested by a likelihood ratio test.

Table 6.4 The OLS estimated VECM for bond futures and spot prices

$\Delta S_t = \alpha_s \widehat{z}_{t-1} + \theta_{ss}\Delta S_{t-1} + \theta_{sf}\Delta F_{t-1} + \epsilon_{st}$			
Variables	Coefficients	Standard deviation	t-value
$\widehat{Z}_{t-1}$	0.022*	0.005	4.307
ΔS_{t-1}	0.086	0.069	1.248
ΔF_{t-1}	0.259*	0.076	3.434
$\Delta F_t = \alpha_f \widehat{z}_{t-1} + \theta_{fs}\Delta S_{t-1} + \theta_{ff}\Delta F_{t-1} + \epsilon_{ft}$			
Variables	Coefficient	Standard deviation	t-value
$\widehat{Z}_{t-1}$	0.002	0.005	0.614
ΔS_{t-1}	0.101	0.061	1.660
ΔF_{t-1}	0.350*	0.066	5.269

Notes: The asterisk * indicates that the statistic is significant at the 5% level, for which the null hypotheses of no linear causality is rejected.

The error correction coefficients α_i $(i = s, f)$ is only statistically significant for bond spot series, but not for the bond futures series, indicating that only the current spot prices respond to the previous period's deviation from long term equilibrium, and all the adjustment between the futures and spots are subject to the adjustment of spot prices. This suggests that the futures can process public available information more efficiently than the spot prices, which is consistent with the fact that the futures should dominate its underlying asset on the price discovery process. It should be noted that the error correction coefficient for the spot, α_s, is positive. If the changes of the spot prices are higher for the previous period and since the error correction term is negative, then the ΔS_t decline in the long-run to its equilibrium, and vice versa. The direction of lead-lag in short term are determined by the causality coefficients θ_{fs} and θ_{sf}. The significance of θ_{sf} indicates bond spot prices response to last period's futures innovations. While the t-tests to the causality coefficient θ_{fs} is statistically insignificant, suggesting changes in spots prices do not Granger cause price changes in the futures prices. This unidirectional of causation also implies that futures prices can improve the predictability of the underlying bond spot prices. In all, the above evidence show that the presence of the equilibrium long-run relationship between the bond futures and spots prices suggests the bond futures prices had the function of providing a hedge against the risks of the underlying bond spot prices. The changes of futures prices may lead the changes of spot prices, and the past information contained in futures prices can improve forecasts of the spot prices.

6.4.2 Conditional volatility co-movements between spot and futures bond market

The results for the bivariate EGARCH estimation are reported in Table 6.5. The method of estimation to maximize the likelihood function is the BHHH algorithm. In order to improve the initial values, the SIMPLEX algorithm has been applied for the optimization procedure. The coefficients γ_i $(i = s, f)$, measuring the degree of volatility clustering, are 0.9011 and 0.9065 for bond futures and spot lagged conditional variances respectively. The statistically significant volatility clustering coefficients suggest that the second moments for both prices series are quite persistent, however, the their unconditional variances are finite.[6] The signs parameters of β_s and β_f are all significantly positive. This is because the change of bond prices are negative to the change of shocks (i.e., interest rates). As the interest rates increases (i.e., positive innovations), the bond price will decline. the degrees of this asymmetry effects, measured by $\frac{1+\beta_i}{1-\beta_i}$ $(i = s, f)$, show that a positive innovation increases volatility 1.41 times more than a negative shock for the futures prices and 1.19 for the spots prices all else being equal, suggesting that the condi-

[6]The implied half-life persistance shocks, calculated by log (0.5) / log γ_i, are 6.66 days and 7.07 days for bond futures and spots respectively.

tional volatilities for both of them increase more when interest rate rises. The parameters η_{ii} $(ii = ss, ff)$ describing the size effect are also significantly positive. Thus, the positive sign of β_i $(i = s, f)$ reinforces such magnitude effects.

Table 6.5 The maximum likelihood estimation for the bivariate EGARCH model

Model estimations						
Coefficients	Futures		Coefficients	Spots		
ω_f	-1.1869*	(-6.23)	ω_s	-1.1254*	(-6.98)	
γ_f	0.9011*	(57.5)	γ_s	0.9065*	(66.4)	
β_f	0.1686*	(5.14)	β_s	0.0860*	(2.50)	
η_{ff}	0.3841*	(5.92)	η_{ss}	0.1342*	(2.94)	
η_{fs}	-0.0268	(-0.92)	η_{sf}	0.1665*	(2.85)	
ρ	0.48*		(14.4)			
Loglikelihood		4454.2				
Model diagnostics						
$Q(6)$ statistics						
$\widehat{\epsilon}_t/\widehat{\sigma}_t$	5.711			4.434		
$(\widehat{\epsilon}_t/\widehat{\sigma}_t)^2$	2.813			2.862		
$\widehat{\epsilon}_{st}\widehat{\epsilon}_{ft}/\widehat{\sigma}_{st}\widehat{\sigma}_{ft}$			5.451			
Aymmetric tests						
Sign bias test	-0.474			0.150		
Negative size bias test	-1.222			-0.100		
Positive size bias test	0.311			1.183		
Joint test	1.040			0.658		

Notes: All results for the second moments are based on bivariate EGARCH model with the assumption of constant correlation between bond futures and spots. Numbers in parentheses are t-statistics. Q

(6) statistics are the Ljing-Box statistic for standardized residuals, squared standardized residuals, and the cross product of the standardized residuals respectively. They are distributed asymptotically as a chi-square distribution with 6 degrees of freedom. For more details of the asymmetric tests see Engle and Ng (1993). The asterisk * denotes significance at the 5% level.

Graphics $(6-1)$ and $(6-2)$ present the news impact surfaces on conditional volatility for the futures and spots. The graphs are drawn in three-dimensional perspectives, with the value of the conditional volatility as height values over grid points which are specified by the x-coordinates and y-coordinates from the vector futures and spots innovations. While the points from 0 to *200* of x- and y-coordinates represent innovations values from maximum negative to zero, the points from *200* to *400* of x- and y-coordinates represent innovations values from zero to maximum positive. Both curves are quadratic functions centered on $\epsilon_{i,t-1} = 0$ $(i = s, f)$. Larger magnitude shocks (both negative and positive) for either futures and spots drive their own conditional volatilities change larger than small magnitude shocks. And both figures clearly show the news impact curves are not symmetric: the positive shocks from either futures or spots have greater impact on its own volatility than the negative shocks of the same size.

The strong existence of asymmetric effects to its own past innovations also cause the cross series information flow asymmetric. The volatility flow effects are determined by parameters η_{ij} $(i, j = s, f$ and $i \neq j)$. The volatility spillover coefficients η_{fs}, measuring the direction of information flow from spots to futures, is statistically insignificant. This implies that the conditional variances in futures series are not influenced by innovations originated from the spots prices. The conditional variances of futures are only subject to their own past shocks, which is similar to the first moments described above. On the contrary, the coefficient η_{sf}, measuring the volatility spillover from futures prices to the spots, is significant, which is also similar to the first moments described above. Further, the futures positive innovations can induce 1.40 times more volatility than negative innovations for the spots conditional volatility.

Again, such results can also be confirmed by comparing Figure 6-1 with Figure 6-

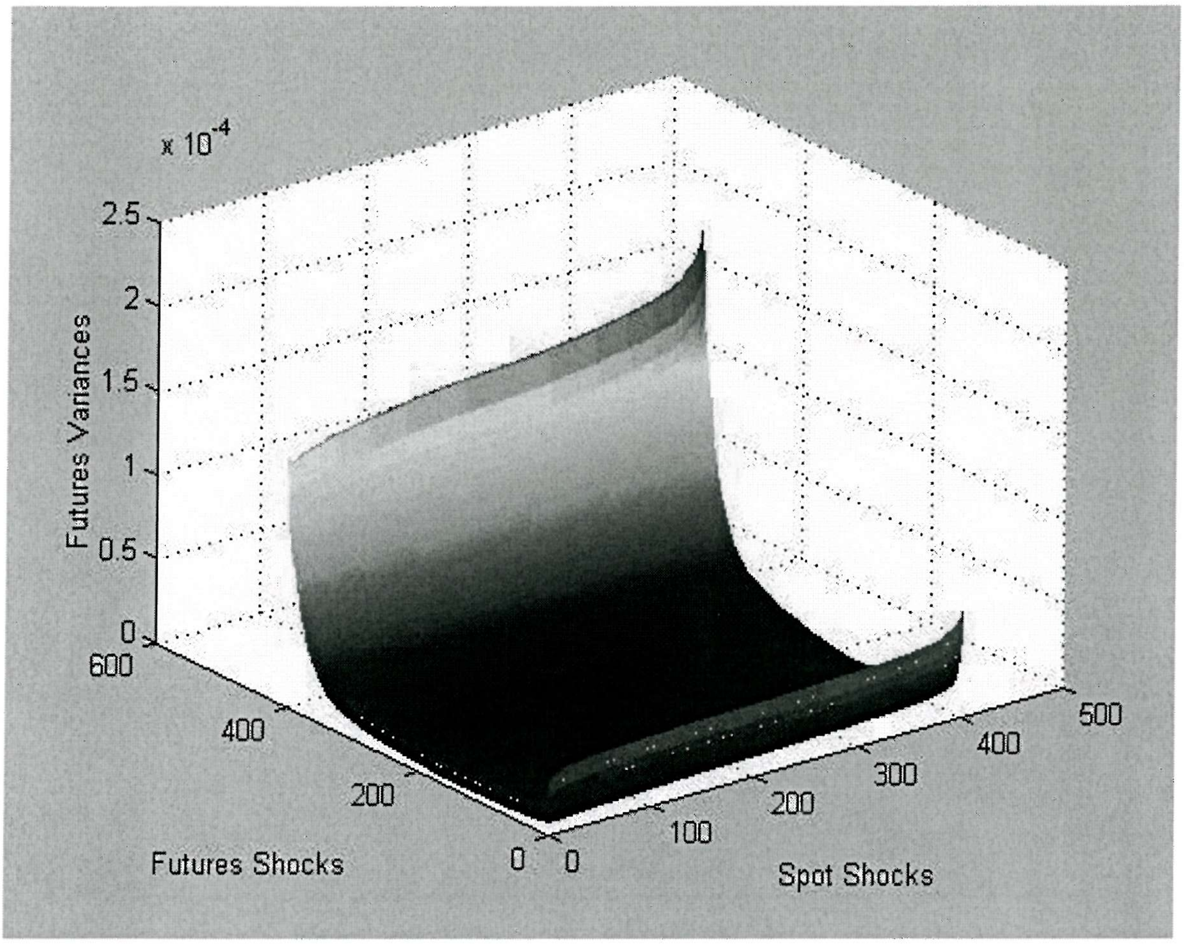

Figure 6-1: Three-dimensional plot showing the bond futures variance as the function of its own shocks and the shocks from the bond market. When comparing with figure $(6-2)$, it can be found that the shocks from the bond market can hardly influence the futures variance.

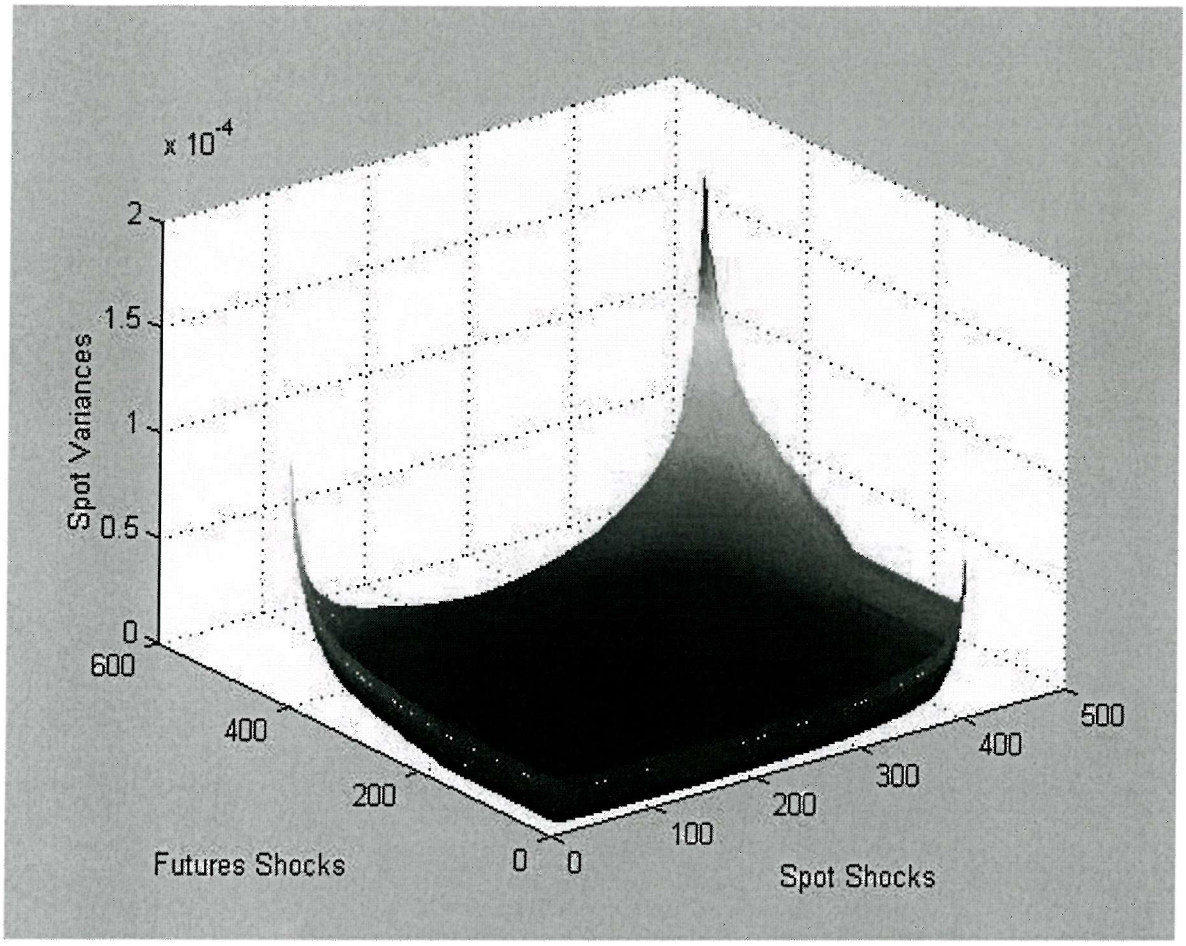

Figure 6-2: Three-dimensional plot showing the bond variance as the function of its own shocks and the shocks from the bond futures. We note that the bond volatility is not only curved to its own shocks but also asymmetrically curved the futures shocks.

2. In figure 1, when the bond spot shocks are moving from centre, the varying spots innovations can hardly influence the futures conditional variances. Whereas the news surface of the spots volatility is bowl-shaped in figure 2, indicating the spots volatility is not only curved to its own past shocks but also asymmetrically curved to futures past shocks of either sign: larger positive futures shocks increase much bigger spots volatility than negative shocks. The strong existed phenomenon of the volatility spillover from the futures prices to the bond prices suggests that the futures could improve the information revealing process for the spot specification. The Ljung-Box statistics for the estimated standardized residuals $(\widehat{\epsilon}_t/\widehat{\sigma}_t)$, squared standardized residuals $(\widehat{\epsilon}_t/\widehat{\sigma}_t)^2$, and cross product residuals $(\widehat{\epsilon}_{st}\widehat{\epsilon}_{ft}/\widehat{\sigma}_{st}\widehat{\sigma}_{ft})$are all insignificant for both futures and spots prices, suggesting no remaining dependence in the estimated standardized residuals nor second moments information existing in squared standardized residuals and cross product residuals. The asymmetric tests also support the adequacy of the EGARCH model specifications.

6.4.3 An non-linear view of Granger Causality and an assessment of the comparative bond volatility performance before and after the closure of the bond futures market

Table 6.6 summarizes the results of non-linear Granger causality between the conditional volatility filtered bond futures and bond spots. Hiemstra and Jones (1994) suggest that suitable values must be chosen for the lead length m, the lag lengths L_s and L_f, and the scale parameter e before implementing the nonlinear Granger causality test. Under their suggestions, these results are estimated by setting $m = 1$, $L_s = L_f = l$, and $e = 1.0 * \sigma$ with $\sigma = 1$. Two major conclusions follow from this table. First, for the whole sample period up to the year 1995, the null hypotheses, ${}_1H_0$, that the bond futures does not Granger cause its underlying spot is rejected at the 5% significance level. The estimates of the difference of the two conditional probability ratio CS are significant for the past information length for the futures and spots up to more than five trading days (This corresponds to $L_s = L_f = 1, \cdots, 5$).

The second main conclusion from Table 6.6 is that the bond spots has little effect on the bond futures for which the null hypothesis ${}_2H_0$ is rejected at the 5% significance level. By applying nonlinear Granger causality tests, the results of the strong effect of the bond futures on the spots but not vice versa are consistent with those conducted by the error correction model. Summarizing these findings, it can be said that inference about the role should be attributed to the bond futures market, that is the futures reflecting new information before the spot market, is invariant with different testing approaches.

Table 6.6 Nonlinear Granger causality test for the conditioanl volatility filtered futures and spots

Number of lags $L_s = L_f = l$	Cs	Test statistic
${}_1H_0$: Bond futures does not Granger cause bond spot		
1	0.0392	3.56*
2	0.0304	3.07*
3	0.0266	2.72*
4	0.0218	2.52*
5	0.0182	2.22*
H_0 : Bond spot does not Granger cause bond futures		
1	0.0050	1.44
2	0.0033	1.09
3	0.0010	0.42
4	-0.0001	-0.05
5	-0.0020	-0.92

Notes: All results for the nonlinear Granger causality are based on Hiemstra and Jones (1994)'s approach with $m = 1$, $L_s = L_f = l$, and $e = 1$. CS denotes the difference between $C1$ and $C2$ calculated by equation (9) and equation (10) respectively. The test statistics in the third column are asymptotically distributed as $N(0,1)$. The asterisk * denotes the null hypothesis of no nonlinear Granger causality is rejected at the 5% significant level.

Given the above empirical findings that the bond futures had positive effects on the spots in that the futures market is a leading information indicator for the spot market and the former can stabilize the latter. Then it is nature and necessary to compare the performance of the bond spot market before and after the cease of trading in the bond futures: how much would the bond volatility have changed if the bond futures were still adopted after the closure? Counter factual simulation can be applied to identify the effects of the changes in the two time regimes before drawing any firm conclusions. The first step of this simulation approach in carrying out the comparison is to identify the volatility estimation procedure for the post futures closure period by using the EGARCH model, then we replace the post-closure residuals $\epsilon_{spot,t}$ by the pre-closure residuals $\epsilon^*_{spot,t}$, and finally, we simulate a new volatility path given by the post-closure volatility estimation procedure.[7]

[7]The EGARCH estimation results of the actural conditional volatilty after the handover is:

$$\sigma_t^2 = \exp\{-1.080 + 0.8984 \ln \sigma_{t-1}^2 + 0.2141 \left[\left| \frac{\epsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} \right| - \sqrt{\frac{2}{\pi}} + 0.7863 \frac{\epsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} \right] \}$$

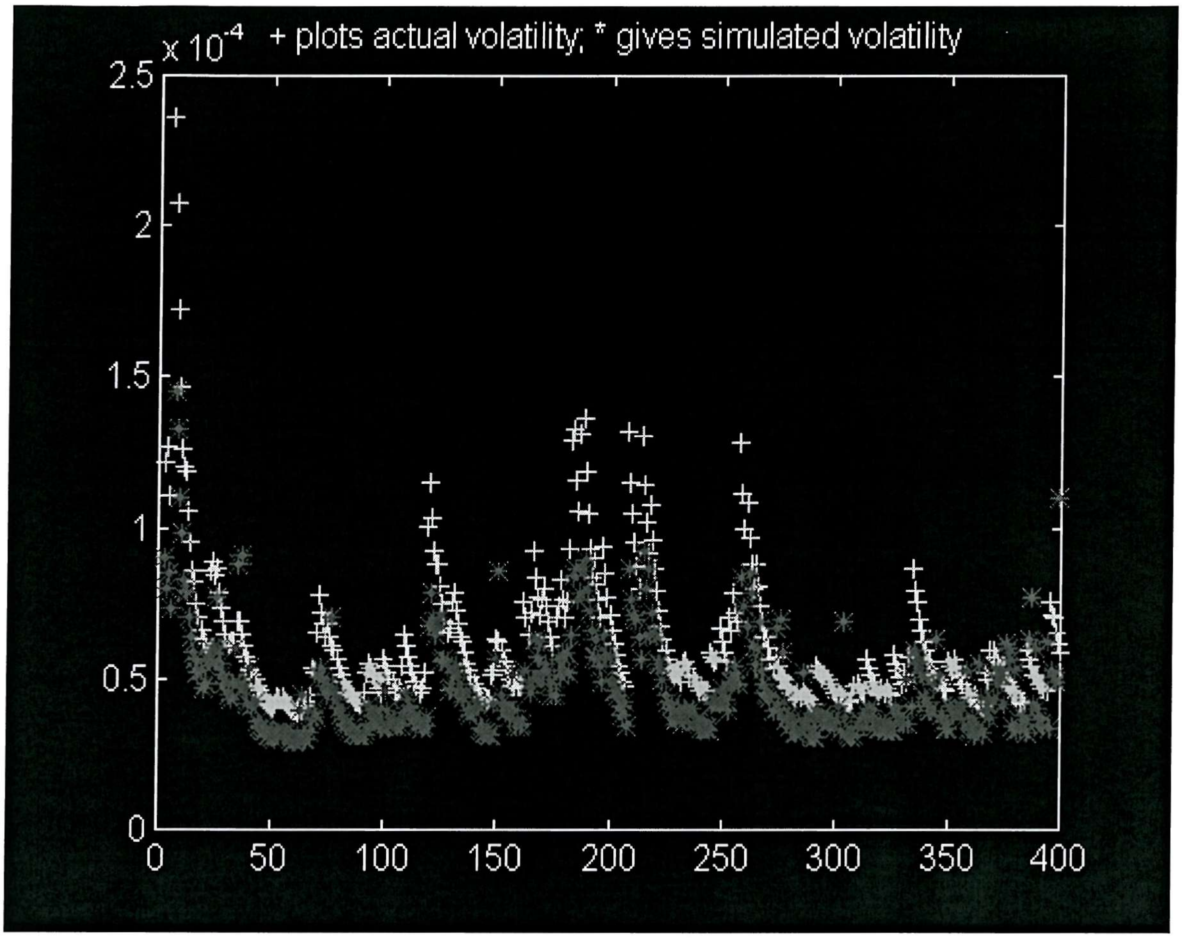

Figure 6-3: This graph presents the results of the counter factual simulations. It is clear from this graph that the magnitudes of the actural volatility are higher than those of the simulated volatility.

The results of the counter factual simulation are presented in Figure 3, which gives the qualifying comparison of the actual volatility to the simulation volatility for the spot bond after the closure of the bond futures. The pluses and open circles represent the actual volatility and the simulated volatility, respectively. It is clear from this graphic that the magnitudes of the actual volatility are higher than those of the simulated volatility. This discrepancy can probably be stem from the closure of the bond futures market. At this junction the questions appears: Does this represent a real increase in the actual volatility, or is it merely an coincidence related to other factors? It is possible that the more volatile bond market is induced by the less liquidity of the bond market after the futures closure.

However, the trading volumes for the bond market do no change significantly during the sample period. The bond investors still stay in the bond market. The reason that closing the bond futures market increased the volatility of the underlying bond prices is that the bond prices are closely tied with the futures prices through the no-arbitrage mechanism. It is conceivable that the suspension of the futures market makes such mechanism disappear, and the underlying bond prices could be too volatile to sustain the theoretical value when confronted with shocks unless such mechanism restored. This line of reasoning is consistent with the result of figure 3.

6.5 Conclusions

This paper is concerned with the relative effects of the closing Shanghai bond futures market on its underlying bond market. In order to improve the reliability of the argument of the closure, numerous measurements have been applied. It had been expected by China policy-makers that the closure of the bond futures market, having had few similar examples, might lead to the bond market being less volatile. However, this does not seem to be the case.

One of the issues which initially prompted this research concerns whether or not the futures market fullfil its functions. Because there is presumably a direct relationship between the futures market and the underlying bond market, the present findings provide some information concerning this original issue. For instance, the bond futures prices are cointegrated with the bond spot prices and the futures prices are an unbiased predictor of the bond prices. These findings suggest that there exists a long equilibrium relationship between the two prices and they would converge to such a relationship in the long term through the no-arbitrage mechanism. The vector error correction model (VECM) further confirms that the bond futures prices can improve the information efficiency for the bond prices. This finding is also consistent with the results conducted by the non-linear Granger causality test. The observation of interests from the second moments examinations show that the bond futures prices could improve the information discovery process for the underlying bond prices. Even under normal conditions, when futures is used as a hedging instrument, there should be a risk reduction.

There is, of course, the possibility of using the futures as a speculation instrument. However, the existence of the long term equilibrium and the volatility spillover mechanism would make bond futures prices and spot prices move together. In addition, the counter-factual simulation shows that the volatility of the bond prices does not become less volatile after the closure of the futures market. In conclusion, these empirical tests show that the Shanghai bond futures market had fulfilled its functions. The bond futures should not be directly responsible for the bankruptcy of those investment companies, which, unfortunately, results in the bond futures market being closed by the government.

Chapter 7

Conclusions to the research

7.1 Overviews of the research

This chapter summaries the major contributions, implications of this research and suggests future research directions. In this thesis, we have touched on two main directions for information integration of the Shanghai financial market by answering four research questions raised in the introduction chapter. They are the impacts of the Hong Kong stock market on the Shanghai stock market and the role of the Shanghai bond futures market for the underlying spot market, and for each one we have discussed the potential contributions of policy concerns.

It is unprecedented that a nation could develop so rapidly like Shanghai from having an obscure stock market only a few years ago to having control of one of the major financial centres in the world — the Hong Kong stock market. So the first question in chapter 1 is whether the Hong Kong stock market has any adverse impacts on the Shanghai stock market after the handover. We have demonstrated that the Shanghai stock market become less volatile after the information from the Hong Kong stock market being integrated with the Shanghai stock prices. Our results are consistent with other empirical findings that the volatility in emerging stock markets declines after market liberalizations.

The information integration is usually thought to be accomplished by taking into account of the US stock market. To aid the completeness of the first research question, the second question is raised in chapter 1, which is to tackle the issue of the US stock market information dominance. Mainly by applying the concept of co-persistence in variances, there is no direct relationship between the US stock market and the Shanghai stock market. Thus, the examination of the US stock market influence on the Shanghai and Hong Kong stock markets proves the validity of the volatility spillover between the latter two stock markets.

In addition to giving detailed insights into the volatility spillover between the Hong Kong and Shanghai stock markets, we answer the third question by examining the potential financial and economic variables which are responsible for two stock markets information integration. Weekly data is collected by conducting a through the structural VAR test. The main finding of the third question study is that the Hong Kong interest rates may be the main factor that influences the Shanghai stock price movements.

The final raised question is with respect to the role of the Shanghai bond futures market for the underlying spot market. We showed the existence of strong long run equilibrium relationship between the bond futures market and its underlying bond market. Furthermore, the futures market could non-linearly Granger cause the underlying spot market, suggesting adequate information could be incorporated into the spot prices as the result of the presence of the futures market. We also examined the influence of the closure of the bond futures market on the spot market. In addition, we also showed the volatility spillover effects from the futures market to the underlying spot market, as the examination of the conditional volatility process can reveal the traces of information movement. As shown elsewhere, the results of this study seem to point to the conclusion that the consequence of the information flow on which the conditional volatility is based can improve financial price discovery processes.

7.2 Contributions of the research

As far as the author is aware, this research is the first to undertake the study of the issue of the handover and the closure effect of the bond futures market in Shanghai, enlarging our understanding of the information influences on financial prices. What are the contributions of the demonstrations of this research? The main arguments against the openness of the emerging financial markets and the introduction of derivative markets have been that the emerging stock markets are insufficiently developed to immune shocks that will be posed. And these un-examined claims have led to the reversal of key policies, notably the prohibition against foreign investors and the abolition of the derivative markets (for example, the Shanghai bond futures market).

Thus the regulation development in the Shanghai stock market has followed the similar and repeating patterns: the Shanghai stock exchange produces rules to encourage foreign investors and introduces derivative markets, and the policy agency directs the regulators to alter the rules. This phenomenon not only means that different bodies have different perceptions towards the development of the Shanghai stock market, but also shows the policy agency's misconceptions about the openness of the stock market and the adoption of the derivative market. Further, there have been concerns that regulation intervention should not undermine the functioning of the financial market structure, or at least not provide policies with no overall improvement in efficiency for the financial markets. This research presented here have clarified the above misconceptions and validated the those concerns by examining the Shanghai stock market and bond futures market.

As the analysis in the previous chapters makes it clear that there is little disagreement on the reduced volatility of the Shanghai stock market after receiving more information from the Hong Kong stock market. Such examinations show that more information flowing into the Shanghai stock market enable the investor to act in an informed and rational manner. The investigation of the unique relationship between the Shanghai and Hong Kong stock markets has added new evidence and supported previous research findings on the importance of inter-market information linkages and interactions.

This research have also brought considerably to the existing knowledge about the information linkages and interactions between stock markets. While many previous work have touched on the issue of information linkages, few have directly addressed the issue of co-persistence in variances applications on the volatility spillover effects. The US stock market factor consideration as noted in chapter four has offered some evidence on areas of application concerning information flow among different financial markets in higher dimensions. Using the concept of co-persistence in variances in this way marks a clear separation from the usual approach. However, this approach is compatible with the usual approach in that it shares the same principles that analyses interactions among stock markets. So the process of applying the concept of co-persistence in variances on the relationships between financial markets is incremental, raising considerations with the application of the concept itself. The research on incorporating the US stock market factor may leave the possibility of further reducing the volatility of the Shanghai stock market by directly allowing the US information to be incorporated into the Shanghai stock prices. Finally, unlike traditional dynamic interactions between spot and futures market analysis we are interested in not only the futures market being as a vehicle for price discovery in the spot market, but also the consequence without futures information considered valuable in improving market efficiency and making investment decisions.

There are still different concerns within the policy agencies on the issue of openness and derivative products towards improving the functioning of the stock markets. Notwithstanding these differences, by using a limited range of well established techniques, the principal unifying thrust of much of this research is attempt to remove the confusion over the markets interactions and to provide a platform on which the policy agency makes policy by helping to improve market efficiency, rather than mechanizing decisions.

7.3 Implications of the research for policy agencies and investors

Since the findings of this research point to the conclusion that, under the usual circumstances, the openness of stock markets and the introduction of derivative markets can help stock price discovery process to be as effective as possible. Therefore, there are several implications for both policy agency and investors.

On the policy agency side, the stock market environment has changed significantly. In particular, the authorities could encounter problems in knowing the sources of variance and causality in different stock market variances that is how conditional variances propagate through economic systems (Granger, Robins, and Engle 1986). In order to avoid policy swinging the authorities from emerging stock markets may benefit from understanding what their policy should aim to achieve, and by what means they should be implemented. This research focused on different stock markets and derivative futures market may shed some light on these concerns with improving the efficiency of policies. The emerging stock market are becoming an integral part of developed markets (see Chan et al. 1992). These changes will reshape the sources of information on which regulators rely in forming regulations. Thus, greater understanding the functions of information (measured by variance) by policy agencies in emerging stock markets may be needed to make more realistic and credible policies.

Indeed, the new sources of information, provided by the Hong Kong stock market, could provide a relatively stable environment that balance out the extremes of stock price oscillations which are common in emerging stock markets. The information and the no-arbitrage mechanism provided by the futures market are also of sufficient to ensure that authorities form appropriate expectations on price movements. These inseparable inter-market effects that could improve stock price discovery processes have led to a far greater emphasis on the openness of stock markets. Of course, this may suffer from adverse effects such as the increased reliance on the stock price movement of the dominated stock markets, i.e., the contagion effects (the issue which is still quite controversial, however.).

On the investors side, understanding the relationship among different markets and its consequences allows investors to find direct application in their investment decisions. Therefore, this research could be particularly beneficial to investors. This research points to the tendency that the unopened Shanghai stock market receives information from the Hong Kong stock market after the handover, which could reduce the volatility of the Shanghai stock market considerably. Then investors can improve their trading strategies rationally and allocate their limited capitals more efficiently by using the information gained from outside the domestic market. Without using information flow both foreign market and domestic market investors may become increasingly and enormously disadvantaged throughout their trading.

Targeting the information flow could also provide investors some scope for controlling their portfolio risks. This may be of particular importance in the Shanghai stock market, where investors are insufficiently experienced in identifying potential risks worthy of consideration due to the limited available information. In order to determine the efficiency of the portfolio decisions, the interaction within the Hong Kong and Shanghai stock markets must be taken into consideration. Thus, this research may raise investors awareness of the impacts of the Hong Kong stock market on the Shanghai stock market as a valuable and important source of information for them.

There is another implication of this research may render to public companies. The knowledge of how information affect stock markets may enable companies to secure more favorable finance. Since more external information may reduce the volatility of stock price movements and increase the liquidity of the stock market, the companies do not have to lower their initial offering prices to investors for the high risks and low liquidity markets. Clearly, understanding more information linkages among different markets may provide additional insight for all relevant parties.

Bibliography

[1] Ahn, S., and Reinsel, G. (1990) Estimation for Partially Nonstationary Autoregressive Models, *Journal of the American Statistical Association*, **85**, 813-823.

[2] Akgiray, V. (1989) Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts, *Journal of Business*, **62**, 55-80.

[3] Amisano, G., and Giannini, C. (1997) Topics in Structural *VAR* Econometrics, Springer Verlag, New York.

[4] Amisano, G., M.Cesura, C., Giannini and Seghelini, M. (1995) The Transmission Mechanism among Italian Interest Rates, working paper, University of Brecia, Italy.

[5] Antoniou, A., and Foster, A. (1994) Short-Term and Long-Term Efficiency in Commodity Spot and Futures Markets, *Financial Markets, Institutions and Instruments*, **3**, 17-35.

[6] Bae, K.H. and Karolyi, G. A. (1994) Good news, Bad News and International Spillovers of Stock Return Volatility Between Japan and the US, *Journal of Pacific-Basin Finance*, **2**, 405-438.

[7] Baillie, R. (1987) Inference in Dynamic Models Containing 'Surprise' Variables, *Journal of Econometrics*, **35**, 101-117.

[8] Baillie, R., and Bollerslev, T. (1989) Common Stochastic Trends in a System of Exchange Rates, *Journal of Finance,* **44**, 167-181.

[9] Baillie, R. and Bollerslev, T. (1990) Multivariate GARCH Processs and Models of Time Varying Risk Premia in Foreign Exhange Markets, *Journal of International Money and Finance*, **6**, 309-324.

[10] Baillie, R., and Bollerslev, T. (1994) Cointegration, Fractional Cointegration, and Exchange Rate Dynamics, *Journal of Finance,* **49**, 737-745.

[11] Banerjee, A., J. Dolado, J. Galbraith, and D. Hendry (1993) Co-integration, Error Correction, and The Econometric Analysis of Non-Stationary Data, Oxford University Press.

[12] Bekaert, G., and Harvey, C. (1997) Emerging Equity Market Volatility, *Journal of financial Economics*, **43**, 29-77.

[13] Berndt, E., Hall, B., Hall, R., and Hausman, J. (1974) Estimation and Inference in Non-linear Structural Models, *Annals of Economic and Social Measurement*, **4**, 653-666.

[14] Bernake, B., and Mihov, I. (1995) Measuring Monetary Policy, NBER Working Paper Series, 5145.

[15] Berry, T., and Howe, K. (1994) Public Information Arrival, *Journal of Finance*, **49**, 1331-1346.

[16] Bierens, H. (1994) Topics in Advanced Econometrics: Estimation, testing, and specification of cross-section and time series models, Camberidge University Press.

[17] Black, F. (1976) Studies in Stock Price Volatility Changes, Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section. *American Statistical Association*, 177-181.

[18] Blanchard, O. (1989) A Traditional Interpretation of Macroeconomic Flucturations, *The American Economic Review*, **79**, 1146-1164.

[19] Blanchard, O., and Quah, D. (1989) The Dynamic Effects of Aggregate Demand and Supply Disturbances, *The American Economic Review*, **79**, 655-673.

[20] Blanchard, O., and Watson, M. (1986) Are Business Cycles All Alike? The American Business Cycle, University of Chicago Press.

[21] Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, **31**, 307-327.

[22] Bollerslev, T., R. Engle, and Wooldridge, J. (1988) A Capital Asset Pricing Model with Time Varying Covariances, *Journal of Political Economy*, **96**, 116-131.

[23] Bollerslev, T. (1990) Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized *ARCH* Approach, *Review of Economics and Statistics*, **72**, 498-505.

[24] Bollerslev, T. and Wooldridge, J. (1992) Quasi-maximum Likelihood Estimation Dynamic Models With Time Varying Covariances, *Econometric Reviews*, **11**, 143-172.

[25] Bollerslev, T., and Engle, R. F. (1993) Common Persistence in Conditional Variances, *Econometrica*, **61**, 167-186.

[26] Bollerslev, T., R. Engle, and Nelson, D. (1994) ARCH Models, in R. Engle and D. MaFadden (eds.) *Handbook of Econometrics*, Vol. IV, Elsevier, Amsterdam.

[27] Bollerslev, T., and Melvin, M. (1994) Bid-Ask Spreads and Volatility in The Foreign Exchange Market, *Journal of International Economics*, **36**, 355-372.

[28] Boswijk, H. (1994) Testing for an Unstable Root in Conditional and Structural Error Correction Models, *Journal of Econometrics*, **63**, 37-60.

[29] Brenner, R., and Kroner, D. (1995) Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets, *Journal of financial and quantitative analysis*, **30**, 23-42.

[30] Brooks, C. (1998) Predicting Stock Index Volatility: Can Market Volume help? *Journal of Forecasting*, **17**, 59-80.

[31] Cai, C. (1994) A Markov Model of Switching-Regime $ARCH$, *Journal of Business and Economic Statistics* **12**, 309-316.

[32] Campbell, J. and Shiller, R. (1988) The Dividend-price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies,* **1**, 195-228.

[33] Campbell, J. and Shiller, R. (1991) Yield Spreads and Interest Rate Movements: A Bird's Eye View, Review of Economic Studies **58**, 495-514.

[34] Campbell, J. (1993) Intertemporal Asset Pricing Without Consumption Data, *American Economic Review*

[35] Campbell, J. S. Grossman, and Wang, J. (1993) Trading Volumes and Serial Correlation in Stock Returns, *Quarterly Journal of Economics*, **108**, 905-939.

[36] Canova, F. (1995) The Economics of VAR Models, Macroeconometrics: Developments, Tensions, and Prospects, Kluwer Academic Publishers.

[37] Caporale, G., and Pittis, N. (1997) Causality and Forecasting in Incomplete Systems, *Journal of Forecasting*, **16**, 425-437.

[38] Chan, K., K. Chan, and Karilyi, G. (1991) Intraday Volatility in the Stock Index and Stock Index Futures Markets, *Review of Financial Studies*, **4**, 657-684.

[39] Chan, K., and Karolyi, G. (1991) The Volatility of the Japanese Stock Market: Evidence from 1977-1990, *Japanese financial market research,* Elsevier Science Publishers, Amsterdam.

[40] Chen, N. (1991) Financial Investment Opportunities and The Macroeonomy, Journal of Finance, **46**. 529-554.

[41] Cheung, Y., and Lai, K. (1997) Finite-Sample Sizes of Johansen's Likelihood Ratio Test for Cointegration, *Oxford Bulletin of Economics and Statistics*, **55**, 313-328.

[42] Cheung, Y., J. He, and Ng, L. (1995) Common Predictable Componets in Regional Stock Markets, *Journal of Business and Economic Statistics,* **15**, 35-42.

[43] Christie, A. (1982) The Stochastic Behavior of Common Stock Variance: Value Leverage and Interest Rate Effects, *Journal of Financial Economics*, **10**, 407-32.

[44] Cochrane, J. (1994) Shocks, *Carnegie-Rochester Confernce Series on Public Policy,* **41**, 295-364.

[45] Cooley, T, and LeRoy, S. (1985) Atheoretical Macroeconomics: A Critique, *Journal of Monetary Economics*, 293-308.

[46] Conrad, J., M. Gultekin, and Kaul, G. (1991) Asymmetric Predictability of Conditional Variances, *Review of Financial Studies*, **4**, 579-622.

[47] De Santis, G., and Improhoroglu, S. (1997) Stock Returns and Volatility in Emerging Financial Markets, *Journal of International Money and Finance*, **16**, 561-79.

[48] Dickey, D. A., and Fuller, W. A. (1979) Distribution of the Estimators for Autoregressive Time Series with Unit Root, *Journal of the American Statistical Association*, **74,** 427-431.

[49] Diebold, F. (1986) Modeling the Persistence of Conditional Variances: Comment, *Econometric Reviews,* **5**, 51-56.

[50] Diebold, F., J. Gardeazabal, and Yilmaz, K. (1994) On Cointegration and Exchange Rate Dynamics, *Journal of Finance,* **49**, 727-735.

[51] Diebold, F., and Lopez, J. (1994) Modelling Volatility Dynamics, in Kevin D. Hoover (ed.), *Macroeconometrics: Developments, Tensions, and Prospects*, Kluwer Academic Publishers.

[52] Domowitz, I., Glen, J., and Madhavan, A. (1997) International Cross-listing and Order Flow Migration: Evidence from an Emerging Market, *Journal of Finance*,

[53] Driskill, R., and McCafferty, S. (1982): Spot and Forward Rates in a Stochastic Model of the Foreign Exchange Market, *Journal of International Economics*, **12**, 313-331.

[54] Duffie (1996) Dynamic Asset Pricing Theory, Princeton University Press.

[55] Ederington, L., and Lee, J. (1993) How Markets Process Information: News Release and Volatility, *Journal of Finance*, **48**, 1161-1191.

[56] Ederington, L., and Lee, J. (1996) The Creation and Resolution of Market Uncertainty: The Impact of Information Releases on Implied Volatility, *Journal of Financial and Quantitative Analysis*, **31**, 513-519.

[57] Engle, R. F., (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation, *Econometrica*, **50**, 987-1008.

[58] Engle, R. F., D. Hendry and Richard, J. (1983) Exogeneity, *Econometrica*, **51**, 277-304.

[59] Engle, R. F., and Bollerslev, T. (1986) Modelling Persistence of Conditional Variances, *Economic Review* **5**, 1-50.

[60] Engle, R. F., and Granger, C. W. J. (1987) Co-Integration and Error Correction: Representation, Estimaton and Testing, *Econometrica*, **55**, 251-276.

[61] Engle, R. F., and Yoo, B. (1987) Forecasting and Testing in Co-integrated Systems, *Journal of Econometrics*, **35**, 143-159.

[62] Engle, R. F., V. Ng, and Rothschild, M. (1990) Asset Pricing with a Factor$ARCH$ Covariance Structure: Empirical Estimates for Treasury Bills, *Journal of Econometrics*, 45, 213-238.

[63] Engle, R. F., Ito, T. and Lin, W. (1990) Meteor Shower of Heat Waves, Heteroskedastci Intra-daily Volatility in the Foreign Exchange Market, *Econometrica*, **55**, 525-542.

[64] Engle, R. F., and Ng, V. (1993) Measuring and Testing the Impact of News on Volatility, *Journal of Finance*, **48**, 1749-778.

[65] Engle, R. F., and Kroner, K. (1995) Multivariate Simultaneous Gerneralized ARCH, *Econometric Theory* **11**, 122-150.

[66] Fama, E. (1965) The Behaviror of Stock Market Prices, *Journal of Business*, **38**, 34-105.

[67] Fama, E., and French, D. (1988) Dividend yields and expected stock returns, *Journal of Financial Economics,* **22**, 3-25.

[68] Figlewski, S. (1984) Hedging Performance and Basis Risk in Stock Index Futures, *Journal of Finance*, **39**, 657-669.

[69] French, K., and Roll, R. (1986) Stock Return Variances-The Arrival of Information and The Reaction of Traders, *Journal of Financial Economics,* **17**, 5-26.

[70] French, K., G. Schwert and Stambaugh, R. (1987) Expected Stock Returns and Volatility, *Journal of Financial Economics*, **19**, 3-30.

[71] Fuller, W. (1996) Introduction to Statistical Time Series, Willey Series in Probability and Statistics.

[72] Gallant, A. P. Rossi, and Tauchen, G. (1992) Stock Price and Volume, *Review of Financial Studies*, **5**, 199-242.

[73] Gali, J. (1992) How Well does the $IS-LM$ Model Fit Postwar US Data, *Quarterly Journal of Economics*, 709-738.

[74] Garolyi, G., (1995) A Multivariate *GARCH* Model of International Transmission of Stock Returns and Volatility: The Case of the United States and Canada, *Journal of Business & Economic Statistics*, **13**, 11-25.

[75] Garratt, A., K. Lee, M. Peasran, and Shin, Y. (1999) A Structural Cointegrating *VAR* Approach to Macroeconometric Modelling, in S. Holly and M. Weale (eds.) *Econometric Modelling; Technique and Applications,* Cambridge University Press.

[76] Geweke, J. (1984) Inference and Causality in Economic Time Series Models, in Z. Griliches and M. Intriligator (eds.) *Handbook of Econometrics* Vol. II, Elsevier Science.

[77] Giannini, C., Lanzarotti, A., and Seghelini, M. (1995) A Traditional Interpretation of Macroeconomic Fluctuations: The Case of Italy, *European Journal of Economy,* **11**, 131-155.

[78] Glosten, L., Jagannathan, R., and Runkle, D. (1993) On The Relation Between The Expected Value and The Volatility of The Nomial Excess Return on Stocks, *Journal of Finance*, **48**, 1779-801.

[79] Granger, C. W. J. (1969) Investigating Causal Relations by Econometric Models and Cross-spectral Methods, *Econometrica*, **37**, 424-439.

[80] Granger, C. W. J. (1981) Some Properties of Time Series Data and Their Use in Econometric Model Specificating, *Journal of Econometrics*, **26**, 1045-1066.

[81] Granger, C. W. J. (1986) Developments in the Study of Cointegrated Economic Variables, *Oxford Bulletin of Economics and Statistics*, **48**, 213-228.

[82] Granger, C. W. J., Robins, R., and Engle, R. (1986) Wholesale and Retail Prices: Bivariate Time-series Modelling with Forecastable Error Varainces, in D. Belsley and E. Kuh (eds.), *Model Reliability*, Cambridge, Mass.

[83] Granger, C. W. J. (1988) Some Developments in the Concept of Causality, *Journal of Econometrics*, **39**, 199-211.

[84] Granger, C. W. J. (1991) Cointegration and Long-run Relationships, in R. F. Engle and C. W. J. Granger (eds.), *Long-run Economic Relationships Readings in Cointegration*, Oxford University Press.

[85] Hamao, Y. (1991) The Effect of The 1987 Stock Crash on International Financial Integration, in Ziemba, Bailey, and Haomao (eds.), *Japanese Financial Market Research*,

[86] Hamao, Y., R. Masulis, and Ng, V. (1990) Correlations in Price Changes and Volatility Across International Stock Markets, *Review of Financial Studies*, **3**, 281:307.

[87] Hamilton, J. (1994) *Time Series Analysis*, Princeton University Press.

[88] Hansen, B. (1995) Regression With Nonstationary Volatility, *Econometrica*, **63**, 1113-1132.

[89] Harvey, C. (1988) The Real Term Structure and Consumption Growth, *Journal of Financial Economics,* **22**. 305-333.

[90] Hausman, J. (1983) Specification and Estimation of Simultaneous Equations Models, in Griliches and Intriligator (eds.), Handbook of Econometrics vol. I, North-Holland Publishing Company.

[91] Heimstra, C., and Jones, J. (1994) Testing for Linear and Non-linear Granger-Causality in the Stock Price-volume Relation, *Journal of Finance*, **49**, 1639-1664.

[92] Hendry, D. (1995) *Dynamic Econometrics*, Oxford University Press.

[93] Hentschel, L. (1995) All in The Family Nesting Symmetric and Asmmetric $GARCH$ Models, *Journal of Financial Economics*, **39**, 71-104.

[94] Ho, M., and Sorensen, B. (1996) Finding Cointegration Rank in High Dimensional Systems using the Johansen Test, *Review of Economics and Statistics*, **78**, 726-732.

[95] Johansen, S. (1988) Statistical Ananlysis of Cointegraton Vectors, *Journal of Economic Dynamics and Control*, **12**, 231-234.

[96] Johansen, S. (1991) Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, **59**, 1551-1580.

[97] Johansen, S. (1992) Cointegration in Partial Systems and the Efficiency of Single Equation Analysis, *Journal of Econometrics*, **52**, 389-402.

[98] Johansen, S. (1995) Likelihood-based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press.

[99] Johansen, S. (1996) Likelihood-based Inference for Cointegration of Some Nonstaionary Time Series, in Cox, Hinkley and Nielsen (eds.) *Time Series Models in econometrics, finance and other fields*, Champan & Hall.

[100] Johansen, S. and Juselius, K. (1990) Maximum Likelihood Estimation and Inference on Cointegration — with Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics*, **52**, 169-210.

[101] Jorion, P. (1995) Predicting Volatility in the Foreign Exchange Market, *Journal of Finance*, **50**, 507-528.

[102] Karatzas, I., and Shreve, S. E. (1991) *Brownian Motion and Stochastic Calculus.* New York: Springer-Verlag.

[103] Karolyi, G. (1995) A Multivariate $GARCH$ Model of International Transmissions of Stock Returns and Volatility: The Case of the United States and Canada, *Journal of Business and Economic Statistics*, **13**, 11-25.

[104] Kasa, K. (1992) Common Stochastic Trends in International Stock Markets, *Journal of Monetary Economics*, **29**, 95-124.

[105] Kawaller, I., Koch, P., and Koch, T. (1987) The Temporal Price Relationship Between S&P 500 Futures and the S&P 500 Index, *Journal of Finance,* **42**, 1309-1329.

[106] King, R., C. Plosser, J. Stock, and Watson, M. (1991) Stochastic Trends and Economic Fluctuations, *The American Economic Review*, **81**, 819-840.

[107] King, M. and Wadhawani, S. (1989) Transmission of Volatility Between Stock Markets, *The review of Financial Studies*, **3**, 5-33.

[108] Koutmos, G. and Booth, G. (1995) Asymmetric Volatility Transmission in International Stock Markets, *Journal of International Money and Finance*, **14**, 742-62.

[109] Kupiec, P., and Sharpe, S. (1991) Animal Spirits, Margin Requirements and Stock Price Volatility, *Journal of Finance*, **46**, 717-732.

[110] Lamoureux, C., and Lastrapes, W. (1990) Persistence in Variance, Structural Change, and the *GARCH* Model, *Journal of Bussiness and Economic Statistics*, **8,** 225-34.

[111] Lee, C., A. Schleifer and Thaler, R. (1991) Investor Sentiment And The Closed End Puzzle, *Journal of Finance*, **46**, 75-109.

[112] Lo, A., and Mackinlay, A. (1995) Maximizing Predictability in the Stock And Bond Markets, *NBER Working Paper, #5027.*

[113] Longin, F., and Solnik, B. (1995) Is The Correlation in International Equity Returns Constant: 1960-1990 ?, *Journal of International Money and Finance*, **14**, 3-26.

[114] Lumsdaaine, R. (1995) Finite Sample Properties of the MAximum Likelihood Estiamtor in GARCH (1,1) and IGARCH (1,1) Models: A Monte Carlo Investigation, *Journal of Business and Economic Statistics*, **13**, 1-10.

[115] Lütkepohl, H. (1982) Non-causality Due to Omitted Variables, *Journal of Econometrics*, **19**, 367-378.

[116] Lütkepohl, H. (1990) Asymptotic Distribution of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models, *The Review of Economics and Statistics*, **72**, 116-125.

[117] Madhavan, A. (1996) Security Prices and Market Transparency, *Journal of financial intermediation*, **5**, 255-283.

[118] Magnus, J., and Neudecker, H. (1986) Matrix Differential Calculus with Applications in Statistics and Econometrics, Wiley Series in Probability and Mathematical Statistical.

[119] McQueen, G. (1992) Long-horizon Mean-reverting Stock Prices Visited, *Journal of Quantitative Analyses*, **22**, 1-18.

[120] Mitchell, M., and Mulherin, J. (1994) The Impact of Public Information on the Stock Market, *Journal of Finance*, **49**, 923-950.

[121] Nelson, D. (1989) Commentary: Price Volatility, International Market Links, and Their Implications for Regulatory policies, *Journal of Financial Services Research,* **3**, 247-254.

[122] Nelson, D. (1990) Stationary and Persistence in The $GARCH\,(1,1)$ Model, *Econometric Theory*, **6**, 318-334.

[123] Nelson, D. (1991) Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, **59**, 347-370.

[124] Park, J., and Phillips, P. C. B. (1988) Statistical Inference in Regressions with integrated Regressions: part 1, *Econometric Theory*, **4**, 468-497.

[125] Parkinson, A.,(1980) The Extreme Value Method for Estimating the Variance of the Rate of Return, *Journal of Business*, **53**, 61-65.

[126] Patelis, A. (1997) Stock Return Predictability and The Role of Monetary Policy, *Journal of Finance,* **52**, 1951-1972.

[127] Penman, S. (1987) The Distribution of Earning news Over Time and Seasonalities in Aggregate Stock Returns, *Journal of Financial Economics* 18, 199-228.

[128] Pesaran, H., and Y. Shin (1998) Generalized Impulse Response Analysis in Linear Multivariate Models, *Economics Letters*, 58, 17-29.

[129] Phillips, P. C. B. (1987) Time Series Regression With a Unit Root, *Econometrica*, **55**, 277-301.

[130] Phillips, P. C. B. (1988) Weak Convergence of Sample Covariance Matrices to Stochastic Intergrals via Martingale Approximations, *Econometric Theory*, **4**, 528-533.

[131] Phillips, P. C. B., and Perron, P. (1988) Testing for a Unit Root in Time Series Regression, *Biometrika*, **75**, 335-346.

[132] Phillips, P. C. B. (1995) Fully Modified Least Squares and Vector Autoregression, *Econometrica*, **63**, 407-436.

[133] Poterba, J., and Summers, L. (1986) the Persistence of Volatility and Stock Market Fluctuations, *Ameerican Economic Review*, **76**, 1142-1151.

[134] Richards, A. (1995) Comovements in National Stock Market Returns: Evidence of Predictability, But Not Cointegration, *Journal of Monetary Economics*, **36**, 631-654.

[135] Roll, R. (1988) R^2, *Journal of Finance,* **43**, 541-566.

[136] Roll, R. (1989) Price Volatility, International Market Links, and Their Implications For Regulatory Policies, *Journal of Financial Services Research*, **3**, 211-246.

[137] Ross, S. A. (1989) Information and Volatility: the No-arbitrage Martingale Approach to Timing and Resolution Irrelevancy, *Journal of Finance*, **44**, 1-17.

[138] Rozeff, M., and Kinney, W. (1976) Capital Market Seasonality: The Case of Stock Returns, *Journal of Financial Economics,* **3**, 379-402.

[139] Serletis, A., and Scowcroft, D. (1991) International Efficiency of Commodity Futures Prices, *Applied Financial Economics*, **1**, 185-192.

[140] Shastri, K., J. Sultan and Tandon, K. (1996) The Impact of the Listing of Options in the Foreign Exchange Market, *Journal of International Money and Finance*, **15**, 37-64.

[141] Shephard, N. (1996) Statistical Aspects of ARCH and Stochastic Volatility, in Cox, Hinkley and Nielsen (eds.) *Time Series Models in econometrics, finance and other fields*, Champan & Hall.

[142] Silber, W. L. (1985): The Economic Role of financial Futures, in Futures Markets: Their Economic Role (ed. A. E. Peck*),* American Enterprise Institute for Public Policy Research, Washington, DC.

[143] Sims, C. (1972) Money, Income, and Causality, *American Economic Review*, **62**, 540-552.

[144] Sims, C. (1980) Macroeconomics and Reality, *Econometrica,* **48**, 1-48.

[145] Stein, J. (1961) The Simultaneous Determination of Spot and Futures Prices, *American Economic Review*, **51**,1012-1025.

[146] Stock, J., and Watson, M. (1988) Testing for Common Trends, *Journal of the American Statistical Association*, **83**, 1097-1107.

[147] Stoll, H., and Whaley, R. (1990) The Dynamics of Stock Index and Stock Index Futures Returns, *Journal of Financial and Quantitative Analysis*, **25**, 441-468.

[148] Stulz, R. (1997) International Portfolio Flows and Security Markets, Unpublished working paper, *Dice Centre for Financial Economics*, The Ohio State University.

[149] Susmel, R. and Engle, R. F. (1994) Hourly Volatility Spillover Between National Equity Markets, *Journal of International Money and Finance*, **13**, 3-25.

[150] Thornton, D., and Battern, D. (1985) Lag Length Selection and Tests of Granger Causality Between Money and Income. *Journal of Money, Credit, and Banking*, **17**, 164-178.

[151] Toda, H., and Phillips, P. C. B. (1993) Vector Autoregressions and Causality, *Econometrica*, 61, 1367-1393.

[152] Tsay, R. (1987) Conditional Heteroscedastic time Series Models, *Journal of the American Statistical Association*, **82**, 590-604.

[153] Turnovsky, S., and Campbell, R. (1985) The Stabilizing and Welfare Properties of Futures Markets: A Simulation Approach, *International Economic Review*, **26**, 277-303.

[154] Urbain, J. (1992) On Weak Exogeneity in Error Correction Models, *Oxford Bulletin of Economics and Statistics*, **54**, 187-207.

[155] Watson, M. (1994) Vector Autoregressions and Cointegration, in R. Engle and D. McFadden (eds.) *Handbook of Econometrics* Vol. IV, Elsevier Science.

[156] White, H. (1980) A Heteroskedasticity-Consistent Covariance Matrix Estimator and A Direct Test for Heteroskedasticity, *Econometrica,* **48**, 817-838.

[157] White, H. (1982) Maimum Likelihood Estiamtion of Misspecified Models, *Econometrica,* **50**, 1-25.

[158] Wooldridge, J. (1990) A Unified Approach to Robust, Regression-based Specification Test, *Econometric Theory,* **6**, 17-43.

[159] Wooldridge, J. (1991) On the application of Robust, Regression-based Diagnostics to Models of Conditional Means and Conditional Variances, *Journal of Econometrics*, **47**, 5-46.