FRACTIONAL INTEGRATION AND LONG MEMORY MODELS OF STOCK PRICE VOLATILITY : THE EVIDENCE OF THE EMERGING MARKETS

By

Jorge Cláudio Cavalcante de Oliveira Lima

Department of Economics

McGill University

Montreal, Quebec, Canada

Montreal, January 2002

A Thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of degree of Doctor of Philosophy in Economics

0-612-78658-7

ABSTRACT

Following the important work on unit roots and cointegration which started in the mid-1980s, a great deal of econometric works has been devoted to the study of the subtleties and varieties of near nonstationarity and persistence that characterize so many economic and financial time series. In recent years research activity has gained importance with outstanding contributions made on estimation and testing of a wide variety of long memory processes, together with many interesting and imaginative applications over a wide variety of different fields of economics and finance. For these reasons, this study provides empirical evidence to an aspect of fractional differencing and long memory processes, or the long memory of volatility. Evidence of long memory persistence is explored using stock price indices for eight emerging economies in both Asian and Latin American markets. The concern with the presence of long memory in higher moments of return series was first drawn by Ding, Granger and Engle (1993), using asset returns. Baillie, Bollerslev and Mikkelsen (1996) developed the fractionally integrated GARCH, or FIGARCH, process to represent long memory in volatility. The measure of long-memory persistence in the volatility is employed either using the original rescaled range statistic by Hurst (1951) and its modified version proposed by Lo (1991). Further analysis of the presence of long memory persistence is conducted using autocorrelation analysis. All the findings point in the same direction, that is, the existence of long memory in volatility irrespective of the measure chosen. Estimation of different models of volatility is undertaken beginning with the ARCH specification and until the FIGARCH model. The results show the effects to be higher in Latin American countries than in the Asian ones. This result seems consistent with the degree of intervention in the Latin American markets, known to be much higher.

Other possible explanations for the occurrence of long term persistence are also pursued such as the Regime Switching modelisation proposed first by Hamilton and Susnel (1994) with the SWARCH approach. Results show that this approach can bring another possible explanation for persistence, specially in economies like Brazil that, have very different regimes for the period covered in this study.

RÉSUMÉ

Depuis des nombreux travaux sur le thème des racines unitaires et de cointégration, une grande diversité de papiers se sont interessés à étudier les sutilités et varietés des modéles qui s'adressent à la quasi-stationarité et persistence qui semblent décrire plusiers séries économiques et financiéres. L'éstimation et le test d'un grand nombre de modéles de longue

memóire ensemble avec d'autres innovations aussi créatives ont occupé des différents champs d'intérêt de l'économie et des finances. À cause de tout cela. cet étude fournit de l'évidence empiríque pour la persistence de la volatilité. L'évidence de l'existence de longue mémoire est fouillé en utilisant des données des marchés boursiers de huit différents économies en émergence, et en particulier pour les marchés de l'Asie et de l'Amérique Latine. L'attention pour ce genre de modële a été attiré par Ding, Granger et Engle (1993) en utilisant des donnés pour les prix actifs. Baillie, Bollerslev et Mikkelsen (1996) ont developé une approche qui a été nommé de Fractionally Integrated ARCH, ou tout simplement FIGARCH pour représenter la longue memoire en volatilité. Plusiers mesures de persistence sont alors employés pour identifier l'occurrence ou pas de persistence dans la volatilité des rendements boursiers. La mesure de Hurst (1951), ensuite celle de Lo (1991) aboutissent au même résultat, c'est-à-dire, l'effet de longue mémoire est présent dans la volatilité des marchés boursiers des pays en émergence. L'analyse d'autocorrélation est aussi employé et encore une fois les résultats précedents sont confirmés.

Ensuite, plusiers modèles sont estimés en débutant avec le modële ARCH jusqu'à la especification de FIGARCH. Il faut mantioner que les marchés Latino-Américains présentent plus d'effets de longue mémoire que ceux de l'Asie.

D'autres possibilités pour expliquer la persistence sont cherchés comme les modéles de changement de régimes appliqués á la volatilité (SWARCH). On poursuit des estimations sous des différentes hipothéses et on constate que ces modèles peuvent expliquer la persistence et en particulier pour des pays comme le Brésil où le degré d'inteventions est assez elevé.

ACKNOWLEDGMENTS

A project is never possible without the aid given by different individuals and institutions that were crucial to the development and production of such a challeging work. First of all, I would like to thank the financial support from CAPES and UFRRJ that enabled me to entertain the study in Canada and at McGill University.

Many individuals played a major role in my academic formation and background. Much of my econometric knowledge and passion were encouraged by professors John W. Galbraith and Victoria Zinde-Walsh that besides being generous and comprehensive in various periods of this thesis, always gave me advice and comments and unconditional support. Professor Jean Mercenier became a great friend of mine and inspired me in the research work and taught me how to be rigorous with methodology. Many other professors of McGill University, especially in the Department of Economics gave me continued support, in particular Franque Grimard. I owe to McGill Economics Department the environment that is important in the achievement of any academic degree. A lot of colleagues were also important for advices, comments or only for their friendship. Ata Assaf, Greg Tkackz, Oladi Gholamreza, were also support for difficult moments.

Finally, my family gave the love that is the origin of all that is good in the world. They gave me all one human being needs. To my father, to my mother, brother and sister, this is the result of four years of life during which I was absent, I do hope this has worth a while. To my wife, the good, dedicated and comprehensive love of my life, gave me the support a man must have in order not to give up. To my son, the most precious gift that life has ever gave me, and for the years of his youth that I could not see him growing.

Also my gratitude to Coleman Nee and Elaine Garnham, both of the Economics Department that were also part of this work.

TABLE OF CONTENTS

Chapter 1

Introduction

1.1. Purpose and Scope of the Thesis

The world economy has undergone significant changes during the decade of the nineteen-nineties. National economies have become increasingly open to international trade and investment, and new patterns have emerged in global financial markets. Many countries have abolished capital controls, leading to tremendous growth in international financial flows. Less developed countries in particular have been the recipients of unprecedented amounts of foreign savings, which many have used to finance current-account deficits. Most striking of all has been the growth of capital markets relative to the investment banks that traditionally supplied most funds to less-developed economies. Since the so-called Black September crisis of 1982, when Mexico defaulted on its debt servicing payments, many investment banks have been unwilling to lend to developing countries. Faced with this situation, developing country governments and companies have increasingly relied on capital markets to get access to funds. This development has been encouraged by recent advances in information technology, which have allowed individuals to transfer money from one bank to another in a matter of seconds. Technology has also made it easier for investors to participate in markets in two or more countries at the same time. It follows, then, that if something disturbs markets in country A, it is very likely to have an effect in country B as well, since individuals may have to sell their positions in one market in order to make up for a loss in another. The result has been increased volatility in financial markets over time.

We have observed international financial markets becoming more integrated as reduced information and transaction costs have opened up new opportunities for investors. A consequence of these increased linkages between markets is that we can no longer talk about isolated effects. Of particular interest is the fact that market disturbances today appear to last much longer than those observed in the past. In other words, there is evidence of long memory in volatility. The traditional ARCH methodology has not done a satisfactory job of accounting for these changes in market volatility that persist for a long time after the initial movement of stock prices. Consequently, a general objective of this thesis is to compare the volatility of established and emerging markets in order to confirm the existence of long memory in volatility and, if such evidence is found, to see whether it is more pronounced in emerging markets. This chapter will discuss the definition of volatility, review the existing literature on the subject, and examine the history of volatility in US markets in order to establish a benchmark for posterior comparisons with emerging markets.

Volatility is an important characteristic of markets for most financial instruments, and it plays a central role in many areas of finance, foreign exchange, etc. It is crucially important in asset pricing models and dynamic hedging strategies, as well as in the determination of option prices. From an empirical standpoint, it is therefore very important to carefully model any temporal variation in volatility. The ARCH model and its various extensions have proven to be effective tools along these lines, and as a result the literature on ARCH has expanded dramatically since the seminal paper by Engle (1982). The question of whether markets have become more volatile in recent years has been the subject of much debate, especially in the aftermath of events such as the Asian Crisis and the recent collapse of technology stocks in the NASDAQ market in the United States. The development of derivatives in all markets is also seen to have contributed not only to liquidity but also to increased volatility.

Currency devaluations, failed economic plans, regulatory changes, coups and other national financial "shocks" are notoriously difficult to predict and may have disasterous consequences for global portfolios. Despite these difficulties, researchers have managed to identify several regular features of emerging markets, including: high average returns, high volatility and low correlations both across emerging markets and with developed markets. Indeed, the lesson of volatility was learned the hard way by many investors in December 1994 when the Mexican stock market began a fall that would reduce equity value in U.S. dollars by 80% over the next three months.

This thesis takes the well-established literature on financial volatility in developed economies and applies it to the so-called emerging markets that occupy an increasingly important place in the menu of choices for investors. The objective is to determine whether some of the findings on volatility for more established markets are still true, and to use these findings as benchmarks to evaluate differences between markets.

In doing so, the hope is to determine the effects of the globalisation of the world economy on a sub-sample of countries, the "emerging" economies. The analysis is restricted to the most important emerging economies, as described by the International Monetary Fund (IMF) and the World Bank. Despite the fact that many countries can be classified as emerging markets, concerns about the accuracy of information and the availability of data have forced us to focus on the following: Argentina, Brazil, Mexico, Thailand, Taiwan, South Korea, Malaysia and Hong-Kong. For certain countries (e.g. Brazil) data are available from 1968 through 1999, but in most cases the period of available data extends from the end of the nineteen-seventies until the middle of the nineteen-nineties.

This thesis could have focused on a variety of themes, but what attracted our attention most was the question of the behaviour of volatility over time. Other researchs have suggested that developed countries have experienced considerable changes in market volatility during this period, so it is natural to ask whether this has also occurred in emerging economies. However, as we will make clear later in this chapter, volatility is a rather vague term that needs to be defined more precisely. In order to make comparisons easier and also more precise, the focus of analysis will be the huge field of the heterocedasticity models. The homocedasticity hypothesis is central to many econometric models, but since the development of the ARCH model a substantial amount of applied work on heteroskedasticity in time-series has been done. We will apply this methodology and the models derived from it to the available data from emerging economies.

The thesis is organised as follows. Chapter 1 defines emerging markets and characterises the data and indices used in the subsequent analysis. Some important differences between developed and emerging economies are higlighted, especially with reference to the United States. Chapter 2 consists of an extensive review of the literature on volatility, with the aims of showing the growing importance of this issue and identifying work that remains to be done with regard to less developed or "emerging" economies. Chapter 3 addresses the question of long memory volatility. A variety of methods are used to arrive at the conculsion that emerging economies do indeed exhibit this long memory property. Chapter 4 deals with the estimation of a great deal of time-series models, from ARCH models through to FIGARCH models. Comparisons between developed and emerging markets are also examined. Chapter 5 contains a discussion of alternative methods of estimating long memory models.

1.2 - EMERGING MARKETS[1]: Definition and Descriptive Remarks

Equity market returns in emerging economies differ substantially from those in developed economies. The term "emerging market" can be defined in various ways. On the one hand, "emerging" implies that a market has begun a process of change, growing in size and in sophistication compared to smaller markets that give little appearance of change. Alternatively, "emerging" can refer to any market in a developing economy, with the implication that all have the potential for development. The International Finance Corporation (IFC) follows a definition that considers most of low- and middle-income countries to be developing, regardless of their particular stage of development, and all stock markets in these countries are considered to be emerging. IFC follows the criteria of the World Bank in classifying economies as low-income, middle-income or high-income:

* low-income economies are those with a GNP per capita of $695 or less in 1993;
* middle-income economies are those with a GNP per capita of $696- $8.625 in 1993;
* high-income economies are those with a GNP per capita of $8.626 or more in 1993

[1] Broadly defined, an emerging market is a financial market in a country making an effort to change its economy with the goal of raising its performance to the level of the world's more advanced nations.

Table 1.1

WORLD BANK CLASSIFICATION OF ECONOMIES BY INCOME AND REGION 1994-95

Low Income	India
Middle Income (Lower)	Indonesia Philippines Thailand
Middle-Income (Upper)	South Africa South Korea Malaysia Argentina Brazil Mexico
High-Income (Lower)	Singapore Taiwan Hong-Kong Chile

Source: The International Finance Corporation

Table 1.1 shows the classification used by the International Finance Corporation, which will be adopted in the remainder of this thesis. Based on this definition, we have chosen a sample of countries that should give a representative view of the emerging markets as a whole. Tables 1.2 and 1.3 show the relative importance of each market. In Table 1.2, countries are listed in order of world capitalisation. We can easily see that Hong Kong, South Africa, Malaysia, Taiwan, India and Brazil are the most important markets among the emerging markets by this criterion. The growing importance of Asia relative to Latin America is obvious and impressive (see Brazil and Mexico, for instance). During the seventies and the eighties the majority of Latin American countries were governed by dictatorships which were often very closed to foreign capital. The Mexican debt crisis of 1982 also discouraged the flow of capital to Latin America, which was then redirected to Asian countries. Recently, in the aftermath of the Asian Crisis, this trend has been reversed to some extent.

Table 1.3 shows world value traded, and again we can see the increased importance of the south-Asian countries vis-à-vis Latin America. Taiwan is the leader, followed by South Korea, Hong-Kong, Malaysia and finally Brazil. The same pattern found before is reproduced here. Movements in both tables indicate the growing relative importance of south-Asian markets, especially in the 1980-1994 period. The exact ranking differs according to the classification used, but the same countries are present in Tables 1.2 and 1.3. For our purposes, this is the really important issue in the analysis we are pursuing. As mentioned before, the relative importance of Latin American markets has increased after the Asian Crisis.

Table 1.2
WORLD MARKET CAPITALIZATION 1986-1999 (US$ BILLIONS)

Emerging Markets	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1999
Argentina	1.6	1.5	2.0	4.2	3.3	18.5	18.7	43.7	36.9	37.8	53.5
Brazil	42.0	17.0	32.1	44.4	16.3	42.8	45.3	99.4	189.3	147.6	155.1
Chile	4.1	5.3	6.8	9.6	13.6	28.0	29.6	44.6	68.2	73.9	58.2
India	13.6	17.1	23.6	27.3	38.6	47.7	65.1	98.0	127.5	127.2	155.2
Korea	13.9	32.9	94.2	140.9	110.6	96.4	107.4	139.4	191.8	182.0	229.7
Malaysia	15.1	18.5	28.3	39.8	48.6	58.6	94.0	220.3	199.3	222.7	131.7
Mexico	6.0	8.4	13.8	22.6	32.7	98.2	139.1	200.7	130.2	90.7	119.0
Philippines	0.7	2.0	2.9	4.3	12.0	5.9	10.2	13.8	40.3	55.5	45.8
South Africa	102.3	128.7	126.1	131.1	137.5	168.5	103.5	171.9	225.7	280.5	221.9
Taiwan	15.4	48.4	120.0	237.0	100.7	124.9	101.1	195.2	247.3	187.2	338.9
Thailand	1.9	2.9	0.5	8.8	25.6	23.4	35.8	58.5	130.5	131.5	44.9
Hong Kong	53.4	54.1	74.4	77.5	83.4	122.0	172.1	385.2	269.5	303.7	305.7
Totals	238.6	319.7	483.2	738.1	611.3	854.8	883.5	1,586.7	1,912.4	1,895.7	2,521.4

Source: The International Finance Corporation

The Emerging Markets Factbook 1994, published by the International Finance Corporation, presents the share of emerging markets based on several criteria, including capitalisation, value, and performance among others. We show these shares in tables 1.4 and 1.5 below. Again we see the importance of the Asian market during the period in question. The Latin American market has also an outstanding performance during these periods. Both markets summarise the behaviour of the emerging markets since the nineties and for this reason we will be analysing these markets in the thesis.

These markets have, however, very different behaviours. The Asian markets are more market-oriented and as such show a lower degree of intervention compared to the Latin American markets. These markets have passed through an opening and liberalizing process that is more recent, since the beginning of the nineties.

Table 1.3
WORLD VALUE TRADED 1986-1999 (U$S BILLIONS)

Emerging Markets	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1999
Argentina	**0.6**	**0.3**	**0.3**	**0.6**	**1.9**	**0.9**	**4.8**	**15.7**	**10.3**	**11.4**	**12.0**
Brazil	**21.5**	**28.9**	**9.6**	**18.0**	**16.8**	**5.6**	**13.4**	**20.5**	**57.4**	**109.5**	**120.0**
Chile	**0.06**	**03**	**0.5**	**0.6**	**0.9**	**0.8**	**1.9**	**2.0**	**2.8**	**5.3**	**18.0**
India	**5.0**	**10.8**	**6.7**	**12.2**	**17.4**	**21.9**	**24.3**	**20.6**	**21.9**	**27.3**	**13.3**
Korea	**4.2**	**10.9**	**24.9**	**79.2**	**121.3**	**75.9**	**85.5**	**116.1**	**211.7**	**286.1**	**63.3**
Malaysia	**2.3**	**1.2**	**3.8**	**2.6**	**6.9**	**19.9**	**10.7**	**21.7**	**153.7**	**126.5**	**44.4.**
Mexico	**2.4**	**3.8**	**15.6**	**5.7**	**6.2**	**12.2**	**31.7**	**44.6**	**624.5**	**83.0**	**51.6**
Philippines	**0.1**	**0.6**	**1.5**	**0.9**	**2.4**	**1.2**	**1.5**	**3.1**	**6.8**	**13.9**	**19.2**
South Africa	**2.8**	**5.0**	**9.6**	**4.9**	**7.1**	**8.2**	**8.1**	**7.8**	**13.0**	**16.0**	**6.1**
Taiwan	**5.0**	**18.9**	**84.1**	**275.6**	**965.8**	**715.0**	**365.2**	**240.7**	**346.5**	**711.3**	**1070.4**
Thailand	**0.6**	**1.1**	**4.6**	**5.6**	**13.5**	**22.9**	**30.1**	**72.1**	**86.9**	**80.2**	**43.2**
Hong-Kong	**9.7**	**15.3**	**47.6**	**23.4**	**34.6**	**34.6**	**38.6**	**78.6**	**131.6**	**147.2**	**150.0**
Totals	**45.2**	**82.9**	**164.7**	**408.6**	**1,165.5**	**894.4**	**605.5**	**612.3**	**1,068.9**	**1,640.1**	**2852.4**

Source: The International Finance Corporation - Emerging Markets Factbook 1995

Table 1.4

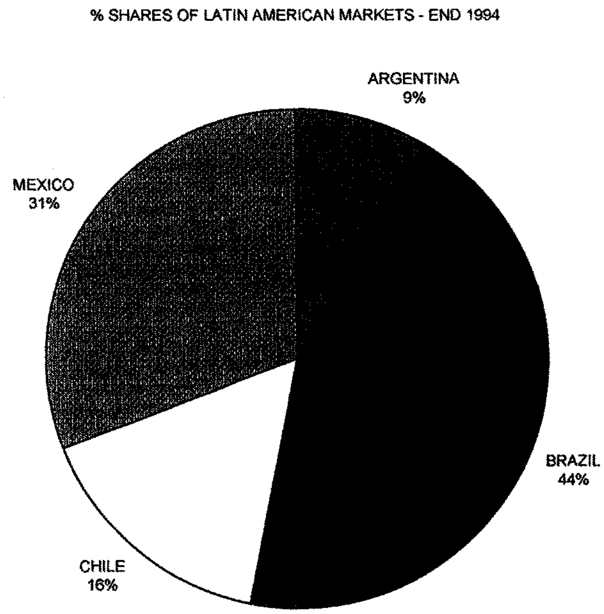
Shares of the Emerging Markets

End of 1994	% of Emerging Markets	% other Latin America	% other Emerging Markets
Argentina	**1.9**	**21.8**	**76.2**
Brazil	**9.8**	**13.9**	**76.2**
Chile	**3.5**	**20.2**	**76.2**
Mexico	**6.8**	**17.0**	**76.2**

Source: The International Finance Corporation - Emerging Markets Factbook 1995

The Latin American country with the largest stock market is Brazil, followed by Mexico, Chile and Argentina. Together, these markets are responsible for some 23% of emerging markets as compiled by the IFC in 1994. If we exclude the former socialist countries (Poland, the Czech Republic, etc.) and also ignore China, this percentage increases to 38%, as is depicted in Figure1.1 below. This confirms our assertion above that these are, together with the Asian markets the more important representative of the emerging economies.

Figure1.1



Source: International Finance Corporation

Table 1.5

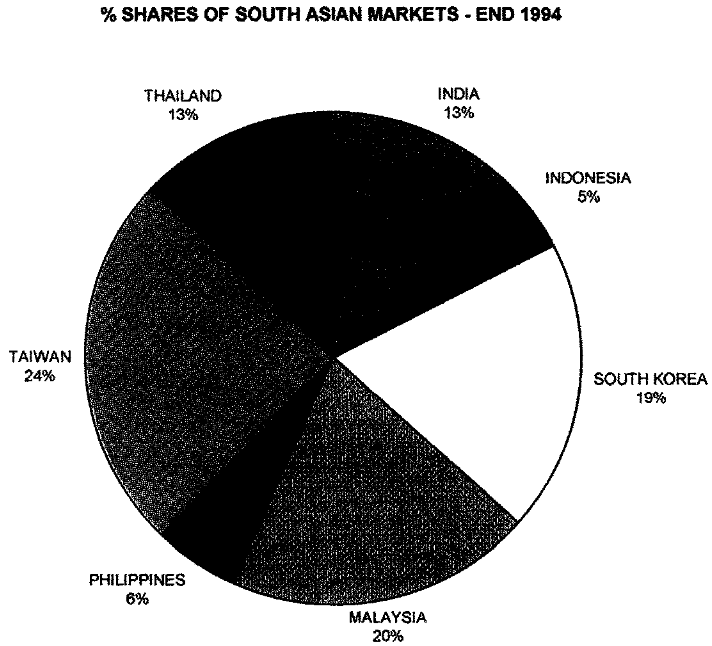
Shares of the Emerging Markets

End of 1994	*% of Emerging Markets*	*% of other South Asia*	*% of other Emerging Markets*
India	6.6	70.4	28.5
Indonesia	2.4	24.6	28.5
Korea	9.9	18.0	28.5
Malaysia	10.3	16.7	28.5
Philippines	2.9	25.0	28.5
Taiwan	12.8	15.1	28.5
Thailand	6.7	20.2	28.5

Source: The International Finance Corporation - Emerging Markets Factbook 1995

Looking at the Asian markets, the most important of these at the end of 1994 was Taiwan, followed by Malaysia, South Korea, Thailand, India, the Philippines and Indonesia (this list excludes Hong-Kong). Together, these markets are responsible for some 72% of overall emerging markets (see Figure 1.2).

Figure 1.2



Source: International Finance Corporation

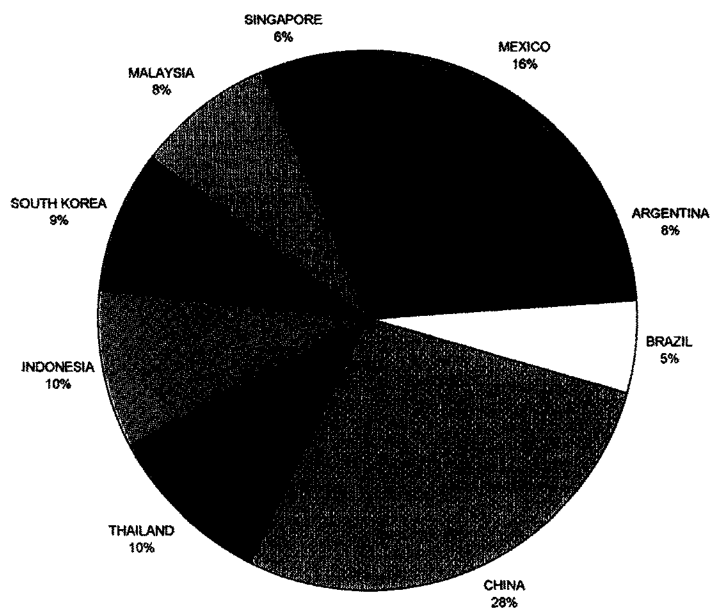
A fundamental factor explaining the return of private capital to emerging markets has been the process of deregulation of financial markets in developed economies since the 1980s, especially in the United States, United Kingdom and Japan. In these countries, deregulation increased the number of investment opportunities for banks, insurance companies and institutional investors (e.g. pension funds) that were previously operating in relatively segmented markets. Deregulation also brought with it increased mobility of capital between markets, a sharp expansion of liquidity, and consequently an increase in the prices of financial assets that has generated an accelerated growth of net financial wealth.

Secondly, the liberalisation of the capital accounts of several countries, along with the reduction and maintenance of low real interest rates in developed economies to offset the effects of recession, have amplified these movement of capital at the international level. The immediate consequence of the expansion of international capital flows has been increased interdependence of national economies. Specifically, for economies with less-developed capital markets, the sudden and unprecedented increases in liquidity caused sharp rises in the prices of financial assets. We have also the change in the regime of financial repression – the practice of very high real interest rates. The financial liberalisation made possible the quick expansion of the so-called " emerging markets" with a big increase in the potential returns of capital in these economies. Also, the consensus view of Washington was an important reference point for policy-makers in implementing policies that were largely based on the idea of liberalisation.

We have already seen that capital flows into Asian countries have far exceeded those into Latin America in recent decades, until 1997. This can be explained by the fact that the process of financial liberalisation was quickier in the Latin America than it was in Asia. Moreover, this happened in the context of a change in financial regime, where fiscal deficits began to be financed by the market and not by the printing of new money. If we add to this the liberalisation of capital controls and the privatisation of many state-owned companies, we can appreciate the impressive growth of portfolio investment in Latin American countries, particularly Mexico, Argentina and Brazil. However, the strategies pursued by these regions are different. In the case of Latin America, it appears that short-term capital has been attracted to the region. In the decade of the nineteen-nineties Asia has confirmed to be a magnet for long-term capital, especially due to (i) the development strategy implemented in the last 15 years, and (ii) the cautious approach of Asian governments towards liberalizing financial markets[2]. This approach has created direct and indirect "filters" to the flow of short-term capital. (See Figure 1.3.)

Figure 1.3

FLOWS OF CAPITAL TO LATIN AMERICA AND SOUTH ASIA IN 1994



Source: Unctad, 1994

[2] Some argue that it is not true that capital movements to Ásia were predominantly long-term. In fact, of the fast growing countries of Asia which depended on foreign capital (South Korea, Thailand, Indonesia and Malaysia), all depended on mostly bank-based short term capital (and got into trouble in 1997). On the other hand, Taiwan, Hong-Kong and Singapore which were net exporters of capital had less trouble.

The rates of economic growth observed in Latin American countries have been slow and erratic, and economic policies have not shown any indication of leading to sustainable growth. Foreign capital has been attracted mostly by the interest rate differentials (i.e. differences between internal and external interest rates) caused by the change in the regimes, and by capital gains from speculation activity related to the privatisations that occurred in many countries. However, we have already stressed that the flows changed dramatically after the Asian crisis, with Brazil being by far the most desired country for foreign investors. This is an important point, but it does not change the fact the emerging markets are an important destination for foreign investors. From the investor's stand-point, the only difference is the geographic destination. Table 1.6 shows the magnitude of the flows of capital towards the emerging economies in 1995.

Table 1.6

Flows of Capital to Latin America and South Asia in 1995

Region/Country	*Total US$ billions*	*Bank Credits*	*Portfolio Investments*	*Foreign Direct Investment*
Latin America (1)	137,4	18,6	67,1	52,0
Mexico	63,0	9,5	32,1	21,4
Argentina	31,4	2,5	14,8	14,1
Brazil	21,4	1,7	13,7	6,0
South and South Asia (2)	295,0	103,0	56,3	136,9
China	110,2	19,9	14,3	77,1
Thailand	39,0	24,2	6,2	8,9
Indonesia	37,7	24,5	5,3	8,0
South Korea	34,0	13,3	17,7	3,1
Malaysia	32,2	11,3	2,3	18,4
Singapore	25,6	6,1	1,8	17,9

Source: Unctad, Trade and Development Report 1996.

The Mexican balance-of-payments crisis was accompanied by a substantial drop in that country's stock market, the consequences of which were not limited to Mexico. The so-called "tequila effect" caused drops in stock markets in other Latin American countries, especially Brazil and Argentina. Several ministers of finance went to New York to explain that their countries were not like Mexico and so should not be penalised for that country's difficulties. The Latin American emerging markets have become an important alternative destination for investors in the developed countries, the United States in particular. The openness of Latin American markets to foreign investors at a time of historically low American interest rates created an enormous flow of capital from North to South in the Americas. Among the largest recipients were Argentina, Mexico and Brazil. According to the World Bank, Latin America received 30 percent of foreign direct investment in the world between 1989 and 1993 (World Bank, 1995). The volume of securities issued by developing countries reached an astonishing $59 billion U.S. at the end of this period. Among the biggest issuers were again the major Latin American countries, espeically Brazil, Mexico and Argentina. In terms of stocks in the international market, the total issued was around some US$12 billions U.S. in 1993. Nearly half of the original amount came from Latin America (IMF, 1995). The investment in securities and stocks coming from Latin America became very popular among the American investors. Recently a lot of Brazilian companies have launched ADR's (American depositary receipts) that are stocks traded in the New York Stock Exchange.

1.3 - Stock Price Indices and their definitions

The stock price indices used here are indicators of the performance of the stock prices in each market, showing the behaviour of the principal shares traded on each stock exchange. For each market used we have relied on the most representative index for that market. As our analysis is based on the evolution of the emerging markets, we have restricted our focus to these representative indices for each country. The indices used are as follows: the MERVAL (MERcado de VALores - Argentina), IBOVESPA (Indice da Bolsa de Valores do Estado de Sào Paulo - Brazil), IPC (Indice de Precos de Cotizaciones - Mexico), KCSPI (Korea Composite Stock Price Index -Korea), SET (Stock Exchange Thailand -Thailand), Kuala Lumpur Stock Exchange (KLSE - Malaysia) , TSEWSI (Taiwan Stock Exchange Weighted Stock Index -Taiwan) and Hang-Seng Index for Hong-Kong (HIS – Hong Kong). We discuss briefly some of the characteristics of these indices below.

1.3.1 - Definition of a Weighted Stock Price Index (used by all the countries in the sample)

An index represents the current value, in domestic currency, of a portfolio made up on a specific date, varying from country to country and starting from a hypothetical investment. Only the reinvestment of the dividends received, the amount resulting from the sale of subscription rights and the maintenance in the portfolio of shares received as bonuses are considered.

Indices should be reliable and should use a methodology that is easily followed by the market. They should represent not only the average behaviour of the prices of the main stocks, but also the profile of trading carried out during each trading session.

A stock market index aims to give an indication of average market behaviour. As such, it seeks to reflect as closely as possible the real configuration of cash tradings (round lot) on the stock market in order for the investor to have an idea of how a particular market is changing over time.

a) Portfolio

The portfolio of an index is composed of stocks which jointly represent at least 70-90% of the amount of cash transacted during the twelve months preceeding the establishment of the index for each of the countries under consideration. This can vary from stock market to stock market. As an additional criterion, it is required that a stock be traded on 70-90% of the trading sessions of the reference period. The percentage varies from market to market, but at least 70% of the market should be included in the established index. This rule of thumb has been found to be effective when considering the coverage of an index.

The share of each stock in the portfolio is directly related to the significance of this security on the cash market in terms of number of trades and the amount in domestic currency, adapted to the size of the sample.

In order to maintain the representativeness of the Index, a revaluation of the market is carried out every four months, always based on the twelve preceding months. Again these figures can vary from country to country, but this is the procedure found in most stock exchanges. After changes in the relative participation of each stock have been identified, a new portfolio is formed and a new weight is given to each security according to the market distribution as assessed by the revaluation study. An example of this is the recent importance of shares of companies related to information technology, the so-called 'new economy' stocks that have led to changes in most of indices around the world.

1.3.2 - The Indices as Representative of Market Capitalisation

The issuing companies whose stocks are part of the portfolios of the indices are responsible, on average, for approximately 70-90% of the total market capitalisation of all companies listed on any particular stock market. Each stock market calculates its index in real time, taking into account all trades on the cash market involving stocks in its portfolio. Information is then disseminated through each stock market's network and also by a series of vendors. Consequently, it is possible to follow the behaviour of an index on-line throughout the day, anywhere in the world.

a) Transparency

An easy calculation methodology and limited changes in that methodology over time are highly desirable. These elements, together with easy access to data, ensure the usefulness of the index. This can be confirmed by the fact that, for several stock markets, these indices are the sole indicators of the performance of stocks to be traded on a liquid futures market.

Each stock market is responsible for the management, calculation, diffusion and maintenance of the index. This responsibility ensures strict observance of regulations and technical procedures in its methodology.

1.3.3 - Calculation of the Indices

The weighted index stands for the summation of the weights of the stocks (i.e. the hypothetical amount of the stock multiplied by its last price) that compose the theoretical portfolio. It can be calculated, at any moment using the following formula:

$$INDEX_T = \sum_{i=1}^{n} P_{i,T} Q_{i,T} \quad (1.1)$$

where:

$INDEX_T$ = the stock market index at time T

n = the total number of the stocks that compose the theoretical portfolio

P = the most recent price of stock i at time T

Q = the theoretical quantity of stock i in the portfolio at time T (equally value weighted)

1.3.3.1 Adjustment of the theoretical amount due to distribution of benefits

The mechanism of change is similar to the one used for the adjustment of the portfolio as a whole. That is, one considers that the investor sold the shares for the latest closing price at the beginning of the distribution of benefits and used the resources in the purchase of the same shares without the benefits distributed.

Formula for Changing in the Theoretical Quantity (at the time of distribution of benefits)

$$Q_n = \frac{Q_0 P_0}{P_{ex}} \quad (1.2)$$ where:

Q_n = the new quantity, Q_0 = the previous quantity, P_0 = last closing price prior to the beginning of distribution of benefits, P_{ex} = ex-benefit theoretical price, calculated based on P

General Formula of Calculation of Ex-benefit theoretical Price

$$P_{ex} = \frac{P_c + (S.Z) - DIV}{1 + B + S} \quad (1.3)$$

where:

P_{ex} = the ex-theoretical price

P_c = the last closing price prior to the beginning of distribution of the benefit

% S = the subscription percentage

Z = the issue price in domestic currency of each share to be subscribed

Div = the value in domestic currency received by each share as dividend

% B = the bonus percentage

a) Criteria for Inclusion of Stocks in the Portfolio

In order to have a stock included in an index, it is necessary that it simultaneously fulfil the following requirements, always in relation to the preceding 12 months (the period may vary slightly from one stock market to another, but this is the figure most commonly used):

- It must be among the list of stocks comprising 80% of the value traded in the stock market. This list results from the accumulation of the value traded for each security, disposed in decreasing order, with the cut-off limit set at 80.

- Its share of total value traded must be greater than 0.1%.

- It must have been traded in more than 70% of the total trading sessions during the period.

Each stock market calculates the index of negotiability for each of the stocks traded on the exchange in the last twelve months according to the formula given below. These indices are placed on a chart in decreasing order, and one column shows the sum of these indices as one reads the chart from the highest to the lowest. The participation of each index in the summation is then calculated, listing the stocks until the total number of accumulated participations reaches 70-80%. The stocks so chosen will enter the portfolio of the index if they fulfil the two other criteria. If a stock does not fulful the criteria it is replaced by the next stock on the list that does fulful them.

When the minimum of 70% of the sum of the index of negotiability is reached, one has the list of stocks which are going to compose the index for the next four months. This process is repeated constantly. The indices of these chosen stocks are listed again and the participation of each of these indices, with respect to the sum of the indices of all the securities in the portfolio, is calculated. The result is multiplied by the original index of negotiability and the adjusted participation is then obtained. The adjusted participation of each stock, applied to the value of the index of the last day of the preceding four-month period, will determine the composition of the portfolio for the next four-month period and so on.

The theoretical amount of each stock will remain the same for the period of validity of the portfolio. It will be altered only in case of a distribution of benefits (dividends, cash and bonus shares, subscriptions) by a company. On the other hand, once a stock is chosen to participate in the index portfolio it will only be withdrawn when it cannot fulfil at least two of the indicated criteria.

1.3.4 - Stock Negotiability Index

$$\sqrt{\frac{n}{N} \cdot \frac{v}{V}} \quad (1.4)$$

where:

n = the number of trades involving the stock carried out on the cash market (round lot) in the last 12 months

N = the total number of trades on the cash market (round lot) in the last 12 months

v = the value in domestic currency obtained with trades with the stock on the cash market (round lot) in the last 12 months

V = the value in domestic currency of the total amount traded on the cash market (round lot) in the last 12 months.

1.4 - Data Used in the Thesis

According to Harvey (1995), we have learned some regular features of emerging markets. First, we need to be careful in interpreting the average performance of these markets as the International Finance Corporation (IFC) has backfilled some of the index data, resulting in a survivor bias in average returns. This problem arises when we choose to work on an average index for the market as a whole like the IFCG (Global Index). For this reason we decided to work with each country individually. Second, the countries that we have currently chosen are the ones the have proven track records. This choice of "winners" introduces a selection bias. Third, some markets have long histories beginning in the last half of the 19^{th} century. At one point in the

1920's, Argentina's market capitalization exceeded that of the U.K. However, this market declined in subsequent years. We will only consider returns in the re-emergence period since the 1960s.

Argentina

The best known index is the Merval index (the most widely used, it is weighted by the traded volume of shares). This index is prepared at the Buenos Aires Stock Market. The index is computed daily based on constituent stocks representating the commercial-industrial, property, mining, and oil sectors. The Merval index comprises only 20 companies. It began to be computed on a daily basis in 1989. This index is revised every three months on up to 80 per cent of the volume of shares traded during the preceding six months.

The data used in the thesis are daily quotes for the period from January 2, 1989 through December 31, 1997. The base date used in the index is September 30, 1994.

Brazil

The IBOVESPA is the most representative Brazilian indicator, due to its continuity - it has not been subject to any methodological changes since its implementation on January 2, 1968 - and also to the fact that BOVESPA is responsible for up to 85 % of the total business transactions carried out by all Brazilian stock exchanges.

Index Subgroups

Many stock markets disseminate information about the performance of subgroups of the principal index in daily bulletins. Group I is composed of stocks with greater importance in the portfolio of the index for a period of four months, chosen according to the criterion of dispersion. Group II evaluates the average performance of other stocks that compose the IBOVESPA index. The existence of these indices helps investors adequately follow the behaviour of the market. Recently (May, 2000) a new group, Group III, has been introduced in order to take account of the so-called 'new economy' stocks.

The calculation methodology used by the IBOVESPA index is published by the stock market in its Methodological Issue (most recently in May 2000). Weights and sources used in preparing the index are revised periodically, with statistics depending on the volume of shares traded in the stock exchange.

IBOVESPA is a composite index of share price movement. The index is computed daily, based on the traded volume of shares of constituent stocks representative of the commercial-industrial sector, property, mining, and oil sectors. The method of computation is based on a comparison of the trade volume of shares of constituent stocks on a given day with the total traded volume on the previous day.

Weighting for each stock varies daily as traded volume changes. The base date used in the index is September 30, 1994. The data used are quoted daily from January 2, 1968 through December 31, 1999. IBOVESPA has 1994=100 as its base year, and uses 56 shares quoted on the São Paulo Stock Exchange.

Malaysia

The Kuala Lumpur Composite Index (KLCI), has its base year in 1977=100, and uses 100 shares quoted on the Kuala Lumpur Stock Exchange (KLSE). Quotes are for market closing prices. The data start on January 2, 1975, and we have daily observations through December 31, 1997. It is a composite index that comprises industrial, financial, mining and other sectors. Various sub-indices by industry also exist to help investors follow movements in the market.

The index is computed daily, based on the market capitalisation of constituent stocks representative of the commercial-industrial, property, mining, and oil sectors. The method of computation is called the "Weighted Market Capitalisation Method," where the composite index for the day is computed by comparing the total market capitalisation of constituent stocks for the day with the total market capitalisation on the previous day.

Mexico

The "Indice de Precios y Cotizaciones" (IPC), or Price and Listing Index, is calculated using the shares of enterprises making up the sample designed for this purpose, which has national coverage.The Mexican Stock Exchange (BMV), located in Mexico City, is responsible for preparing the IPC.

"Indicadores Bursátiles" reports indices that allow statistical cross-checks and that provide assurance of reliability of the IPC by dividing it into several sub-indices: industrial enterprises, public utilities, and transportation. The index comprises 45 companies. The base date used is October 1978=100. The data coverage is from November 1, 1978 through December 31, 1997.

Thailand

The share price index used here is from the Stock Exchange of Thailand (SET). The SET index is calculated from all common stocks listed on the stock exchange. The base period is April 30, 1975. The SET index is defined as the current market value divided by the base market value, multiplied by 100.

The index is computed daily, based on the market capitalisation of constituent stocks representative of the commercial-industrial, property, mining, and oil sectors. The method of computation is the "Weighted Market Capitalisation Method" where the composite index for the day is computed by comparing the total market capitalisation of constituent stocks for the day with the total market capitalisation on the previous day. The base date used in the index is 4th January 1980 and uses 39 companies begining on the 30 April 1975 and going until 30 December 1997.

Korea

The Korea Composite Stock Price Index (KCSPI) starts on January 4, 1977 and ends on December 31, 1997. The base period is January 4, 1980. It comprises 14 industry sectors including some 743 companies.

The index is computed daily, based on the traded volume of shares of constituent stocks representative of the commercial-industrial, property, mining, and oil sectors. It is a composite index of share price movements in the stock market. The method of computation is the "Weighted Trade Value Method," where the composite index for the day is computed by comparing the trade volume of shares of constituent stocks for the day with the total traded volume on the previous day.

The KCSPI disseminates information about the performance of subgroups of the principal index in its daily bulletin. These indices are divided into first and second section stocks. There are also indices by capital size: Large-Sized Capital stocks are those whose capital is more than W$15 billion, Medium-Sized Capital stocks are those with more than W$5 and less than W$15 billion, and finally, Small-Sized Capital stocks are those with less than W$5 billion.

Taiwan

We utilize the share price index quoted on the Taiwan Stock Exchange Weighted Stock Index. The data we use start on January 4, 1975 and finishes on December 31, 1997. The base period is January 5, 1981 = 100. The index comprises 67 companies and is computed daily, based on the traded volume of shares of constituent stocks representative of the commercial-industrial, property, mining, and oil sectors. The method of computation is the "Weighted Trade Value Method" where the composite index for the day is computed by comparing the trade volume of shares of constituent stocks for the day with the total traded volume on the previous day.

Hong-Kong

We utilize the share price index quoted on the Hang Seng Exchange Weighted Stock Index. It starts on the January 4, 1975 and we use observations up to December 31, 1997. The base period is January 5, 1981 = 100. The index comprises 27 companies and is computed daily, based on the traded volume of shares of constituent stocks representative of the commercial-industrial, property, mining, and oil sectors. The method of computation is the "Weighted Trade Value Method" where the composite index for the day is computed by comparing the trade volume of shares of constituent stocks for the day with the total traded volume on the previous day.

It is a well-established fact that there is a fairly low correlation among the world's equity markets, and this is frequently presented as evidence in support of the portfolio gains to investors from international diversification. Some studies have documented a significant increase in the correlation between national equity market co-movements during the 1987 international equity market crash and other periods of high volatility, reducing the benefits of international diversification.

For the period covered in our thesis. the data used for Hong Kong cover the period during which it remained under British protection, so that for much of this time it was seen as an extension of part of the London market. We have not captured the passage from Hong Kong to China that took place in 1997, and the attendant effects on the volatility. While the period covered was one in which Hong Kong's status as an emerging market might be questioned because of the links with London, we have at least avoided mixing the pre- and post- 1997 market movements in one data set.

1.4.1 The Choice for the Returns

The choice of nominal returns rather than another possible specification for the returns is based on a number of observations gathered from the literature. For example, some studies of volatility have used daily inflation-adjusted returns instead of the raw returns. These studies have suggested that for closed economies (i.e., closed to foreign investors) one could use either real returns (adjusted by inflation) or simply the US$ dollar returns.

However there are problems with each of these procedures. In the case of real returns, there is no estimate of the inflation on a daily basis, so some ad-hoc procedures have to be used, such as that described by Harvey (1995). Harvey took annualized monthly inflation rates for each country obtained from the *International Financial Statistics* database; each monthly inflation rate was then divided by 240 to estimate daily inflation (assuming 12 months with 20 trading days in each month as an approximation.)

The daily returns are then computed as the difference between the nominal return and the inflation rate computed for that day, using the methodology above. This is an approximation that can be extremely rough during hyperinflationary times; hyperinflation does occur in a number of periods covered by the analysis of this thesis, so that this is a serious shortcoming for our purposes, although it might be an acceptable approximation in other circumstances.

The other alternative identified above is to transform the returns in local currency into US$ using the prevailing exchange rate. However, for these countries the exchange rate has been fixed for some part of the sample period of interest; again, this transformation would be misleading for a part of the sample period. Calculating the returns in U.S. dollars eliminates local inflation, but retains the U.S. inflation that may be a problem for longer periods of data coverage. Perhaps more importantly, the series under study will be affected by the relatively long periods in which the exchange rate does not reflect purchasing power.

In Brazil, during periods of hyperinflation, new standard lots for transactions were defined as stock prices went up to reflect the inflation. For example, on February 28, 1986, when the new currency was introduced through a division of the old monetary standard by dividing by 1000, the unit became round lots of 1000 shares instead of one single share (see, e.g., Barry et al. 1998).

Many local and foreign researchers have studied the Brazilian stock market, one of the largest emerging markets in the world. A few examples are Aggarwal and Tandon (1994), Aggarwal and Leal (1996), Lemgruber et al. (1998), Costa Jr. (1990), Costa Jr. and Lemgruber (1993), Almeida et al. (1993) and Leal and Sandoval (1994). All of these studies used nominal returns.

For Argentina, by contrast, the data used in this thesis begin in 1989 with the Austral Plan, and so do not suffer from the problem of the potential effects hyperinflation.

One final point about returns could be related to the choice of returns and not the excess returns (gross returns minus the risk free rate). We have been influenced in this decision by the fact that for most of the emerging economies, it is difficult to define precisely what would be a risk free rate; the usual Treasury bonds are not viewed as risk-free assets, as for example in the U.S.

In the absence of an acceptable measure of the risk free rate, we have chosen to work with the returns themselves and their power transformations.

We do not remove possible calendar effects, such as the well-known January effect, in this thesis. While such effects may well exist, our primary purpose it to compare patterns in emerging markets with those in developed markets: we prefer minimal pre-filtering of data, so that any differences in markets are observed in the data chosen for analysis.

1.5 - Descriptive Statistics and Features of the Emerging Economies

It has been argued that in the nineteen-nineties we have seen a return of voluntary capital flows to South-Asian and Latin American countries, with important differences between the two regions. Both regions have received significant amounts capital in the last five years, but some of the Asian newly industrialised countries (NICs) have succeeded in attracting greater volumes of high quality capital with longer maturities. Table 1.7 shows some descriptive statistics for the Emerging economies.

Table 1.7
Descriptive Statistics from Monthly Returns in percent (Returns in US$) (January 1982 to April 1995)

Country	*Average*	*Standard Deviation*	*Coefficient of Variation*	*Kurtosis*	*Asymmetry*	*Minimum*	*Maximum*
Argentina (Merval)	3.64	24.11	6.62	6.6	1.78	-60.86	119.6
Brazil (Ibovespa)	4.11	23.72	5.77	.97	.58	-65.9	281.79
Chile (IGPA)	1.59	7.93	4.99	-.24	-.10	-19.93	20.76
Mexico (IPC)	3.00	16.02	5.34	2.83	-.69	-60.60	53.00
USA (S&P 500)	1.00	4.32	4.32	4.63	-.68	-21.76	13.18
Developed Markets	1.04	4.23	4.07	1.83	-.43	-17.12	11.57
Latin America	2.81	9.64	3.43	.60	-.17	-25.49	24.22
Emerging Markets	1.62	6.61	4.08	.18	-.43	-18.21	19.20

Source: The International Finance Corporation

The renewed interest in Latin American markets, as well as the openness of these markets to foreign investors, draws our attention to their relations to American and world markets. The American crisis of 1987 suggests that these markets have changed in some fundamental respects, so it may be interesting to re-examine the behaviour of returns in Latin American markets. An investigation of the ways in which American, Latin American and other markets influence each other would be particularly worthwhile.

It would be especially interesting to determine if there is any causality among the Latin American markets and the developed country markets. This may be important in the event that investors could exploit the relationships to explain the behaviour of these markets. The emerging markets are more predictable than the developed markets in the sense that they exhibit strong serial correlation (Harvey, 1995). The research of Aggarwal and Leal (1995), Harvey (1995), Mullin (1993), and Divecha, Drach and Stefek (1992) has shown low correlations between these markets, but recent events suggest that this may be changing. The low correlations show important gains from diversification for foreign investors. The ability to predict well may generate abnormal risk-adjusted gains.

Causality among developed markets has been extensively examined by Eun and Shim (1989) and Hamao, Masulis and Ng (1980), among others. The term "causality" as it is used here follows the usage of Granger. That is, we examine which series precedes the other and which follows if they are or not contemporaneous (Granger and Newbold, 1977). This definition does not imply that one series determines another, since both can be affected by a third non-observed series. We say that one market "Granger causes" another if the first series precedes the other series. It has been observed that the American market precedes or "causes" several other developed country markets, for example Japan, Canada and the United Kingdom.

The number of studies of emerging markets has been increasing rapidly. However, the vast majority limit themselves to the study of Asian markets and especially to correlation among these markets. In general, studies of this type assume implicitly that relationships between markets are linear, and that markets are completely integrated, which is obviously not the case. An incomplete list of recent works in this area would include Aggarwal and Leal (1995), Mullin (1993), Divecha, Drach and Stefek (1992). Speidell and Sappenfield (1992), Cheung and Ho (1991) and Bailey and Stulz (1990). The general conclusion from this literature is that correlations between emerging markets and markets in the United States and Japan have been increasing, but these increases have not been sufficient to allow for important gains in diversification. It is clear that larger correlations do not necessarily imply greater integration, since common factors can affect two different markets in similar ways. One possible explanation is increased capital flows between countries. Despite growing interest in the emerging markets, investment in international assets by large American pension funds has been restricted to 5% of their portfolios. This indicates enormous potential for growth in these capital flows in the future (Errunza, 1994).

Other authors suggest that low observed correlations may be due to non-linearity in the relationships between markets (Mullin, 1993: Cheung, 1993). The correlations are not stable over time (Agggarwal e Leal, 1995; Cheung, 1993) and the transmission mechanisms are affected by the volatility of the developed markets (Bekaert e Harvey, 1995; Aggarwal, Inclan e Leal, 1995).

Corhay, Rad and Urbain (1994), Chan, Gup and Pan (1992). DeFusco, Geppert and Tsetsekos (1994), Leal and Austin (1996) use unit root and co-integration to test the efficiency of several Asian markets. In general they find support for the hypothesis of independence between Asian markets and developed markets. The developed market indices cannot be used to predict the behaviour of these markets. Other authors, using the methods of ratio of variances and vector auto-regression, conclude that no significant relationships exist between Asian and developed markets. See Lo, Fung, Chen, Lai (1993), and Park and Fatemi (1993).

It seems that it is impossible to use an index for a developed market to predict the behavior of an emerging market. If there is any relationship, either it is non-linear or intra-day. There is also evidence that these markets only move together in periods of extremely volatility, as was the case during the Gulf War and the Crash of 1987 (Aggarwal e Leal, 1995).

In the case of Latin American markets, the empirical evidence is much more limited. Aggarwal, Inclan and Leal (1995) study the relationships between several emerging markets by examining variances and sudden changes in these variances. They conclude that the variance does not seem to change smoothly over time, but rather institutional and economic changes can generate sharp changes in its behaviour. After taking this into account in an ARCH model, variance can be considered to be constant. After controlling for sudden changes in variance, there does not seem to be any strong relationship between these markets. Aggarwal and Leal (1995) arrive at a similar conclusion using much simpler econometric models. It has also been observed that the emerging market returns are more predictable than those in developed markets. Harvey (1995) points out that the source of this predictability could be time-varying risk exposures and/or time-varying risk premiums. This predictability could also be the result of fundamental inefficiencies.

The American market affects other developed country markets through the transmission of innovations in the market index. The other markets, on the contrary, do not transmit their innovations to the American market. Hamao, Masulis and Ng (1989) find evidence of the transmission of volatility from the American market to the Japanese and British markets, but not in the reverse direction.

Other authors have documented the fact that the American market affects Asian markets, whereas other developed markets, including the Japanese market, do not seem to affect Asian emerging markets in any significant way (Liu, Pan and Hsueh, 1994; Leal and Austin, 1996). Roll (1992) shows that many of the relationships between markets can be explianed by developments in the international sector. There are high correlations between data for enterprises of the same sector in different countries. However, Heston and Rouwenhorst (1994) have suggested that factors affecting sectors of the economy are dominated by domestic factors in every country.

Some studies have tried to create models to explain the variabilty of returns and volatility across countries. Harvey (1995), for example, showed that volatility of emerging markets can be explained by the concentration of certain economic sectors in markets. However, he was unable to build a model that would predict changes in returns in any significant way. Harvey used several factors, including petroleum prices, returns from American markets, international interest rates, and unexpected changes in inflation and exchange rates in an arbitrage model and concluded that the only significant variables were American market returns and exchange rates.

Even so, the explanatory power of the model is low. Mullin (1993) suggests that returns in Asian countries appear to be related to the performance of exports, and volatility seems to be related to changes in exchange rates and the inflation.

In terms of the best way to model the time series in these markets, there is a trend towards considering the ARCH/GARCH family of models to be most appropriate. This type of model permits autocorrelation in the series. allowing the variance to change in a smooth way. It also allows asymmetries in the distribution of returns (EGARCH), and for the distribution to be leptokurtic with fat tails by modelling returns with a Student's t distribution instead of the normal distribution. The process of resolution of these models through numeric iterative methods is dynamic and non-linear. Among the many papers that favour the use of these models in examining emerging markets is Errunza, Hogan, Kini and Padmanabahn (1994). Others studies for specific markets include Woo, Lai and Cheung (1995) for the Thailand and Hong Kong, Nicholls and Tonuri (1995) for Australia, Hargis (1994) for Latin America. and Sewell, Stansel, Lee and Pan (1993) for Taiwan and Korea.

A representative sample consisting of the biggest markets in terms of capitalisation in Latin America are Argentina, Brazil, Chile and Mexico. We present some statistics for emerging countries below. According to the International Finance Corporation (1995), these countries were collectively responsible for 94.3 percent of the $450 billion US total capitalisation of the Caribbean and Latin America markets in 1994. Tables 1.8 and 1.9 below present a summary of descriptive statistics for these markets. Brazil was the largest market in the region, followed by Mexico. The Brazilian market was 20 times smaller than the American market. In 1994, the Brazilian market was the 15th largest in the world and the Mexican market was 21st.

The capitalisation of the emerging markets in Latin America and the Caribbean was 3.0 percent of world capitalisation and 23.3 percent of the capitalisation of emerging markets, also according to the International Finance Corporation (1995). The level of concentration of a market is related to volatility (Harvey,1995b). Chile was the most concentrated and Mexico the least concentrated. Brazil seemed to be the most under-valued market, not only in terms of price-profit but also in terms of market value relative to patrimonial value. The 1994 figures show that Brazil was the cheapest among the Latin America markets. Tables 1.8 and 1.9 also show summary statistics for monthly returns in US$ dollars. Looking at the coefficient of variation for monthly returns, we can see that Argentina's market was the one that presented the most risk to investors for each percentage point increase in returns, and that the Chilean market was the safest for the foreign investor. That is, the degree of volatility was higher in Argentina and lower in Chile for the period 1982-1995. Brazil also showed a very high degree of volatility, and as the returns are in U$S dollars, the coefficients of variation are not that bad. Taking into account only the standard deviation as a measure of volatility, the of US$ dollar returns for Argentina's and Brazil's markets were at least 5 times more volatile than that those for the United States and other developed countries. The volatility of the Mexican market was nearly 4 times greater than that of the US market, while the Chilean market was only 1.8 times more volatile. In dollar terms, the Latin American markets are more than 2 times more volatile than the developed markets and about 1.5 times more volatile than the other emerging markets as measured by Morgan Stanley Capital International data. This can be partly explained by the growing inflow of capital from developed countries to emerging markets. This flow of capital is inherently extremely volatile, and with improvements in information technology making the international transfer of funds much easier, the result seems to be an increase in volatility.

Below we have some figures. Values are for U.S. dollar returns and are based on monthly data from January 1976 to June 1992 from International Finance Corporation. Arithmetic and geometric mean returns have an important difference. The arithmetic average is the return to a strategy that requires equal investment in each period. That is, the gains are not reinvested in the market. The geometric mean has a more appealing portfolio interpretation. The geometric average represents the average return to a buy-and-hold strategy. In this strategy, a fixed amount is invested in the first period, and the portfolio is held until the end of the sample period.

Mean (in terms of the US$ dollar returns) in these emerging markets range from 72 per cent in Argentina to -6 percent in Indonesia. This range sharply contrasts with the range of average returns in the industrialised country markets. No developed market has an arithmetic mean that exceeds 25 per cent. In the emerging markets sample, nine markets have returns that average above 25 per cent.

The emerging markets returns are characterised by high volatility, which produces large differences between the arithmetic and geometric mean returns. These differences are especially evident for Argentina, where the arithmetic mean is 72 per cent and the geometric mean is 27 per cent. Volatility, as measured by standard deviations, ranges from 18 per cent to 105 per cent. In contrast, volatilities for the industrial markets range from 15 to 33 percent. 13 emerging markets have volatilities greater than 33 percent.

Prices and returns reveal some of the most crucial differences between emerging markets and mature markets. Many markets in developing countries offer yields far in excess of developed-market returns and offer low-to-negative correlation with world markets. Both of these facts suggest that unexploited profit opportunities may exist. High autocorrelation in returns, characteristic of speculative inefficiency, indicates that lagged prices may contain information about future returns; volatile stock prices also indicate inefficiency. High price volatility in emerging markets may stem from small market effects or from informational imperfections. With few trades occurring, information about stock values (and therefore stock prices) tends to be noisy.

Table 1.8

Means and Standard Deviations of International Equity Returns, 1976-92 (%)
Industrial Markets based on U.S. dollar returns from monthly data

Industrial Markets	Arithmetic Mean	Geometric Mean	Standard Deviation
Australia	15.95	12.17	26.34
Austria	15.20	12.31	24.21
Belgium	18.03	15.80	20.97
Canada	12.44	10.39	19.93
Denmark	14.98	13.13	19.08
Finland	-9.66	-12.17	22.15
France	17.78	14.51	25.26
Germany	15.17	12.73	21.81
Hong Kong	25.45	19.25	33.88
Ireland	12.61	9.72	24.28
Italy	14.68	11.11	26.84
Japan	17.97	15.20	23.38
Netherlands	18.95	17.30	17.53
New Zealand	-1.98	-5.18	26.12
Norway	16.60	12.49	28.41
Spain	10.32	7.32	24.47
Sweden	18.65	15.87	23.24
Switzerland	14.18	12.37	18.74
United Kingdom	19.20	16.50	22.90
United States	14.17	13.00	15.46

Source: The World Bank

Moreover, limited reporting requirements in many markets mean that investors typically have less information about firms and receive less frequent updates than do investors in industrialised country markets. Uncertainty about the financial condition of firms may introduce high variance in expected returns.

The low correlation between returns in emerging markets and industrial markets suggests that the global investor would benefit from diversification in emerging markets. New interest in international investing has been partly caused by the emerging equity markets, which are attractive because of their high average returns and low correlations with industrial markets. So, we have seen from simple comparisons of data from emerging and developed economies that volatility tends to be higher in emerging economies. We will be investigating this point further throughout the thesis.

Table 1.9

Means and Standard Deviations of International Equity Returns, 1976-92 (%)
Emerging Markets

Emerging Markets	*Arithmetic Mean*	*Geometric Mean*	*Standard Deviation*
Argentina	71.66	27.02	105.06
Brazil	22.69	4.71	60.83
Chile	38.65	30.90	39.84
Colombia	45.60	40.27	32.57
Greece	9.75	3082	36.27
India	21.45	17.88	26.87
Indonesia	-6.29	-12.35	34.95
Jordan	10.14	8.53	18.04
Korea	20.02	15.15	31.97
Malaysia	13.56	9.81	26.90
Mexico	30.44	19.02	45.00
Nigeria	2.18	-6.36	37.20
Pakistan	25.65	23.21	22.38
Philippines	51.16	43.23	38.79
Portugal	40.85	29.00	51.43
Singapore	16.72	13.05	26.21
Taiwan	39.93	25.37	54.06
Thailand	47.89	18.11	25.69
Turkey	21.55	22.04	76.71
Venezuela	37.92	26.23	47.52
Zimbabwe	10.92	4.33	34.30

Source: The World Bank

1.6 - Stock Market Volatility definition and principles

As stated above, one of the most important measures of risk in financial markets is volatility, usually understood as the standard deviation of returns in annualised percentage terms. The volatility of an asset can vary greatly over time, and recent volatility in global markets has risen substantially, particularly since the Asian Crisis. However, over the long-term, the volatility of a financial asset tends to hover around an average figure. A general indication of volatilities of various financial investments can be seen in Table 1.10.

Table 1.10 - Volatility of different Investments in 1997

Investment	*Long-Term Volatility*
Gold	7.5 %
US Stocks	10.5 %
US Dollar	11.5 %
French Stocks	15.5 %
Japanese Stocks	19 %
Hong Kong Stocks	22 %
Silver	22 %
IBM Stock	28 %
Mexico Stocks	30 %
Russian Stocks	52 %
Netscape Stocks	75 %

Source: Bloomberg

Note: The above volatility estimates are crude estimates based on historical analysis over periods of about 1 year. Volatility is defined as the annualised standard deviation of the natural logarithms of the consecutive price relatives of the financial instrument.

In practice, volatility is measured by the degree of fluctuation in share prices during the previous 12 months based on the last 52 weekly values. Volatility is then calculated as the standard deviation of the price, and is a measure of its dispersion around the 12-month average.

In this common definition, the absolute value of the returns is preferred to the more usual squared value or more generally to any power. This is because the former quantity better captures the autocorrelation and the seasonality of the data (Granger and Ding, 1993; Müller et al., 1990 and Taylor, 1986). This greater capacity to reflect the structure of the data can also be easily derived from the non-existence of a fourth moment in the distribution of the price changes.

Although this definition is the most appropriate for the assessment of risk and for forecasting, one might prefer other definitions of volatility that would give more weight to the tails of the distribution. For instance, the cubed root of the third moment could be used for the evaluation of extreme downside risk in portfolio optimisation, as in Roy (1952). One might also prefer the use of conditional volatility such as defined by the option model (Cox and Rubinstein, 1985) or the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model (Bollerslev, 1986).

However, although both approaches might hold some appeal in the case of daily frequency, their use in cases of intra-daily frequency presents important drawbacks. On the one hand, the implicit volatility cannot be computed at very high frequency since options are not quoted at such frequencies. On the other hand, as a consequence of the heterogeneity of the stock market, intra-daily stock indices cannot be described by one homogeneous GARCH model (Guillaume et al., 1994a).

There are two kinds of volatility: historical volatility and implied volatility. Historical volatility is a statistical measurement of past price movements, and is what people are usually refering to when they use the term 'volatility'. Implied volatility measures whether option premiums are relatively expensive or inexpensive. This type of volatility is calculated based on currently traded option premiums. Ideally, what traders would like to know is what volatility is going to be in the future. If we knew what future volatility would be, we could make a fortune quite easily. Because we do not in fact have this information, we make an educated guess. The starting point for this guess is historical volatility. One would ask, "What has been the volatility of this stock or security over a cetain period of time?" Black and Scholes (1973) have built their model based on the idea of implied volatility.

When evaluating volatility, we may look at several different periods. We may look at what the volatility has been for the past week, or we may choose some other period. A longer time period will give a better idea of average volatility. Stocks or other securities that are volatile on a daily or weekly basis usually remain that way over time. When evaluating the purchase of an option, it is the historical volatility of the underlying security we are looking at.

However, there is a different interpretation of volatility that is not associated with the underlying security. This is implied volatility. However, what if we use our historical volatility formula and come up with a price that is very different from the range where the option is currently trading? We are all using the same inputs. We all use the price of the underlying security, the time until expiration, the strike price, dividends to be paid by the stock, the current risk free interest rate, and volatility. The only factor that is not known, and for which we have to take a guess, is volatility. What has happened is that the marketplace is assuming a different volatility than historical volatility. The way to solve for this implied volatility is to use the option pricing model in reverse. We know the price of the option and all the other variables except the volatility the marketplace is using. Therefore, instead of using the equation to solve for the option's price, we use the model to solve for the option's volatility. If we insert the price into the model, leaving out volatility (which is what we are looking for) and keeping the other variables constant, we will find the level of volatility that would yield the current market price.

The first thing that one thinks about when trying to evaluate historical volatility is that the standard deviation should be used. If a person is looking for a simple way to measure volatility, the simple standard deviation will work well enough. However. use of the standard deviation assumes that there is a normal distribution. If stock prices were normally distributed, the implication would be that there could be negative prices. This assumption is a very strong one. In reality, the furthest a stock's price can fall is to zero, but it can rise infinitely. To account for this asymmetry we can take the standard deviation of the logarithmic price changes measured at regular intervals.In what follows we will discuss the treatment of volatility in a more precise way.

$$s' = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\ln X_i - l n\bar{X})^2} \qquad (1.6)$$

However, in the next section we will be using the definition of volatility as defined in (1.6), above.

1.7 - Facts about Volatility in the US Market and Emerging Economies

According to Schwert (1990) investors, regulators, brokers, dealers, and the press have all expressed concern over the level of stock market volatility. However, the perception that prices move a lot, and have been moving a lot more in recent years, is in part a reflection of the historically high levels of popular indices. The drop in stock prices on October 13, 1989, while large in terms of a point decline, was not even among the 25 worst days in NYSE history in terms of percentage changes, and this is also the case with the impact of the Asian crisis. While a 6 percent drop in prices is not inconsequential. neither is it a rare event when considered within the context of the behaviour of stock returns over the 1885-2000 period.

Apart from October 1987, October 1989, September 1998, October 1997 and March 2000, volatility was not particularly high in the 1980s and 1990s. Moreover, the growth in stock index futures and options trading has not been associated with an upward trend in stock volatility. There is little evidence that computerised trading per se increases volatility, except perhaps within the trading day.

On October 13, 1989, all major networks flashed reports on the market decline. The ability of investors to track stock prices on a continuous basis has raised a public question of a volatility problem. As intra-day data on stock prices are simply unavailable, do the large but extremely brief price drops that have characterised recent market declines also occurred in the past? The growing interest in addressing this question can be seen in the profusion of recent papers on the subject.

The evidence to date as to whether trading halts or circuit breakers can reduce volatility in a beneficial way is inconclusive, in spite of the fact that these mechanisims are widely used. They were introduced in Brazil in 1996 and have been used quite a few times. Even if circuit breakers do reduce volatility, the important question to ask is whether the benefits of stability are greater than the inefficiency costs associated with trading halts. Despite the fact that this question remains unanswered, all major stock markets have adopted similar mechanisms.

The stock market crash of October 19, 1987, and the drops in stock prices of October 13, 1989 and October 1997 left many people wondering whether stock prices have become too volatile. Since the 1987 crash, numerous studies have looked at the effects of modern investment techniques on the volatility of stock prices. Various means of limiting volatility have been proposed since then. The evidence so far indicates that the volatility of rates of return to broad portfolios of NYSE common stocks has been unusually high in the 1990s.

Volatility has appeared high to many people because the level of stock prices is much higher than it has ever been. Thus, while there have been large absolute changes in the level of the Dow Jones Industrial Average (DJIA), in percentage terms these changes have only been moderate.

There is little evidence that the level of stock return volatility has increased since the start of index futures and options trading in the early 1980s. Although high volatility has been associated with high levels of trading in stocks, futures and options, it is unclear whether the large volume of trading has caused the high volatility, or whether the high volatility and trading volumes reflect the arrival of important information.

The remarkable technological advances in the computer and communication industries (the so-called 'new economy') and the globalisation process have made it much easier for large numbers of people to learn about and react to information more quickly. They have also made it possible for financial markets to provide liquidity for investors. These changes have had two imporant by-products. First, there are large incentives for investors to obtain and act on new information. Second, because new information spreads more quickly, the rate at which prices change in response to information has also accelerated. The liquidity of organised securities markets plays an important part in supporting the value of traded securities, but it also means prices can change quickly. From this perspective, volatility is a symptom of highly liquid securities markets.

On April 14, 2000, the Dow Jones Industrial Average (DJIA) fell from 10,923.6 to 10,305.8, over 617 points. This was the largest one-day drop since Dow Jones began computing index numbers in 1885, but the 5.66 per cent drop was not the largest in percentage terms. Nevertheless, most public attention focused on the absolute size of the drop. The 508 point drop on October 13, 1987 also caused a large public reaction, as it represented 22 per cent drop in value.

By focusing on the absolute level of the DJIA, we can exaggerate the severity of recent volatility. For example, the DJIA reached 509.76 for the first time on March 19, 1956; prior to that date; it would have been impossible for the index to drop 508 points. Alternatively, the DJIA fell "only" 38 and 31 points on October 28 and 29, 1929, yet these are the second and the third largest daily percentage drops in the history of the NYSE.

Keeping this distinction in mind, the reaction to the decline on October 13, 1989, which was not even among the 25 largest percentage drops in stock prices, would not appear to be justified. Table 1.11 shows the 25 highest and lowest daily returns to the DJIA between January 1900 and April 2000. If we also compute the 25 highest and lowest monthly returns from January 1900 through April 2000, we see that many of these extremes occurred during the Great

Depression. October 1987 is only the fifth lowest return in the sample, and the recent episodes of the Asian crisis appear but are surpassed by earlier events.

The highest and lowest returns tend to be clustered in brief sub-periods over the whole period, indicating an increase in stock price volatility during these periods. The recent episode of the Asian crisis does not even appear in the best/worst list, as we might expect *a priori*. From a visual inspection of table 1.11, we can see that the recent episodes of volatility, including the market crashes in Thailand (1997), Russia (1998) and Brazil (1999) have not been among the most severe in history despite the large amount of public attention and media coverage they have received.

As stated, the most commonly used measure of stock return volatility is standard deviation. This statistic measures the dispersion of returns. Financial economists find it to be useful because it summarises the probability of seeing extreme values of returns. When the standard deviation is large, the chance of a large positive or negative return is large also.

Table 1.11

THE 25 HIGHEST AND LOWEST DAILY PERCENTAGE RETURNS TO MARKET 1900-2000

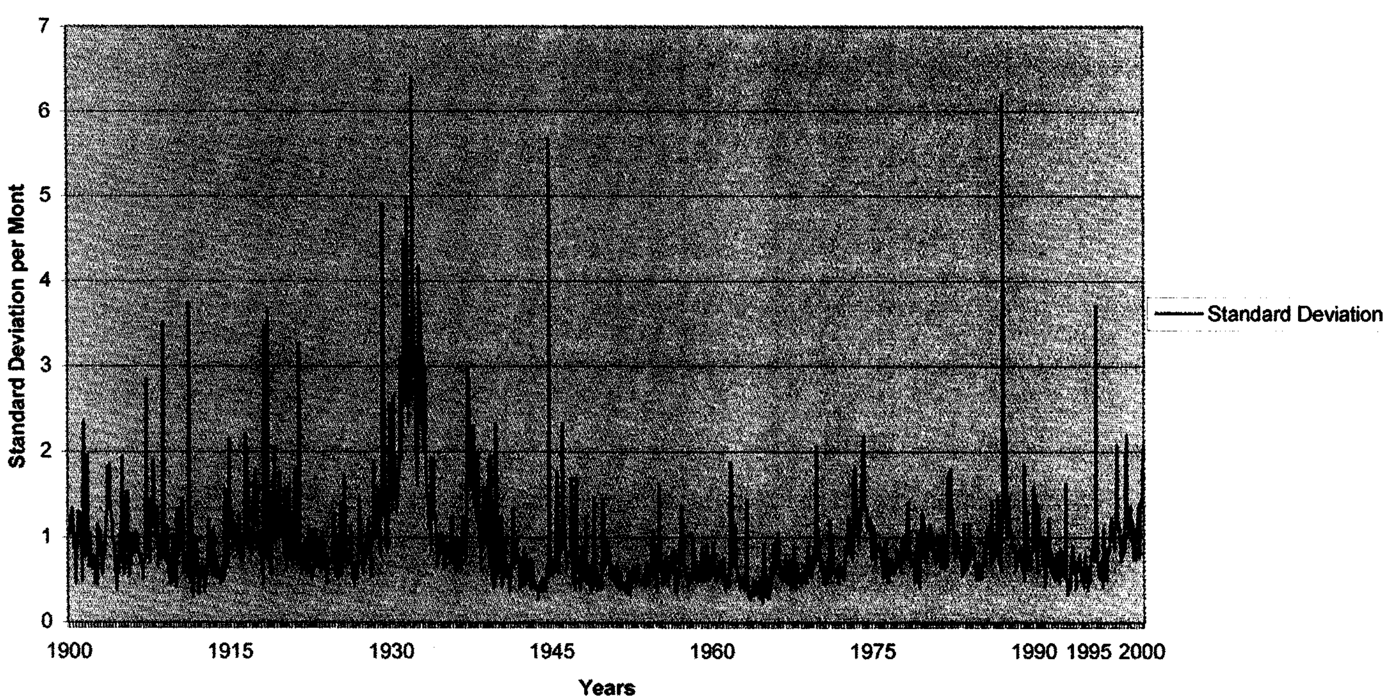
Date	Percentage Change	Date	Percentage Change
October, 19 1987	-22.61	March, 15 1933	+ 15.34
October, 28 1929	-12.82	October, 30 1929	+12.34
October, 29 1929	-11.73	June, 12 1996	+11.83
June, 11 1996	-10.89	September, 21 1932	+11.36
November, 6 1929	-9.92	October, 21 1987	+10.14
October, 10 1937	-9.27	August, 3 1932	+ 9.52
July, 20 1933	-8.88	December, 18 1931	+ 9.35
July, 21 1933	-8.70	November, 14 1929	+ 9.35
October, 5 1932	-8.28	April, 20 1933	+ 9.03
March, 5 1907	-8.28	October, 6 1931	+8.70
October, 26 1987	-8.04	October, 8 1931	+ 8.59
August, 12 1932	-8.02	February, 13 1932	+ 8.37
May, 31 1905	-7.84	February, 11 1932	+ 8.27
July, 26 1934	-7.83	July 24, 1933	+ 8.14
May, 14 1940	-7.47	June, 10 1932	+ 7.66
September, 24 1931	-7.29	June 3, 1931	+ 7.54
October, 27 1997	-7,18	November, 10 1932	+ 7.51
September, 12 1940	-7.18	October 20, 1937	+ 7.48
May, 9 1901	-7.02	June 19, 1933	+ 7.23
June, 15 1933	-6.97	May, 6 1932	+ 7.21
January, 8 1988	-6.95	April 19, 1933	+ 7.21
January, 26 1907	-6.91	August 15, 1932	+ 7.20
October, 16 1933	-6.78	October, 11 1932	+ 7.17
September, 3 1946	-6.73	January 6, 1932	+ 7.02
May, 28 1962	-6.68	October,1 1932	+ 6.90

Source: The Wall Street Journal

Figure 1.4 plots the standard deviations of monthly returns to the DJIA from 1900-2000. Daily returns are used to calculate the standard deviation for each month. There are 1200 standard deviation estimates and the average value for the whole period was 1.1831%. Months like October 1929, October 1987 and November 1997 appear as periods of high volatility. It is also clear that, except for 1987 and 1997, the 1980s and 1990s have not been a period of unusually high volatility. Except for the last three months of 1987 and November 1997, the 1980s and 1990s do not stand as being a period of very high volatility. October 1989 has a lower deviation than the 1973-74 bear market for example. We must stress that October 1998 was also a period of instability. We can also see quite easily the period of turmoil caused by the Great Depression.

Figure 1.4

Volatility of Monthly Returns using Daily Returns within the Month



Source: Wall street Journal

The industrial average rocketed to a then-record high of 2722.42 in 1987, crashed 508 points on Oct.13, 1987 then clambered upward by 32.5% in 33 months to a new two-day peak of 2999.75 in mid-July 1990. The transportation average crackled to life as take-over fever spread to the airline industry, reaching a record 1532.01 on Sept. 5, 1989. In April 2000 the DJIA crashed, falling by 617 points.

Because of take-overs, and partly because of a host of new computer-guided trading techniques, the stock market was often described as volatile in the late 1980s and 1990s. However, this description of the DJIA is off the mark in at least one respect. Ten years ago, a one-day move of 10 points in the DJIA was noteworthy; a 20-point change made headlines and a 50-point move was unheard of. Today, 50 points up or down from one day to the next means nothing. Many people conclude from these changes that the stock market is becoming more volatile.

The perception of what constitutes a big market move simply has not kept up with the rise of the industrial average to higher and higher levels. And yet the impression that stocks are increasingly volatile seems to be supported by a different set of statistics. For example, nearly everyone would agree that a 2% change in the Dow Jones in one day would qualify as volatile, and such moves are occurring more and more frequently.

According to many investors, the first quarter of 1996 was a volatile period for the US stock market. The 50-point circuit breaker in the Dow Jones Industrials was trigered 31 times in the first quarter alone, more than in all of 1995. And in 1997 this happened 86 times! However, does this necessarily mean that the US stock market is more volatile today than in the past? Despite the apparently wild swings that we have seen this year, the market's volatility is actually quite normal. Coming off a period of historically low volatility in 1995-97, today's market swings may appear to be quite high. However, in a historical context they are actually in-line with the average for most of the post-war period, and are in fact smaller than those in earlier periods.

It is easy to see why investors argue that the US market is more volatile today than previously. Still according to Schwert (1990) with the 50-point circuit breaker kicking in 31 times in 1995, or nearly every other day, and with 171-point drops in the Dow being followed by 110-point gains, investors assume that the market is becoming increasingly volatile. However, Schwert reminded that it is important to keep in mind that the 50-point rule today has a completely different meaning with the Dow at 5600 than it did when it was instituted in 1990 and the Dow at 2900. At that time, a 50-point move translated into a change of 1.7%. Today, it is less than 1%. To put the Dow's wild tumble and subsequent rebound in March into perspective, although they were among the largest point movements ever, they didn't even rank in the top 100 in percentage terms.

With this in mind, it is useful to look at the market a little more closely to see if volatility has actually increased as much as investors and market watchers have contended. As a basis, both daily absolute percentage changes and annualised standard deviation in the Dow Jones Industrials going back nearly 50 years were analysed. Interestingly, volatility was lower during the 1993-1995 period than it has been at any time in nearly 30 years. 1993 marked a low point, with the market moving up or down only 0.40% each day on average. In 1994, volatility did increase to 0.50%, but the three year period between 1993-1995 was still one of the least volatile periods in recent years, with the market experiencing absolute percentage changes of only 0.44% on average.

In terms of standard deviations, the volatility during this three-year period was 9.31%, with 1993 and 1995 exhibiting relatively low levels of only 8.51%. However, this can be considered an anomaly. The last time that the market exhibited such low risk levels was during the 1964-65 period when volatility was a mere 6.81%.

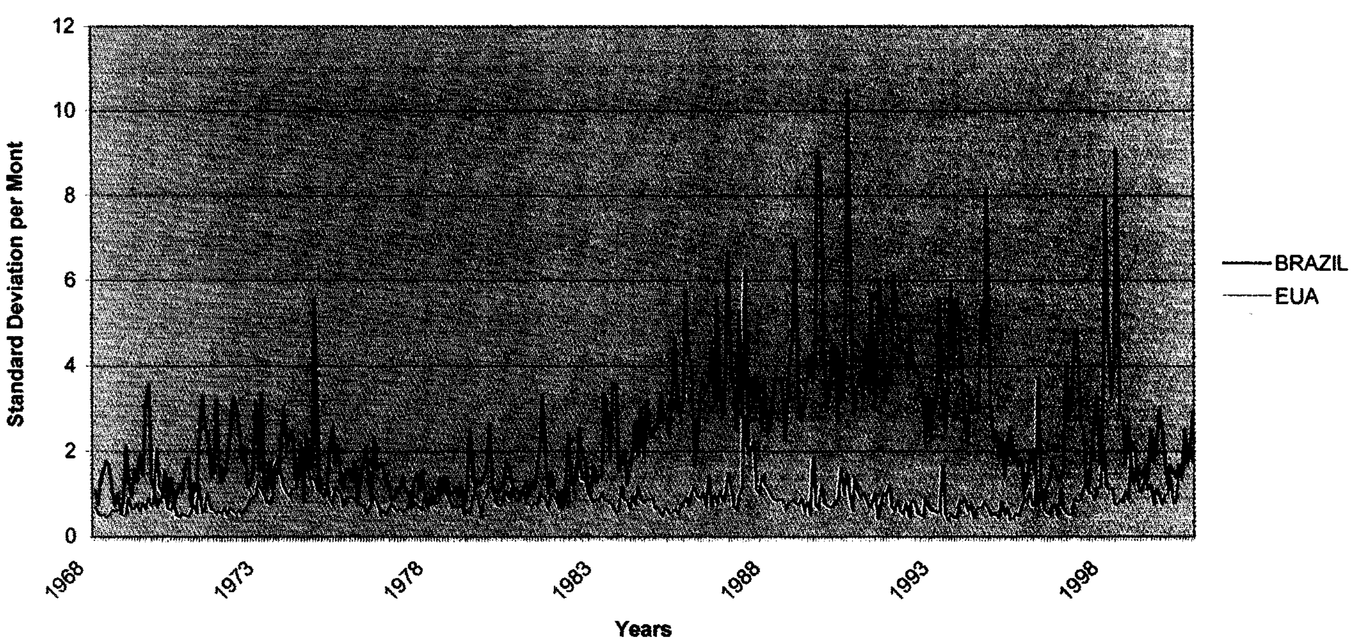
So where do we stand today? Granted, the market's volatility was higher in the first quarter of 1996 than in 1995, with the market moving up or down 0.67% each day in 1996 on average compared to 0.41% in 1995. However, this is only slightly higher than its long-term average of 0.60%, and is roughly in line with the most recent 10-year average of 0.64%.

In terms of annualised standard deviations, the US stock market's volatility in 1996 is in line with historic norms. During most of the post-war period (1950-96), the standard deviation of returns averaged 13.49% compared to 13.75% in 1996. This is actually less than the average (16.44%) over the last ten years.

As we can see from Figure 1.5, which displays the volatility of returns in Brazil (one of the more important emerging markets), volatility as measured by the standard deviation is much larger in the later periods than in the earlier ones. This is quite easily seen in the case of Brazil. Especially in the 1980s, due to the risk of hyperinflation, Brazil's volatility far exceeded that of the United States. This volatility was maintained through several of the Brazilian government's economic plans, and was only reduced after the introduction of the Real Plan in 1994.

Figure 1.5

Volatility of Monthly Returns using Daily Returns within the Month



Source: Bovespa (Brazil) and DJIA (USA)

To summarize the perspective of this thesis, we are interested in using econometric tools in order to measure the effects of volatility in the emerging economies. A result that stands out from an inspection of the data, and one that will be explored further in chapter three, is the fact of clustering of returns. This striking fact was observed for every economy under consideration. A possible explanation for it is the sucession of financial crises that occurred in the 1990s. We will elaborate more on this in later chapters.

What is most striking is the fact that volatility in these markets exceeds by far the volatility of markets in the US, independent of the time period (18 years for Brazil, 7 years for Argentina) or the period of computation (daily or weekly). This is consistent with other studies that found the volatility of Latin American markets to be higher than US volatility (see World Bank Review).

1.8 - Stock Market Volatility- How to explain it? – Econometric Tools

1.8.1 – A First Attempt to treat it – An Alternative View of Volatility

The volatility of stock prices is a well-known phenomenon to all investors. However, why is this volatility frequently so pronounced?

It is apparent that there are extremely wide day-to-day variations in the prices quoted on most stock exchanges. Many people have tried to put forward theories to explain this phenomenon, and more still have tried to use these theories in order to predict future changes in prices. However, most economic theorists ignore the fact that there is no universally accepted body of work explaining what lies behind day-to-day price changes. Instead, they have concentrated on market details in the mistaken belief that the question has already been answered. In fact, theorists cannot agree whether economic or psychological factors are the most important causes of price fluctuations in stock markets. This is an important issue in the study of investment analysis, as it brings into question the whole realm of fundamental and technical analysis, something on which millions of dollars are spent every year.

Among the literature most relevant to the whole issue of volatility is Robert Shiller's *Market Volatility*. Shiller is a firm advocate of the popular model explanation of stock market volatility. Popular models represent a qualitative explanation of price fluctuations. Briefly, the theory proposes that investor reactions due to psychological or sociological beliefs exert a greater influence on the market than rational economic reasoning. It should be noted, however, that Shiller does not totally disregard the work of earlier economists who proposed the Efficient Market Hypothesis (EMH). In fact, he admits that the EMH can be substantiated by statistical data, but he also believes that investor attitudes are of great importance in determining price levels. He claims that substantial price changes can be explained by a collective change of mind on the part of the investing public, which can only be explained by thoughts and beliefs about future events, i.e. psychology.

Schiller's popular model theory proposes that people act inappropriately to information that they receive. Thus, freely available information is not necessarily already incorporated into a stock market price, as the EMH would have one believe. He says that investor behaviour depends on ex-post values, which is the value of an asset taking into account the future dividends. By definition though, ex-post values cannot be known ahead of the payment of dividends, so if future dividends are expected to be high then the ex-post value today will also be high. If investors already knew the future dividend, then forecasting the future price (P_t) would present no problem under the EMH, using ex-post values (P_t^*):

$$P_t = E_t P_t^*$$

In other words, the price is equal to the best possible forecast, or expectation, of ex-post values. Efficient markets theorists claim that the EMH can be used to explain sudden movements in price. For example, new information about dividends could be released. Shiller's argument is that fluctuations are far too big to be accounted for by mere changes in information. He provides statistical evidence suggesting that fluctuations in dividends, due to their nature of being calculated on a moving average, would have to be quite substantial, both in terms of size and length of trend, to resemble observed fluctuations in price. Even during the great depression of the 1930s, dividends were only slightly below their growth path.

In the conclusion of his article, Shiller shows that volatility in stock market prices is five to thirteen times higher than the volatility that could be explained by the EMH and new information. Some efficient markets theorists try to attribute this excess volatility to changes in expected real interest rates. However, Shiller claims that the movements in expected real interest rates needed to explain this excess volatility are far larger that the movements in nominal interest rates over a sample period. The other argument in support of the EMH is that perhaps the fears of investors are greater than the actual changes in price.

Stephen Le Roy and Richard Porter conducted a study that came to virtually the same conclusions as those of Shiller. They published their work in the May 1981 issue of *Econometrica* under the title "The Present-Value Relation: Tests Based on Implied Variance Bounds." In this paper, which was an in-depth statistical study of excess volatility, Le Roy and Porter observed that stock prices based on aggregated and dis-aggregated data are more volatile than the efficient capital markets model would suggest. This conclusion differs greatly from all previous work on the subject by such noted economists as Fama. The importance of the conclusions of this article, the authors suggest, lies in their similarity to results on stock price movements found by Shiller, so neither article's findings can be dismissed as statistical accident.

Not all investors are equally well informed, so insider information can be used to one's advantage as long as no one else is in receipt of the information. This is at complete odds with the strong form of the EMH, which claims that all information, both publicly and privately held, is incorporated into the stock price. Investors may also react differently to the same information.

Risk averse investors might sell as the market becomes bearish, whereas more speculative investors might sell short to gain high profits. Finally, long term 'buy and hold' investors might see a market downturn as a buying opportunity.

In his article "Meltdown Monday or Meltdown Money: Consequences of the Causes of a Stock Market Crash," Mullins claims that the Meltdown Monday explanation is incorrect. This is the name given to the theory that technology and herd behaviour of investors was the cause of the crash. On the contrary, Mullins suggests that investors were reacting rationally to changing economic circumstances. He also cites government intervention in the financial markets as a cause of the crash. Mullins points out that. leading up to 1987, there had been a five-year bull market due to the economic recovery in the US and falling interest rates. From August 25th 1987, it can be seen that prices began to fall at a slow rate until they accelerated into a sheer drop on October 19th. The important point to note is that is that Mullins sees this crash as a rational reaction. This is in contrast to the Meltdown Monday approach.

The author, however, says that it would be inaccurate to view the crash as a bubble bursting. Price movements in the U.S. from 1982 to 1987 have been shown to be random, so there was no bubble to burst. If a bubble were to have burst, surely it would have occurred on August 25th when prices were highest. Thus, this argument has no grounding in statistical evidence. However, an important point to note is that Mullins accepts the possibility that the crash could be explained by investors believing that a bubble existed and fearing that it was about to burst! Indeed, 38% of normal investors felt that the market was overvalued in the week October 12th to 16th, 1987. Thus, we can see that. although Mullins is a supporter of the EMH, Shiller's popular model argument can be applied here. Mullins defends the EMH by claiming that the assumptions about investor rationality, preferences and ability to act are too restrictive. If one relaxes these assumptions then one can examine the fundamentals within the context of the EMH.

1.8.2 – Treatment of Volatility – Econometric Tools

An agent must make decisions based upon the distribution of a random variable, for instance a stock price, at some time in the future. For a risk averse agent, a measure of dispersion would be also be of primary importance, as stated above. Much of the work in conventional econometrics on deriving measures of risk and uncertainty has not addressed this question adequately. In these models, the variance of the disturbance term is assumed to be constant, a convenient but often implausible assumption.

However, many economic series do not fit this description, since they exhibit periods of unusually large volatility followed by periods of relative tranquillity. In such circumstances, the assumption of constant variance clearly does not apply. Also, as an asset holder you would be interested in forecasts of the rate of return and its variance over the holding period. The unconditional variance (i.e. the long-run forecast of the variance) would be unimportant if you were planning to buy the asset at t and sell it at t+1.

The analysis of financial series data usually requires the study of the first and possibly second conditional moments of the series in order to characterise the dependence of future observations on past values. These two conditional moments are often identified with important economic concepts. For example, if we consider the stochastic process for stock market returns, we would have the first two conditional moments given the information set as being μ_t and σ_t^2. In this context, μ_t is usually associated with the risk premium as a whole and σ_t^2 with its volatility. From an empirical standpoint, the first step often consists of estimating of the conditional mean and variance.

It is possible to simultaneously take account of the mean and variance of a series, giving a much more reasonable description of stock prices. The first to introduce this technique was Engle (1982) with the so-called autoregressive conditional heteroskedastic (ARCH) class of models. In the ARCH model, the conditional variance is allowed to change through time as a function of both current and past information. Although this new class of time series models allows for a much wider class of non-linear dynamic econometric models, the linear ARCH(p) has been found to be a particularly useful parameterisations in the modelling of monetary and financial data. We have then:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots\dots\dots + \alpha_q \varepsilon_{t-q}^2 \qquad \text{or} \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \qquad (1.7)$$

A practical difficulty with ARCH models is that, with large p, estimation often leads to violations of the non-negativity constraints that are needed to ensure that the conditional variance is always positive. This has led to the imposition of a rather arbitrary declining lag structure to ensure that these constraints are met. To obtain more flexibility, we can use the GARCH(p,q) model, a generalised form of ARCH that posesses a richer conditional variance function.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots\dots\dots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots\dots\dots + \beta_p \sigma_{t-p}^2 \quad (1.8)$$

A great deal of work using this methodolgy has been applied to both emerging markets and developed economies. Hargis (1994) examines several Latin American markets using various forms of GARCH models, and concludes that these models do have some explanatory power in the period after the liberalisation of market access to foreign investors. Hargis, Aggarwal and Leal (1995) also conclude that when there is an important variation in the American market, such as during the Gulf War in 1991, all markets tend to move together. These joint movements are strongest when the events that cause them are perceived to be negative. Several other papers examining the relationships between emerging markets using ARCH techniques or models of co-integration suggest that links between markets are weak, particularly for Latin American markets, and that several of these markets appear to be segmented.

There is a trend towards considering the models of the ARCH/GARCH family to be the most appropriate for modelling volatility. This type of model allows for autocorrelation in the series, smooth changes in variance over time, asymmetries in the distribution of returns (e.g. EGARCH) and leptokurtic distribution, modelling fat tails with a Student's t distribution rather than using the Normal. The process of resolution of these models is dynamic and not linear through numeric iterative methods.

1.9 Conclusion

The stock market crash of October 1987 was a traumatic event, raising fears of financial collapse and depression, and reviving the spectre of the 1930s. Little wonder then that the volatility of stock markets subsequently came under intense scrutiny by governments, market professionals and academics. Of central concern was the issue of what causes volatility: if volatility were predictable, then steps might be taken to reduce it. We have seen that the market volatility of the 1980s and 1990s has not been unusually high by historical standards. We have also found that volatility increases during recessions and major financial panics, that it is lower in bull markets than in bear markets, and that shocks to returns have a persistent influence upon volatility. The evidence also shows sub-periods for which volatility was markedly higher than in other periods. It is possible to characterise part of this volatility, the deterministic part. This predictable component of volatility, taken to be the conditional variance of stock returns, can be identified and calibrated using a variety of models of the ARCH/GARCH type. Modelling predictable volatility allow us to distinguish periods of unusual volatility. These models show that volatility was noticeably higher during various financial crises. the Depression years of the 1930s, both World Wars, and for much of the past two decades.

The first thing we notice when looking at the data is that the market volatility in recent years has not been higher than it was in the past, although we can see some periods where volatility increased substanstially. This has already been extensively stated for the United States, and is a fairly good description of what has been happening. The second notable feature is that the emerging market returns are indeed more volatile than those in the US, a phenomenon that deserves careful examination. Hardly any work has been done to characterise the behaviour of returns in these markets, and any additional knowledge would represent important progress in our understanding of stock markets.

This first chapter has shown that the conventional wisdom that volatility has increased in recent years cannot be confirmed on a non-technical basis. No matter what conclusion we ultimately reach, we have seen that price movements in emerging markets have been much more pronounced than those observed in US markets. Bearing this observation in mind, we will measure this effect in a more technical fashion in subsequent chapters. In particular, we will make use of the FIGARCH approach that has been used as a tool to assess the long memory of volatility. We hypothesize that if researchers have succeeded in finding significant evidence for long memory in developed markets, then similar evidence should also be found for emerging markets.

Appendix to Chapter One

The distribution of the Stock Price Index P_t

We can perform a simple experiment to determine what kind of disitribution we should expect from the price P_t. The experiment consists of listing the deviations of prices (increases or decreases of prices) observed during a given time period and drawing a histogram. We could do it for tomorrow's prices, for prices over the next 30 days, or any number of days we care to choose.

It is wrong to suppose that the average deviation of stock prices should be an absolute value for any level of prices. For example, a stock that varies US$ 10 per day on average when its price is US$ 100 cannot be expected to vary the same US$ 10 when its price is US$ 1, since prices can never be negative.

A second experiment involves examining relative changes in prices (as opposed to the absolute changes considered previously) to see if they follow some specific distribution. It is not unreasonable to suppose that a 1% rise and a 1% fall in prices might be equally likely to occur. However, it is difficult to imagine what magnitude of fall would have the same probability of occuring as a 120% increase in nominal terms.

The only way to solve for the negative price problem is to suppose that prices P_t have the same probability of varying upwards or downwards by some factor, 1.01 for example. In this case, the price P_t has equal probability of rising to 1.01 times its initial value or falling 1.01 times. A rise of 100% (2 times intial value) might have the same probability as a fall of 50% (a reduction by 2). Even for very big factors we would never predict negative prices. A simple mathematical expression for a distribution in which multiplication and division by a factor are equally likely is the distribution of the difference in the logarithm of the prices. An increase by a factor of 1.15 (i.e. an increase of 15%) has the same probability as a fall by 1.15 (13,04%). This happens because the distance between the logarithm of 100 and 100 / 1.15 = 86.96 is:

$$\text{Ln}\,(115) - \text{Ln}\,(100) = 0.1397$$
$$\text{Ln}\,(100) - \text{Ln}\,(115) = 0.1397$$

It is this difference that we want to represent as a random variable. We can find the same number by computing the logarithm of the returns (or factors):

$$\text{Ln}\,(115/100) \quad = 0.1397$$

$Ln (100/86.96) = 0.1397$

In order to derive the equivalent factor from this logarithmic difference, we need to raise the number e (2.71881...) to the difference. To get the difference from the factor, we take the neperian logarithm of the factor. For very small numbers (differences up to 0.05), the factor is approximately the logarithmic difference plus 1, and the percentual change in P_t upwards or downwards is apprximately equal to the difference itself.

Table 1

Logarithmic Difference	Factor	Fall Change (%)	Hike Change (%)
0.01	1.010	1	1
0.02	1.020	2	2
0.05	1.051	4.9	5.1
0.10	1.105	9.5	10.5
0.20	1.221	18.1	22.1
0.50	1.649	39.3	64.9
0.70	2.014	50.3	101.4
1.00	2.718	63.2	171.8

Log-Normal Distribution

From a series of logarithmic differences of P_t (or of a series of logaritmic returns of P_t), we may arrive at the conclusion that this series is normally distributed. In this event we would say that stock prices P_t follow a log-normal distribution. In that case, what measure of dispersion is to be used? This measure is the standard deviation of the distribution. If we assume that the returns of P_t are not serially correlated and that the variance of daily returns is constant, our experiment is simplified. Instead of taking the differences observed in periods equally spaced over thirty days, we can take the differences from one day to the next, as the variance computed for 30 days will be equal to 30 times the variance for one day.

If we have R_{mt} = the monthly return of the Price P_t counting from date t and assuming that we have 21 working days (and 21 daily observations) within a month, $Ln (R_{mt}) = Ln (P_{t+21}/ P_t) = Ln[(P_{t+1}/ P_t) \times (P_{t+2}/P_{t+1}) \times \ldots\ldots\ldots \times (P_{t+21}/P_{t+20})$. If we use the operative properties of the logarithm, we would have, then:

$Ln(R_{mt}) = Ln(P_{t+21}/P_t) = Ln(P_{t+1}/P_t) + Ln(P_{t+2}/P_{t+1}) + \ldots\ldots + Ln(P_{t+21}/P_{t+20})$. If we write the daily return of the stock price P_t in date t as R_{dt}, we would have that $R_{mt} = R_{dt} + R_{dt+1} + R_{dt+2} + \ldots + R_{dt+21}$. The variance of the monthly returns can then be written as $Var(R_{mt}) = Var(R_{dt} + R_{dt+1} + R_{dt+2} + \ldots + R_{dt+21})$. If we assume that daily returns are not serially correlated, then $Var(R_{mt}) = Var(R_{dt}) + Var(R_{dt+1}) + Var(R_{dt+2}) + \ldots + Var(R_{dt+21})$ and also under the hypothesis that daily variances are constant and equal to s^2, we would have finally that $Var(R_{mt}) = s^2 + s^2 + s^2 + \ldots + s^2 = 21\, s^2$ and the standard deviation would be equal to the square root of the variance $\sqrt{21s^2} = s\sqrt{21}$. That is, the standard deviation of the distribution of differences for a period of 30 days would be $\sqrt{30}$ = 5,48 times the standard deviation of the daily distribution for the differences. If the 30 day standard deviation of the logarithmic difference of the stock price is 0.1397, there is a probability of 68% that the price P_t would be between $e^{0.1397}$ = 1.15 times below and $e^{0.1397}$ = 1.15 times above its most likely value (between 100/1.15 = 86.96% and 100 x 1.15 = 115%) or between a fall of 13.04% and an increase of 15%.

Volatility

The standard deviation of a log-normal distribution of prices is defined as its volatility, represented by σ and indicated in percent terms. A volatility of 20% means that the standard deviation of the logarithmic differences is 0.20. A price P_t that has volatility for 30 days equal to 20% means that there is a probability of 68% of it being between $e^{-0.20}$ = 0.82 times (or 18% below) and $e^{0.20}$ = 1.22 times (or 22% above) its most likely value.

There are 22 dates, but only 21 returns. By definition the variance of a random variable X, of an infinite population, is:

$$\sigma^2 = V(X) = E[X - E(X)]^2 = E(X - \mu)2$$

Dividing the sum of square of deviations by the number of observations, we have the variance of the stock price:

$$Var(X) = \frac{\sum e_i^2}{N} = \frac{\sum (X_i - \bar{X})^2}{N}$$

If X_1 , X_2 ,......, X_N are the observed values in a sample, the non-biased estimator of VAR(X) is:

$$s^2 = \frac{\sum e_i^2}{N-1} = \frac{\sum (X_i - \bar{X})^2}{N-1}$$

But $\sum (X_i - \bar{X})^2 = \sum X_i^2 - \frac{1}{N}(\sum X_i)^2$

The standard deviation is by definition the square root of the variance, so the historical volatility (standard deviation) can be computed as:

$$\sigma(\text{day}) = \left(\frac{\sum \ln_i^2 - \left(\frac{\sum \ln i}{N} \right)^2}{N-1} \right)^{1/2} = \left(\frac{0.010771 - \left(\frac{0.067101}{22} \right)^2}{22-1} \right)^{1/2} = 0.022975$$

In this case the annual volatility is σ(year) = $\sqrt{252} x \sigma(day)$ = 0.3647 = 36,47%, where 252 is the number of working days in a year.

Table 2 – Historical Series of Telebras PN

Date	Day	Stock Price	$1+R_{dt}$	$Ln(1+R_{dt})$	$Ln(1+R_{dt})^2$
02/18/1997	1	102.30	-	-	-
02/19/1997	2	103.09	1.0077	0.0077	0.0001
02/20/1997	3	99.20	0.9623	-0.0385	0.0015
02/21/1997	4	100.20	1.0101	0.0100	0.0001
02/24/1997	5	104.10	1.0389	0.0382	0.0015
02/25/1997	6	106.10	1.0259	0.0256	0.0007
02/26/1997	7	103.60	0.9700	-0.0304	0.0009
02/27/1997	8	101.60	0.9807	-0.0195	0.0004
02/28/1997	9	102.60	1.0098	0.0098	0.0001
03/03/1997	10	105.20	1.0253	0.0250	0.0006
03/04/1997	11	104.60	0.9943	-0.0057	0.0000
03/05/1997	12	106.80	1.0220	0.0218	0.0005
03/06/1997	13	108	1.0103	0.0102	0.0001
03/07/1997	14	108.70	1.0065	0.0065	0.0000
03/10/1997	15	112.30	1.0331	0.0326	0.0011
03/11/1997	16	115.10	1.0249	0.0246	0.0006
03/12/1997	17	111.90	0.9722	-0.0282	0.0008
03/13/1997	18	112	1.0009	0.0009	0.0000
03/14/1997	19	115	1.0268	0.0264	0.0007
03/17/1997	20	111.40	0.9687	-0.0318	0.0010
03/18/1997	21	110.40	0.9910	-0.0090	0.0001
03/19/1997	22	109.40	0.9909	-0.0091	0.0001
Σ				0.0067101	0.010771

Chapter 2

Review of the Literature of Long Memory and Volatility

2.1 - Introduction

In this chapter we will be reviewing a variety of models that have become commonplace in the literature dealing with volatility (i.e. the second moment or variance). In spite of the fact that these developments are relatively recent, having been introduced only in the last twenty years, a great deal of research has already been conducted. An extensive review of this literature is included here, with the objective of showing the evolution and the variety of these models. The first model dealing with variance, ARCH, is discussed, including its inability to account for persistence in variance. Other developments, up to and including the recent FIGARCH model, are also discussed. Our objective is to review and define the models that will be estimated in the next chapter. As we will see, the vast majority of the results have been established for the developed markets, but little evidence has been shown for emerging markets. The review is structured so that we begin with the simpler models of variance, and then build up to the more elaborated ones that constitute the focus of this thesis. By proceeding this fashion, we will show the shortcomings of the earlier models, as well as the ways in which professionals and researchers have overcome them.

This review does not pretend to add anything new, but rather aims at being a comprehensive survey of the literature on volatility models from a more technical standpoint than that of the overview supplied in the first chapter. We will see that there are in fact many different ways of defining volatility beyond the definition used thus far. The review also demonstrates that, while is useful to understand what has happened with volatility in the past, it is equally important to be able to forecast it in order to anticipate future movements in stock prices. We will also attempt to show that the concern about what is going on in the stock market is not limited to the developed world. Our ultimate objective is to learn more about emerging markets by discovering the extent to which their behaviour conforms with existing findings for developed markets. In this sense, the review provides hints about how one should proceed when dealing with data from emerging markets.

We begin with the earliest developments in this area: the autoregressive conditional heteroskedasticity model (ARCH) and an extension of this methodology, the GARCH model (G standing for "generalised"). We then discuss the inability of ARCH/GARCH models to deal with persistence in volatility, a result established by several authors. We show that the development of long memory models in volatility follow the development of long memory in the first moment through the ARFIMA specification.

2.2 - Autoregressive Conditional Heteroskedasticity (ARCH) and Generalised ARCH (GARCH)

The history of ARCH is a very short one. The technique was introduced by Robert Engle in 1982, and the ARCH literature has grown in spectacular fashion since then. As we have seen, much attention has been paid to its application to stock markets. The development of the ARCH methodology defies the general trend in scientific advancements, where applications usually lag behind theoretical developments. Engle's original ARCH model and its various generalisations have been applied to numerous economic and financial data series in many countries, but the models have experienced relatively few theoretical refinements. Nowadays, with the increased availability of high-frequency financial series, interest in ARCH has been renewed.

Bachelier (1900) was the first to conduct a rigorous study of speculative behaviour in stock markets. After Bachelier, it appears that the issue was completely forgotten for some time. Mandelbrot (1963a,b, 1967) revived interest in the time series properties of asset prices with his theory that "random variables with an infinite population variance are indispensable for a workable description of price changes" (cf. 1963b, p. 421). His observations, such as the fact that unconditional distributions have thick tails, that variances change over time and that large (small) changes tend to be followed by large (small) changes of either sign, are 'stylised facts' for many economic and financial variables. Even today, too few models can convincingly account for these observations.

The first noticeable thing about time series data dealing with stock markets is that the mean appear to be constant while the variance change over time. Prior to the introduction of ARCH, researchers were very much aware of changes in variance but they only used informal procedures to deal with it. Mandelbrot (1963a) used recursive estimates of the variance over time to account for changes in variance. Klien (1977) took five period moving variance estimates around a ten period moving sample mean. Engle's (1982) ARCH model was the first formal model that seemed to capture the stylised facts mentioned above without the improvisation of the approaches just described. The ARCH model is useful not only because it captures the salient features of these stylised facts, but also because it can be applied usefully in a variety of areas. It has been used to measure the term structure of interest rates; to develop optimal dynamic hedging strategies; to examine how information flows across countries, markets and assets; to price options; and to model risk premia. The literature on ARCH is so vast that it is almost impossible to provide a comprehensive review. A few survey papers have already been written on this topic. In particular, the reader may wish to refer to Engle and Bollerslev (1986) and Bollerslev, Chou and Kroner (1992). The latter noted several hundred papers that apply the ARCH methodology to various financial markets. Some recent references to this rapidly growing bibliography include Bekaert (1992), Bollerslev and Hodrick (1992), Duffee (1992), Koedijk, Stork and De Vries (1992) and Ng and Pirrong (1992), to name just a few.

Agents sometimes must make decisions based upon the distribution of a random variable some time in the future (e.g. the exchange rate). In many rational expectations models it is assumed that only the mean of the conditional distribution affects the decision. Of course this a simplification, but for more general utility functions and risk averse agents, a measure of dispersion should also be very important. Many standard econometric methods have not been responsive to the need for quantitative measures of risk and uncertainty. An example is price dispersion. In conventional econometric models, the variance of the disturbance term is assumed to be constant, a convenient but often implausible assumption of linear covariance stationary models with finite unconditional second moments and time invariant conditional variances and covariances. However, many economic series exhibit periods of unusually large volatility followed by periods of relative tranquillity. In such circumstances, the assumption of constant variance seems inadequate. It is easy to imagine situations in which one might want to forecast the conditional variance of a series. For example, an asset holder might be interested in forecasts of the rate of return and its variance over the holding period. The unconditional variance (i.e. the long-run forecast of the variance) would be unimportant if the investor planned to buy the asset at time t and sell it at time t+1. One approach to forecasting the variance is to explicitly introduce an independent (exogenous) variable that would help to predict volatility (e.g. $y_{t+1} = \varepsilon_{t+1} x_t$).

The analysis of economic and financial time-series data usually involves the study of the first and possibly the second conditional moments of the series in order to characterise the dependence of future observations on past values. These two conditional moments are often identified with important economic concepts. For example, consider a univariate stochastic process for stock market returns, whose first two conditional moments given the information set are μ_t and σ_t^2. In this context, μ_t is usually associated with the risk premium as a whole and σ_t^2 with its volatility, and μ_t / σ_t^2 with the market price of risk. From an empirical standpoint. the first step often consists of estimating the conditional mean and variance. This is not a simple task in practice, since both the mean $\mu(..)$ and variance $\sigma_t^2(..)$ are generally unknown functions of the information set and are thus unobservable. The most common approach employed in practice is to assume a particular functional form for $\mu(..)$ and $\sigma_t^2(..)$, which are characterised by certain unknown parameters that need to be estimated. Although non-parametric and semi-parametric techniques have received a lot of attention recently, the parametric approach is still dominant in practice.

Instead of using ad-hoc variable choices for x_t (e.g. $y_{t+1} = \varepsilon_{t+1}x_t$), it is possible to simultaneously model the mean and variance of a series, which results in a more reasonable description of stock prices. The first to use this approach was Engle (1982) when he introduced the so-called autoregressive conditional heteroskedastic (ARCH) class of models. In the ARCH model the conditional variance is allowed to change through time as a function of both current and past information. Although this group of time-series models provides scope for a much wider class of non-linear dynamic econometric models, the linear ARCH(q) model has been found to be a particularly useful parameterisation for modelling monetary and financial data.

Detailed discussions of the ARCH(q) model, setting out further technical conditions that do not concern us here, may be found in Engle (1982), MilhØj (1985) and Weiss (1986a). A practical difficulty with ARCH is that, with q large, estimation (to be discussed later) will often lead to the violation of the non-negativity constraints that are needed to ensure that the conditional variance is always positive. This has led to the imposition of a rather arbitrary declining lag structure to ensure that these constraints are met. To obtain more flexibility, a further extension to the ARCH(q) model has been proposed by Bollerslev, (1986, 1988). The GARCH(p,q), or Generalised ARCH, includes a richer conditional variance function.

The distinguishing feature of the model is not simply that the conditional variance is a function of the conditioning information set, but rather that the particular functional form allows us to reproduce some stylised facts. Episodes of volatility are generally characterised by clusters of shocks to the dependent variable. The conditional variance function specified by Engle is formulated to mimic this phenomenon. In the regression model, a large shock is represented by a large deviation from its conditional mean or equivalently, a large positive or negative value of the error term. In the ARCH regression model, the variance of the current error conditional on realised values of lagged errors is an increasing function of the magnitude of the lagged errors, irrespective of their signs. Hence, large errors of either sign tend to be followed by more large errors of either sign. Similarly, small errors of either sign tend to be followed by small errors of either sign. The order of the lag q (ARCH(q)) determines the length of time for which a shock persists in conditioning the variance of subsequent errors. The larger the value of q, the longer the episodes of volatility will tend to be. This permits the ARCH specification to model a lot of financial and monetary series.

In the first empirical application of ARCH, which studied the relationship between the level and the volatility of inflation, Engle (1982, 1983) found that a large q was required in the conditional variance function to avoid the problem of a negative variance parameter. This would necessitate estimating a large number of parameters subject to inequality restrictions to avoid a negative conditional variance. To reduce the computational burden, Engle (1982, 1983) parameterised the conditional variance as a linear combination of past squared errors, where the weights declined linearly and were constructed so that they would sum to one. With this parameterisation (which is not unique) a large lag can be specified, yet only two parameters need to be estimated in the conditional variance function. Although linearly declining weights are plausible, the formulation puts too many undue restrictions on the dynamics of the ARCH. With this in mind, it seemed particularly important to extend the ARCH class of models to allow both longer memory and a more flexible lag structure.

The availability of long duration high frequency time series on returns from speculative assets led researchers to devote more time to studying the long run behaviour of financial data. A common finding in much of the empirical literature was that absolute returns (or squared returns) are significantly correlated over long lags. This suggested that the financial data might require many ARCH lags in order to be fully described, but this approach is potentially burdensome. In fact, the models estimated in Engle (1983), Engle and Kraft (1983), and Engle, Lilien and Robins (1985) imposed linearly declining weights in the errors so that the only free parameters were q and the sum of the weights. The choice of q can be addressed by model selection techniques, but since this can be somewhat awkward another way of dealing with a longer lag structures was highly desirable.

Bollerslev (1986) developed an extension of the conditional variance function proposed by Engle, which he called generalised ARCH (GARCH), and this model has proven to be very useful in empirical work. The GARCH model was also independently proposed by Taylor (1986), who used a different acronym. Both suggested that the conditional variance be specified as a linear combination of past squared errors (as in Engle) while also allowing past conditional variances to appear in the current conditional variance equation. The process is defined as GARCH(p,q). Some additional restrictions must also be imposed in order to ensure that the conditional variance is positive.

The motivation of the GARCH(p,q) process can be seen by expressing it as:

$$h_t = \alpha_0 + \alpha \ (L)\varepsilon_t^2 + \beta(L)h_t \quad (2.1)$$

where $\alpha(L)$ and $\beta(L)$ are polynomials in the backshift operator L. If the roots of $1-\beta(L)$ lie outside the unit circle, we can invert (2.1) and rewrite this as:

$$h_t = \alpha_0^* + \sum_{i=1}^{\infty} \delta_i \varepsilon_{t-i}^2 \qquad (2.2)$$

Hence, expression (2.2) reveals that a GARCH(p,q) process is an infinite order ARCH process with a rational lag structure imposed on the coefficients. The generalisation of ARCH to GARCH is similar to the generalisation of an MA process to an ARMA process. The intention is for GARCH to parsimoniously represent higher order ARCH processes. Instead of working with an ARCH (q) with a very high q, which is burdensome, we can achieve the similar results using GARCH(p,q) representation.

Using the law of iterated expectations, we can easily derive the fundamental properties of ARCH/GARCH processes. First, a GARCH process has zero mean. This property is observed in many monetary and financial series.

For the general GARCH (p, q) process, Bollerslev (1986) gave the necessary and sufficient condition

$$\alpha(1) + \beta(1) = \sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{q} \beta_i < 1 \qquad (2.3)$$

for the existence of the variance. When this condition is satisfied, the variance is

$$\sigma_\varepsilon^2 = E(\varepsilon_t^2) = \frac{\alpha_0}{1 - \alpha(1) - \beta(1)} \qquad (2.4)$$

Although the variance of the error term (conditional on past realised values) changes with the elements of the information set, the ARCH process is unconditionally homoskedastic. So the visual appearance of an ARCH/GARCH series conveys the impression that the unconditional variance changes over time. This false perception results from the clustering of large deviations. A major contribution of the ARCH literature is the finding that apparent changes in the volatility of economic time series may be predictable and result from a specific type of non-linear dependence rather than an exogenous structural change in the variance.

The nature of the unconditional density of an ARCH process can be analysed by higher order moments. If the error term is conditionally normal[1], the skewness coefficient is immediately

---

[1] Engle's (1982) ARCH model assumes $y_t = x_t' \xi + \varepsilon_t$
where y_t is the dependent variable and x_t is a k x 1 vector of exogenous variables, which may include lagged values of the dependent variable, and ξ is a k x 1 vector of regression parameters. The ARCH model characterises the distribution of the

seen to be zero. Since the error is continuous, this implies that the unconditional distribution is symmetric. Higher moments indicate further properties of the ARCH process. An expression for the fourth moment of a general GARCH (p, q) process is not available, but Engle (1982) gave it for the ARCH(1) process and Bollerslev (1986) for the GARCH(1,1). Recently Terasvirta (1999) derived the necessary and sufficient condition for the existence of an unconditional fourth moment of a GARCH (p,q) process, as well as an expression for the fourth moment itself and the aucorrelation function of the centred and squared observations.

The ARCH(1) process has tails heavier than a normal distribution. This property makes the ARCH attractive because the distributions of asset returns frequently display tails heavier than the normal distribution. Although no known closed form for the unconditional density function of an ARCH process exists, Nelson (1990b) demonstrated that under suitable conditions, as the time interval goes to zero, a GARCH (1,1) process approaches a continuous time process whose stationary unconditional distribution follows a Student's t distribution. Nelson's result indicates why heavy tailed distributions are prevalent in high frequency financial data rather than the normal distribution. That the parameterisation of the ARCH process does not impose *a priori* the existence of unconditional moments is an important characteristic of the model. It has long been suggested, at least as early as Mandelbrot (1963b) that the distribution of asset returns is such that the variance may not exist. In empirical applications of GARCH, estimated parameters frequently do not satisfy (2.3).

The fact that ARCH models admit an infinite variance is desirable because such behaviour may be a characteristic of the data generating process that should be reflected in the estimated model, especially when dealing with persistence. Fortunately, as will be noted later, even for GARCH models with infinite variances, standard results on consistency and asymptotic normality are still be valid.

The GARCH process is serially uncorrelated with a constant mean of zero. The process is also weakly stationary as long as the variance exists. The lack of serial correlation is a characteristic of the ARCH process that makes it suitable for modelling financial time series. The efficient markets hypothesis (EMH) discussed in chapter 1 asserts that past rates of return can not be used to improve the prediction of future rates of return. Suppose that y_t is the rate of return on an asset and that there is no regression component in the model. Then y_t is identical to the error term and becomes a pure GARCH process. The optimal prediction of the return is the expectation of

---

stochastic error ε_t conditional on the realised values of the set of variables ψ_{t-1} = { y_{t-1} . x_{t-1} , y_{t-2} , x_{t-2} , ...}. So we can standardise the error term divinding by the standard deviation to get it normal.

the return conditional on any available information. But because the GARCH model specifies that the expected value of y_t conditioned on the information set is equal to zero, the past observations do not alter the optimal prediction of the rate of return. Therefore, the presence of ARCH does not represent a violation of market efficiency. Of course, the lack of serial correlation does not imply that the error terms are independent. We suggested above that the qualitative appearance of data generated from an ARCH process arises from a particular type of dependence. Bollerslev (1986) gave a representation for the GARCH(p,q) process that reveals the nature of this dependence. Thus, the GARCH model appears to be a natural and simple generalisation of ARCH, and empirical evidence suggests that GARCH models fit the data as well as or better than ARCH models with linearly declining weights and roughly the same mean lag. See Bollerslev (1986) for more details.

One reason for the GARCH model's popularity is its convenience in implementation. One can find a counterpart in Box-Jenkins' ARMA technique in modelling means.

Taylor (1986) gives the autocorrelation function for the error term of an ARCH (p) process, and shows that it follows the same Yule-Walker equation for the associated zero-mean AR (p) process provided the fourth moment of the error term exists.

Sastry Pantula (1986) and Bollerslev (1986) show that the squared error has an ARMA (m,p) representation with m= MAX {p,q}[2]. In this representation, the autoregressive parts of the process are given by the sum of the $\alpha's$, and the moving average part is characterised by the $\alpha + \beta$ coefficients. Such an analogy has been used by Bollerslev (1988) to motivate the use of the autocorrelogram and partial autocorrelogram for ε_t^2 in model identification in assessing the number of lags (p).

The autocorrelation and partial autocorrelation functions of the squared process will display the familiar patterns of an ARMA process. Bollerslev (1988) has suggested that these autocorrelation functions may be used to identify the orders p and q of the GARCH process. In practice, the identification of the order of a GARCH(p, q) has not posed much of a problem, at least in comparison with the earlier modelling experience with ARMA(p,q) processes. In applied work, it has been frequently demonstrated that the GARCH(1, 1) process is able to represent the majority of financial time series. A data set that requires a model of order greater than GARCH (1,2) or

---

[2] Let $v_t = \varepsilon_t^2 - \sigma_t^2$, then in the conditional variance expression, $\sigma_t^2 = \varpi + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$, we can rewrite it as $\varepsilon_t^2 = \varpi + \sum_{i=1}^{m} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{i=1}^{p} \beta_i v_{t-i} + v_t$

GARCH (2, 1) is very rare, and may be difficult to estimate given the restrictions that may have to be imposed.

In empirical research, the most often used GARCH model is GARCH(1,1). It is interesting to see the theoretical autocorrelation functions for GARCH(1,1) processes. For ease of exposition, it is assumed that the distribution is conditional normal and covariance stationary. The autocorrelation function still decreases exponentially like those of an ARMA process. This result is also given in Bollerslev (1988), along with an excellent discussion on the structure of a GARCH(1,1) model when its fourth moment exists. Much of the above discussion is valid even when the fourth moment does not exist.

A natural question that has been asked concerns temporal aggregation. Does a high frequency (e.g. fitted to daily data) ARCH/GARCH process aggregate to a low frequency (e.g. fitted to weekly data) ARCH/GARCH process? Drost and Nijman[3] (1993) considered this issue in detail and showed that it is possible to have algebraic expressions between the parameters corresponding to low and high frequencies. However, as the frequency decreases the aggregate process behaves more like a conditional homoskedastic model, as pointed out by Diebold (1988).

Going from low to high frequency, in the limit the process will be an integrated ARCH as noted by Nelson (1990). Also, the distribution is no longer normal. Therefore, from a distributional point of view an ARCH process is not closed under aggregation, but this still remains an open area for research.

For practical purposes, if we specify an ARCH model only in terms of moments it is possible to estimate the low frequency parameters from an estimation of a high frequency model, and vice versa. Drost and Nijman (1993) demonstrated this using the empirical results of Baillie and Bollerslev (1989), who fitted a GARCH(1, 1) model to several exchange rates. They found that, for the Swiss franc, estimates of α_1 and β_1 from the daily data were 0.073 and 0.907. Using the relationship between the parameters of high and low frequency data. Drost and Nijman (op.cit.) showed that the implied weekly estimates were 0.112 and 0.792. Baillie and Bollerslev's estimates using actual weekly data were 0.121 and 0.781, close to the ones obtained previously by Drost and

[3] Consider ARCH(1) with $\varepsilon_t | \psi_{t-1} \sim N(0, \sigma_t^2)$ and $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$, t = 1,2,....T and we want to find the corresponding for ε_t, when t = m,2m,.....,T. Drost and Nijman showed that $E(\varepsilon_t | \psi_{t-(m)}) = 0$ and $E(\varepsilon^2{}_t | \psi_{t-(m)}) = \alpha_0 \dfrac{1-\alpha_1^m}{1-\alpha_1} + \alpha_1^m \varepsilon_{t-m}^2$. Therefore in terms of the first two moments, na ARCH process is closed under temporal aggregation and we have na algebraic relationship between the parameters corresponding to high and low frequency data.

Nijman. Except for the Japanese yen, Drost and Nijman found that direct estimates were very close to the implied weekly estimates.

We have noted that ARCH takes account of the clustering of large and small errors and the fat tails observed in the distributions of many financial data series. One of the major considerations in the introduction of ARCH by Engle (1982, p.989) was that econometricians' ability to predict the future varies from one period to another. Predictions are usually done using a conditional mean model. This is easy to see if we take an AR(1) process as an example. In this case, neither the one-step ahead nor k-step ahead forecast variance depends upon the information set. Thus, all forecast variances will be constant over the sample period. This is not very interesting and even not very descriptive for asset prices. Also for this reason, the ARCH/GARCH model seems more appropriate as it allows the forecast to change as new information is incorporated. This is especially important in the context of implied volatility, as discussed before.

2.3- Other Models dealing with Variance - Persistence in Variance

A common finding in most studies of the ARCH model concerns the possible presence of an approximate unit root in the estimated autoregressive polynomial for the conditional second moment. In this kind of model shocks to the variance are persistent in the sense that current information remains important for forecasts of all horizons. Although many economic or financial time series may exhibit persistence in their conditional variances, as previously noted by Engle and Bollerslev (1993) a non-trivial linear combination of such variables may have no persistence. This discussion puts in perspective the possibility of shocks affecting volatility forever. There is a practical interest nowadays in assessing to what extent developments in financial market have a permanent effect on volatility.

Even though many time series may exhibit persistence in variance, it is likely that several different variables share the same common long-run component. In such a situation, the variables are naturally defined to be co-persistent in variance, and the co-persistent linear combination is interpretable as a long-run relationship. Conditions for co-persistence to occur in the multivariate linear GARCH model are presented. These conditions parallel the conditions for linear co-integration in the mean, as developed by Engle and Granger (1987). The presence of co-persistence has important implications for asset pricing relationships and in optimal portfolio allocation decisions.

We have already stressed that the use of an autoregressive model (AR) with variance σ_ε^2 gives a conditional variance that depends on the forecast horizon but not on the available information set and therefore does not change over time. ARCH/GARCH models have been developed to account for this change in variance. This class of models has been designed so that the conditional variance and the conditional mean depend on the available information set.

If generated by a GARCH(p,q) model, the forecast of the conditional variance is a non-trivial function of the information set today, just as in the case of the ARCH(1) model. This allows a solution to a rational expectations model to be found, which gives a recursive solution for the conditional expectations. In particular, for the GARCH(1,1) model with horizon greater than two, if all roots of the autoregressive polynomial lie outside the unit circle then the sum of the $\alpha's$ and the $\beta's$ is less than one (the stationary case).

When the ARCH effect is present, current information is useful for assessing the accuracy with which a process can be forecast, and this can be important for producing forecasts of volatility. It is interesting to consider how the available information affects forecast uncertainty as the forecast horizon s increases. For s > p, the conditional variance of the innovation to the forecast error reduces to a linear difference equation for the sequence $E(\varepsilon_{t+s}^2|\psi_t)$ for s = p+1 to infinity. If the roots lie outside the unit circle, the solution converges to:

$$\lim_{s\to\infty} E(\varepsilon_{t+s}^2|\psi_{t-1}) = \frac{\alpha}{1-\alpha_1-...-\alpha_q-\beta_1-...-\beta_p} \qquad (2.5)$$

which is the unconditional variance of the innovation. In this case, as the forecast horizon becomes very large, the conditioning set provides no information about the variance. If, however, the roots lie on or inside the unit circle, this will not be the case we would have the denominator in (2.5) to sum to zero. For example, consider a GARCH(1, 1) process with $1-\alpha(L)-\beta(L)$ having a unit root, implying $\alpha_1+\beta_1=1$. Then the solution reduces to

$$E(\varepsilon_{t+s}^2|\psi_t) = s\alpha + E(\varepsilon_t^2|\psi_t) \qquad (2.6)$$

Therefore the conditional variance grows linearly with the forecast horizon and the dependence on the information set persists forever.

Engle and Bollerslev (1986) were the first to consider GARCH processes with $\alpha_1+\beta_1=1$ as being a distinct class of models that deserved attention. They termed it integrated GARCH (IGARCH) because of the similarity between IGARCH processes and processes that are integrated in the mean (like ARIMA). A process that is integrated in the mean is one that must be differenced to induce stationarity, and in which a shock in the current period affects the level of the series into the indefinite future. In an IGARCH process. a current shock persists indefinitely in conditioning the future variances. The IGARCH model is important because a remarkable empirical regularity, repeatedly observed in applied work, is that sums of estimated coefficients of GARCH conditional variances are very close to one. Baillie and Bollerslev (1989) estimated GARCH(1,1) models for six U.S. exchange rates and found $\hat{\alpha}+\hat{\beta}$ ranging between 0.94 and 0.99 for these six series.

Bollerslev and Engle (1991) considered multivariate IGARCH processes and defined a concept of co-integration in variance that they called co-persistence. Sets of univariate IGARCH processes are co-persistent if there exists a linear combination of the processes that is not integrated in variance. Here again we can see the similarity with the concept of co-integration. However, Nelson (1990) has cautioned that drawing an analogy with processes that are integrated in the mean may be somewhat misleading. Nelson (1990a) also demonstrated that although IGARCH models are not weakly stationary. because they have infinite variances they could be strongly stationary. Processes that are integrated in the mean are not stationary in any sense. The consistent finding of very large persistence in variance in financial time series is perplexing because currently no theory predicts that this should be the case. Lamoureux and Lastrapes (1990) have argued that large persistence may actually represent misspecification of the variance and result from structural change in the unconditional variance of the process, as represented by changes in α . A discrete change in the unconditional variance of a process produces clustering of large and small deviations that may show up persistence in a fitted ARCH model. Lamoureux and Lastrapes used 17 years of daily returns on the stocks of 30 random selected companies and estimated GARCH(1,1) models holding α constant and allowing α to change discretely over sub-periods of the sample. For the restricted model, in which α is constant, the average estimate of α for 30 companies was 0.978, while for the unrestricted model, in which α is allowed to change, the average estimate fell to 0.817. Lamoureux and Lastrapes present Monte Carlo evidence demonstrating that the MLE of $\alpha_1 + \beta_1$ has a large positive bias when changes in the unconditional variance are ignored.

In the IGARCH model, the conditional variance s steps ahead is the same as the conditional variance one-step ahead for all horizons s. The model is obviously closely related to the traditional random walk, which has a unit root in the conditional mean rather than a unit root in the conditional variance. The effect on future conditional variances is permanent: shocks are not forgotten and their effects are permanent. Consequently, information today is important for forecasting. As in the case of the conditional mean with a unit root, shocks to the conditional variance in integrated GARCH(1,1) models are not forgotten. The integrated GARCH(1,1) model which restricts $\alpha_1 + \beta_1$

= 1 was introduced by Engle and Bollerslev (1986) and was an attempt to model long run volatility persistence. The IGARCH(1,1) model is always related to the random walk process in mean.

A plot of the first differences of the logarithm of these series shows that even though the series seems to be uncorrelated over time, the observations are clearly not independent. There is a tendency for large changes to be followed by large changes but of unpredictable sign. The modified Box-Pierce test statistic can be used to test for serial correlation and this would be the case in an IGARCH model.

The first autocorrelations and partial autocorrelations die out fairly slowly. The estimate for the GARCH(1,1) suggests consideration of a particular class of GARCH models that have the property that the multi-step forecasts of the variance do not approach the unconditional variance. A necessary condition is that the $\alpha's$ and $\beta's$ sum to one. In this way, the integrated GARCH(p,q) models are part of a wider class of models with a property called "persistent variance," in which current information remains important for forecasting conditional variances for all horizons.

The empirical plausibility of IGARCH models has been established by findings in Engle, Lilien and Robbins (1985) and Bollerslev, Engle and Wooldridge (1988) that ARCH and GARCH models of interest rates typically exhibit parameters that are not in the stationary region. Mills (1994) presents some evidence on the dollar/pound exchange rate. He asserts that "since the changes in the dollar/sterling exchange rate are uncorrelated, but the squared changes are correlated, an alternative possibility for modelling this series is an ARCH process. Formal tests for the presence of ARCH are provided by the Q statistics and the LM test statistic from the regression of x on lagged values of itself. Using p = 4 lags, Mills finds the statistic 18.71(chi-square), clearly indicating the presence of ARCH. If the PACF dies out fairly slowly, an (I)GARCH(1,1) process seems a reasonable candidate for modelling x. Estimation of such a process obtains $\alpha_1 + \beta_1$ = 0.921, which is close to one. Multi-step forecasts from the model will approach the unconditional variance quite slowly. The estimated mean lag of this variance expression, is found to be 7.75, or about eight weeks, which shows persistence of variance in such markets where the changes are daily.

It certainly seems plausible to conclude that changes in exchange rates for many countries, although serially uncorrelated, have a time-varying conditional variance that can be modelled as an IGARCH process. In fact, this conclusion relies on many studies of exchange rates that find ARCH models to provide a satisfactory representation of the dynamic behaviour of such series: see, for example, MilhØy (1987), Diebold and Nerlove (1989), Hsieh (1988, 1989a, 1989b) and Baillie and Bollerslev (1989). It has also been found in other empirical research that the sum of the estimated ARCH and GARCH parameters in a GARCH(1,1) model is very close to one. For example, Taylor

(1986) estimated GARCH(1,1) models for 40 different financial time series. The results show that, for all but six of the 40 series, the estimated value of $\alpha_1 + \beta_1$ is greater than or equal to 0.97. In Ding et al. (1993), the estimated value of $\alpha_1 + \beta_1$ for daily S&P 500 returns is equal to 0.97. As has already been remarked, this regularity is widely considered to be a characteristic of volatility persistence.

Assuming conditional normality the usual procedure for estimating the IGARCH model is by estimating the GARCH model by maximum likelihood (MLE) and imposing the restriction for integration in the variance ($\alpha_1 + \beta_1 = 1$). Most of the applied work on ARCH/GARCH (including IGARCH) models use the Berndt, Hall, Hall and Hausman (1974) algorithm (BHHH) to maximise the log likelihood function. In most applications it is very difficult to justify conditional normality, and therefore the log likelihood function is misspecified. Weiss (1986) studied the asymptotic properties of quasi-maximum likelihood estimators (QMLE) to estimate the log likelihood when normality is not imposed. His results were extended by Lumsdaine (1991), who established the consistency and asymptotic normality of the QMLE of the GARCH(1,1) and IGARCH(1,1) models under a different set of assumptions. He showed that the QMLE for the IGARCH(1,1) model has the same asymptotic distribution as that of the GARCH(1,1). This result is important because it asserts that the difficulties of the unit root model are not encountered with IGARCH.

As pointed out in Ding et al. (1993), the sample autocorrelation function for absolute returns and squared returns remains significantly positive for all of these lags, while most sample autocorrelations are not significantly different from zero. In this paper, Ding et al. provide new evidence of long-term dependence in speculative returns series. Five speculative returns series from different markets are examined. Absolute returns and their power transformations are found to have long, positive autocorrelations. Usually this property is strongest for absolute returns. One exception is the exchange rate return, which has the strongest property when taken to ¼ power. (This property is examined in more detail in a later paper). The theoretical autocorrelation functions for various GARCH(1,1) models are derived and found to be exponentially decreasing, which is rather different from the sample autocorrelation function for the real data. A general class of long memory models with no memory in returns but long memory in absolute returns and power transformations is proposed. The issue of estimation and simulation for this class of models is discussed. The estimated results show that the proposed models give much better descriptions of the real data.

In spite of the fact that it is common to view volatility persistence as being best represented by ARCH/GARCH with parameters summing very close to one, Ding et al. (1994) prove that the

ACF for an IGARCH (1,1) process is exponentially decreasing and very different from the sample autocorrelation function found for various speculative returns. The IGARCH(1,1) process is not at all persistent in volatility in the sense that the autocorrelation function for the returns dies out quickly.

Ding et al. have proved, under the assumption that the distribution is conditional normal, that the theoretical autocorrelation functions for GARCH(1,1), IGARCH(1,1) and IGARCH (1,1) with drift processes are all exponentially decreasing, which is different from what we have found in real data. This is probably due to the fact that the effect of a lagged squared error on the conditional variance is exponentially decreasing. This can be generalised to a situation where the distribution is not normal, and the conditional variance equation is not a linear function of the lagged residuals. They simulate a covariance stationary model where the fourth moment exists, so the theoretical autocorrelation function is precisely defined. We can see that the sample autocorrelations are very close to the theoretical ones. As predicted by the theory, they decrease exponentially and very quickly.

The first negative sample autocorrelation occurs at lag 47, and the sample autocorrelations after this lag are not significantly different from zero. We still see that the theoretical autocorrelation function fits the sample data quite well. The sample autocorrelation decreases too fast to account for the long memory property found in the real data. The authors also plot the simulated sample autocorrelations and their theoretical approximations for IGARCH(1,1) process. One can see that the pattern of the sample autocorrelation is very different from previous ones. It decreases linearly for about the first 400 lags then collapses to zero. Thus we see that the IGARCH(1,1) process without drift is not at all persistent in volatility. Also the sample and theoretical autocorrelations for IGARCH(1,1) process with drift is shown. The shape of the simulated sample autocorrelation is quite similar to that of a covariance stationary GARCH (1,1) process. The theoretical exponentially decreasing autocorrelation function provides a very good approximation.

From the discussion above, we see that the patterns of sample autocorrelations for various speculative returns are quite different from those of the theoretical autocorrelation functions given for GARCH(1,1) or IGARCH(1,1) processes. Usually, the real data has a longer memory in volatility than the GARCH(1,1) model would suggest. The autocorrelation from a GARCH(1,1) process decreases exponentially while the sample autocorrelation usually decreases faster than exponentially at first, then much more slowly, remaining significantly positive over long lags. Usually a GARCH(1,1) process can describe short run effects better than long run effects. It is quite clear from the sample autocorrelation that there are different volatility components that will dominate in different time periods. Some volatility components may have large short run effects but will die out very quickly. Some of them may have relatively smaller short run effects but will last for long periods of time.

2.4 - Long-Memory models

Much of the analysis of financial time series considers the case when the order of differencing, d, is either 0 or 1. If the latter, x_t is I(1) its ACF declines linearly. If it is the former, x_t is stationary (I(0)) and its ACF will exhibit an exponential decay: observations separated by a long time span may, therefore, be assumed to be independent, or at least nearly so. As it is well known, I(l) behaviour of financial time series is an implication of many models of efficient markets.

However, many empirically observed time series, although satisfying the assumption of stationarity (perhaps after some differencing transformation), seem to exhibit dependence between distant observations that, although small, is by no means negligible. Such series are commonly found in hydrology, where this 'persistence' is known as the Hurst phenomenon, (see, for example, Mandelbrot and Wallis (1969), and Hosking (1989)), but many economic time series exhibit similar characteristics. This may be characterised as a tendency for large values to be followed by large values of the same sign so that the series may seem go through a succession of 'cycles,' including long cycles whose length is comparable to the total sample size.

2.4.1- Some Preliminary Ideas and Definitions

Interest in long memory processes in economics appears to have been stimulated by developments in data analysis from the physical sciences, especially in hydrology. These developments preceded the recent work done by economists. The first original work in the field of physical sciences has been found in Hurst (1951). He analysed 900 geophysical time series including tidal flows and inflows into reservoirs in the hope of understanding the determinants of persistence in streamflow data and to improve the design of reservoirs. The idea that economic time series exhibit long-range dependence has been the subject of many early theories. For an example, see Granger (1966).

There are several possible definitions of the so-called property of 'long memory'. Given a discrete time series process y_t with autocorrelation function ρ_j at lag j, then according to McLeod and Hipel (1978), the process possesses long memory if the quantity

$$\lim_{n\to\infty} \sum_{j=-n}^{n} |\rho_j| \tag{2.7}$$

is non-finite. In other words, the spectral density $f(\omega)$ will be unbounded at low frequencies.

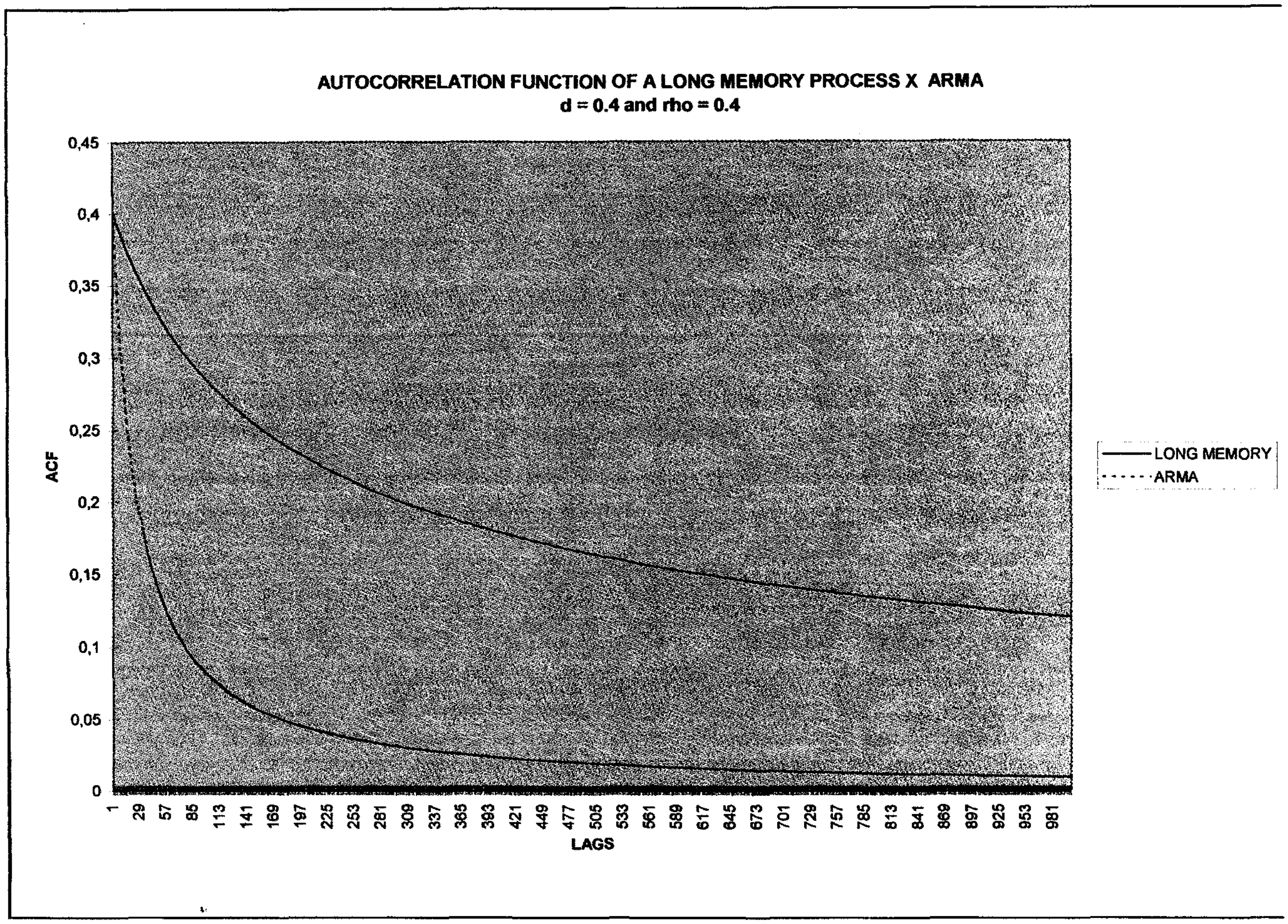
The correlogram plot of the estimated correlation between y_t and y_{t-j} against j is sometimes very useful to describe some of the linear properties of a single time series. In cases where the plot declines either exponentially or very slowly, Box and Jenkins (1970) suggested two models as possible candidates for the generating mechanism of the series: the stationary ARMA or the integrated ARIMA. However, not all correlogram plots have exactly these shapes, so new classes of models have been proposed as possible generating mechanisms. In some cases the correlogram declines steadily but not exponentially, or may start with a small estimated autocorrelation (for example ρ_1 = 0.45), and decline only slowly from this value. Among the models that have been derived as possible explanations to generate series with such properties are the fractionally integrated I(d) models or 'long memory' models. In other words, a stationary process y_t is said to have long memory if the covariance between y_t and y_{t-j} declines slowly as j increases. More specifically, the autocorrelation function ρ_j at lag j can be approximated as follows:

$$\rho_j \cong Kj^{\partial-1} \qquad \text{as} \quad j \to \infty \tag{2.8}$$

for some nonzero constant K and a positive ∂. That is, the autocorrelation function decreases at a slow hyperbolic rate rather than the fast exponential rate that characterises short memory ARMA processes. As we can see in Graph 2.1, this is indeed the case when we compare the long memory ARFIMA with ARMA. The correlogram of these two series shows that the autocorrelation for the long memory process exhibits a clear pattern of slow decay and persistence.

The Efficient Market Hypothesis (EMH) states that, because current prices reflect all available public information, price changes can only be brought about by the arrival of new information. With all prior information already reflected in prices, markets should follow a random walk and price movements on any given day should be unrelated to the previous day's activity. EMH implicitly predicts that all investors react immediately to new information, and therefore future prices will be unrelated to past or present values. The presence of long memory components in asset returns has important implications for modern financial economics, since recent tests of investor rationality and the "efficiency" of markets hinge on the presence or absence of long memory.

Graph 2.1



Do people really make decisions in this manner? Generally speaking, some do react to information as soon as they receive it. However, most people wait for additional information to confirm their suspicions and do not react until a trend is clearly established. The amount of confirming information necessary to validate a trend varies, but an uneven assimilation of information may cause a biased random walk. In the 1970s and 1980s Mandelbot used the term "fractional brownian motion" to describe biased random walks. Regular Brownian motion is a continuous-time stochastic process, denoted as B(r), and is composed of several independent Gaussian increments. Mandelbrot and Van Ness (1968) also note that in a certain way, fractional Brownian motion, $B_H(r)$, can be regarded as the approximate (1/2 - H) fractional derivative of regular Brownian motion,

$$B_H(r) = [1/\Gamma(H+1/2)]\int_0^r (r-x)^{H-1/2}dB(x), \text{ for } \quad r \in (0,1) \qquad (2.9)$$

where $\Gamma(.)$ is the gamma function, B(x) is regular Brownian motion with unit variance and H is the Hurst exponent, originally due to Hurst (1951) and to be introduced later on. When H = 1/2, $B_H(r)$ reduces to regular Brownian motion, B(r). The autocovariance function Brownian motion is given by

$$E|B_H(t) - B_H(s)|^2 = |t - s|^{2H}$$

and

$$\gamma_j = |j|^{2H-2} \tag{2.10}$$

so that for high lags hyperbolic decay occurs in the autocovariance function exhibiting the long memory property. Continuous time fractional Brownian noise is denoted by $B_H(t)$ and is the derivative of fractional Brownian motion. The (1/2 - H) fractional derivative of continuous time white noise reduces to white noise when H = 1/2.

2.4.1.1 - The Hurst Exponent[4]

As has already been mentioned, Hurst was a hydrologist who began working on the design of reservoirs and inflows and outflows of rivers. especially in Egypt. Since an ideal reservoir should never overflow, a policy should be put in place to ensure the discharge of a certain amount of water each year. On the other hand, if inflows into a reservoir are too small then water levels can become dangerously low. In this context, the important question for policy makers is "What level of discharge should be chosen so that a reservoir will never overflow or be empty?"

Many variables must be considered in building a model that can address this question. One of the most important is clearly uncontrollable, namely the amount of water coming from rainfall. Hurst made the reasonable assumption that inflows from rainfall were a random walk, but naturally he wanted to test this hypothesis. This is the origin of the Hurst exponent (H). Among its desirable properties are its high degree of robustness and its broad applicability to time series analysis. It allows us to characterize time series while making only a few underlying assumptions. It can distinguish a random series from a non-random series, even if the random series is non-Gaussian (i.e. not normally distributed). Hurst found that most natural systems do not in fact follow a random

[4] It has been widely established that Hurst exponent is a biased estimator. We just want to recover the discussion of long-memory and the need to develop a better estimator. This discussion is taken in next chapters and we will see that the Geweke-Portar-Hudak (GPH) and minimum distance estimators do much better.

walk.

Hurst measured how the reservoir level fluctuated around its average level over time. As might be expected, the range of this fluctuation changed depending on the length of the time used for measurement. If the time series were random, the range would increase with the square root of time ($T^{0.5}$). This is the $T^{0.5}$ rule. To standardise the measure over time, Hurst created a dimensionless ratio by dividing the range by the standard deviation of the observations. Hence the analysis is called re-scaled range analysis (R/S analysis). Hurst found that most natural phenomena, including river discharges, temperatures, rainfall, and sunspots, follow a "biased random walk"- a trend with noise or fractional Brownian motion. The level of noise could be measured by how the R/S is linked to time, or by how far H is above 0.50. A natural extension of Hurst's study is to see how we can apply it to capital market time series. Among the first to have considered the possibility and implications of persistent statistical dependence in asset returns was Mandelbrot (1971). Since then, several studies have lent further support to Mandelbrot's findings. For example, Greene and Fielitz (1977) claim to have found long-range dependence in daily returns for many securities listed on the New York Stock Exchange. More recent investigations have uncovered anomalous behaviour in long-horizon stock returns; alternately attributed to speculative fads and to time-varying conditional expected returns, these long-run swings may be further evidence of the Joseph effect (long-memory).

The original statistical measurement of long memory due to Hurst (1951) is the re-scaled range or R/S statistic. The rescaled range statistic R_T / S_T is defined as

$$R_T = \frac{1}{S_T}[\underset{1\leq j\leq T}{Max}\sum_{k=1}^{j}(y_k - \bar{y_T}) - \underset{1\leq :\leq T}{Min}\sum_{k=1}^{j}(y_j - \bar{y_T})] \qquad (2.11)$$

where R is the range (the difference between the maximum and the minimum levels), $\bar{y}$ is the sample mean and S_T is the sample variance:

$$S_T = \{1/T)[\sum_k (y_k - \bar{y}_T)^2]^{1/2} \qquad (2.12)$$

As we can see, the R/S statistic is the range of partial sums of deviations of a time series from its mean. The first term in brackets in (2.11) is the maximum (over j) of the partial sums of the first j deviations of y_t from the sample mean. Since the sum of all T deviations of the y_t from their mean is zero, this maximum is always non-negative. The second term is the minimum (over j) of the same sequence of partial sums: hence it is always non-positive. The difference between the two quantities, called the 'range' for obvious reasons is therefore always non-negative, hence $R_T \geq 0$.

In order to compare different types of time series, Hurst divided this range by the standard deviation of the original observations. This 'rescaled range' would increase with time. Hurst and Mandelbrot (1972, 1975) formulated the following relationship:

$$R / S = a.(T)^{H} \tag{2.13}$$

where R/S = rescaled range

T = number of observations

a = constant

H = Hurst exponent

Or written differently (2.13), we would have then:

$$\log[R_T / S_T] \approx a + H[\log(T)] \tag{2.14}$$

and then the Hurst exponent is then estimated as $\log[R_T / S_T] / [\log(T)]$, or alternatively by taking the slope coefficient of a regression of $\log[R_T / S_T]$ on log(T), for different values of T.

According to statistical mechanisms (fractional brownian motion), H should equal 0.5 if the series is a random walk. In other words, the range of cumulative deviations should increase with the square root of time, T. Since a short memory process would have a value H equal to 1/2, an estimated value of H that exceeds 1/2 is interpreted as evidence of long memory. Mandelbrot (1972,1975) focused on the relation of the R/S statistic's logarithm to the logarithm of the sample size as the sample size increases without bound. For short-range dependent time-series, the ratio approaches 1/2 in the limit, but converges to quantities greater or less than 1/2 according to whether there is a positive or negative long-range dependence. The limit of this ratio is also called H, which is precisely the Hurst coefficient. For example, the fractionally differenced process satisfies the simple relation: H = d + 1/2, where d is the fractional differencing parameter. Mandelbrot and Wallis (1969) suggest estimating H by plotting the log of R/S against the log of the sample size. Beyond some large T, the slope of such a plot would settle down to H. However, although H = 1/2 across general classes of short-range dependent processes, the finite-sample properties of the estimated H are not invariant to the form of short-range dependence. When Hurst applied his statistic to the Nile River discharge record, he found H = 0.90! He then tried other rivers, and H was usually found to be greater than 0.50. He finally tried different natural phenomena. In all cases, he found H to be greater than 0.50. What did it mean?

When H differed from 0.50, the observations were not independent. Each observation carried a 'memory' of all the events that preceded it. Furthermore, this was not a short-term memory, which is often called 'Markovian', but a different kind of memory altogether. This memory was long-term and. theoretically at least, lasted forever. More recent events had a greater impact than distant events, but there was still residual influence from the distant past. On a broader sense, a system that exhibits this kind of memory is the result of a long stream of interconnected events. What happens today influences the future, and events in the past are important for explaining we are today. The impact of the present on the future can be expressed as a correlation

$$C = 2^{(2H-1)} - 1 \tag{2.15}$$

where C = correlation coefficient

H = Hurst exponent

There are three distinct classifications for the Hurst exponent (H):

(1) $H = 0.50$

(2) $0 \le H < 0.50$

(3) $0.50 < H < 1.00$

H equal to 0.5 denotes a random series. Events are random and uncorrelated. Equation (2.15) equals zero when H = 1/2. In this case the present does not influence the future. Its probability density function is normal, but it does not have to be. R/S statistics can classify an independent series, no matter what the shape of the underlying distribution. In statistics courses we are usually taught that natural processes follow a normal distribution. Hurst's findings refute this notion, since H is typically greater than 0.5 implying that the probability distribution is not normal. In the case, $0 \le H < 0.50$ is an antipersistent series. In other words, if the system has been up in the previous period, it is more likely to be down in the next period. The strength of this antipersistent behaviour depends on how close H is to zero. The closer it is to zero, the closer C in (2.15) moves towards -0.50. In other words, this negative correlation implies an indirect relationship between past and future values of the variable. If it has been high (up) in the past it will be more likely to be down in the future (i.e. there is a negative relationship). This relation is more volatile than a random series because of the 'mean reverting' property and frequent reversals that characterise the process.

When $0.50 < H < 1.00$, we have a persistent, or trend-reinforcing, series. If the series has been up (down) in the last period, then the changes will continue to be positive (negative) in the next period. The strength of the trend-reinforcing behaviour, or persistence, increases as H approaches 1.0, or 100 percent correlation in equation (2.15). That is, if past values of the series are moving in a positive (negative) direction, future values will tend to move in the same direction. The closer H is to 0.5, the series will be noisy and trends will not be well defined. Persistent series are examples of fractional Brownian motion, or biased random walks. The strength of the bias depends on how far H is above 0.50. Persistent time series are the most interesting class because, as Hurst found, they are plentiful in nature.

In equation (2.14), finding the slope of the log/log graph of R/S versus T will therefore give us an estimate of H. This estimate of H requires no assumptions about the nature of the underlying distribution. For very large T we would expect H to converge to 0.50, because the memory effect diminishes over time and is expected to exhibit properties similar to regular brownian motion, or a pure random walk. The regression referred to above would thus give us an indication of the process and would be performed on the data prior to the convergence of H to 0.50.

It is important to remember that the correlation measure in (2.15) is not related to the Autocorrelation function (ACF) of gaussian random variables. The ACF works well in determining short-run dependence, but fails to capture long-run correlations for non-gaussian series. Graph 2.2 shows the log/log plot of R/S versus T for H = 0.5, using data from a pseudo-random number generator in the Gauss programming language. The estimation used by the Hurst methodology shows H = 0.47. This estimate is a little lower than expected, but these are pseudo-random numbers generated by a deterministic algorithm. In this case, rescaled range analysis seems to have captured this bias. It is important to note that the R/S statistics is an extremely useful tool since it does not require the assumption that the underlying distribution is gaussian. However, finding H = 0.50 does not prove a gaussian random walk. It only proves that there is no long memory process. In other words, any independent process, gaussian or otherwise, would produce H = 0.50.

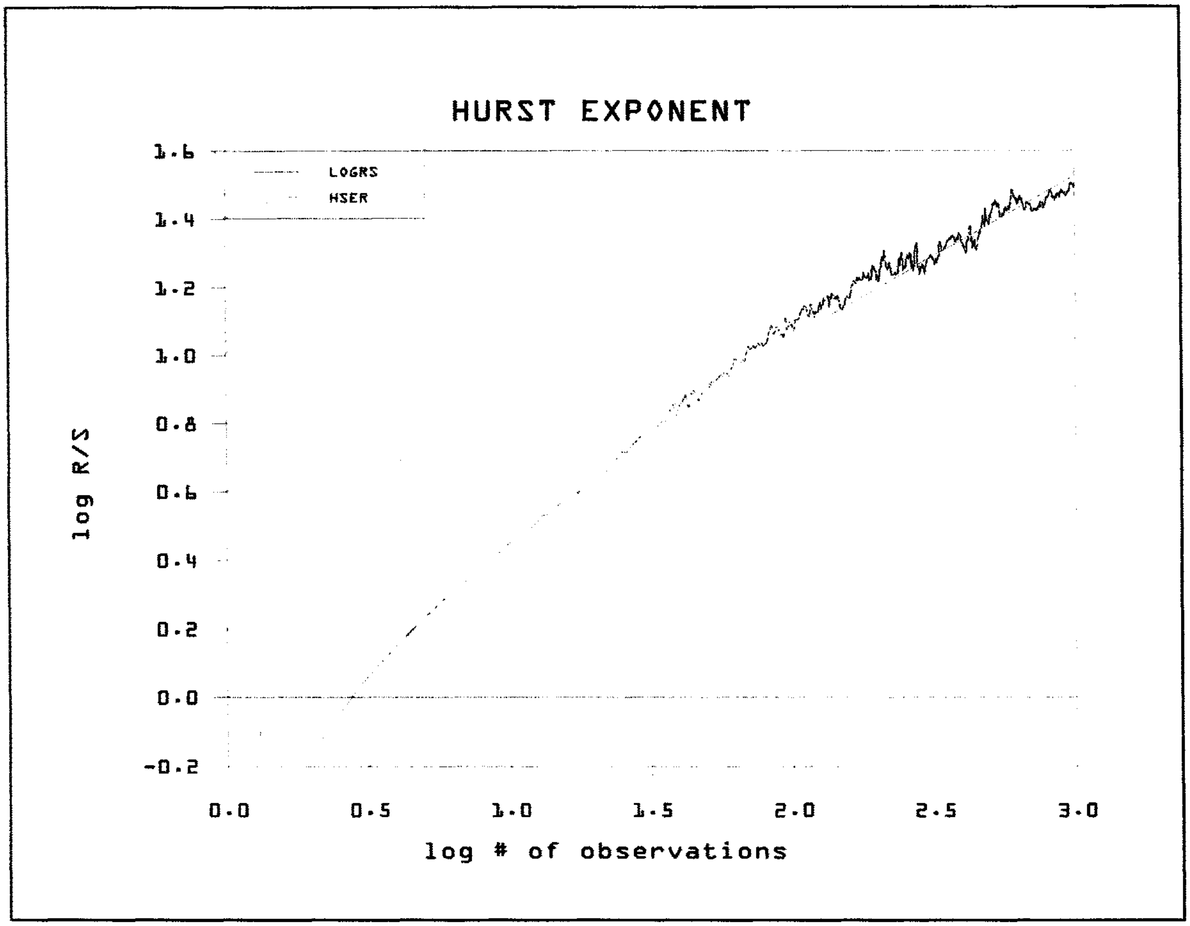
Even if we find a value for H that is significantly different from 0.50, we should not automatically accept this estimate as valid. It may be that the span of the data is not sufficient, or there may still be a problem with the R/S statistic. An estimate of H that is significantly different from 0.50 has two possible explanations:

a) There is a long memory component present in the time series being studied. In this case each observation of the process is correlated with the observations that follow.

b) The analysis is flawed, and an anomalous value of H does not mean that there is a long memory component in the time series.

Sometimes we do not know if have enough data to perform a valid test, since there are no clear guidelines as to what constitutes a sufficient amount. We can test the validity of our results by randomly scrambling the data so that the order of the observations is completely different from that of the original time series. The frequency distribution of the reordered data would remain unchanged since the observations are the same. If we then calculate the Hurst exponent using the scrambled data, H should remain virtually unchanged, since there would be no long memory in the reordered data. Consequently, scrambling the series would have no effect on the qualitative aspects of the data. However, if a long memory component is present the order of the data is important. In scrambling the observations we break the structure of the system, so the H estimate should be very close to 0.50 even though the frequency would not change. The rule would be as follows:

Graph 2.2



a) Take the original series and compute H. Scramble the series from which the H was computed and plot the log/log of the scrambled and unscrambled series. If there is virtually no qualitative difference between the two, then the original series does not have long memory.

b) Take the original series and compute H. Scramble the series from which H was computed and plot the log/log of the scrambled and unscrambled series. Different values of H for the scrambled and unscrambled data would indicate the presence of long memory in the original series. This is because the scrambled series would be independent despite having the same non-normal frequency distribution as the original series.

Applying the R/S statistic is simple and straightforward, but we do need a considerable amount of data in order to calculate it. When analysing financial markets, we generally use the logarithm of returns, defined as follows:

$$S_t = \log(P_t / P_{t-1}) \tag{2.16}$$

where S_t = return in logarithm

P_t = stock price index at time t

For the R/S statistic, the logarithm of returns is more appropriate than the more commonly used percentage change in prices. The range used in R/S analysis is the cumulative deviation from the average, and log returns sum to cumulative returns while percentage changes do not. The first step, then, is to convert the price series into log returns. The second step is to apply equation (2.11) for various increments of time. We should start with a reasonably fine increment. For example, a monthly time series covering 40 years of data could be converted into 480 log returns according to (2.16). We could then split the series into 80 independent six-month increments, being careful that they do not to overlap. Because these are non-overlapping six-month periods the observations should be independent, but this might not always be the case. For instance, they would not be independent if there were a short-term Markovian-type dependence lasting longer than six months. We can again use equation (2.11) to calculate the range of each six-month period, then rescale each range by the standard deviation of the observations in the respective period according to (2.12) to obtain 80 separate R/S observations. By averaging the 80 observations, we obtain the R/S estimate for the series with T = 6 months.

We continue in this manner for T=7,8,9,..., 240. The stability of the estimate can be expected to decrease as T increases since we have fewer observations to average. Eventually we would run a regression of log(T) versus log(R/S) for the full range of T, taking the slope of T as the estimate of H according to (2.14). A natural question arises here concerning the amount of data. How much data should we consider? Simulated data with fewer than 2,500 observations are not to be considered but however there is no clear indication of how many data points should be considered. In our work we will consider several daily data series beginning in the nineteen seventies with some 4,500 observations. This should be sufficient for our purposes. It is worthwhile that this procedure is a sort of bootstrap procedure. This procedure has been proposed by Peters (1991) in order to produce a formal test for the Hurst statistic by means of using a bootstrap distribution. We do not pursue the discusion about the procedure itself as the reader should refer to Peters (1991). We will use other tests as well in order to verify for the existence of long-memory effects, so at this point the procedure described above should be seen as an earlier attemp and an example of testing for long-memory.

This point has been persuasively argued by Mandelbrot (l969, 1972) in extending his work on non-Gaussian (marginal) distributions in economics, and in particular on financial prices (1963), to an exploration of the structure of serial dependence in economic time series. While Mandelbrot considered processes that took the form of discrete time fractional gaussian noise, attention has recently focused on extensions of the ARIMA (p,d,q) class for modelling long-term persistence. However, if d is not an integer then x_t is said to be fractionally integrated. Such models are termed ARFIMA by Diebold and Rudebusch (1989), with the F standing for 'fractionally.'

In the analysis of stationary ARMA, models the ACF will exhibit an exponential decay as the lag increases. Thus, observations separated by a long time span can be taken to be independent, or at least nearly so. However, many observed time series are seen to exhibit dependence between distant observations that is not negligible. This is not taken into account in traditional ARMA models.

The notion of fractional differencing seems to have been proposed independently by Hosking (1981) and Granger and Joyeux (1980). Further references include Hosking (1982), Granger (1980) and Geweke and Porter-Hudak (1983). ARIMA processes with non-integral d are often referred to as long memory models in this literature. The reason for this terminology may be found in the property exhibited by the simplest member of the class, the ARIMA (0,d,0) process (fractional integrated white noise). Hosking (1981a) shows that, if $|d| < 1/2$, then

(i) x is stationary and invertible.
(ii) The ACF declines monotonically and hyperbolically to zero as the lag increases. This is a much slower rate than the exponential decay of an ARMA (d = 0) process.
(iii) The PACF is independent of d.

ARFIMA models possess interesting and potentially useful long memory properties. Non-integer d values provide for flexible modelling of low frequency variation, which gives better long-run forecasts than conventional models. The ARFIMA(p,d,q) offers much greater flexibility in simultaneous modelling of short-term and long-term behaviour of time series.

How does the ARFIMA model incorporate long memory behaviour? For $0 < d < 1/2$, it can be shown that its ACF declines hyperbolically to zero, i.e. at a much slower rate than the exponential decay of a standard ARMA (d=0) process. For $d > 1/2$, the variance of x_t is infinite, so the process is non-stationary. Examples of how autocorrelations vary with d are provided in Hosking (1981), Diebold and Rudebusch (1989) and Lo (1991). Typically, autocorrelations from ARFIMA processes remain noticeably positive at very long lags. even after the autocorrelations from I(0) processes have declined to near zero.

So when $0 < d < 1/2$ the autocorrelations have an infinite sum and the ARIMA (0,d,0) is said to have a long memory. As such it should be useful for modelling long-term persistence. When $-1/2 < d < 0$, all auto and partial correlations are negative, and even though the autocorrelations have an infinite sum such negative correlations do not allow for long-term persistence. The ARIMA (0, d, 0) process then has short memory, a property which is also referred to as "antipersistence". When $d > 1/2$ the variance of x is infinite and the process is nonstationary. When $d = 1/2$ will be invertible but not stationary, whereas the converse applies when $d=-1/2$. Thus, the ARFIMA(0, d, 0) process has special long memory properties that make it useful for modelling long-term persistence. However, there is still a need for a family of models flexible enough to account for both the short term and long-term behaviour of time series.

The effect of the d parameters on distant observations declines hyperbolically as the lag increases, whereas the effects of the AR and MA parameters declines exponentially. It is important for modelling purposes that d be chosen to describe the high lag correlation structure of a time series, while the other parameters should be chosen describe the low-lag correlation structure. Indeed, the long-term behaviour of an ARIMA (p, d, q) process may be expected to be similar to that of an ARIMA (0, d,0) process with the same value of d, since the effects of the AR and MA parameters will be negligible.

Since it is the fractional difference parameter d that allows long persistence to be modelled, the value chosen for d is obviously crucial in any empirical application. Typically, this value will be unknown and therefore must be estimated.

A number of different methods have been proposed, but there are no clear guidelines to suggest which techniques are superior. However, this state of affairs will surely be remedied before long. Early suggestions are discussed in Mills (1990, chapter 11.7), Pagan and Wickens (1989) and Sowell (1992a). To summarize, four suggestions have been made in the literature:

(i) Granger and Joyeux (1980) use a grid search of d values, using a measure of h-step ahead forecastibility to determine the chosen value. They stress that this method is clearly arbitrary and sub-optimal, but they argue that it appears to work quite well in their, admittedly limited empirical experience with the method.

(ii) Hosking (1981a) suggests that a maximum likelihood estimate (MLE) of d may be obtained by the methods of McLeod and Hipel (1978), although no details are given.

(iii) Hosking (1981a) also suggests that d could be estimated by the rescaled range or R/S exponent. This statistic was explicitly developed to measure long-term persistence, and its use in the analysis of economic time series has been established by Mandelbrot (1972). A more recent survey of R/S analysis and applications may be found in Mandelbrot and Taqqu (1979). The biggest problem associated with this method appears to be bias.

(iv) Geweke and Porter-Hudak (1983) propose an estimator of d based on a simple linear regression of the logarithm of the periodogram of a time series on an associated deterministic regressor. The estimator is the OLS slope parameter in this regression, formed using only the lowest frequency ordinates of the log periodogram. An alternative frequency domain estimator has recently been proposed Kashyap and Eom (1988).

Geweke and Porter-Hudak (GPH) has been used widely due to its ease of implementation. They propose estimating d by regressing the periodogram at different frequencies against a constant and the frequency. As pointed out by Pagan and Wickens (1989), Kunsch (1986) shows that frequencies around the origin need to be excluded to obtain consistent estimates. He also shows that k should expand with sample size, and setting $K = g(T) = \sqrt{T}$ has been found to work well.

Having obtained an estimated of d, x can be transformed by the long memory filter, truncated at each point to the available sample. The transformed series is then modelled as an ARMA process. Further details of this procedure and a discussion of its properties may be found in Geweke and Porter-Hudak (1983) and Diebold and Rudebusch (1989). Sowell (1992a, 1992b) discusses joint maximum likelihood estimation of d and the ARMA parameters, and presents Monte Carlo experiments which show that maximum likelihood gives more accurate estimates than Geweke and Porter-Hudak's method when the correct specification is known. However, Sowell emphasises that, when the specification is uncertain (as is usually the case in practice), which method is superior remains an open question. GPH has several merits as well as some drawbacks:

1) The GPH estimation method is a very simple procedure. Under regularity conditions on g (.), the consistent and asymptotically normal estimate of d is just the negative of the OLS estimate of the slope coefficient in a simple linear regression.

2) This method does not require a large amount of data. The experimental results of Geweke and Porter-Hundak (1983) suggest that the effect of sample size on the reliability of the confidence

interval is negligible compared to the Granger and Joyeux (1980) grid-search method, where an AR(50) is used to form ten-step forecasts.

3) The GPH method provides consistent and asymptotically normal estimates of d regardless of the orders and parameterisations of the polynomials underlying the stationary process.

Alternative procedures, such as maximum-likelihood methods for simultaneous estimation of d and the parameters of the polynomials, have desirable properties if the model is correctly specified, but may be inconsistent otherwise (Brockwell and Davis (1987) and Sowell (1987)).

One drawback of the GPH method is the difficulty of choosing k = g(T), i.e. the number of low-frequency periodogram ordinates used in the regression. Alternative choices of g(T) either produce biased estimates if g(T) is too large, or imprecise estimates (large standard errors) if g(T) is too small.

Diebold and Rudebusch (1989b) suggest that, since d is estimated from the long-run dynamics of the time series, economic considerations can suggest a reasonable GPH sample. Their choices are K = T/5 for annual series and K = T/20 for quarterly series. However, their estimates have large associated standard errors.

Based upon theoretical considerations and Monte Carlo simulation, Geweke and Porter-Hundak (1983), Brockwell and Davis (1987) and Shea (1989) recommend using g(T)= T^{α} and obtain good results with α = 0.5. However a lot of doubts remains why we should use α = 0.5. The estimates of d of U.S. GNP series are quite robust, but the standard errors for these estimates are also quite large. Thus, the confidence we can have in the estimates of d is quite low.

Experience with each of these methods is presently too limited to offer any firm guidance as to which is most useful. Nevertheless, the empirical experience of forecasting economic time series and fractionally differenced series is quite encouraging. Both Granger and Joyeux (1980) and Geweke and Porter-Hudak (1983) find that such models perform better out-of-sample with long forecast horizons than do conventional ARMA models, and Granger (1980a) finds that fractionally differenced models arise naturally in economics through component aggregation. One approach to detecting the presence of long memory in time series is to use the "range over standard deviation", or "rescaled range", statistic originally developed by Hurst (1951).

Semiparametric estimates of long memory are useful in the analysis of financial time series because they are consistent under much broader conditions than parametric estimates. However, recent large sample theory on semiparametric estimation forbids conditional heteroskedasticity. Robinson and Henry(1999) show that a leading semiparametric estimator, the Gaussian or local White estimate, can be consistent and have the same limiting distribution under conditional heteroskedasticity as under Robinson's (1995) assumption of conditional homoskedasticity. Actually, noting that long memory has been observed in the squares of financial time series, they allow (under regularity conditions) for conditional heteroskedasticity of the general form introduced by Robinson (1991). This may include long memory behaviour for the squares, such as the fractional noise and autoregressive fractionally integrated moving average form, and also standard short memory ARCH and GARCH specifications.

2.5 - Long Memory volatility processes

Due to the fact that long-memory has recently attracted a great deal of attention, there has been growing interest in the second moment of a process following the example of the ARFIMA model. Many applications of long memory have emerged from studies of financial market data. The availability of large amounts of data exhibiting this phenomenon has justified the development of theoretical tests and models of long memory in volatility. However, in many cases the choice of model has generally been restricted by mathematical tractability. The first contribution was Taylor (1986), who observed that autocorrelations in the absolute values of stock returns tend to decay very slowly. Ding, Granger and Engle (1993) note the same stylized fact for the power transformations of daily returns, and Dacorogna, Muller, Nagler, Olsen and Picet (1993) find similar evidence for squared exchange rate returns recorded every twenty minutes over a four-year period.

A long-memory conditional variance process can be based on the ARCH model of Engle (1982). Baillie, Bollerslev and Mikkelsen (1996) have considered a long memory process in the conditional variance, known as Fractionally Integrated Generalised AutoRegressive Conditional Heteroskedasticity, or FIGARCH. This process implies a slow hyperbolic rate of decay for lagged squared innovations and persistent impulse response weights. The cumulative weights also tend to zero, a property shared by weakly stationary[5] process or stable GARCH processes. However, the impulse response weights of the FIGARCH process decay at a very slow hyperbolic rate. The FIGARCH(p,d,q) is defined as

[5] A recent paper by Giraitis, Kokosaka and Leipus (2000) study a broad class of nonegative ARCH(∞) models. Sufficient conditions for the existence of a stationary solution are established and an explicit representation of the series is found.

$\phi(L)(1-L)^d \varepsilon_t^2 \ = \omega + \{1-\beta(L)\}v_t$ (2.17)

where all the roots of $\phi(L)$ and $\{1-\beta(L)\}$ lie outside the unit circle. The FIGARCH process can also be represented as

$\{1-\beta(L)\}\sigma_t^2 = \omega + \{1-\beta(L)-\phi(L)(1-L)^d\}\varepsilon_t^2$ (2.18)

and as

$\sigma_t^2 = \omega\{1-\beta(1)\}^{-1} + \lambda(L)\varepsilon_t^2$ (2.19)

where $\lambda(L) = \{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^{d}\}$ (2.20)

A necessary and sufficient condition for the FIGARCH (1,d,0) process to have nonnegative impulse response coefficients, $\lambda_j \geq 0$ for positive integer j is for $0 \leq d \leq \beta$. Following Baillie et al (op.cit.), the polynomial in the lag operator of the impulse response coefficients is denoted by $\gamma(L)$, where $\gamma(L) = \sum_{k=0}^{\infty} \gamma_k L^k$. Then,

$(1-L)\varepsilon_t^2 = \omega + \gamma(L)v_t$, and $\gamma(L) = (1-L)^{1-d}\phi(L)^{-1}\{1-\beta(L)\}$ (2.21)

The impact of past shocks on the volatility process is given by the limit of the cumulative response weights,

$\gamma(1) = \lim_{k\to 0} \lambda_k = \sum_{j=0}^{\infty} \gamma_j$ (2.22)

For the FIGARCH process and for a value of d > 1, then $\gamma(1)$ will be infinite, while for the FIGARCH(1,d,0) process,

$\lambda_k = [\Gamma(k+d-1)/\{\Gamma(k)\Gamma(d)\}][(1-\beta)-(1-d)/k]$ (2.23)

The cumulative effect of a shock will be zero on the volatility process since $\gamma(1) = 0$; and from Stirling´s approximation,

$$\lambda_k \approx [(1-\beta)/\Gamma(d)]k^{d-1} \, , \quad (2.24)$$

so that hyperbolic decay occurs in the response of the conditional variance to past shocks. Since $\lambda(1)=1$, it follows that E(ε_t^2) is undefined, and hence the second moment of the unconditional density of ε_t is infinite. The FIGARCH process is then clearly not weakly stationary, a common feature with the IGARCH process. Approximate maximum likelihood estimates of the parameters of the FIGARCH(p,q,d) process can be estimated through maximisation of Quasi-Maximum Likelihood which realises $T^{\frac{1}{2}}$ consistent estimates of the FIGARCH parameters. Then,

$$T^{\frac{1}{2}}(\hat{\theta}_T-\theta_0) \Rightarrow N\{0, A(\theta_0)^{-1}B(\theta_0)A(\theta_0)^{-1}\} \quad (2.25),$$

where A(.) and B(.) represent the Hessian and outer product gradient, respectively, and θ_0 denotes the true parameter values. Simulation evidence indicates that the limiting distribution theory works well in sample sizes of 1500 and 3000. Baillie et al (op.cit.) also report the effects of estimating stable GARCH process where the true data generating process is FIGARCH. The sum of the estimated GARCH(1,1) parameters is always close to one, which implies integrated GARCH or IGARCH behaviour and suggests that the apparent widespread IGARCH property so frequently found in high frequency financial data (see Bollerslev, Chou and Kroner, 1992) may well be spurious.

Fiorentini et al. (1995) have employed analytic derivatives to compute GARCH estimates. They show that, in the context of univariate GARCH models, these derivatives can be successfully used for estimation purposes. They argue that this approach is better than maximum likelihood GARCH estimation, which relies on numerical approximation of log-likelihood derivatives since exact analytic differentiation burdensome.

Following Nelson's (1991) Exponential ARCH model, which allows for asymmetries, Bollerslev and Mikkelsen (1996) have extended the FIGARCH process to FIEGARCH, E standing for "exponential." The FIEGARCH(p,q,d) model is then

$$\log(\sigma_t^2) = \omega + \phi(L)^{-1}(1-L)^{-d}[1-\lambda(L)]g(\xi_{t-1}) \quad (2.26)$$

where $\quad g(\xi_t) = \theta\xi_t + \gamma[|\xi_t| - E|\xi_t|] \quad (2.27)$

and all the roots of $\phi(L)$ and $\lambda(L)$ lie outside the unit circle. When d=0, we have the EGARCH process and when d=1, the process becomes IEGARCH. They also present evidence on the efficacy of QMLE applied to estimate the parameters of the FIEGARCH process and illustrate its application to the pricing of options.

Bailie et al. (1995) consider the application of long-memory processes to describing inflation for ten countries using GARCH-type conditional heteroscedasticity. They find strong evidence of long-memory, and mean-reverting behaviour for all countries except Japan, which appears to be stationary.

Breidt, Crato and de Lima (1993) and Harvey (1993) propose a different modelling of persistence in volatility. The model is then;

$$y_t = \xi_t \sigma_t \quad (2.28) \quad \text{and} \quad \sigma_t^2 = \sigma^2 \exp(h_t) \quad (2.29)$$

where ξ_t is NID(0,1). It is common to specify that h_t is an AR(1) process, which implies an ARMA(1,1) representation for $\log(y_t^2)$. If it is assumed that h_t is the fractional white noise process,

$$(1-L)^d h_t = \varepsilon_t \quad (2.30)$$

where ε_t is $NID(0,\sigma_t^2)$ then the previous model generates a long memory stochastic volatility process. The usual procedure for regular stochastic volatility models has been the estimation through the state space representation and uses QML via the Kalman filter. Since a state space representation does not exist for long memory processes, estimation of the long memory stochastic volatility process is correspondingly difficult. Breidt, Crato and de Lima (op.cit.) use frequency domain approximate MLE to estimate an ARFIMA(0,d,1) model for $\log(y_t^2)$, while Harvey (1993) uses the GPH estimator to obtain an estimate of d in a fractional white noise model for $\log(y_t^2)$.

Mahieu et al. (1997) study the empirical performance of stochastic volatility models using twenty years of weekly exchange rate data for four major currencies. They focus on the effects of innovations, both on estimates of parameter and on estimates of the latent volatility series. The density of the log squared exchange rate innovations is modelled as a flexible mixture of normals. They find that explicitly incorporating fat-tailed innovations increases estimates of the persistence of volatility dynamics. Another finding is that estimates of the errors in the volatility time series are very large. In fact, they are so large that calculated option prices are rarely significantly different from those in models with constant volatility.

Watanabe (1998) develops a new model for the analysis of stochastic volatility (SV). Since volatility is a latent variable in SV models, it is difficult to evaluate the exact likelihood. A non-linear filter, which yields the exact likelihood, is constructed and a smoothing algorithm for volatility estimation is also used. The model is first tested in Monte Carlo experiments, where it performs well. It is then used to analyze daily stock returns on the Tokyo Stock Exchange, where the results confirm the earlier findings.

Liesenfeld and Jung (1999) compare the SV model, which assumes that the conditional distribution of returns given the latent volatility process is normal, to ones based on conditional heavy-tailed distributions, especially Student's t and the generalized error distribution. Their methodology is based on a simulated maximum likelihood approach. The results are based on daily data for exchange rates and stock prices. They reveal that the SV model with a conditional normal distribution does not adequately account for the leptokurtik distribution of returns and the low but slowly decaying autocorrelation functions of the squared returns. These empirical facts are better explained by SV models with conditional heavy-tailed distributions. They also argue that the choice of conditional distribution has systematic effects on the parameter estimates of the volatility process.

One of the most exciting current applications of long memory processes concerns the volatility of asset prices. The work of Ding, Granger and Engle (1993) promises an additional stylised fact in asset pricing. They suggest an Asymmetric Power ARCH (A-PARCH), model to describe the long memory properties encountered in returns data. The model imposes a power transformation on the conditional standard deviation and the asymmetric absolute innovations. This still implies an exponential decay of the volatility process. In a recent paper, McCurdy and Michaud (1996) extend the A-PARCH model to the class of Fractionally Integrated A-PARCH by the introducing the FIAPARCH process.

Loudon et al. (1999) present empirical evidence on the effectiveness of eight different parametric ARCH models in describing daily stock returns. They use twenty-seven years of UK daily data on a broad-based value weighted stock index for the period 1971-1997. The results demonstrate the utility of parametric ARCH models for describing time-varying volatility in this market. The parameters proxying for asymmetry in models that recognise the asymmetric behaviour of volatility are highly significant in each and every case. However, the authors find that the various parameterizations often perform similarly, with the exception of the multiplicative

GARCH model. This model performs qualitatively differently on several dimensions of sample, suggesting that the optimal choice of a model is period specific. Performance is not consistent as we change from in-sample inference to out-of-sample inference within the same period.

Greene and Fielitz (1977) and Aydogan and Booth (1988) both use the original R/S analysis of Hurst to test for the long memory in common stock returns. Lo (1991) uses the modified re-scaled range statistic on returns from value and equally weighted CRSP indices from July 1962 through December 1987. Lo (1991) obtains significant results using the regular re-scaled range statistic and insignificant results using his modified rescaled range statistic. He attributes the difference in the test statistics to short-term persistence within the return series. He also reports finding a lack of evidence for long-range persistence in annual returns from 1872 through 1986.

Hamilton and Lin (1996) investigate the joint time series behaviour of monthly stock returns and growth in industrial production. They find that stock returns are well characterised by year-long episodes of high volatility, separated by longer quiet periods. Real output growth is subject to abrupt changes in the mean associated with economic recessions. They propose a bivariate model in which these two changes are driven by related unobserved variables, and conclude that economic recessions are the primary factor that drives fluctuations in the volatility of stock returns. This framework can then be used for forecasting stock volatility and for identifying turning points.

Fornari and Mele (1996) develop two conditionally heteroscedastic models that allow asymmetric reactions to arrival of news. Such reactions in conditional volatility are related to both the sign and size of past shocks. They propose a volatility-switching ARCH model that differs from existing asymmetric reaction models in its ability to capture a particular aspect of volatility, namely the reversion of asymmetric reactions to news. Empirical evidence from stock market returns in six countries shows that such a model outperforms traditional asymmetric ARCH equations.

Rydén et al. (1997) show that a mixture of normal variables with zero mean can generate series which display long-memory properties. In this case, the temporal higher-order dependence observed in return series can be described by a hidden Markov model. They estimate the model for ten sub-series of the well-known S&P series. which contains around 17.000 daily observations. The results reproduce the stylized facts of long-memory series quite well, but the parameter estimates sometimes vary considerably from one sub-series to the next.

Dijk et al. (1998) investigate the properties of the Lagrange Multiplier (LM) test for ARCH and GARCH in the presence of additive outliers (AOS). They show that both asymptotic size and power are adversely affected if AOS is neglected. The test rejects the null hypothesis of homoscedasticity too often when it is in fact true, and it also has difficulty detecting genuine GARCH effects. The authors then design and implement a robust test with better size and power properties than the conventional test in the presence of AOS. The tests are applied to a number of US macroeconomic time series, illustrating the danger of routinely using nonrobust tests for ARCH to diagnose misspecifications.

Wright (1999) proposes a test for non-stationarity of volatility processes by testing for a unit root in the log-squared time series. This strategy has many advantages, but is not followed in practice because these unit roots tests are known to have very poor size properties. He shows that newer tests that are robust to negative MA roots allow a reliable test for a unit root in the volatility process to be conducted. In applying these tests to exchange rate and stock returns, strong rejections of non-stationarity in volatility are obtained.

Perron (1999) suggests that the conditional variance of financial returns may exhibit sudden jumps. This was first proposed by Delgado and Hidalgo (1996), who used a non-parametric procedure to detect discontinuities in otherwise continuous functions of a random variable. Perron extends this procedure to higher moments, in particular the conditional variance. His results provide a method to identify the locatation and number of jumps.

2.6 – Conclusion

We have seen that the first attempt to model changes in volatility, which is a stylised fact in finance, was the ARCH model proposed by Engle in 1982. This technique was a belated response to a phenomenon that had been detected early in the century, namely the clustering of volatility. A natural extension of ARCH was the GARCH model proposed by Engle and Bollerslev (1986). However, although both could deal with volatility in a reasonable way, they failed to take account of persistence in volatility, which had been discovered in several studies dealing with stock prices and foreign exchange.

The challenge was to model this persistence in volatility. The first attempt was the IGARCH model, which incorporated a unit root for the second moment as another variant of ARCH had done for the mean. Many studies were produced for the advanced industrialised countries (North-America, Europe, etc.), but very little work was done on the emerging markets of Latin America and Asia.

The literature reviewed in this chapter will allow us to pursue extensive estimations in chapter four and to compare our findings with results found elsewhere. We will begin by estimating ARCH models, before turning to the GARCH specification. We will finish our examination of traditional models by using the IGARCH specification.

Next, we will estimate related models such as the asymmetric power ARCH (A-PARCH) that embodies the flexibility of these different competing models. We will estimate the Taylor/Schwert model that uses the absolute value of the returns in estimating volatility. This model stimulated other researchers to investigate alternative definitions of volatility, for example Ding et al. (1993).

We have seen that the phenomenon of persistence was first observed in hydrology, where the R/S statistics was proposed by Hurst (1951) and later revised by Lo (1981). We will use these to investigate the occurrence of long memory in volatility. Another approach would be to use the autocorrelations of returns to determine whether long memory is present in stock markets in different countries. These results would be interesting since they would not depend on any assumption about the underlying distribution of stock prices. We will conduct this analysis in chapter three, where we will be interested in R/S and autocorrelations analysis for the emerging markets.

We will introduce in chapter three as well the idea that keeping in mind the existence of an unit root in the mean of a process that conducts to persistence in the first moment, this idea can be extended to higher moments.

As it will become clearer later we have chosen to use power transformations of the returns (log-squared and so on) following Ding et al. (1993), who applied this methodology to the U.S. equity market. The suggestion of working with power transformations of returns was first made by Mandelbrot (1963) and later by Taylor (1986). Both recognized that there may be more correlation among power transformations of the returns than among the returns themselves. Since our goal has been to try to replicate such results for the emerging markets and make comparisons between developed and developing markets, it is clearly advantageous to do so using the same types of transformations that have already been studied in the developed markets.

Wright (1999) introduced a methodology in order to test for the existence of an unit root in the log of the squared returns that will enable us to identify the occurrence of persistence in volatility. This will be also pursued in chapter three.

After having tested for the existence of long memory effects, in chapter four we will turn to the estimation of long memory volatility models (FIGARCH), and will compare our results to the those obtained for well-known market indexes such as the S&P 500 (USA). We hope to determine whether the standard results for these markets are reproduced in less developed countries and especially the emerging markets.

Finally, in chapter five we will pursue the investigation of other possible explanations for the occurrence of long memory in volatility.

Chapter 3

Inference About Long Memory in Volatility in the Emerging Economies

3.1 - Introduction

The issue of long-memory in financial time series has been the subject of much debate in the econometrics literature. This discussion has produced a number of proposed measures to confirm the presence of this property. These procedures belong to one of two groups: parametric and non-parametric. These will be presented in this chapter and the next, in order to apply them to our data for emerging markets.

We will begin with a brief description of the data set used in this thesis, which will form the basis for the more thorough examination to follow.

The first empirical measure to be discussed is the Hurst exponent. We will employ the methodology developed in the previous chapter and use it to estimate Hurst exponents for each of the countries in our sample. We will then compare our findings with earlier results for developed markets.

We will then proceed to R/S analysis, which has received some criticism of late (see Lo (1991)). The R/S statistic represents an attempt to correct for the bias inherent in Hurst's original formulation. We will estimate this index for all of the countries in our sample and compare the results to established findings for developed country markets.

We will also use a methodology proposed by Wright (1999), which tests for the presence of a unit root in squared and absolute returns. This is yet another method of confirming persistence in the volatility of returns.

Finally, we will use autocorrelation analysis in order to characterise the long-memory properties of our data, following the approach used by Ding (1993). The autocorrelation function (ACF) will be employed to verify slow decay in the ACF. This will be done both for returns and powers of returns. As it will become clearer later we have chosen to use power transformations of the returns (log-squared and so on) following Ding et al. (1993), who applied this methodology to the U.S. equity market. The suggestion of working with power transformations of returns was first made by Mandelbrot (1963) and later by Taylor (1986). Both recognized that there may be more correlation among power transformations of the returns than among the returns themselves.

We have found that the same results are obtained regardless of which measure is chosen to identify long-memory in the series for the emerging economies. Our findings all point in the same direction, confirming the presence of long-memory in the data for emerging markets. This result is consistent with existing studies of markets in developing countries. What is more striking is that the magnitudes of our results are greater than those found elsewhere.

3.2 - Characterisation of the Data used

The data used in this paper consist of daily closing prices for stock market indices in several emerging markets. All indices are weighted, with the weights given by shares of stocks in transactions over a 12-month period. The primary data sources were the Economática and PACAP databases.

If we denote P_t as the price index t=0,.....,T, we can define the compounded return or logarithm return as follows, we will also be considering different definitions of volatility understood as power transformations of the absolute returns :

$$r_t = \ln P_t - \ln P_{t-1} \qquad (3.1)$$

As stated in chapter 1, for Argentina, we have 2501 daily observations from the Merval index from 1989 to 1997. For Hong-Kong. we have 6138 daily observations from 1975 to 1997. For Taiwan, we have 6124 observations from 1975 to 1997. For Brazil, we have 7829 daily observations from the Bovespa index from 1968 to 1999. For Korea, there are 5722 daily observations for 1977 through 1997. For Thailand, we have daily data from 1975 to 1997, for a total of 6133 observations. For Mexico, we use data from the IPC index from 1978 through 1997, for a total of 5523 data. Finally, there are 6144 daily observations for Malaysia from 1975 through 1997.

3.3 - R/S Statistics and the Capital Markets

Peters (1981) applies R/S analysis to the S&P 500 return's using monthly data over a 38-year period, from January 1950 to July 1988. He finds H to be 0.78, which clearly indicates that the stock market is fractal (i.e. it has long-memory) and is not characterised by a random walk. He applies a scrambling test to the series of monthly returns and the log-log plots of the two series are clearly different. H is equal to 0.51 for the reordered series, so scrambling destroys the long memory property of the original series. The sequence of price changes is important in preserving the scaling feature of the series. Changing the sequences of returns by scrambling has changed the character of the time series.

The independence assumption (EMH) seems seriously flawed. Time series of market returns are persistent, with an underlying fractal probability distribution, and seem likely to follow a biased random walk as noted by Hurst. Table 3.1 shows the results for the S&P 500 and some individual stocks identified by Peters. In this study, stocks grouped by industry tend to have similar values of H. Industries with higher levels of innovation, such as the technology industry, tend to have higher values for H. In contrast, utilities have lower levels of innovation and consequently lower values for H. If the value of H for a series is low, then the series is noisy its behaviour more random. For example, Consolidated Edison's time series is less persistent and more jagged than Apple's, which has a value of H = 0.68.[1]

Because both stocks have H values greater than 0.5 they are both considered fractal (i.e. they have long memory). A final observation is that the S&P 500 has a higher value of H than any of the individual stocks in the index. The high H value shows that diversification in a portfolio reduces risk by decreasing the noise factor and reducing randomness. Variances in the case of long-memory are frequently undefined or infinite, possibly making volatility a misleading indicator of risk. H values are associated with less noise, more persistence and clearer trends in series. It is frequently suggested that larger H values imply less risk since there is less noise in the data.

Table 3.1 - Analysis of R/S of Individual Stocks

	HURST EXPONENT (H)
S&P 500	0.78
IBM	0.72
XEROX	0.73
APPLE COMPUTER	0.75
COCA-COLA	0.70
ANHEUSER-BUSCH	0.64
McDONALD'S	0.65
NIAGARA MOHAWK	0.69
CONSOLIDATED EDISON	0.68

Source: Chaos in Capital Markets, Edgar Peters, 1991

[1] As we have already argued Hurst Exponent is a biased estimator and our intent here is that even using Hurst Exponent there is a strong evidence towards 'long memory'. But however we will contrast further these results with other estimators like GPH.

We can also calculate Hurst statistics for international markets. Table 3.2 shows Hurst exponents (H) for the U.K., Japan, and Germany using the Morgan Stanley Capital International (MSCI) index for each country. The monthly MSCI data used by Peters are for January 1959 to February 1990. If we take the S&P 500 to be representative of U.S. markets, the all four countries have different H values. The U.K. has the lowest H (0.68), followed in ascending order by Japan (0.68), Germany (0.72) and U.S. (0.78). Market efficiency can be judged by the amount of noise in the data. Because the United States has the highest H, it is the most efficient" market, since it has less noise than the others do.

Table 3.2 - Hurst Exponent for Several Countries

	HURST EXPONENT (H)
S&P 500	**0.78**
MSCI GERMANY	**0.72**
MSCI JAPAN	**0.68**
MSCI U.K.	**0.68**

Source: Chaos in the Capital Markets, Edgar Peters, 1991

We now analyse the Hurst Exponents for the emerging markets of Argentina, Brazil, Mexico, Taiwan, Korea, Thailand, Malaysia, Hong-Kong and Indonesia, where the returns shown are actually logarithmic of returns. The results are shown in Graphs 3.1 to 3.8, below. We use the same methodology proposed by Hurst (1951). That is, we use the slope of the log/log graph of R/S against T to give us an estimate of H. We have developed a procedure in RATS to compute the H exponent.

Graph 3.1 displays the results found for Argentina. We find the estimated Hurst coefficient to be equal to 0.63, which possibly indicates a long-memory[2]. Scrambling the data produces a Hurst coefficient of 0.54, which is a big change from the previous result and seems to confirm the existence of long-memory using the methodology of Hurst. Compared with data shown in Table 3.1 above, we can see that this finding for Argentina is moderate compared to markets in other countries. It is possible that we do not have enough observations to perform a valid test, so this result should be treated with caution. We must also point out that the Merval index only started in 1989, replacing the General Index that was used before.

In Graph 3.2, we present the H estimate for Hong-Kong. We found Hong-Kong's value to be 0.48. This would mean that the series is mean-reverting and possibly contains no long-memory component. Scrambling the data does not change the H estimate (0.45), an indication

[2] We should bear in mind throughout that an estimator is not a statistic for inference per si. We cannot conclude that an estimate of , let's say 0.65, shows that there is long memory. This is an estimate based on the assumption of long memory which needs to be tested by Lo/MacKinlay and actually we do perform these tests below.

that H is indeed less than 0.5 and that the market does not have long-memory. We should keep in mind that, according to the IFC, Hong-Kong is a developed market, so this may be the result of efficiency linked to development and the fact that investors exploit all gains from available information.

Graph 3.3 displays the results we found using data for Taiwan. In this case, we found the estimate of H to be equal to 0.59, an indication of long-memory. Scrambling the data gives an H value of 0.47 and seems to indicate that the series is persistent. This is because changing the order of the observations caused a big impact on the value of the H Exponent, indicating persistent behaviour of the series. The number of observations for Taiwan is very large, which gives us more confidence in our estimate of H. Details for the Taiwan stock price index can be found in the statistical appendix in chapter one. However, the degree of persistence in the Taiwan data seems to be smaller than that of Argentina as expressed by the values of the Hurst exponents (0.63 for Argentina vs. 0.59 for Taiwan).

Graph 3.4 displays the Hurst exponent for the Brazilian market using Bovespa data. The estimate for Brazil (0.83) is higher than any of the countries reported by Peters (1991) in Table 3.1 and 3.2. This strongly indicates that the series is persistent with a considerable long-memory effect present. Scrambling the data changes the estimate of H from 0.83 to 0.88, which confirms the pattern exhibited by the data. Perhaps it is not surprising to find such a large number for a country such as Brazil, where stabilisation plans have frequently been required. Also, periods of high inflation have been common due to the widely used system of indexation. Details for the methodology of the Bovespa index can be found in the statistical appendix.

In Graph 3.5 we show the results of Hurst's methodology for the Korean stock market. The estimate we found for H was 0.66, not terribly high by international standards. For example, the H values for all of the developed countries in Table 2 (U.K, USA, Germany and Japan) exceed the value found for Korea. Using the technique of scrambling the data, we found H to 0.55, giving an indication that changing the order of the observations is important and suggesting that the series has a long-term component. However, Korea's H is not as high as the value found for Brazil. Since $H > 0.5$ with both scrambled and unscrambled data, we are lead to think that this series is persistent, as stated by Hurst-Mandelbrot.

Graph 3.6 shows the estimate for Thailand. We found the value of H to be equal to 0.55, which suggests that the series is persistent. We proceed as usual by using the scrambling technique. In this case we found H to be equal to 0.65 after scrambling. Since the change is "big" we accept that the series has long memory effects. The same pattern was found for Brazil, Argentina, Taiwan and Korea. Thailand's value for H is moderate compared to Brazil's and Argentina's but large compared those for the other countries. Actually, the value of H for Thailand would be the lowest among the countries listed in Table 3.2.

Graph 3.7 shows the results for Mexico. H was found to be 0.73, a clear indication of possible long-memory effects. Scrambling the data changes the value to 0.64, suggesting that the series indeed has long-memory. Mexico's H = 0.73 is high compared to the developed countries in Tables 3.1 and 3.2. However, the effect is lower than that found for Brazil and higher than those of Argentina, Taiwan, Thailand, Korea, Hong-Kong. The Mexican market has suffered from frequent speculative attacks and the risk of crisis is almost permanent. However, even under these circumstances we do observe persistent behaviour. Perhaps the existence of a indexation in the economy can explain this result.

Finally, Graph 3.8 shows the data used for Malaysia. H is calculated to be 0.51 and scrambling the data indicates that the series is anti-persistent since the new value is 0.50, representing only a small change from the original value. One interesting thing to note is that the Latin American countries (Argentina, Brazil and Mexico) generally have higher estimates of H than those from Asia. We will return to this point later. However, we should be sceptical about the Hurst exponent since we are not certain of what constitutes a "big" change in H after scrambling. As we will see shortly, there is not a strong sampling theory provided by it to help to decide the statistical significance of our findings when applying this technique.

Tests for long-range dependence have been developed as a simple generalisation of the statistic first proposed by the English hydrologist Hurst (1951), called "rescaled Range" or "range over standard deviation" or "R/S" statistic. This statistic has been refined by Mandelbrot (among others) in some important ways. However, these refinements were not designed to distinguish between short-range and long-range dependence. This is a shortcoming in the application of R/S to returns data, since Lo and McKinlay (1988,1990) show that such data display substantial short-range dependence. Therefore, to be of current interest, any empirical investigation of long-term memory in stock returns must first account for the presence of higher frequency autocorrelation. That is, although it has been established that R/S has the ability to detect long-range dependence, this statistic is also sensitive to short-range dependence. Thus, any incompatibility between the data and the predicted behaviour of the R/S statistic under the null hypothesis of no long-run dependence need not come from long-memory, but could be merely a symptom of short-term autocorrelation. By modifying the rescaled range appropriately, Lo (1991) constructs a test statistic that is robust to short-range dependence. He also derives its limiting distribution under both short-range and long-range dependence. Contrary to the findings of Greene and Fielitz (1977) and others, when this statistic is applied to daily and monthly stock returns data over different sample periods, there is no evidence of long-range dependence once the effects of short-range dependence are accounted for. In several papers, Mandelbrot, Taqqu and Wallis demonstrate the superiority of R/S analysis to more conventional methods of determining long-range dependence, such as analysing autocorrelations, variance ratios, and spectral decompositions. For example, Mandelbrot and Wallis (1969a) can detect long-range dependence in highly non-Gaussian time series with large skewness and kurtosis. Further aspects of the R/S statistic's robustness are derived in Mandelbrot and Taqqu (1979). Although these claims may attest to the fact that long-range dependence can indeed be detected by the "classical" R/S statistic, perhaps the most important shortcoming of the rescaled range is its sensitivity to short-range dependence. This implies that any incompatibility between the data and the predicted behaviour of the R/S statistic under the null hypothesis need not come from long-term memory, but may merely be a symptom of short-term memory. As we have seen and stated before the Hurst estimator is a biased estimator, if the value moves away from 0,5 by randomising it through a sort of bootstrap technique ("scrambling"), it may only mean that we have got an unfortunate single sample with short memory that makes the estimates biased.

To see this specifically, Lo (1991) takes the example of an AR(1) stationary process. He argues that it is well known, by a Central Limit Theorem, that as T increases without bound, the R/S converges in distribution to a well defined random variable V when properly normalised, i.e.,

$$\frac{1}{\sqrt{T}}(R/S)_T \Rightarrow V \tag{3.2}$$

where "$\Rightarrow$" denotes weak convergence and V is the range of a Brownian bridge on the unit interval. Supposing that y_t is an AR(1) stationary process, it yields a R/S that does not satisfy (3.2). Actually, the limiting distribution in this case would be ξV, where $\xi \equiv \sqrt{(1+\rho)/(1-\rho)}$ and ρ is the autoregressive parameter of the AR(1) process. So if ρ is 0.5, the mean of R/S may be biased upward 73 percent ($\sqrt{(1+0.5)/(1-0.5)} = 1.73$) and since the mean of V is $\sqrt{\pi/2} \approx 1.25$ (see Feller (1951), Kennedy (1976), and Sidiqui (1976)), the mean for the classical R/S would be 2.16 (1.73*1.25) for such an AR(1) process what exceeds by far the critical values (1.862) and this would yield a rejection of the null hypothesis at any conventional significance level , so that we would be accepting the hypothesis of long-term memory. This bias is the problem with the R/S statistics. It would be possible to correct for this bias by dividing R/S by ξ, so that convergence may be restored. But this requires knowledge of ξ, and also of ρ. Moreover if y_t follows a process other than an AR(1), ξ would change so that we would have to correct it case-by-case what becomes impracticable. The estimated Hurst coefficient is not invariant to the form of short-range dependence. Davies and Harte (1987) show that even though the Hurst coefficient of a stationary Gaussian AR(1) is precisely 1/2, the 5 percent R/S statistics rejects this null hypothesis 47 percent of the time for an autoregressive parameter of 0.3. Ideally, we would like to have a form to correct for short-range dependence without requiring too many restrictions. Its limiting distribution should also be invariant to many forms of short-range dependence and still sensitive to long-range dependence.

To distinguish between long-range and short-range dependence, the R/S statistic must be modified so that its statistical behaviour is invariant over a general class of short memory processes, but that also takes in account long-memory processes. Lo (1991) constructs a test statistic that is robust to short-range dependence. He considers a modified R/S statistic in which the standard deviation becomes (the square root of) a consistent estimate of the variance of the partial sum in

$$\bar{R}_T = \frac{1}{\bar{\sigma}_T(q)}\left[\underset{1\le j\le T}{Max}\sum_{k=1}^{j}(y_k - \bar{y}_T) - \underset{1\le j\le T}{Min}\sum_{k=1}^{j}(y_j - \bar{y}_T)\right] \tag{3.3}$$

where

$$\bar{\sigma}_T^2(q) = S_T^2 + 2\sum_{k=1}^{q}\omega_j(q)\bar{\gamma}_k\ , \qquad \omega_j(q) \equiv 1 - \frac{j}{q+1}, q < n \tag{3.4}$$

and S_T^2 and $\bar{\gamma}_k$ are the usual sample variance and autocovariance estimators of y_t. $\bar{R}_T$ differs from R_T only in its denominator, which is the square root of a consistent estimator of the partial sum's variance. If y_t is subject to short-range dependence, the variance of the partial sum is not simply the sum of the variances of the individual terms, but also includes the autocovariances. Therefore the estimator $\bar{\sigma}_T^2(q)$ involves not only sums of squared deviations of the series, but also its weighted autocovariances up to lag q. The weights $\omega_j(q)$ are those suggested by Newey and West (1987). By allowing q to increase with the number of observations T, but at a slower rate, the denominator of $\bar{R}_T$ adjusts appropriately for general forms of short-range dependence. There is, however, little guidance in selecting a truncation lag q. Andrews (1991) and MacKinlay and Lo (1989) have shown that when q becomes too large relatively to T, we have a radical different finite-sample distribution from the asymptotic limit of the estimator. But, on the other side, q can not be chosen to small since by doing this, we can be skipping autocovariances beyond q that may be very high and that should be considered in the weighted sum. The truncation is then subject to the criticism of data. Andrews (1991) provides a data-dependent rule for choosing q. It is based on an asymptotic mean-squared error criterion but not too much is known about how to best choose q in finite samples.

Greene and Fielitz (1977) were probably the first to apply R/S analysis to common stock returns. Other applications are related to the price of gold (Booth and Kaen,1979), foreign exchange rates (Booth, Kaen and Koveos, 1982) and futures markets (Helms, Kaen and Rosenman, 1984). All of these empirical studies share some features in common:

1) They provide no sampling theory to determine the statistical significance of their findings.

2) They also use the "classical R/S", which is not robust to short-range dependence.

3) They do not focus on the R/S statistic itself but rather on the regression of its logarithm on (sub) sample sizes.

Point (3) is supported by Davies and Harte (1987), who show such regression tests to be significantly biased toward rejection even for a stationary AR(1) process with an autoregressive parameter of 0.3.

Lo (1991) tests for long-memory in stock returns using data from the Center for Research in Security Prices (CRSP) monthly and daily returns observations from July 1962 through December 1987, amounting to 6409 observations for the daily data and 744 observations for the monthly data. For the daily returns, the R/S statistic is significant at the 5% level but the modified R/S is not. While the modified R/S is found to be 1.46, the classical R/S is 2.63. The bias in this case amounts to some 80%! The statistical significance of the modified R/S is consistent with the short-memory null hypothesis. For the monthly returns, none of the modified statistics was found to be significant.

We present the results for the classical R/S statistic and the modified R/S statistics as proposed by Lo (1991) for Argentina, Brazil, Mexico, Thailand, Korea, Taiwan, Malaysia and Hong-Kong. Table 3.3 presents the values for the modified R/S statistics. We used a procedure in Gauss based on the Andrew's (1991) data-dependent formula to choose the truncation lag (q) in the process of computing the modified R/S. Table 3.3 also measures the bias of the classical R/S when compared to the modified R/S. The bias is defined to be: ((classical R/S)/(modified R/S)-1)*100. We also contrast these results with those for the US Market described by the Standard & Poor's index (S&P).

The statistics computed in table 3.3 have a distribution with critical values given in Lo (1991, pg.1288, Table II). Using these values, we can test the null hypothesis at the 95 percent level of confidence by accepting or rejecting according to whether the modified R/S is or is not contained in the interval (0.809,1.862) which assigns equal probability to each tail. The classical R/S is also shown and the Hurst exponent. Tests marked with an asterisk indicate those that are significant at the 95 percent level.

Table 3.3 - Modified R/S Statistics for the Countries under Analysis

	CLASSICAL R/S	MODIFIED R/S	BIAS	HURST EXPONENT
ARGENTINA	2.18*	2.15*	1.4%	0.63
BRAZIL	5.81*	4.36*	33.3%	0.83
MEXICO	3.05*	2.37*	28.7%	0.73
TAIWAN	1.72	1.45	18.6%	0.59
THAILAND	1.75	1.55	12.9%	0.55
KOREA	1.90*	1.74	9.2%	0.66
USA	1.25	1.16	7.76%	0.52
MALAYSIA	1.70	1.43	18.9%	0.51
HONG-KONG	1.15	1.06	8.5%	0.48

Some interesting results can be seen in Table 3.3. The first is the fact that all three Latin American markets show long-range dependence whether measured by the classical R/S, the modified R/S or the Hurst exponent. For Argentina, we found values of 2.18 (classical), 2.15 (Modified) and 0.63 (Hurst Exponent). The bias is relatively small compared to that found by Lo in his 1991 paper. The same pattern is found for Brazil: 5.81 (classical), 4.36 (Modified) and 0.83 (Hurst Exponent). The effects in the case of Brazil appear to be much larger than for Argentina, and the process is highly persistent. The bias, however, is higher: 33.3 percent. Mexico is an intermediate case between Argentina and Brazil. The classical R/S was found to be 3.05, the modified 2.37 and the Hurst Exponent 0.73. All of the values for Mexico are between those for Argentina and Brazil, and there is also evidence of long-range dependence. The bias is 28.7 percent.

For the Asian countries, none of the statistics are significant at the 95 percent level except for Korea, where the classical R/S was found to be 1.90. However, the modified R/S is not significant. For all other countries there is no significance for both statistics. Concerning the Hurst Exponent, it only follows the pattern indicated by the R/S analysis in the case of Hong-Kong and possibly Malaysia. We do find confirmation of the fact that Latin American markets present more long-range dependence than Asian markets. It is also interesting to see that the results for the US markets follow the same pattern as those found for the Asian markets in the sense that we find no significance for both statistics. In our view, this may reveal two different things. The Latin American markets have suffered from many interventions during the period under consideration, which may have created confusion among investors until they understood the nature of each new measure. This means that after new information arrives in the market it takes some time for investors to digest it fully and these markets were far more closed than the Asian/US markets. Foreign investors react much more quickly than domestic investors. This could contribute to long-range dependence, especially in case of Brazil. It could also be explained by the widespread practice of indexation.

In order to put this into perspective and provide some intuition, we show the autocorrelation functions up to 260 lags (approximately one year) for the countries under consideration. These are depicted in Graphs 3.9a and 3.9b below. For the Asian countries, only the lowest order autocorrelation coefficients are statistically significant. For the Latin American countries, the highest order coefficients are still statistically significant.

3.4 Unit Roots in the Volatility of Stock Returns

Another way to test for long-memory in the volatility of returns is to consider a stochastic volatility model where the volatility process is non-stationary.[3] The stochastic volatility model implies that the log of the squared time series is an ARMA process, the largest autoregressive root of which is the same as the largest autoregressive root of the volatility process. It is then possible to test for a unit root in the unobserved volatility process by testing for a unit root in the log of the squared time series. This test is straightforward and does not require a distribution to be specified for the error term. However, it is a well established result that standard unit roots suffer from extreme size distortions in the presence of negative MA roots. Perron and Ng (1996) have proposed modified unit roots tests which are robust. Wright (1999) proposed using Perron's method to test for a unit root in the log of the squared time series and hence to test for a unit root in the volatility process.

Let's start by considering the standard autoregressive stochastic volatility (ARSV) model which specifies that $\{y_t\}_{t=1}^{T}$ is a time series of returns such that

$$y_t = \sigma_t \varepsilon_t$$

The model further specifies that ε is i.i.d. with mean zero and variance 1. $\log(\sigma_t^2) = \mu + h_t$, $a(L)h_t = \eta_t$ and $a(L) = b(L)(1 - \alpha L)$ is a pth-order autoregressive lag polynomial such that b(L) has all roots outside the unit circle. The parameter α is the largest autoregressive root of the volatility process. It is assumed that η_t is i.i.d. with mean zero and variance σ_η^2 and is distributed independently of ε_t. Clearly we can write,

$$a(L)\log(y_t^2) = a(1)\mu + \eta_t + a(L)\log(\varepsilon_t^2)$$

$$a(L)\log(y_t^2) = \varpi + \eta_t + a(L)\zeta_t = \varpi + x_t \qquad (3.5)$$

[3] See Hansen, 1992; Harvey et al., 1994; and Ruiz, 1994.

where $\xi_t = \log(\varepsilon_t^{\,2}) - E(\log(\varepsilon_t^{\,2}))$, $\bar{\varpi} = a(1)(\mu+E(\log(\varepsilon_t^{\,2})))$ and $x_t = \eta_t + a(L)\,\xi_t$. If we are only interested in deciding whether $\alpha = 1$ or not, then we can use an approach that does not require any distributional assumption to be made concerning the error term (unlike for estimation) and that is simple to apply for any value of p. The time series $x_t = \eta_t + a(L)\,\xi_t$ has a Wold representation and, from inspection of its autocovariance funcion, this is non-MA(p) reduced form. It follows from equation (3.5) that $\log(y_t^{\,2})$ is a stationary ARMA(p,q) process if $|\alpha| < 1$, but it is an ARIMA process if $\alpha = 1$. So, one may test the hypothesis that $\alpha = 1$ by testing for a unit root in $\log(y_t^{\,2})$ using, in principle, any one of the unit roots tests available in the econometric literature.

However, it has been pointed out by Harvey (1994) that these unit root tests have very poor size properties, and thus they attach little significance to the rejection of the unit root null that they obtained using exchange rate and stock market data. The presence of a large negative moving average root is known to cause serious distortions in standard unit roots in finite samples. Accordingly, standard unit root tests applied to the log of squared time series on asset returns may in principle be interpreted as tests for a unit root in the volatility process, but should in practice suffer from serious finite sample size distortions as suggested by Wright (1999), and possibly even in the large sample sizes that are available for asset returns data.

Perron and Ng (1996) have proposed modified unit root tests which have much better finite sample properties in the presence of large negative MA roots. Wright (op.cit.) tests the hypothesis that $\alpha = 1$ against the alternative $|\alpha| < 1$ applying these tests to $\log(y_t^{\,2})$. The three statistics are

$$MZ_\alpha = \left[T^{-1}(v_T - \bar{v})^2 - s^2\right]\left[2T^{-2}\sum_{t=1}^{T}(v_t - \bar{v})^2\right]^{-1}$$

$$MSB = \left[s^{-2}T^{-2}\sum_{t=1}^{T}\ (v_T - \bar{v})^2\right]^{1/2}$$

$$MZ_t = MZ_\alpha . MSB$$

where $v_t = \log(y_t^{\,2})$, $\bar{v} = T^{-1}\sum_{t=1}^{T} v_t$ and s^2 is the autoregressive spectal density estimate obtained from the autoregression

$$v_t = a_0 + a_1 v_{t-1} + \sum_{j=1}^{k} a_j \Delta v_{t-j} + e_t$$

where $k = o(T^{1/3})$. These unit roots have power against local alternatives in a T^{-1} neighbourhood of unity. This contrasts with maximum-likelihood tests for a unit root in a GARCH/IGARCH model which use root-T asymptotics and correspondingly have power only in a $T^{-1/2}$ neighbourhood of unity.

The unit root tests are designed to have power against the alternative that the volatility is a stationary autoregression (the ARSV model with $|\alpha| < 1$). The squared, log-squared and aboslute value of stock returns often have correlograms that decay very slowly, as we showed in chapter three. This finding has motivated some researchers to propose a fractionally integrated stochastic volatility model. A simple case of this model specifies that $\{y_t\}_{t=1}^{T}$ is a time series of returns such that

$$y_t = \sigma_t \varepsilon_t$$

where ε is i.i.d. with mean zero and variance 1, $\log(\sigma_t^2) = \mu + h_t$ and $(1 - L)^d(1 - \alpha L)h_t = \eta_t$, $(1 - L)^d$ denotes the fractional differencing operator and η_t is i.i.d. $N(0,\sigma^2_\eta)$ and is independent of ε_t. More generally h_t can be an arbitrary Gaussian fractional ARIMA process. He shows the tests to have power agains the FISV alternatives.

We apply these tests for a unit root in volatility developed by Wright (1999) to the stock returns data used in our thesis. The procedures described above were then applied to testing for a unit root in the volatility of each of the series of stock returns for Argentina, Brazil, Mexico, Taiwan, Thailand, Hong-Kong, Korea and Malaysia. For comparison, the familiar Z_α, Z_t, and ADF statistics were also used. The results are reported in Tables 3.4 to Table 3.19 for both log-squared returns and absolute returns). In obtaining s^2, the autoregressive spectral density estimator, the results are reported for k = 5, 10, 15 and 20. For the $Z_\alpha$and Z_t, statistics s^2 was used as the spectral density estimate. For the ADF test, k lags of the differenced data were added to the Dickey-Fuller regression.

Table 3.4 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Thailand, k = 5	-16,84	-1.579,6	-50,54	-487,84	0,032	-15,61
Thailand, k = 10	-11,52	-1.252,5	-69,73	-160,76	0,057	-8,95
Thailand, k = 15	-9,34	-1.181,8	-87,76	-90,14	0,074	-6,69
Thailand, k = 20	-8,11	-1.157,4	-100,58	-65,67	0,087	-5,71

Note: All test statistics are significant at 1% level

Table 3.5 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Brazil, k = 5	-22,20	-2.850,8	-72,44	-774,13	0,025	-19,76
Brazil, k = 10	-14,50	-2.299,1	-108,96	-222,51	0,047	-10,55
Brazil, k = 15	-10,76	-2.172,4	-156,89	-95,75	0,072	-6,92
Brazil, k = 20	-8,88	-2.133,6	-199,74	-56,94	0,094	-5,33

Note: All test statistics are significant at 1% level

Table 3.6 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Argentina, k = 5	- 9,70	-663,32	-38,41	-147,79	0,058	- 8,56
Argentina, k = 10	-5,86	-548,83	-65,94	-33,46	0,120	-4,22
Argentina, k = 15	-4,68	-533,65	-85,55	-18,29	0,160	-2,93
Argentina, k = 20	-4,07	-528,69	-98,19	-13,34	0,186	-2,48

Note: All test statistics are significant at 1% level

Table 3.7 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Mexico , k = 5	-20,86	-2.280,3	-57,60	-783,49	0,025	-19,79
Mexico, k = 10	-14,61	-1.767,6	-75,94	-270,78	0,043	-11,63
Mexico, k = 15	-12,40	-1.680,3	-87,68	-183,52	0,052	-9,58
Mexico, k = 20	-10,83	-1.619,2	-103,45	-122,40	0,064	-7,82

Note: All test statistics are significant at 1% level

Table 3.8 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Korea , k = 5	-20,03	-2.512,2	-71,96	-610,27	0,029	-17,46
Korea, k = 10	-13,81	-2.112,8	-103,50	-207,87	0,049	-10,18
Korea, k = 15	-11,01	-2.013,4	-136,37	-108,53	0,068	-7,35
Korea, k = 20	-8,91	-1.961,7	-103,23	-56,86	0,093	-5,31

Note: All test statistics are significant at 1% level

Table 3.9 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Taiwan , k = 5	-19,48	-2.690,7	-81,86	-539,07	0,030	-16,40
Taiwan, k = 10	-12,41	-2.291,8	-136,29	-140,33	0,060	-8,35
Taiwan, k = 15	-9,34	-2.215,5	-194,15	-64,08	0,088	-5,62
Taiwan, k = 20	-8,29	-2.197,7	-226,02	-46,25	0,103	-4,76

Note: All test statistics are significant at 1% level

Table 3.10 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Hong-Kong , k = 5	-21,74	-2.860,8	-75,35	-720,75	0,026	-18,98
Hong-Kong, k = 10	-14,27	-2.338,4	-117,39	-198,34	0,050	-9,98
Hong-Kong, k = 15	-11,39	-2.242,9	-156,35	-102,86	0,070	-7,17
Hong-Kong, k = 20	-9,98	-2.211,0	-185,54	-70,96	0,084	-5,96

Note: All test statistics are significant at 1% level

Table 3.11 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Malaysia , k = 5	-21,88	-2.630,4	-64,95	-820,04	0,025	-16,40
Malaysia, k = 10	-14,78	-2.056,1	-92,72	-245,76	0,045	-11,08
Malaysia, k = 15	-11,69	-1.924,2	-127,48	-113,82	0,066	-7,54
Malaysia, k = 20	-10,49	-1.893,3	-146,90	-82,97	0,078	-6,44

Note: All test statistics are significant at 1% level

Table 3.12 – Unit Root test statistics for log-squared stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Thailand, k = 5	-16,44	-1.309,7	-41,27	-503,23	0,032	-15,86
Thailand, k = 10	-11,52	-999,1	-50,87	-192,63	0,051	-9,81
Thailand, k = 15	-9,84	-920,9	-60,79	-114,55	0,066	-7,56
Thailand, k = 20	-8,64	-899,0	-66,00	-92,57	0,073	-6,80

Note: All test statistics are significant at 1% level

Table 3.13 – Unit Root test statistics for absolute stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Brazil, k = 5	-18,01	-2.149,6	-66,44	-523,08	0,031	-16,17
Brazil, k = 10	-12,02	-1.792,8	-98,23	-166,34	0,055	-9,11
Brazil, k = 15	-9,78	-2.172,4	-156,89	-95,75	0,072	-6,92
Brazil, k = 20	-8,88	-1.695,0	-143,77	-69,36	0,084	-5,88

Note: All test statistics are significant at 1% level

Table 3.14 - Unit Root test statistics for absolute stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Argentina, k = 5	-8,97	-498,76	-29,48	-142,83	0,059	- 8,44
Argentina, k = 10	-6,17	-403,30	-41,30	-47,40	0,102	-4,85
Argentina, k = 15	-5,63	-395,32	-44,36	-39,42	0,112	-4,42
Argentina, k = 20	-4,94	-385,99	-49,52	-30,10	0,128	-3,86

Note: All test statistics are significant at 1% level

Table 3.15 – Unit Root test statistics for absolute stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Mexico , k = 5	-24,55	-2.539,9	-40,68	-1948,70	0,016	-31,21
Mexico, k = 10	-18,09	-1.694,3	-36,07	-1.102,8	0,021	-23,48
Mexico, k = 15	-15,00	-1.339,7	-34,63	-748,51	0,026	-19,35
Mexico, k = 20	-12,96	-1.068,6	-34,58	-476,84	0,032	-15,44

Note: All test statistics are significant at 1% level

Table 3.16 – Unit Root test statistics for absolute stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Korea , k = 5	-18,39	-2.020,6	-61,14	-545,82	0,030	-16,52
Korea, k = 10	-13,99	-1.746,4	-74,88	-271,62	0,043	-11,65
Korea, k = 15	-11,96	-1.666,3	-85,06	-191,52	0,051	-9,78
Korea, k = 20	-10,20	-1.594,1	-103,04	-119,32	0,065	-7,72

Note: All test statistics are significant at 1% level

Table 3.17 – Unit Root test statistics for absolute stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Taiwan , k = 5	-14,66	-1.621,6	-62,38	-337,41	0,039	-12,98
Taiwan, k = 10	-9,27	-1.385,7	-97,03	-101,56	0,070	-7,11
Taiwan, k = 15	-7,10	-1.341,9	-124,39	-57,78	0,092	-5,36
Taiwan, k = 20	-6,52	-1.325,9	-144,32	-41,80	0,109	-4,55

Note: All test statistics are significant at 1% level

Table 3.18 – Unit Root test statistics for absolute stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Hong-Kong , k = 5	-19,58	-2.082,6	-58,17	-640,66	0,028	-17,90
Hong-Kong, k = 10	-12,97	-1.641.2	-82,19	-199,23	0,050	-9,98
Hong-Kong, k = 15	-11,12	-1.570,5	-97,88	-128,58	0,062	-8,01
Hong-Kong, k = 20	-10,51	-1.561.	-101,00	-119,35	0,054	-7,72

Note: All test statistics are significant at 1% level

Table 3.19 – Unit Root test statistics for absolute stock returns

	ADF	Z_α	Z_t	MZ_α	MSB	MZ_t
Malaysia , k = 5	-18,90	-1.726,3	-47,58	-657,99	0,028	-18,14
Malaysia, k = 10	-13,02	-1.304,5	-60,00	-236,20	0,046	-13,02
Malaysia, k = 15	-11,78	-1.233,3	-67,86	-164,94	0,055	-9,08
Malaysia, k = 20	-11,54	-1.207,1	-72,41	-138,74	0,060	-8,32

Note: All test statistics are significant at 1% level

The $Z_\alpha$and Z_t statistics yield overwhelming rejections, specially for Mexico. The ADF statistics also yield rejections at all conventional significance levels, though are less extreme than the $Z_\alpha$and Z_t statistics. Using the unit root tests that are robust to a large MA root, the hypothesis of a unit root in the volatility is clearly rejected at all conventional significance levels, regardless of the choice of k, for all countries and irrespective of log-squared stock returns or absolute stock returns. In light of the fact that these tests control size reasonable well, this is strong evidence against the model of a unit root in the volatility process.

This indicates that, while there is considerable persistence in the volatility of returns, a unit root in the stochastic volatility model is too extreme a specification. Models in which the volatility process is modelled as an AR(p) with a large root (but not a unit root) or in which the volatility process is fractionally integrated may provide a better representation of the data.

3.5 - Autocorrelation analysis of Stock Market Returns

With the availability of high frequency long time series from returns on speculative assets, much research has been devoted to the study of long-run behaviour of financial data. A common finding in much of the empirical literature is that, contrary to what was previously thought, the returns themselves carry little serial correlation, which is an indication that the efficient markets hypothesis holds. However, the absolute returns and their power transformations are highly correlated. Taylor (1986) was the first to study this in a systematic way. $|r_t|^d$ has significant positive serial correlation over long lags. These findings were rediscovered by Ding, Granger and Engle (1993). They examined this property for long daily stock market price series. They found it possible to characterise $|r_t|^d$ as having 'long memory', with quite high correlations for long lags. In this paper they found positive autocorrelation for the S&P 500 series for more than 2700 lags with a series of 17054 observations. Similar results were also found for other values of d, and this seems to be strongest when d = 1 compared to both smaller and larger values of d. This result seems to argue against ARCH specifications based upon squared returns.

Table 3.20 gives the summary statistics for r_t (returns itself[4]) for Argentina, Brazil, Mexico, Korea, Taiwan, Thailand, Malaysia, Hong-Kong and US. Table 3.20 can also be used to illustrate some stylised facts that have been noted other researchers and that we reproduce here.

---

[4] In what follows returns will mean the returns itself.

Table 3.20 -Summary Statistics for the Daily Returns

STATISTICS	ARG	BRA	MEX	KOR	TAI	THAI	MAL	US	HK
MIN	-0.313	-0.223	-0.183	-0.159	-0.078	-0.094	-0.156	-0.228	-0.333
MAX	0.336	0.360	0.266	0.053	0.067	0.101	0.144	0.154	0.123
MEAN	0.0026	0.0049	0.0021	0.0005	0.0007	0.0005	0.0008	0.0002	0.0090
STD	0.043	0.029	0.0210	0.0112	0.016	0.0132	0.0148	0.0115	0.0170
KURTOSIS	25.57	8.635	20.42	11.22	2.549	12.28	11.25	25.42	35.73
RANGE/STD	15.09	20.10	21.38	19.02	9.13	7.65	20.27	33.04	26.82
SKEWNESS	0.497	1.231	0.872	-0.365	-0.015	0.681	0.382	-0.487	-1.761
NORMALITY	46663	23858	76812	27042	1512	31031	25928	357788	278604
CLASSICAL R/S	2.18	5.81	3.05	1.90	1.72	1.75	1.70	1.25	1.15
MODIFIED R/S	2.15	4.36	2.37	1.74	1.45	1.55	1.43	1.16	1.06
VARIANCE	3707.4	2341.9	1467.6	346.9	507.8	478.9	580.6	360.2	1385.4
SAMPLE SIZE	2501	7879	5223	6022	6524	5583	5674	17054	6138
RANGE	0.649	0.583	0.449	0.213	0.146	0.101	0.300	0.380	0.456

There exists a wide variety of opinion about the distributions of stock price returns and the data generating processes. Some authors claim the distributions to be Paretian stable (McFarland et al., 1982), some say they follow the Student's t distribution (Boothe and Glassman, 1987), and others reject any single distribution (Calderon-Rossel and Ben-Horim, 1982).

Instead of looking at the centre of the distribution, an alternative way to characterise the distribution is to look at the tails.

(i) Thin-tailed distributions are those for which all moments exist and whose cumulative distribution function declines exponentially in the tails.
(ii) Fat-tailed distributions are those whose cumulative distribution function declines with a power in the tails
(iii) Bounded distributions are those with no tails.

A nice result is that these categories can be distinguished by the use of only one parameter, the tail index for distributions of category (i), for category (ii) and for category (iii).

The empirical estimation of the tail index and its variance crucially depends on the size of the sample. On the one hand, using too many observations introduces a bias in the tail index since some of the observations do not belong to the tail anymore. On the other hand, using too few observations introduces inefficiency in the estimation of the variance. The very large sample sizes available with intra-daily data ensure that enough tail observations will be present.

An important result is that the tails of a fat-tailed distribution are invariant under addition, although the distribution as a whole may vary according to temporal aggregation (see Feller, 1971). That is, if weekly returns are identically and independently distributed Student's t, then monthly returns will not follow the t distribution. However, the tails of the monthly returns distribution will be like the tails of the weekly returns, with the same exponent.[5]

Another important result in the case of fat-tailed distributions concerns the existence of the moments of the distribution. It can be easily seen that only the first k-moments of the distribution are bounded. Finally, the tail index reflects the interaction of different agents in the market. Indeed, the probability of extreme events depends on the presence or absence of certain market participants such as medium-term investors or pure speculators, as well as on changing market conditions.

We also use a kernel estimator to compare the empirical distribution of returns with a Standard Normal distribution. We can easily see from Graphs 3.10a and 3.10b that the distribution for the returns have fatter tails than the standard normal for all countries, which seems to confirm the fact that the distribution for the returns have fat-tails. We can also see in Table 3.20 that the kurtosis for r_t is higher than that of a normal distribution (3.0) for all countries except for Taiwan and Korea. The kurtosis varies from 35.73 for Hong-Kong to 2.549 for Taiwan. The Jarque-Bera test for normality is far beyond the critical values of 5.99 (at the 5 percent level) and 7.38 (at the 2.5 percent level). We reject the hypothesis of normality for the distribution of the returns without exception. The Jarque-Bera test statistic varies from 1512 (Taiwan) to 278604 (Hong-Kong). Table 3.20 indicates that the returns belong to the class of fat-tailed non-stable distributions that have a finite tail index. We also find evidence of 'fat-tailed' behaviour for the United States. Here, the Jarque-Bera statistic is also far beyond conventional critical values, which suggests that returns are far from being distributed normally. This is consistent with our findings for the emerging economies.

Table 3.20 suggests that the variance exists for all stock markets. We also note that the variances for the Latin American countries are higher than those for their Asian counterparts (with the exception of Hong Kong), a fact that we noticed previously in Chapter 1. Again we find that markets in Asian countries behave more like U.S. markets than Latin American markets. The degree of intervention in Latin American economies could help to explain this higher variance.

It appears from Table 3.20 that the third moment exists. Skewness is also computed for all countries, and the distributions are found to be symmetric in every case. We can see this either by looking at Graphs 3.10a and 3.10b or by examining the skewness statistics. We do not reject the null hypothesis for all countries except for Hong-Kong. We observe a small amount of right skewness for Brazil and Argentina, while the US skewness is to the left. What is important to

[5] Actually the kurtosis and the studentized range statistics (which is the range divided by the standard deviation) are used as an index of the tail behaviour compared to a normal distribution. For those more interested in the tail index, a good reference is Koedijk et allii (1990)

remember here is that returns do not follow a normal distribution, and that this is true for emerging markets as well for developed ones.

We can see from Graphs 3.10a and 3.10b and from Table 3.20 that the distributions of returns for Argentina, Taiwan, Hong-Kong, Thailand, Korea and US are leptokurtic. For Brazil and Mexico, they are platykurtic. Finally, Malaysia's distribution it is mesokurtic.

Table 3.20 shows the classical and modified R/S statistics again. Just as a reminder, we have seen that Latin American markets display long-memory effects, whereas this property is not present for U.S. and Asian markets.

In Graphs 3.11a and 3.11b we plot the returns for all countries in our sample. We can see that r_t is stable around the mean in every case. We also observe some clustering in the data, i.e. large (small) returns are more likely to be followed by more large (small) returns (see, Ding et all). For comparison, we also show the series of absolute returns for all countries in Graphs 3.12a and 3.12b. We can clearly see the observation of Mandelbrot (1963) and Fama (1965) that large absolute returns are more likely than small absolute returns to be followed by a large absolute return. Market volatility changes over time, which suggests that a suitable model for returns should have a time varying volatility structure similar to that suggested by the ARCH methodology. Every market shows some periods of increased volatility. For example, volatility was much higher during the Great Depression than in any other period. There was also a sudden drop in share prices after the Black Monday crash of 1987, but unlike the Great Depression, the increased volatility did not last very long. To establish a comparison to emerging markets, Brazil also experienced a period of extremely high volatility during the hyperinflation of 1989-93. Thus, emerging and developed markets both show periods of higher volatility followed by periods of relative calm, and these periods tend to be clustered.

Until now we have been concerned with returns themselves, but we should also consider higher moments as well. It is now an established fact that stock market returns themselves contain little serial correlation (see Fama (1970) and Taylor (1986)). This is a point in favour of the efficient markets hypothesis (EMH). We have already detected a long-memory effect in the Latin American countries, but this standard finding does not appear to apply in the Asian countries. In Graphs 3.13a and 3.13b, where we plot the autocorrelation functions (of the returns) up to 2500 lags, we see that some countries do indeed display persistent correlation over time, as is the case for Brazil and Taiwan.

Taylor (1986) studied the correlations between 40 different transformed returns series and concluded that absolute returns and squared returns are even more highly correlated than the original series. Following this study, as well as Granger (1993), we will examine the autocorrelations for r_t and $|r_t|^d$ with positive d's. Table 3.21 to 3.29 gives the sample autocorrelations for r_t, $|r_t|$ and r_t^2 with lags 1 to 5, 10, 20, 30, 40, 70 and 100 for each country. We also plot the autocorrelograms for r_t, $|r_t|$ and r_t^2 with lags 1 to 100 in Graphs 3.14 to 3.21. The 95 percent confidence interval for the estimated sample correlation if the process r_t is independently and identically distributed (i.i.d.) is $\pm 1.96/\sqrt{T}$. In our data, T ranges from 2501 for Argentina to 7829 in the case of Brazil. There is a lower bound of $\pm 1.96/\sqrt{T}$ = 0.0392 (or 0.015, in the case of US) and an upper bound of $\pm 1.96/\sqrt{T}$ = 0.0222. Bartlett (1946) proved that if r_t is an i.i.d. process then the sample correlation ρ_t is approximately N(0,1/T). In Graphs 3.14 to 3.21 and also in Tables 3.21 to 3.29, about one quarter of the sample autocorrelations within lag 100 are outside the 95 percent confidence interval for an i.i.d process. This is true for all countries including the U.S. The first autocorrelations are significantly positive for all countries. Other researchers have found similar evidence, for example Fama (1976), Taylor (1986), Ding et al. (1993). In some of the countries the first order autocorrelations are very small (e.g. Argentina, Hong Kong, and US), which could suggest that r_t does have some memory albeit very short. Some portion of stock market returns is predictable, even though it might not be important in practice. As a result of these finding we can conclude that the EMH or random walk theory does not hold strictly.

Actually, we can establish an interesting comparison with the results found for US markets. This overall picture would lead us to conclude that emerging markets generally do show higher persistence, taking the ACF into account. However, the degree of persistence differs across countries. For example, in Tables (3.21 to 3.29) Brazil and Taiwan have higher ACF values compared to other countries, an indication of a higher persistence. This is also true for other emerging markets compared to U.S. market. We will return to this result later on. On the other hand, some markets such as Hong-Kong and Korea have ACFs that die out very quickly, even more quickly than the U.S. This is an indication that these markets would have short memory compared to the U.S. market. This can be explained by the fact that these markets are much smaller than the U.S. markets and so the lack of synchronization is much smaller affecting the degree of persistence. In Hong-Kong, for instance, only a few stocks are negotiated. Another interesting characteristic that appears in these tables is that some markets have strong 'mean-reverting' tendencies. This is the case for Thailand and Malaysia.

However, for the other markets the first lag is highly positive, which confirms the unreliability of the efficient market hypothesis. The second lag is significantly negative for some of the countries (Argentina, Mexico, Thailand, Hong-Kong, Korea and US), and this supports the hypothesis of "mean-reversion" in stock market returns. This behavior suggests that the stock market returns in our sample are not realisations of i.i.d. processes.

TABLE 3.21 - AUTOCORRELATION FUNCTION - TAIWAN

	NUMBER OF LAGS										
DATA	1	2	3	4	5	10	20	30	40	70	100
r_t	0.1160	0.0072	0.1117	0.0398	0.0078	0.0402	0.0086	-0.0233	-0.0180	0.0269	0.0190
$\|r_t\|$	0.3325	0.3746	0.3743	0.3626	0.3459	0.3427	0.2989	0.2858	0.2814	0.2568	0.2151
r_t^2	0.3558	0.4010	0.3923	0.3737	0.3691	0.3423	0.3236	0.3010	0.2945	0.2658	0.2023

TABLE 3.22 - AUTOCORRELATION FUNCTION - THAILAND

	NUMBER OF LAGS										
DATA	1	2	3	4	5	10	20	30	40	70	100
r_t	0.1407	-0.0175	0.0281	0.0465	-0.0058	-0.0140	0.0118	0.0306	0.0001	-0.0007	-0.0383
$\|r_t\|$	0.4261	0.3789	0.3502	0.2999	0.2782	0.2395	0.1840	0.1704	0.1613	0.1156	0.0910
r_t^2	0.3248	0.2864	0.2556	0.2045	0.1403	0.1511	0.0872	0.0753	0.1022	0.0472	0.0425

TABLE 3.23 - AUTOCORRELATION FUNCTION - MALAYSIA

	NUMBER OF LAGS										
DATA	1	2	3	4	5	10	20	30	40	70	100
r_t	0.1459	0.0172	0.0544	0.0514	0.0154	0.0489	-0.0133	-0.0036	0.015	-0.0030	0.0102
$\|r_t\|$	0.3438	0.3057	0.2709	0.2048	0.2139	0.1826	0.0862	0.0864	0.1050	0.0469	0.0496
r_t^2	0.3699	0.2856	0.2683	0.1696	0.1246	0.1749	0.034	0.0543	0.0541	0.0057	0.0116

TABLE 3.24 - AUTOCORRELATION FUNCTION - HONG-KONG

	NUMBER OF LAGS										
DATA	1	2	3	4	5	10	20	30	40	70	100
r_t	0.0688	-0.0042	0.0681	0.0129	-0.0129	-0.0006	-0.0091	0.0182	-0.0024	-0.0197	-0.0041
$\|r_t\|$	0.2682	0.2350	0.2199	0.2005	0.1899	0.1736	0.1163	0.0951	0.1035	0.0538	0.0203
r_t^2	0.1415	0.0425	0.00602	0.0395	0.0308	0.0543	0.0181	0.0236	0.0258	0.0029	-0.0016

TABLE 3.25 - AUTOCORRELATION FUNCTION - MEXICO

	NUMBER OF LAGS										
DATA	1	2	3	4	5	10	20	30	40	70	100
r_t	0.2760	-0.0360	-0.0700	0.0200	0.0830	0.030	0.009	0.007	-0.0009	0.014	0.007
$\|r_t\|$	0.397	0.249	0.214	0.220	0.196	0.190	0.147	0.115	0.138	0.088	0.039
r_t^2	0.333	0.212	0.184	0.159	0.131	0.181	0.129	0.077	0.07	0.038	0.012

TABLE 3.26 - AUTOCORRELATION FUNCTION - ARGENTINA

	NUMBER OF LAGS										
DATA	1	2	3	4	5	10	20	30	40	70	100
r_t	0.083	-0.117	0.040	0.105	0.081	0.0246	-0.0229	0.0778	0.00924	0.006	-0.0296
$\|r_t\|$	0.3603	0.3653	0.2943	0.2915	0.3050	0.2534	0.2219	0.2324	0.1524	0.1062	0.0752
r_t^2	0.1848	0.2974	0.0941	0.1091	0.1083	0.0560	0.1552	0.1229	0.0327	0.0055	0.0133

TABLE 3.27 - AUTOCORRELATION FUNCTION - BRAZIL

	NUMBER OF LAGS										
DATA	1	2	3	4	5	10	20	30	40	70	100
r_t	0.1952	0.0705	0.0349	0.0663	0.0767	0.0343	0.0159	0.0474	-0.0072	0.0317	0.0350
$\|r_t\|$	0.3295	0.3163	0.2967	0.3018	0.3014	0.2894	0.2475	0.2208	0.1948	0.2025	0.1970
r_t^2	0.1693	0.1920	0.1189	0.1527	0.1645	0.1686	0.1039	0.0686	0.0565	0.0781	0.0630

TABLE 3.28- AUTOCORRELATION FUNCTION - KOREA

	NUMBER OF LAGS										
DATA	1	2	3	4	5	10	20	30	40	70	100
r_t	0.1233	-0.0340	0.0152	0.0175	0.0073	-0.0002	-0.0050	-0.0252	0.0278	-0.0041	-0.0024
$\|r_t\|$	0.2530	0.2571	0.2503	0.2197	0.1977	0.1378	0.0810	0.0742	0.1045	0.0827	0.0773
r_t^2	0.0916	0.0865	0.0850	0.0591	0.0577	0.0330	0.0050	0.0052	0.0253	0.0155	0.0136

TABLE 3.29- AUTOCORRELATION FUNCTION - US

DATA	1	2	3	4	5	10	20	40	70	100
r_t	0.063	-0.0390	-0.0040	0.0310	0.0220	0.0180	0.0170	0.0000	0.0000	0.0040
$\|r_t\|$	0.318	0.3230	0.3220	0.2960	0.3030	0.2470	0.2370	0.2000	0.1740	0.1620
r_t^2	0.218	0.2340	0.1730	0.1400	0.1930	0.1070	0.0830	0.0590	0.0580	0.0450

Furthermore, if r_t is an i.i.d. process, then any transformation of r_t is also i.i.d., for example $|r_t|$ and r_t^2. The standard error of $|r_t|$ will be $1/\sqrt{T}$ =[0.02396 (Argentina) to 0.011014 (Brazil)] if r_t has finite variance, and the same standard error is applicable for the sample autocorrelations of r_t^2 providing r_t has also finite kurtosis. From Graphs 3.14 to 3.21 we can see that, not only are the sample autocorrelations of $|r_t|$ and r_t^2 outside the 95% confidence interval but they are also positive over long lags in every case. Also, the sample autocorrelations for absolute returns are greater than those for the squared errors at every lag up to 100 and for every country. It is clear that stock market returns for emerging markets are not generated by i.i.d. processes, and this finding also applies to the United States. A long- memory component is present, and the fact that absolute returns are more positive over long lags is indicative of this.

We now examine the sample autocorrelations for the transformed absolute stock market returns $|r_t|^d$ and various positive values of d. Tables 3.30 to 3.38 give $\rho(|r_t|^d, |r_{t+\tau}|^d)$ for various d (d = 0.125,0.25,0.5,0.75,1,1.25,1.5,1.75,2 and 3) at lags 1 to 5, 10, 20, 40, 70 and 100. From these tables and also from Graphs 3.22 to 3.29 we can see that the conclusions found above remain valid. All power transformations of absolute returns have significant positive autocorrelations up to lag 100, which supports the claim that stock market returns have long-memory. In particular, Taiwan, Brazil and Thailand show slow decays in their respective autocorrelation functions. All power transformations for all countries show the same pattern. The power transformations of absolute returns have significant positive autocorrelations at least up to lag 100, supporting the claim that stock market returns have long-term memory. The autocorrelations decrease quickly after the first lag, then decrease very slowly while remaining positive for long lags. We can verify from table 3.30 that the same pattern is found for the US.

We can see that the ACF's for $|r_t|^d$ exceed those for the US for various values of d and various lags, reinforcing our perception that emerging markets do show higher persistence than developed markets[6]. However, this does not appear to be true for Korea and Hong-Kong, two markets with lower ACF values for the different lags and various values for d. This is consistent with our earlier finding that these markets have a shorter memory than the US market.

Perhaps, the most striking finding from the autocorrelations is that $|r_t|^d$ has the largest autocorrelation up to lag 100 when d is close to 1. Thailand, Brazil, Argentina, Mexico, Hong-Kong and Korea have d = 0.75. The exceptions are Taiwan and Malaysia, where d = 1.75. To highlight this, we calculated the sample autocorrelations $\rho_\tau(d)$ as a function of d, d > 0 for τ = 1, 2, 5, and 10 and setting d = 0.125, 0.0, 130,...., 1.745, 1.750,........., 2.0, 2.05, ..., 4.95, 5. Graphs 3.30 to 3.37 show the plots of calculated $\rho_\tau(d)$ for $|r_t|^d$. Here we see that the autocorrelation $\rho_\tau(d)$ is a smooth function of d except in the case of Mexico. There is a saddle point $\underline{d}$ between 2 and 3 (roughly) such that when d > d, $\rho_\tau(d)$ is a convex function of d. Again, Mexico is an exception. There is a unique point d* near 1 where $\rho_\tau(d)$ reaches its maximum point, i.e. $\rho_\tau(d^*)$ > $\rho_\tau(\underline{d})$ for every d other than d*. Ding et al. (1993) report the same findings for the US.

[6] This is only true for Brazil for d lower than 1,25 and for Argentina except for d = 3.

TABLE 3.30 - AUTOCORRELATION OF $|r_t|^d$ THAILAND

	NUMBER OF LAGS										
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.3772	0.3323	0.3153	0.2891	0.2706	0.2617	0.2035	0.2022	0.2011	0.1678	0.1702
d=0.250	0.4044	0.3630	0.3291	0.3037	0.2912	0.2642	0.2068	0.1979	0.1950	0.1606	0.1757
d=0.500	0.4343	0.3924	0.3572	0.3235	0.3112	0.2705	0.2167	0.2020	0.1917	0.1534	0.1364
d=0.750	0.4387	0.3939	0.3627	0.3196	0.3037	0.2605	0.2065	0.1919	0.1795	0.1371	0.1135
d=1.000	0.4261	0.3789	0.3502	0.2999	0.2782	0.2395	0.1840	0.1704	0.1613	0.1156	0.0910
d=1.250	0.4041	0.3562	0.3282	0.2738	0.2437	0.2147	0.1573	0.1442	0.1424	0.0941	0.0725
d=1.500	0.3779	0.3316	0.3030	0.2475	0.2068	0.1906	0.1310	0.1182	0.1259	0.0751	0.0589
d=1.750	0.3509	0.3080	0.2783	0.2241	0.1717	0.1692	0.1074	0.0949	0.1125	0.0595	0.0494
d=2.000	0.3248	0.2864	0.2556	0.2045	0.1403	0.1511	0.0872	0.0753	0.1022	0.0472	0.0425
d=3.000	0.2411	0.2189	0.1880	0.1574	0.0575	0.1038	0.0368	0.0287	0.0794	0.0201	0.0266

TABLE 3.31 - AUTOCORRELATION OF $|r_t|^d$ ARGENTINA

	NUMBER OF LAGS										
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.2003	0.2442	0.1941	0.2094	0.2243	0.2307	0.1768	0.1285	0.1369	0.1483	0.0752
d=0.250	0.2684	0.3023	0.2595	0.2575	0.2849	0.2822	0.2167	0.1696	0.1674	0.1738	0.0891
d=0.500	0.3349	0.3495	0.3124	0.3006	0.3288	0.3104	0.2384	0.2093	0.1844	0.1722	0.0936
d=0.750	0.3635	0.3665	0.3200	0.3105	0.3318	0.2963	0.2361	0.2292	0.1772	0.1454	0.0884
d=1.000	0.3603	0.3653	0.2943	0.2915	0.3050	0.2534	0.2219	0.2324	0.1524	0.1062	0.0752
d=1.250	0.3296	0.3526	0.2464	0.2496	0.2577	0.1951	0.2035	0.2188	0.1177	0.0663	0.0569
d=1.500	0.2822	0.3348	0.1909	0.1966	0.2026	0.1369	0.1858	0.1915	0.0825	0.0352	0.0384
d=1.750	0.2310	0.3159	0.1399	0.1449	0.1508	0.0896	0.1700	0.1572	0.0535	0.0157	0.0234
d=2.000	0.1848	0.2974	0.0991	0.1019	0.1083	0.0560	0.1522	0.1229	0.0327	0.0055	0.0133
d=3.000	0.0742	0.2233	0.0223	0.0200	0.0246	0.0068	0.0966	0.0347	0.0025	-0.011	0.0006

TABLE 3.32 - AUTOCORRELATION OF $|r_t|^d$ BRAZIL

	NUMBER OF LAGS										
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.2531	0.2372	0.2415	0.2225	0.2255	0.2242	0.1721	0.1487	0.1394	0.1481	0.1606
d=0.250	0.2982	0.2687	0.2739	0.2625	0.2560	0.2507	0.2166	0.1974	0.1815	0.1885	0.1884
d=0.500	0.3351	0.3048	0.3068	0.2999	0.2931	0.2784	0.2538	0.2346	0.2118	0.2152	0.2110
d=0.750	0.3415	0.3186	0.3123	0.3103	0.3065	0.2835	0.2609	0.2382	0.2126	0.2170	0.2124
d=1.000	0.3295	0.3163	0.2967	0.3018	0.3014	0.2714	0.2475	0.2208	0.1948	0.2025	0.1970
d=1.250	0.3005	0.3004	0.2636	0.2779	0.2807	0.2464	0.2119	0.1883	0.1641	0.1761	0.1686
d=1.500	0.2611	0.2723	0.2179	0.2415	0.2477	0.2384	0.1811	0.1472	0.1264	0.1428	0.1324
d=1.750	0.2152	0.2345	0.1669	0.1976	0.2069	0.2035	0.1407	0.1051	0.0886	0.1085	0.0952
d=2.000	0.1693	0.1920	0.1189	0.1527	0.1645	0.1686	0.1039	0.0686	0.0565	0.0781	0.0630
d=3.000	0.0528	0.0611	0.019	0.0386	0.0288	0.0721	0.0271	0.0062	0.0040	0.0180	0.0073

TABLE 3.33 - AUTOCORRELATION OF $|r_t|^d$ HONG-KONG

	NUMBER OF LAGS										
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.1386	0.1504	0.1511	0.1529	0.1294	0.1153	0.0940	0.0741	0.0918	0.0525	0.0272
d=0.250	0.1673	0.1791	0.1762	0.1754	0.1548	0.1369	0.1075	0.0848	0.1030	0.0604	0.0320
d=0.500	0.2142	0.2197	0.2096	0.2033	0.1869	0.1639	0.1235	0.0976	0.1140	0.0672	0.0343
d=0.750	0.2496	0.2408	0.2254	0.2128	0.2001	0.1765	0.1271	0.1012	0.1137	0.0648	0.0295
d=1.000	0.2682	0.2350	0.2199	0.2005	0.1899	0.1736	0.1163	0.0951	0.1035	0.0538	0.0146
d=1.250	0.2626	0.1972	0.1901	0.1650	0.1549	0.1532	0.0921	0.0797	0.0852	0.0370	0.0104
d=1.500	0.2314	0.1382	0.1433	0.1160	0.1055	0.1192	0.0616	0.0586	0.0626	0.0204	0.0030
d=1.750	0.1860	0.0816	0.0960	0.0708	0.0607	0.0832	0.0353	0.0386	0.0416	0.0089	-0.0006
d=2.000	0.1415	0.0425	0.0602	0.0395	0.0308	0.0543	0.0181	0.0236	0.0258	0.0029	-0.0016
d=3.000	0.0431	0.0017	0.0093	0.0033	0.0010	0.0093	0.0007	0.0028	0.0031	-0.0005	-0.0007

TABLE 3.34 - AUTOCORRELATION OF $|r_t|^d$ KOREA

	NUMBER OF LAGS										
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.1684	0.1698	0.1896	0.1623	0.1335	0.1302	0.0931	0.0885	0.0755	0.0819	0.0678
d=0.250	0.2049	0.2107	0.2248	0.1981	0.1672	0.1468	0.1097	0.0990	0.0954	0.0923	0.0829
d=0.500	0.2387	0.2482	0.2530	0.2269	0.1967	0.1556	0.1126	0.1011	0.1104	0.0978	0.0917
d=0.750	0.2544	0.2624	0.2602	0.2326	0.2057	0.1517	0.1009	0.0914	0.1123	0.0939	0.0883
d=1.000	0.2523	0.2571	0.2503	0.2197	0.1977	0.1378	0.0810	0.0742	0.1045	0.0827	0.0773
d=1.250	0.2323	0.2318	0.2234	0.1899	0.1738	0.1153	0.0573	0.0531	0.0887	0.0661	0.0614
d=1.500	0.1925	0.1886	0.1433	0.1469	0.1369	0.0866	0.0342	0.0323	0.0617	0.0469	0.0432
d=1.750	0.1411	0.1357	0.1813	0.0996	0.0948	0.0571	0.0161	0.0156	0.0442	0.0290	0.0263
d=2.000	0.0916	0.0865	0.1313	0.0591	0.0577	0.0330	0.0050	0.0052	0.0253	0.0155	0.0136
d=3.000	0.0085	0.0073	0.0850	0.0029	0.0036	0.0013	-0.0016	-0.0014	0.0005	-0.0001	-0.0005

TABLE 3.35 - AUTOCORRELATION OF $|r_t|^d$ MALAYSIA

NUMBER OF LAGS											
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.1990	0.1577	0.1519	0.1162	0.1386	0.1157	0.0715	0.0805	0.0526	0.0538	0.0507
d=0.250	0.2287	0.1900	0.1743	0.1327	0.1597	0.1341	0.0800	0.0795	0.0666	0.0597	0.0580
d=0.500	0.2751	0.2396	0.2126	0.1633	0.1916	0.1594	0.0908	0.0847	0.0880	0.0617	0.0625
d=0.750	0.3134	0.2785	0.2454	0.1883	0.2101	0.1751	0.0935	0.0873	0.1013	0.0567	0.0586
d=1.000	0.3438	0.3057	0.2709	0.2048	0.2139	0.1826	0.0882	0.0864	0.1050	0.0469	0.0496
d=1.250	0.3655	0.3196	0.2867	0.2108	0.2037	0.1839	0.0769	0.0821	0.0993	0.0347	0.0384
d=1.500	0.3771	0.3197	0.2911	0.2057	0.1824	0.1817	0.0622	0.0747	0.0864	0.0227	0.0274
d=1.750	0.3782	0.3072	0.2842	0.1909	0.1544	0.1782	0.0472	0.0650	0.0702	0.0128	0.0183
d=2.000	0.3699	0.2856	0.2683	0.1696	0.1246	0.1749	0.0340	0.0543	0.0541	0.0057	0.0116
d=3.000	0.2945	0.1816	0.1740	0.0817	0.0394	0.1643	0.0062	0.0199	0.0150	-0.0023	0.0013

TABLE 3.36 -AUTOCORRELATION OF $|r_t|^d$ TAIWAN

NUMBER OF LAGS											
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.1679	0.2223	0.2154	0.2132	0.1902	0.2163	0.1734	0.1686	0.1687	0.1412	0.1355
d=0.250	0.2035	0.2567	0.2523	0.2486	0.2251	0.2486	0.2021	0.1964	0.1967	0.1690	0.1573
d=0.500	0.2609	0.3095	0.3090	0.3027	0.2801	0.2952	0.2459	0.2374	0.2372	0.2112	0.1880
d=0.750	0.3036	0.3483	0.3488	0.3396	0.3196	0.3254	0.2775	0.2664	0.2814	0.2397	0.2065
d=1.000	0.3325	0.3746	0.3743	0.3626	0.3459	0.3427	0.2989	0.2858	0.2814	0.2568	0.2151
d=1.250	0.3496	0.3907	0.3888	0.3748	0.3617	0.3505	0.3125	0.2975	0.2909	0.2657	0.2168
d=1.500	0.3576	0.3992	0.3951	0.3793	0.3695	0.3516	0.3201	0.3031	0.2951	0.2689	0.2140
d=1.750	0.3590	0.4021	0.3956	0.3784	0.3714	0.3484	0.3234	0.3041	0.2959	0.2685	0.2087
d=2.000	0.3558	0.4010	0.3923	0.3737	0.3691	0.3423	0.3236	0.3018	0.2945	0.2658	0.2023
d=3.000	0.3229	0.3770	0.3615	0.3387	0.3399	0.3074	0.3090	0.2747	0.2799	0.2445	0.1763

TABLE 3.37 - AUTOCORRELATION OF $|r_t|^d$ MEXICO

NUMBER OF LAGS											
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.2762	0.2178	0.1605	0.1653	0.1649	0.1636	0.1604	0.1471	0.1686	0.1439	0.1391
d=0.250	0.2745	0.1892	0.1340	0.1416	0.1338	0.1236	0.1184	0.1050	0.1295	0.1068	0.0867
d=0.500	0.3206	0.1995	0.1534	0.1641	0.1475	0.1274	0.1117	0.0955	0.1219	0.0916	0.0521
d=0.750	0.3685	0.2290	0.1894	0.1994	0.1781	0.1605	0.1308	0.1081	0.1341	0.0925	0.0450
d=1.000	0.3978	0.2495	0.2149	0.2205	0.1963	0.2108	0.1472	0.1155	0.1390	0.0884	0.0399
d=1.250	0.4054	0.2545	0.2247	0.2231	0.1967	0.2069	0.1547	0.1137	0.1324	0.0778	0.0329
d=1.500	0.3931	0.2463	0.2198	0.2097	0.1818	0.2083	0.1526	0.1045	0.1169	0.0639.	0.0250.
d=1.750	0.3669	0.2304	0.2047	0.1865	0.1578	0.1980	0.1430	0.0915	0.0972	0.0501.	0.0178.
d=2.000	0.3336	0.2124	0.1845	0.1597	0.1313	0.1810	0.1291	0.0778	0.0775	0.0384	0.0120
d=3.000	0.2062	0.1543	0.1054	0.0729	0.0525	0.1104	0.0718	0.0370	0.0256	0.0126	0.0016

TABLE 3.38 - AUTOCORRELATION OF $|r_t|^d$ US

	NUMBER OF LAGS										
	1	2	3	4	5	10	20	30	40	70	100
d=0.125	0.110	0.108	0.102	0.098	0.121	0.100	0.100	0.100	0.095	0065.	0.089
d=0.250	0.186	0.181	0.182	0.176	0.193	0.164	0.164	0.164	0.148	0.120	0.131
d=0.500	0.257	0.255	0.263	0.251	0.259	0.222	0.221	0.221	0.192	0.166	0.165
d=0.750	0.297	0.299	0.305	0.286	0.291	0.246	0.241	0.241	0.207	0.180	0.173
d=1.000	0.318	0.323	0.322	0.286	0.303	0.247	0.237	0.237	0.200	0.174	0.162
d=1.250	0.319	0.326	0.312	0.280	0.295	0.227	0.211	0.211	0.174	0.153	0.138
d=1.500	0.300	0.309	0.278	0.242	0.270	0.192	0.170	0.170	0.136	0.122	0.106
d=1.750	0.264	0.276	0.228	0.192	0.234	0.149	0.125	0.125	0.095	0.088	0.073
d=2.000	0.218	0.234	0.173	0.140	0.193	0.107	0.083	0.083	0.059	0.058	0.045
d=3.000	0.066	0.088	0.036	0.025	0.072	0.019	0.009	0.009	0.004	0.006	0.003

In fact, $|r_t|^d$ has a positive autocorrelation over a much longer lags than 100. Table 3.39 shows the lags (τ^*) at which the first negative autocorrelation of $|r_t|^d$ occurs for various d.

TABLE 3.39 -LAGS AT WHICH FIRST NEGATIVE AUTOCORRELATION OF $|r^d|$ OCCURS

d	0.125	0.250	0.500	0.750	1.000	1.250	1.500	1.750	2.000	3.000
ARG	408	446	450	450	450	450	177	82	78	17
BRA	1799	1621	1621	1621	1595	1588	1574	1415	1378	145
MEX	780	375	156	156	210	119	102	102	102	66
HON	133	133	124	120	120	98	94	63	60	12
THA	547	862	1087	1078	1078	497	485	167	155	73
TAI	956	956	1005	1006	1007	996	990	990	973	682
USA	2028	2534	2704	2705	2705	2705	2705	1685	2598	520
KOR	370	370	408	408	408	325	162	161	123	13
MAL	163	167	167	167	167	129	129	80	76	59

It can be seen from Table 3.39 that $|r_t|^d$ has positive autocorrelations ranging from more than 1799 lags (Brazil) to 12 (Hong-Kong). This seems to show that some countries have a very strong long-memory effect, up to 1799 days or almost 7 years. For d = 1 lags vary from 120 (Hong-Kong) to 1595 (Brazil). To summarize, Brazil, Taiwan, Thailand and Mexico do display long-memory effects, and these effects are strong even with long lags. It is worthwhile to note that the first lag for which a negative autocorrelation for $|r_t|^d$ occurs for the United States is always bigger than no matter which emerging market we consider. Again, this provides support for the view that the US market exhibits a stronger mean reverting behaviour than the less developed markets.

We typically observe slow decays in autocorrelation functions in the presence of long-memory, which we see for Brazil, Taiwan and Thailand in Graphs 3.13a and 3.13b. Many different models have been employed to try to explain this pattern in the sample autocorrelation curve. Among them we can identify:

(1) ρ_t an exponentially decreasing function of τ $(\rho_t = \alpha\beta^{\tau})$ like an ARMA model .

(2) ρ_t as the same in the autocorrelation function of a fractionally integrated process (Granger and Joyeux(1980) , that is $\rho_t = \rho_{\tau-1}\dfrac{(\tau+\beta-1)}{(\tau-\beta)}$.

(3) ρ_t as a polynomially decreasing function of τ $(\rho_\tau = \alpha / \tau^{\beta})$, which is approximately the same as (2) when τ is large.

In the literature it has been found that when we use real data, the fitted autocorrelations using (1) decrease too slowly at the beginning and too fast at the end, whereas the opposite occurs if (2) and (3) are used. A preferred model would have to combine the effects of (1), (2) and (3). The theoretical autocorrelation function proposed by Granger (1993) is then:

$$\rho_\tau = \frac{\alpha\rho_{\tau-1}^{\beta_1}\beta_2^{\tau}}{\tau^{\beta_3}} \qquad (3.6)$$

which we can change into a linear model with just a few transformations:

$$\log\rho_\tau = \log\alpha + \beta_1\log\rho_{\tau-1} + \tau\log\beta_2 - \beta_3\log\tau \qquad (3.7)$$

We can let $\alpha^* = \log\alpha$, $\beta_1^* = \beta_1$, $\beta_2^* = \log\beta_2$, and $\beta_3^* = -\beta_3$, then

$$\log\rho_\tau = \alpha^* + \beta_1^*\log\rho_{\tau-1} + \beta_2^*\tau + \beta_3^*\log\tau \qquad (3.8)$$

We have estimated the relationship (3.8) for some countries that had been identified before as having a long-memory according to the autocorrelation function, i.e. Brazil, Taiwan and Thailand. Ordinary Least Squares estimates are shown below.

TABLE 3.40 OLS ESTIMATES OF EQUATION 3.8

						D.W.
BRAZIL	-0.231841	0.745069	-0.000135	0.039353	0.92	2.632
	(-10.96948)	(43.466296)	(-12.104542)	(5.354717)		
TAIWAN	-0.139905	0.616428	-0.000241	-0.047036	0.871	2.567
	(-4.680133)	(24.304403)	(-9.1825)	(-3.159659)		
USA	-0.049	0.784	-0.000195	-0.057	0.92	2.65
	(-3.9)	(62.9)	(-5.9)	(-9.1)		
THAILAND	-0.0819	0.759177	-0.000014	-0.081688	0.743	2.543
	(-1.993683)	(37.631087)	(-0.5019)	(-3.713581)		

The t-statistics are in parentheses. Nearly all coefficients are significant, with the single exception being β_2^* for Thailand. If we re-transform the equation we find that, for Brazil,

$$\rho_\tau = 0.586\rho_{\tau-1}^{0.745}\ (0.999689)^\tau / \tau^{-0.0339}.$$

Graphs 3.38, 3.39 and 3.40 plot the fitted autocorrelations and the sample autocorrelations using (3.8) above. We see that the theoretical model fits the actual sample autocorrelations quite well.

Other studies have been conducted for the New York Stock Exchange daily price index, producing results similar to those in table 3.40. Granger(1993) also finds support for the existence of long memory in most financial time series.

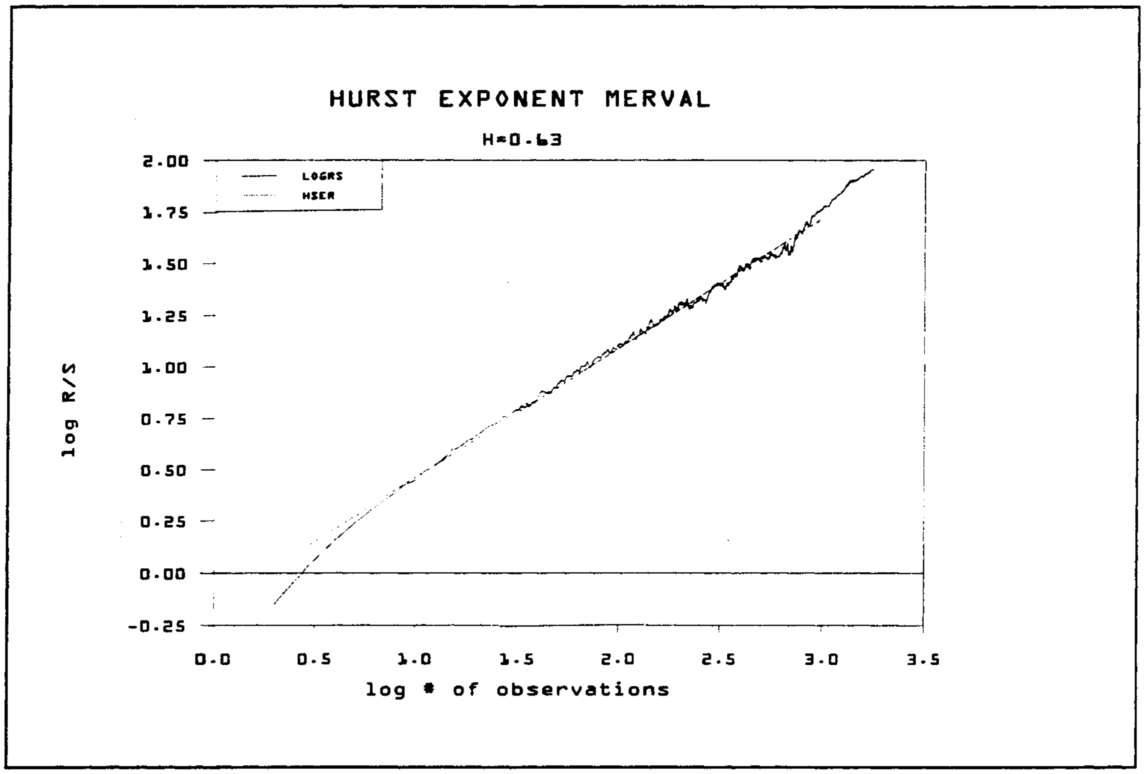
3.6 - Conclusion

In this chapter we investigated the long-memory properties of stock market returns series using data for emerging markets. We found that there are substantially higher correlations among absolute returns series than among non-transformed series. This is also true for power transformation of the absolute returns $|r_t|^d$, which have quite high autocorrelation even for long lags. This result has been established for the US markets using S&P data, but we have managed to show that the same phenomenon is present in emerging markets. It is worthwhile mentioning that the overall picture shows emerging markets to be more persistent than the US market, regardless of the measure used to capture this effect. We have used the Hurst exponent, the classical R/S statistic, the modified R/S and finally the autocorrelation function to arrive at the same conclusion, i.e. that 'long-memory' is more pronounced in emerging markets than in developed ones (using the U.S. as a benchmark).

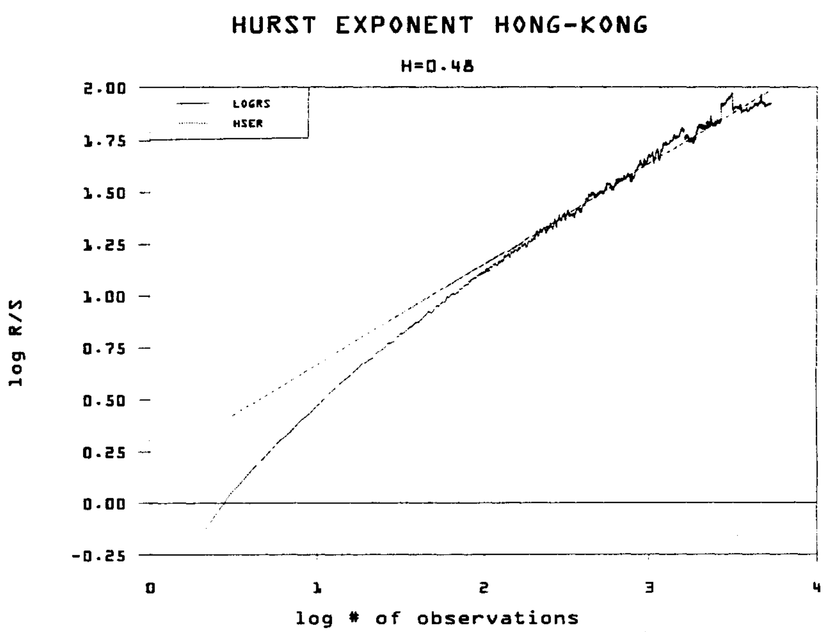
In coming to this conclusion, we must highlight the fact that the degree of persistence appears to vary across countries. On one hand, we have the Korean and Hong-Kong markets displaying even shorter memory than the US market. On the other hand, we see other emerging markets with longer memory than the U.S. market. Longer memory is summarised by higher values of the ACF for different lags and different power transformations. In particular, we observe Brazil and Taiwan as extreme cases, where we consistently have values for the ACF much higher than those for the US. We also found mean reverting behaviour to be stronger in emerging markets, especially in Malaysia.

These results accord well with the idea that investors may react more quickly and use information more efficiently in some markets than in others. In particular, results may differ depending on whether interventions are frequent (e.g. Brazil) or infrequent (e.g. the United States).

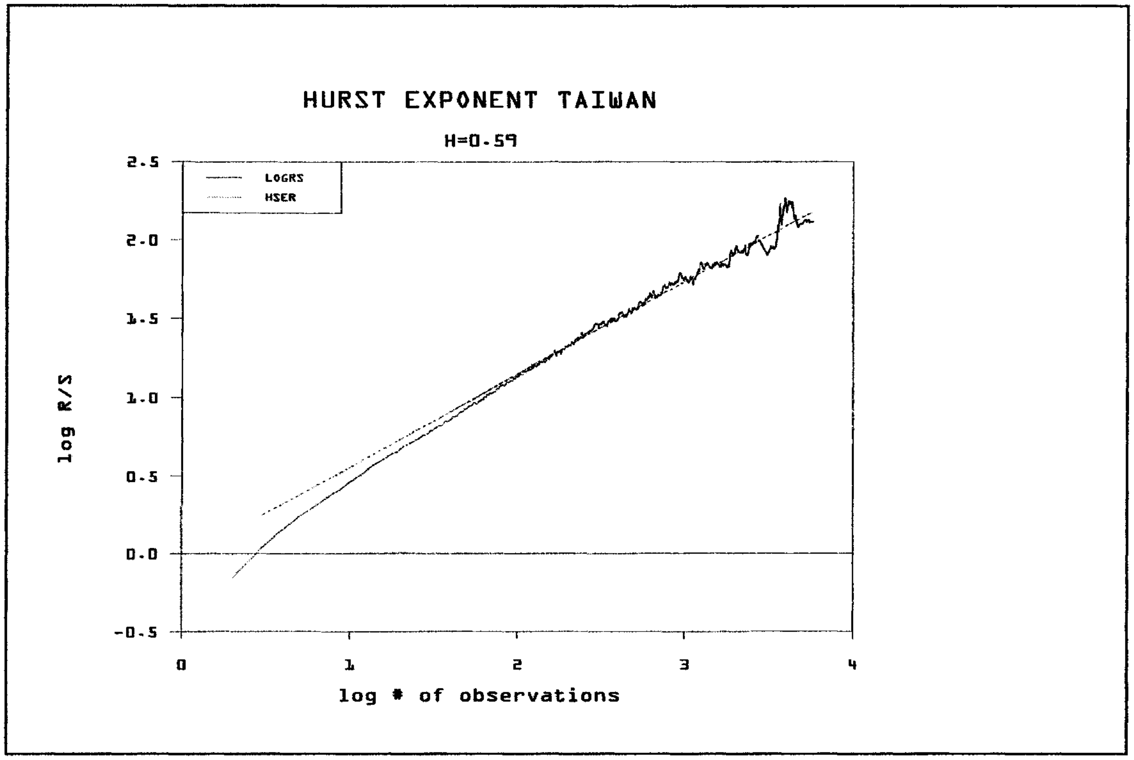
Graph 3.1



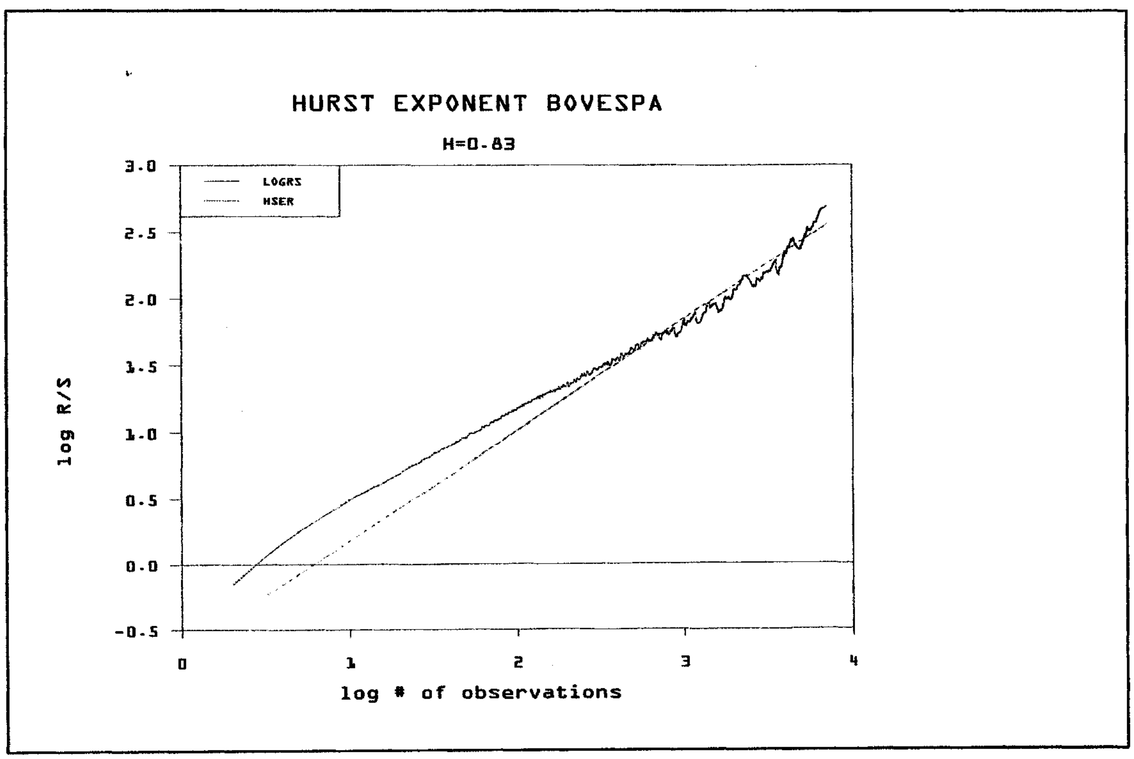
Graph 3.2



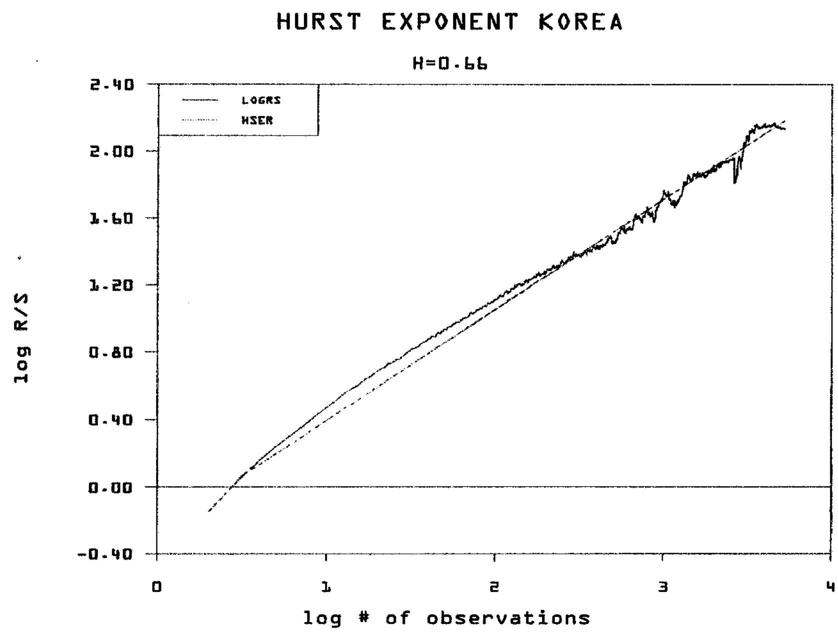
Graph 3.3



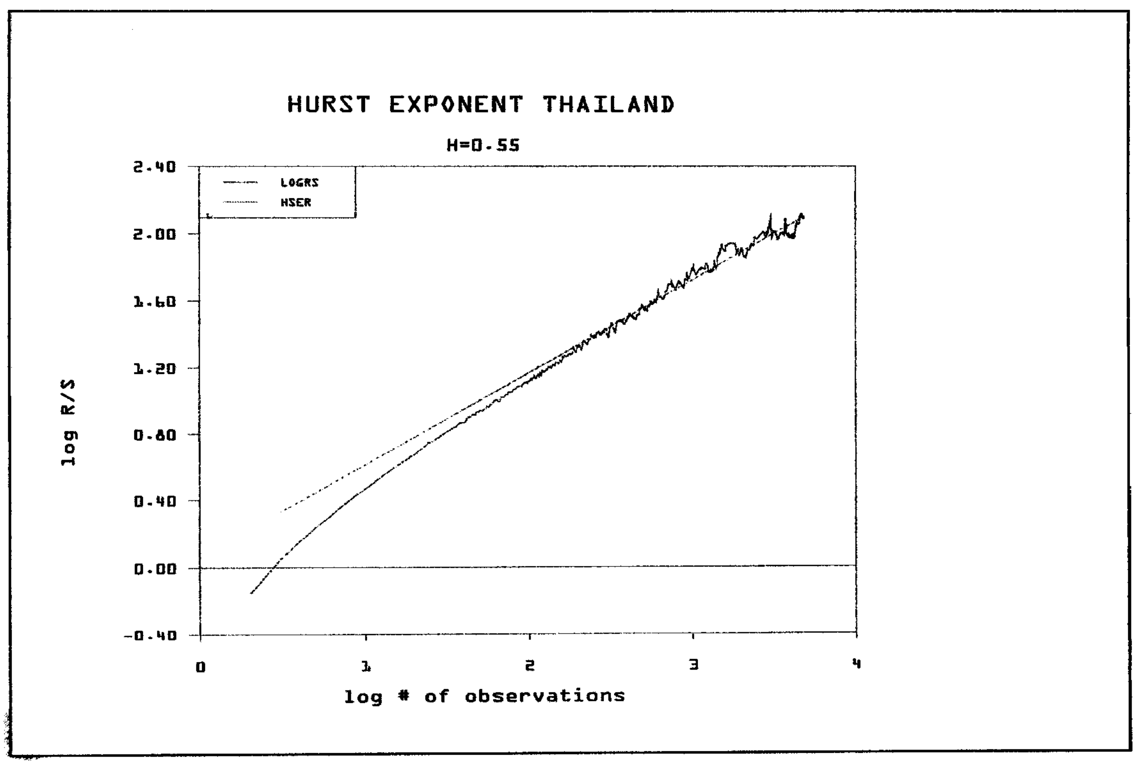
Graph 3.4



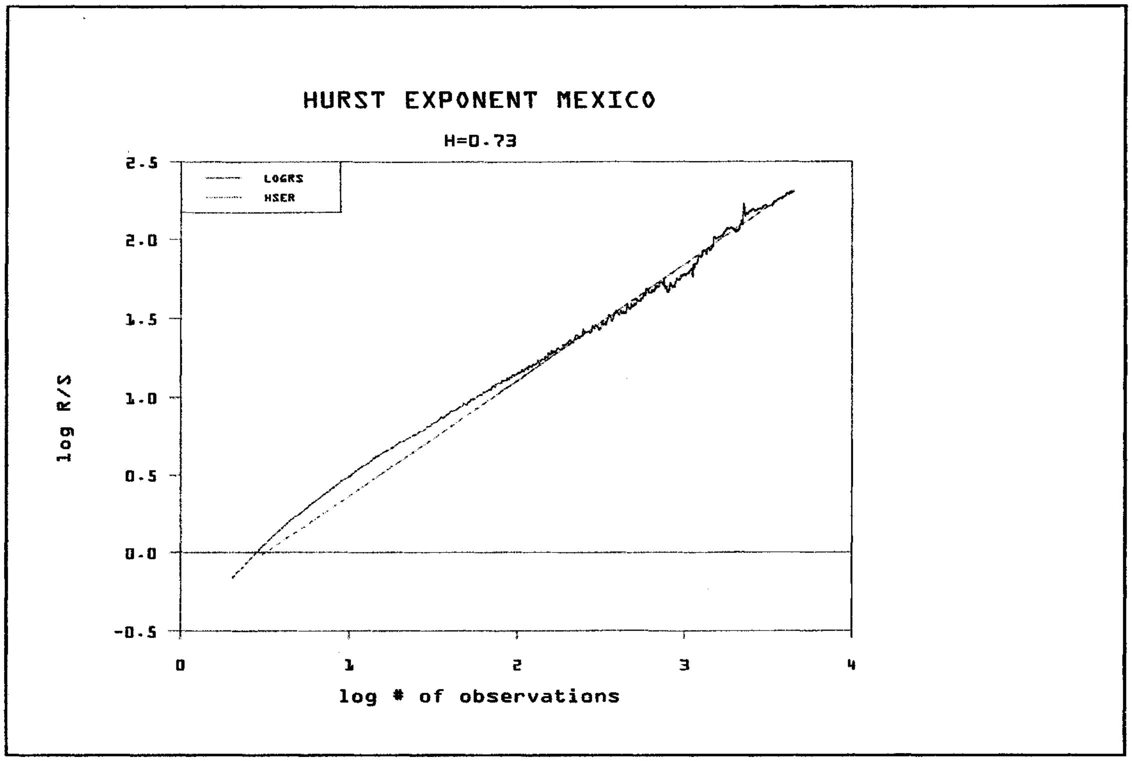
Graph 3.5



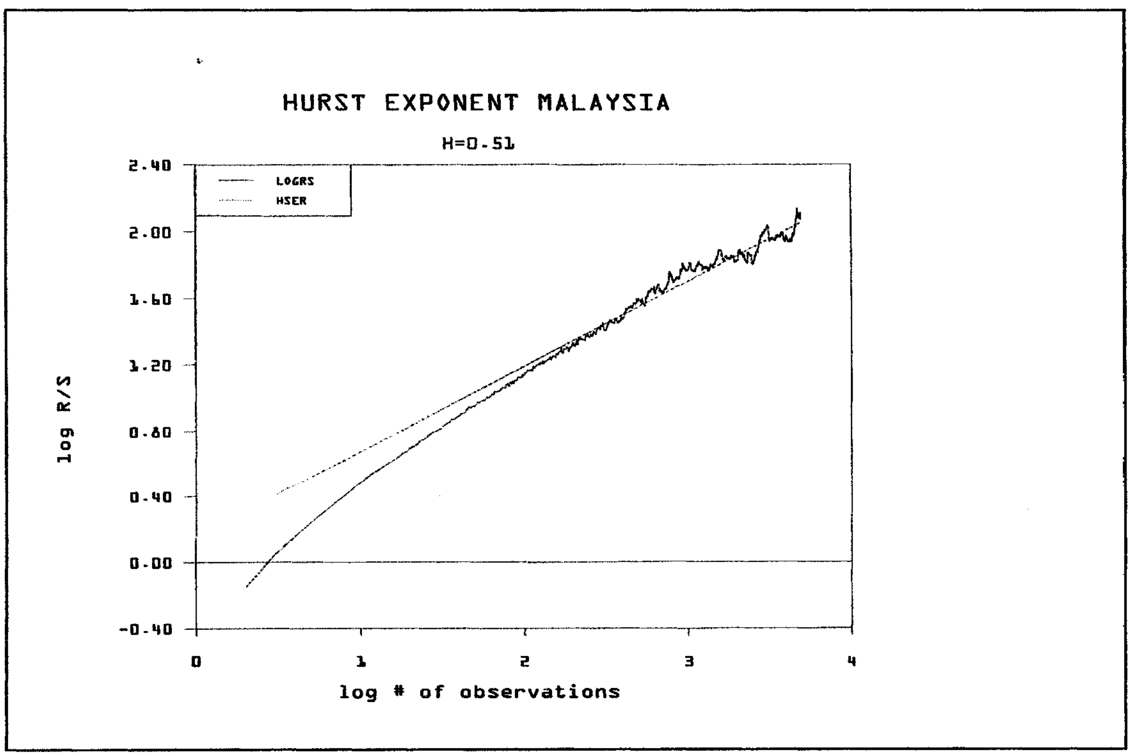
Graph 3.6



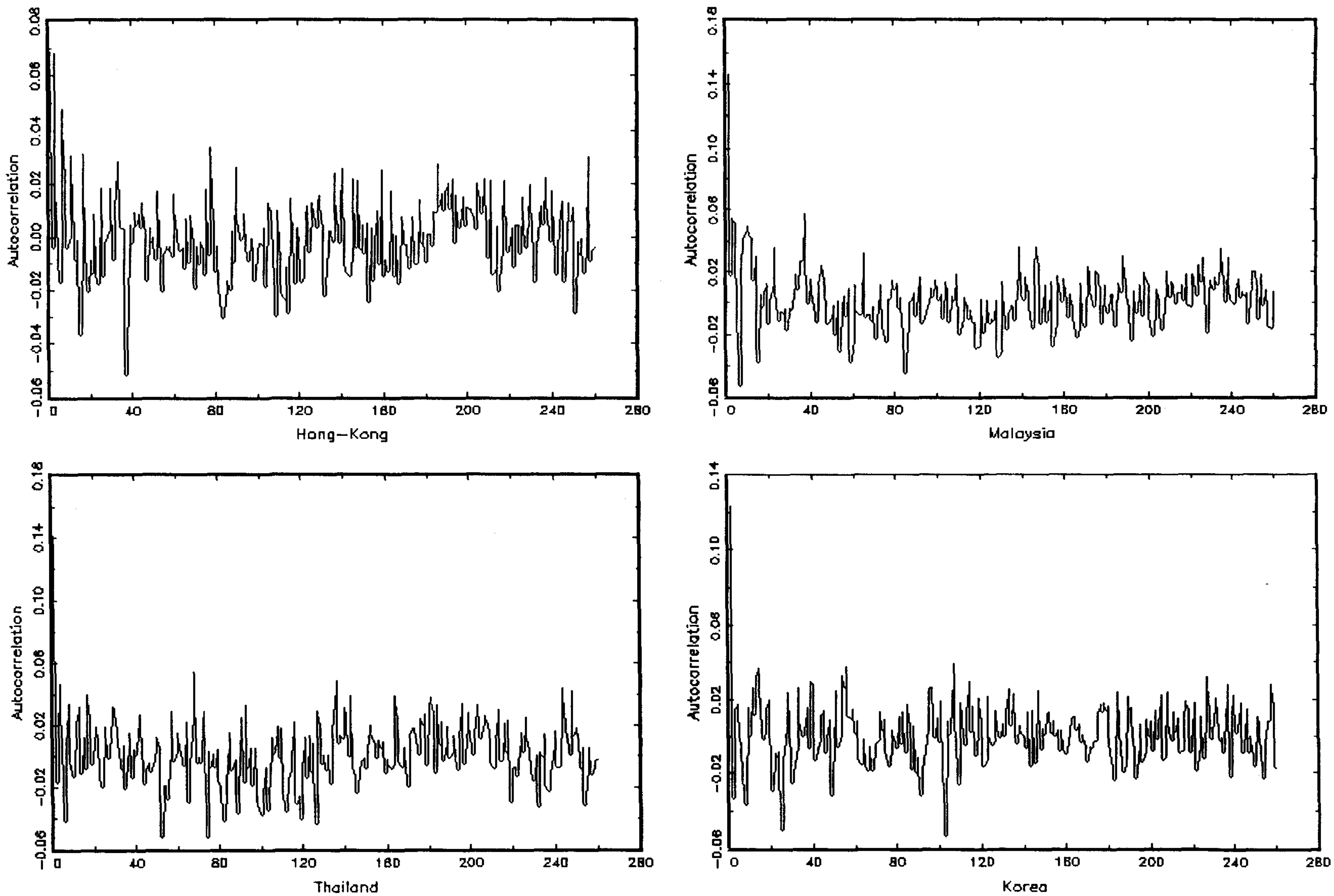
Graph 3.7



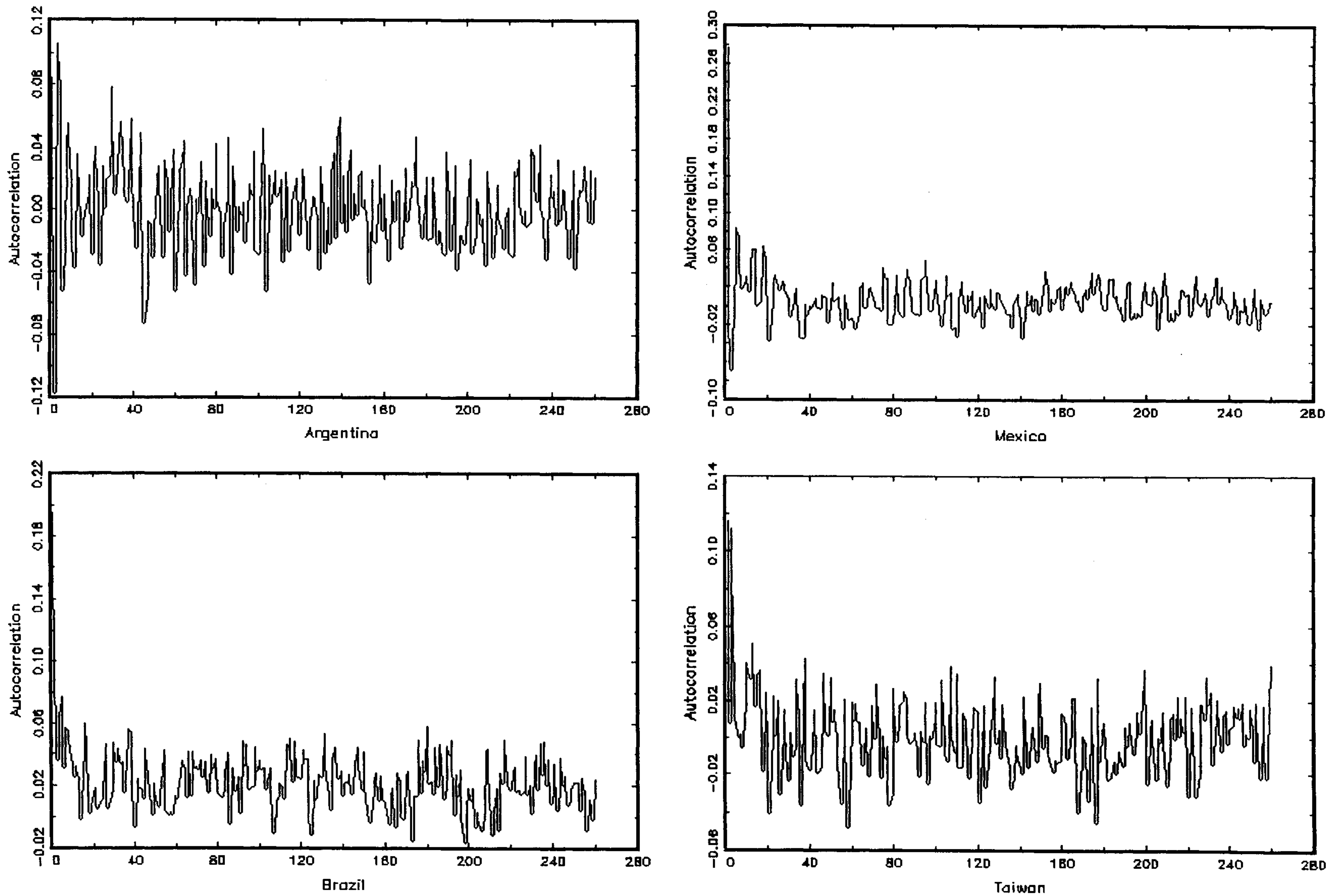
Graph 3.8



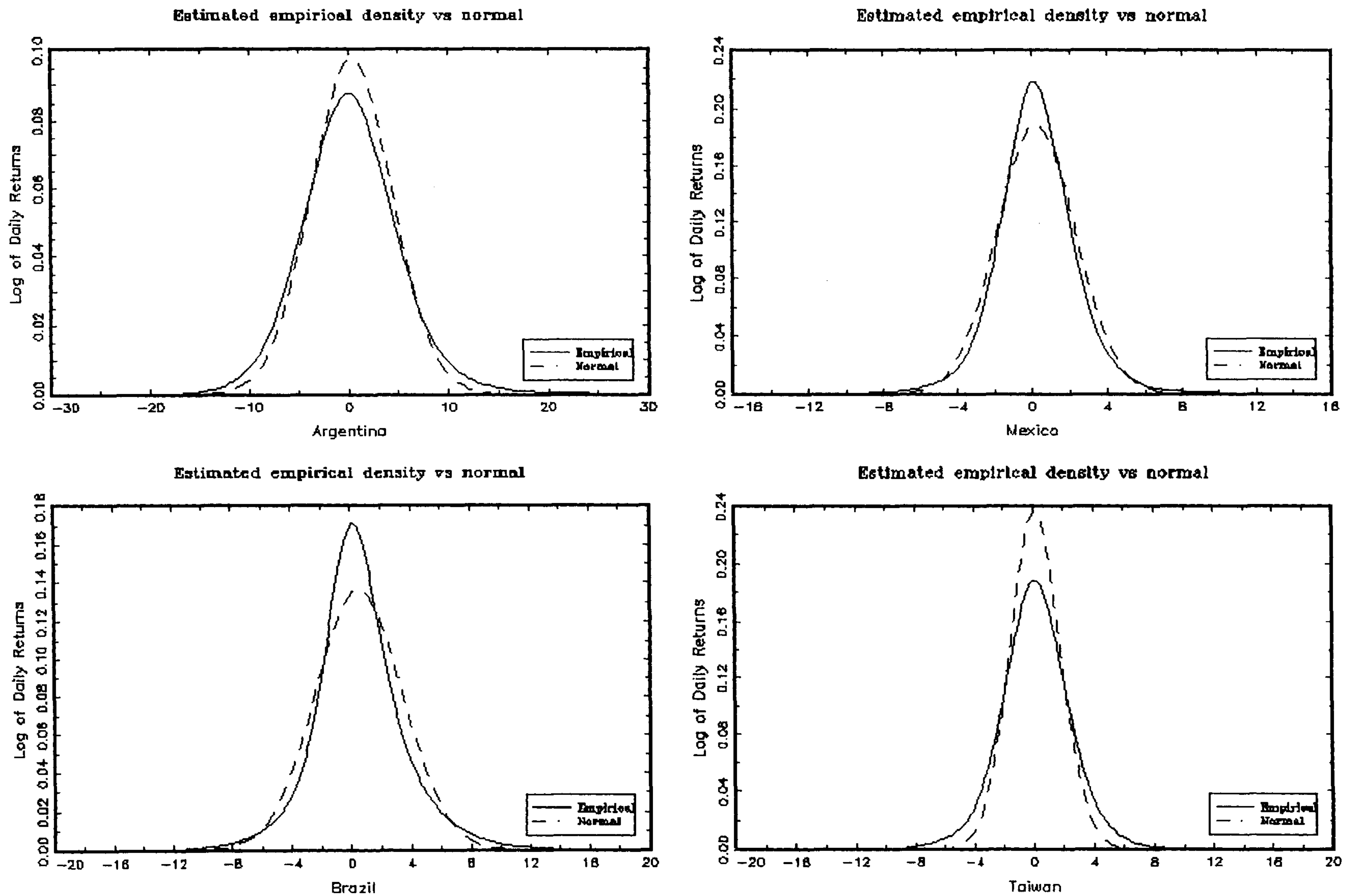
Graph 3.9a – Autocorrelation Function



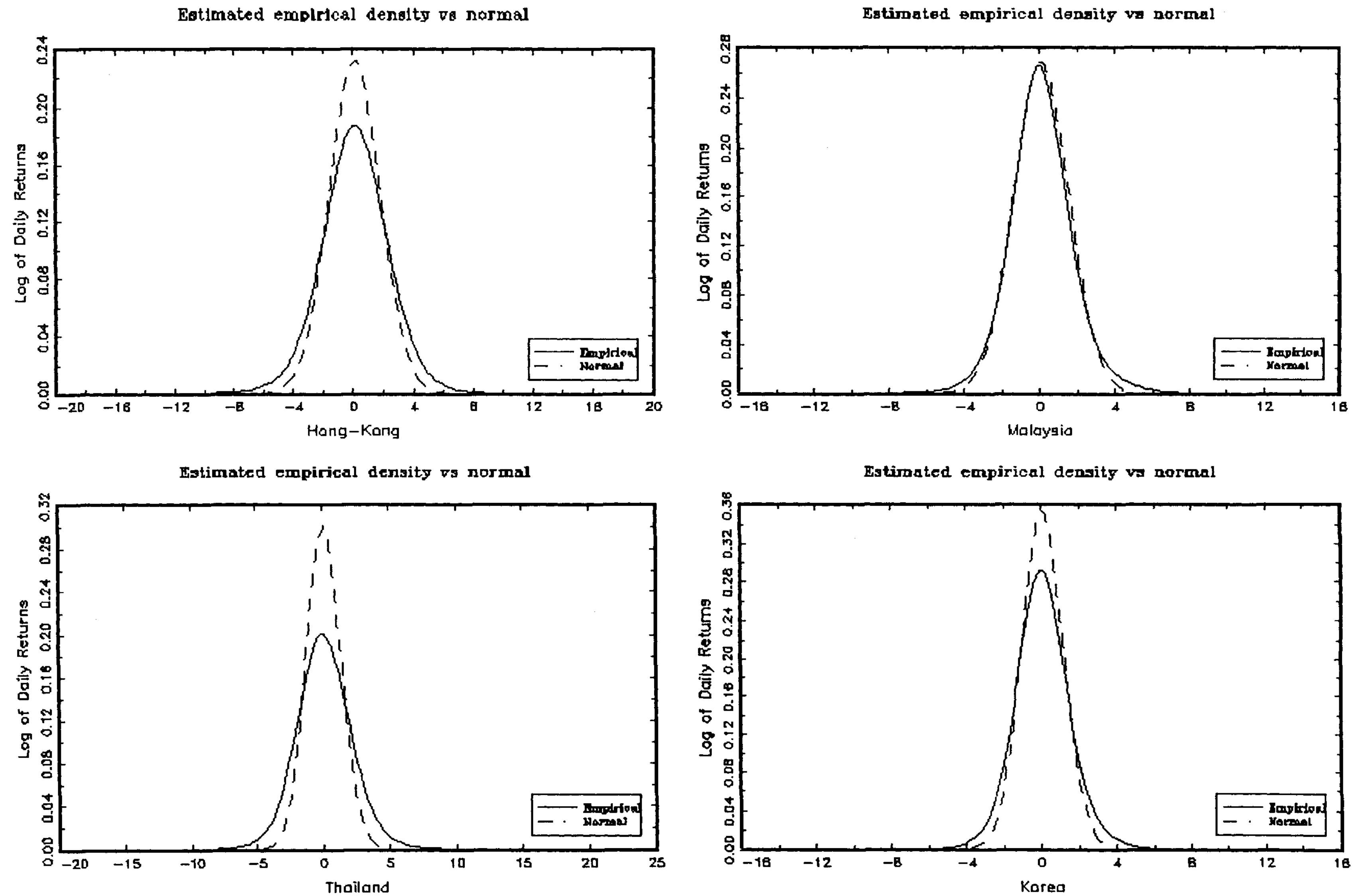
Graph 3.9b — Autocorrelation Function



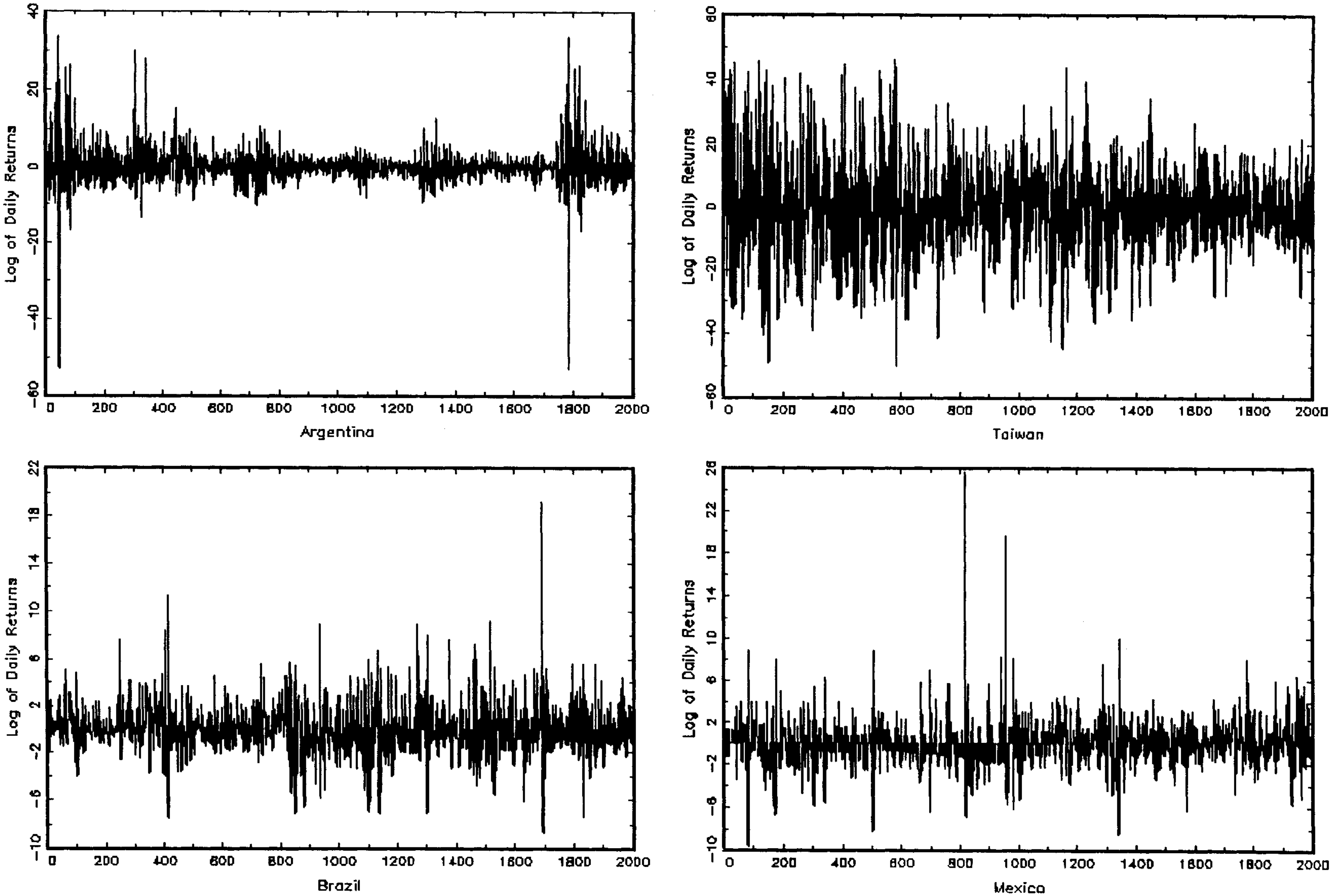
Graph 3.10a – Empirical Density x Normal



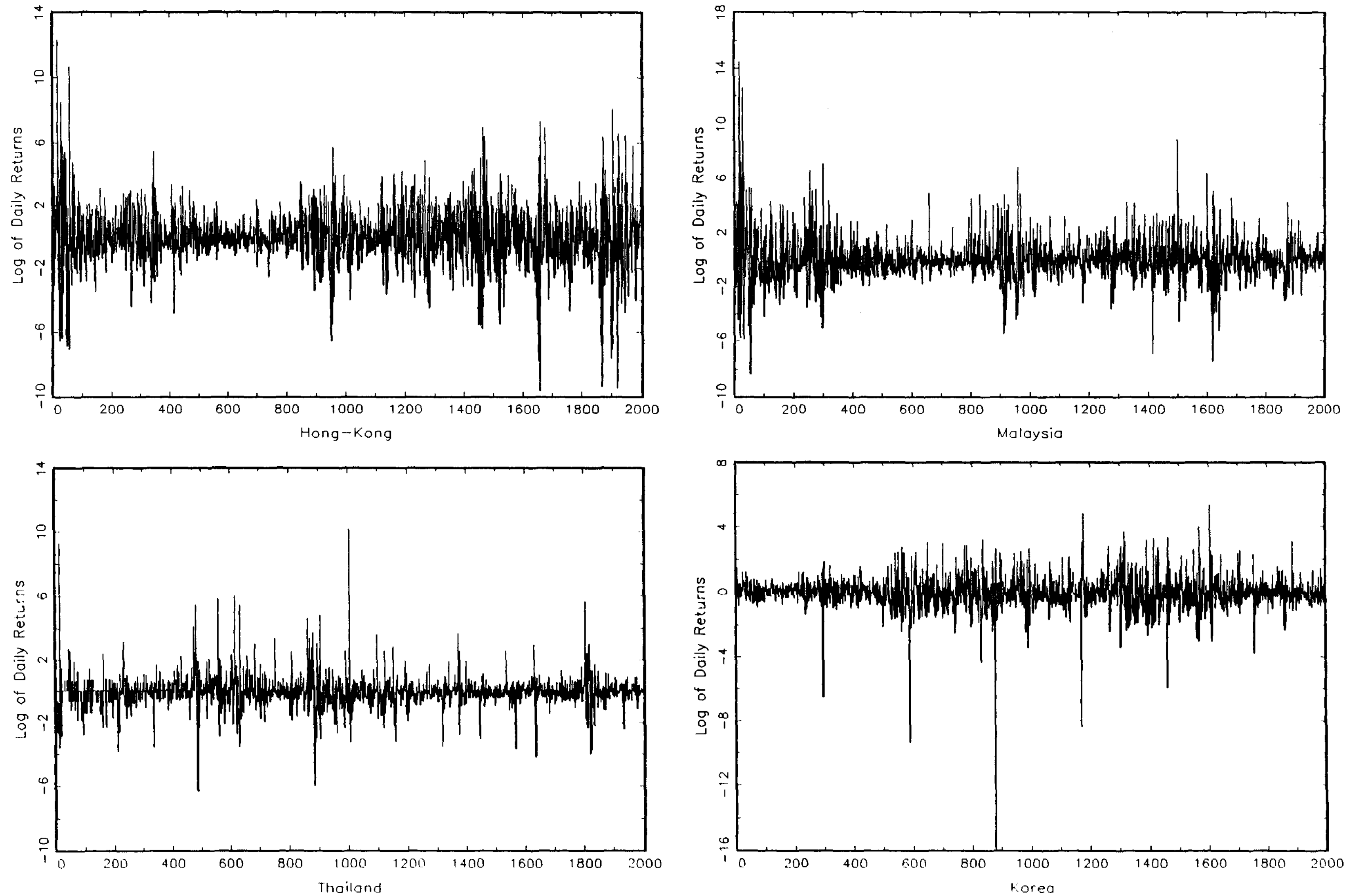
Graph 3.10b — Empirical Density x Normal



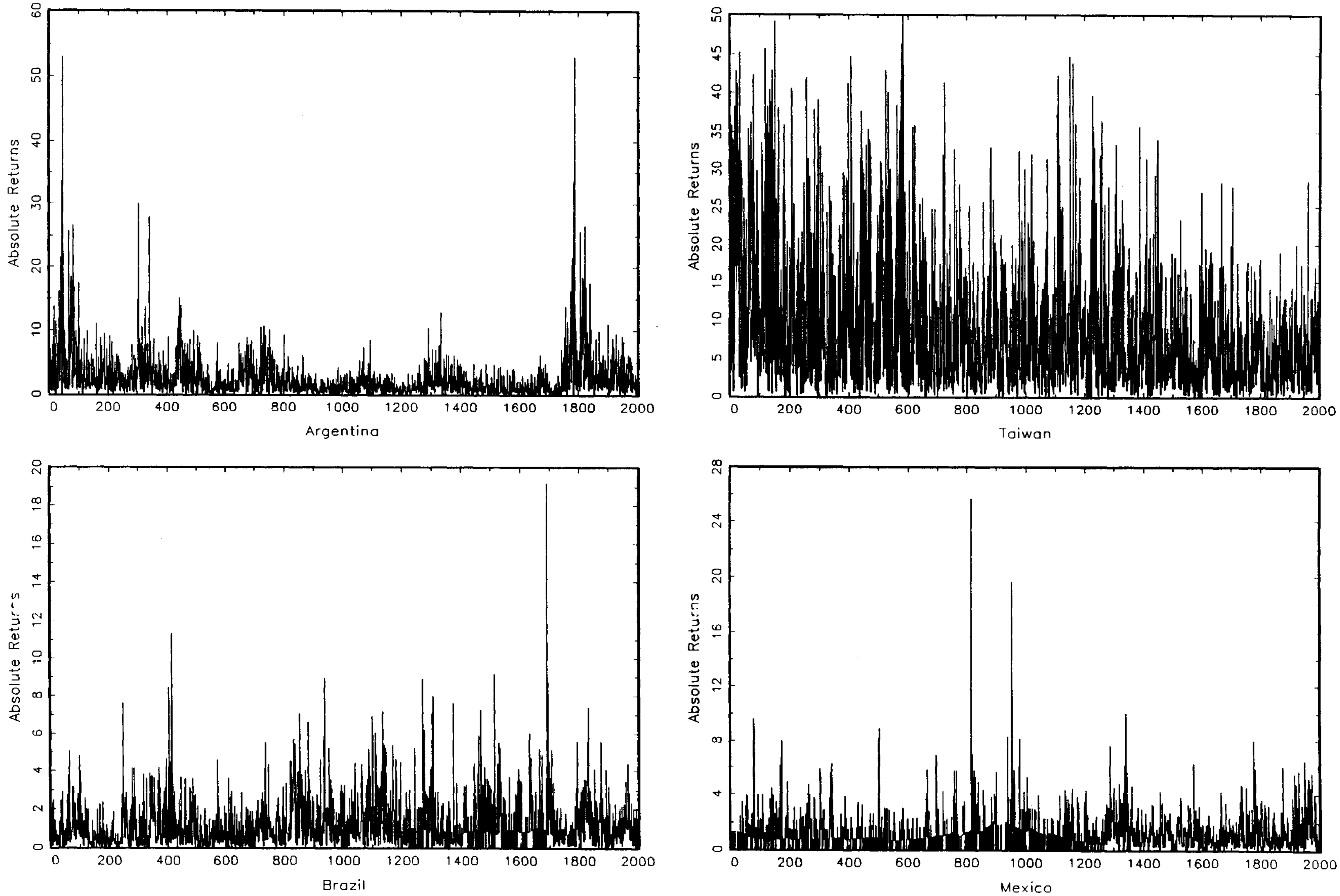
Graph 3.11a – Returns in Emerging Markets



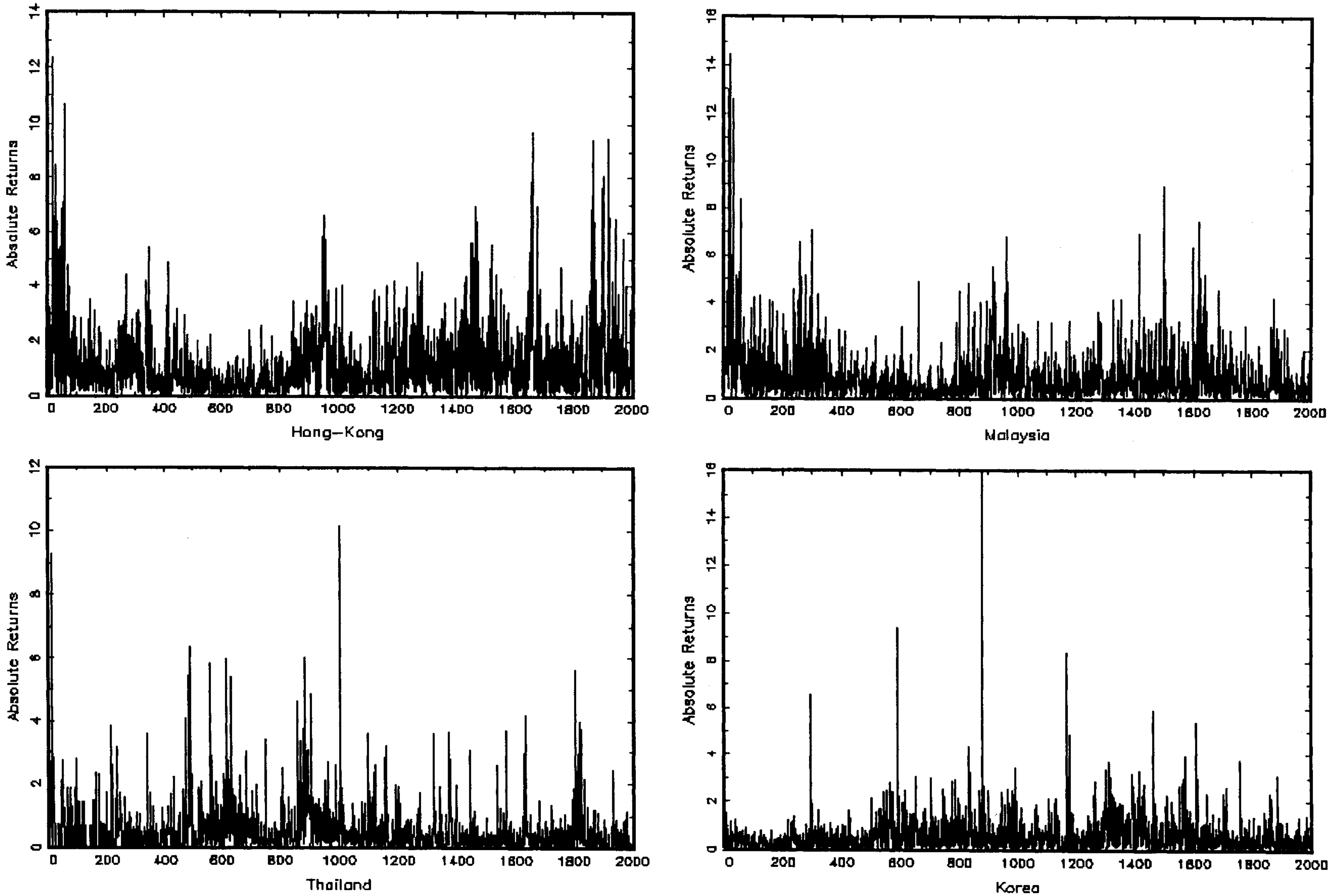
Graph 3.11b – Returns in Emerging Markets



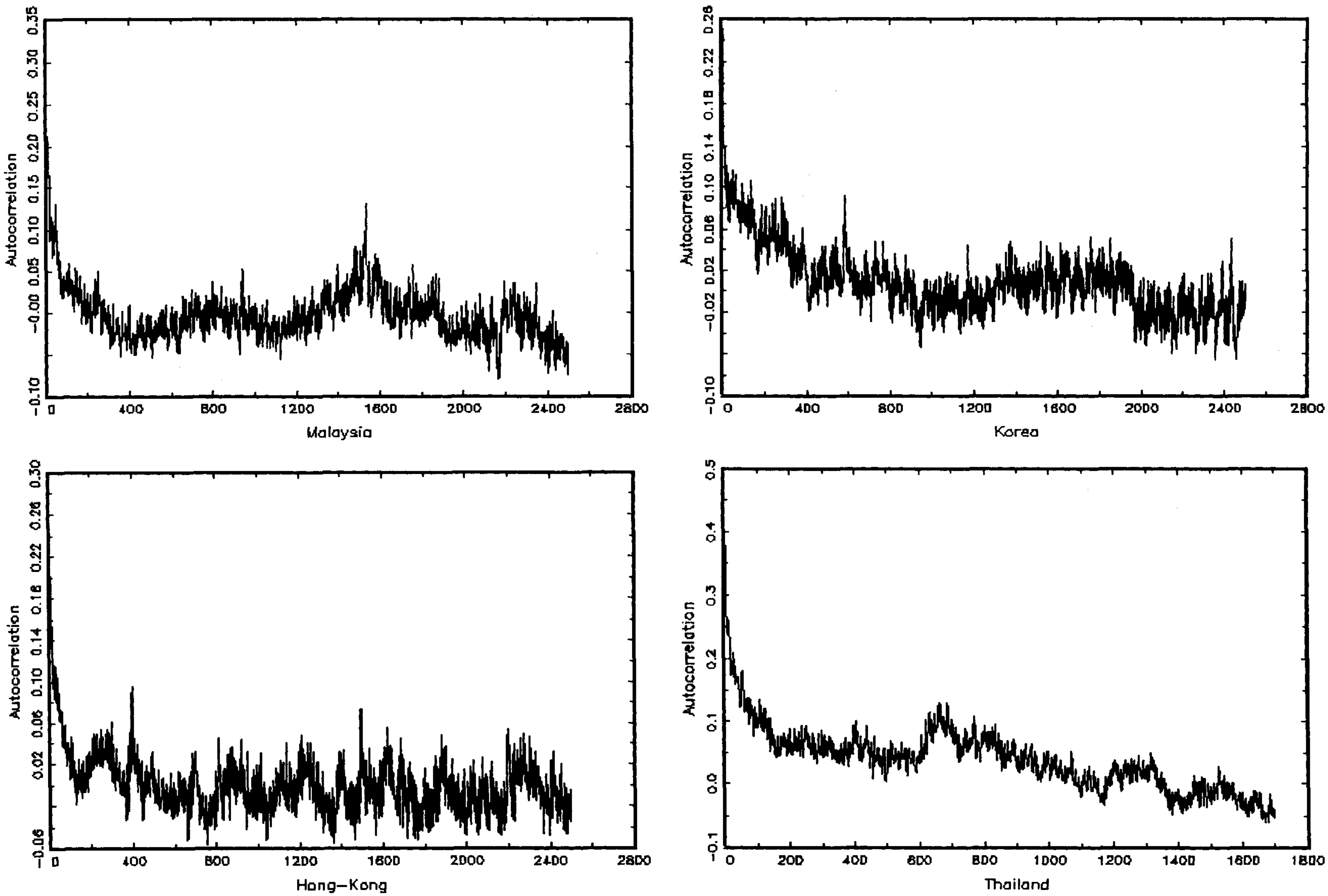
Graph 3.12a – Absolute Returns in Emerging Markets



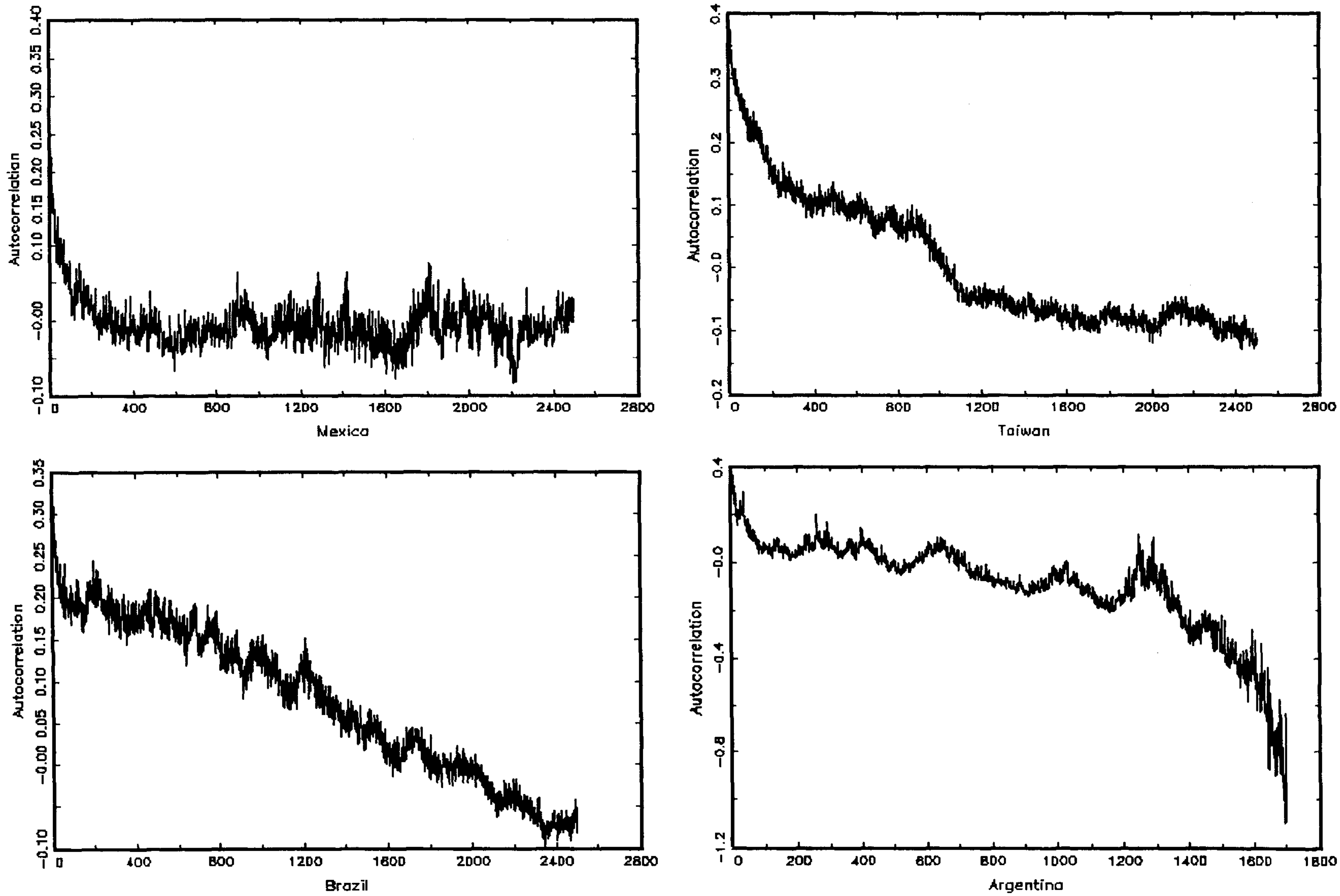
Graph 3.12b – Absolute Returns in Emerging Markets



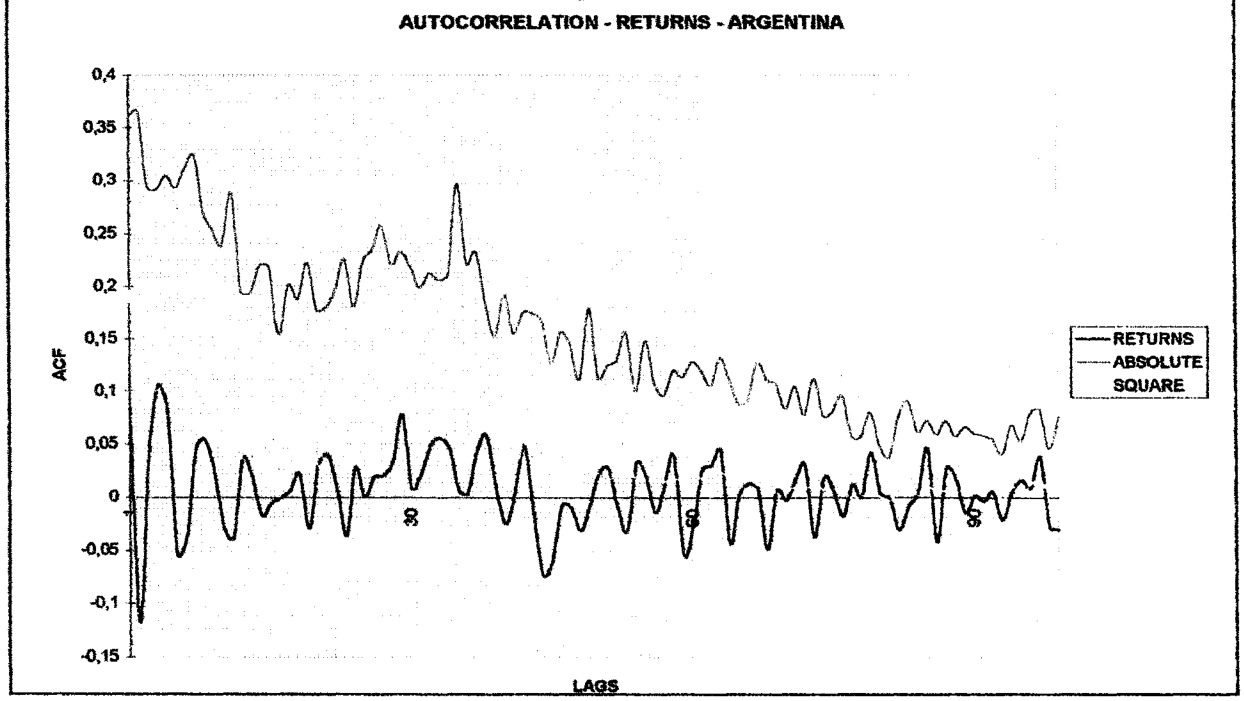
Graph 3.13a – Autocorrelation up to 2500 lags



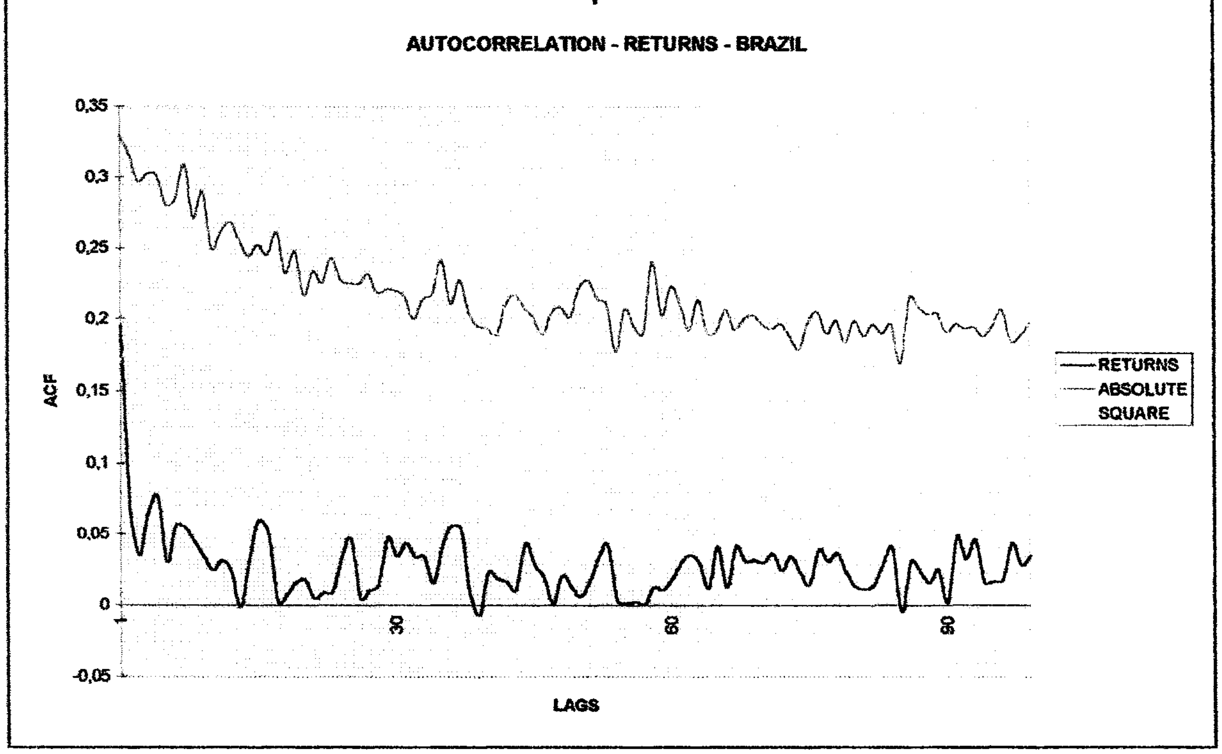
Graph 3.13b - Autocorrelation up to 2500 lags



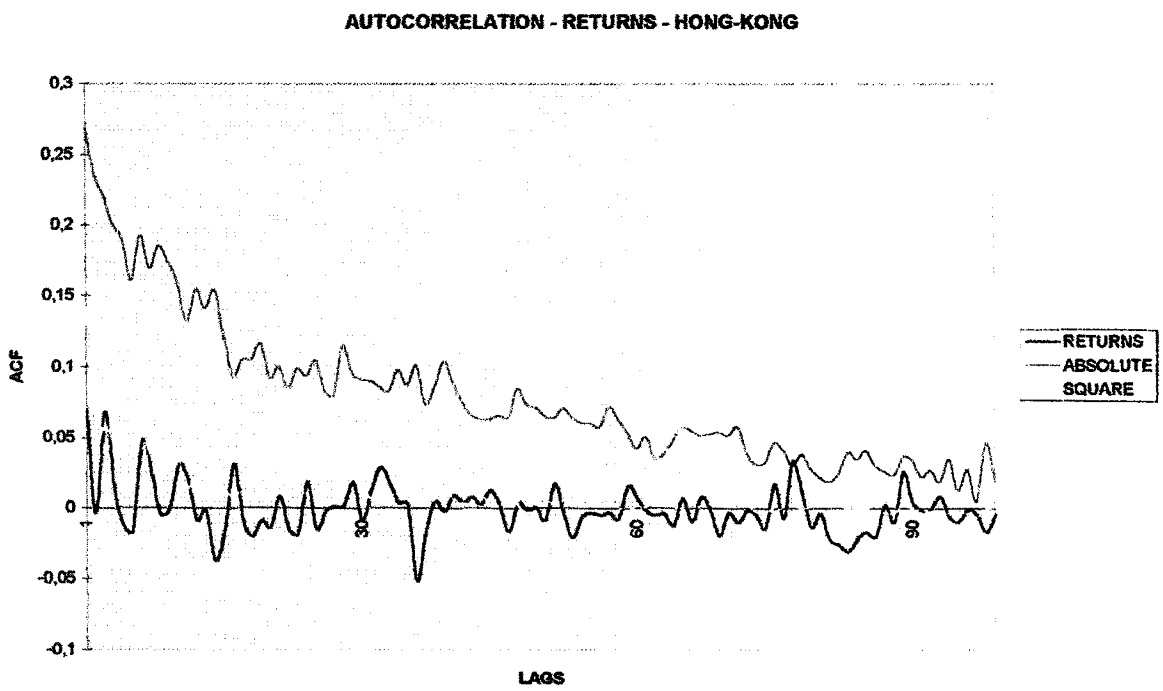
Graph 3.14



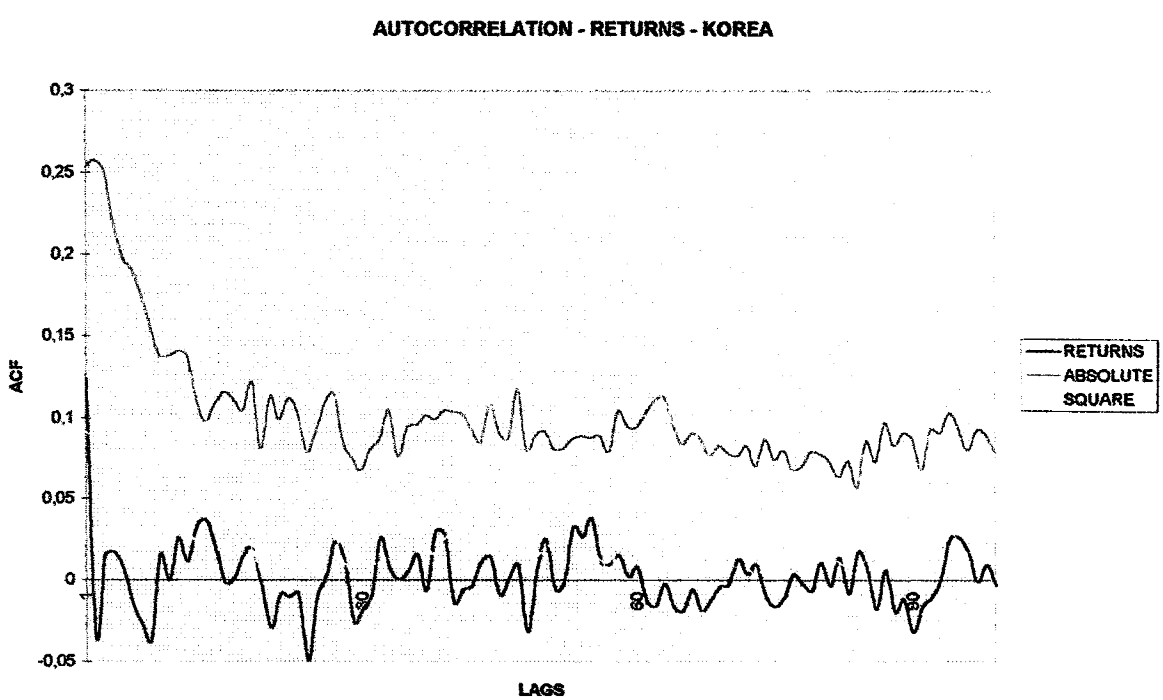
Graph 3.15



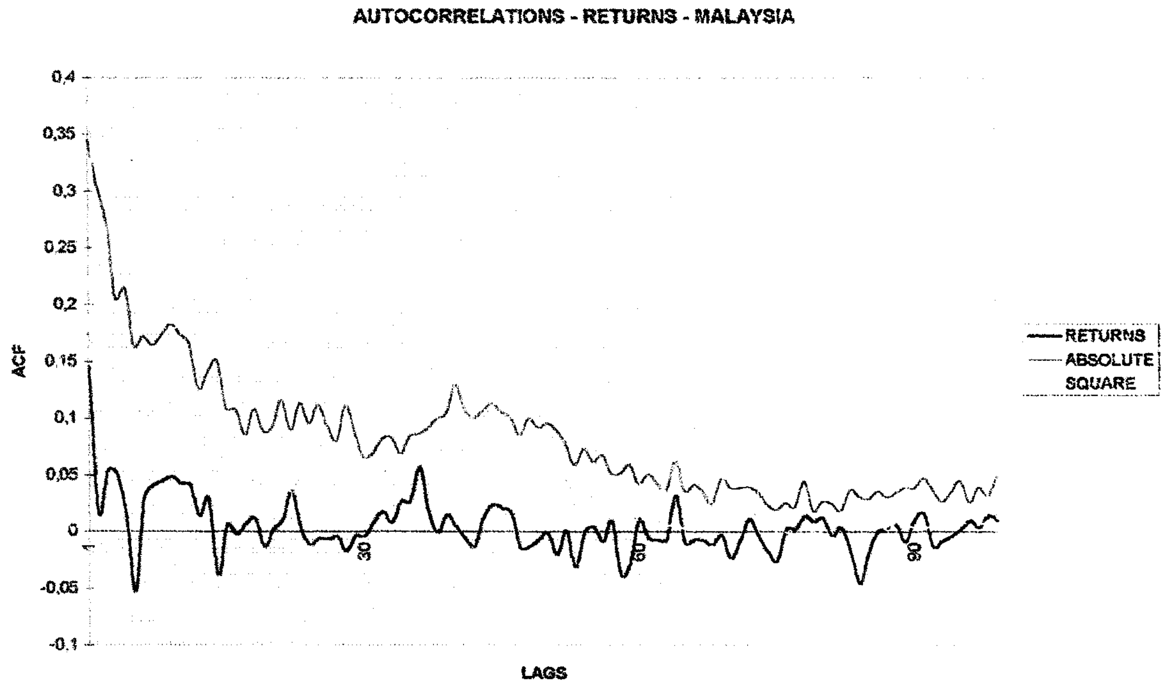
Graph 3.16



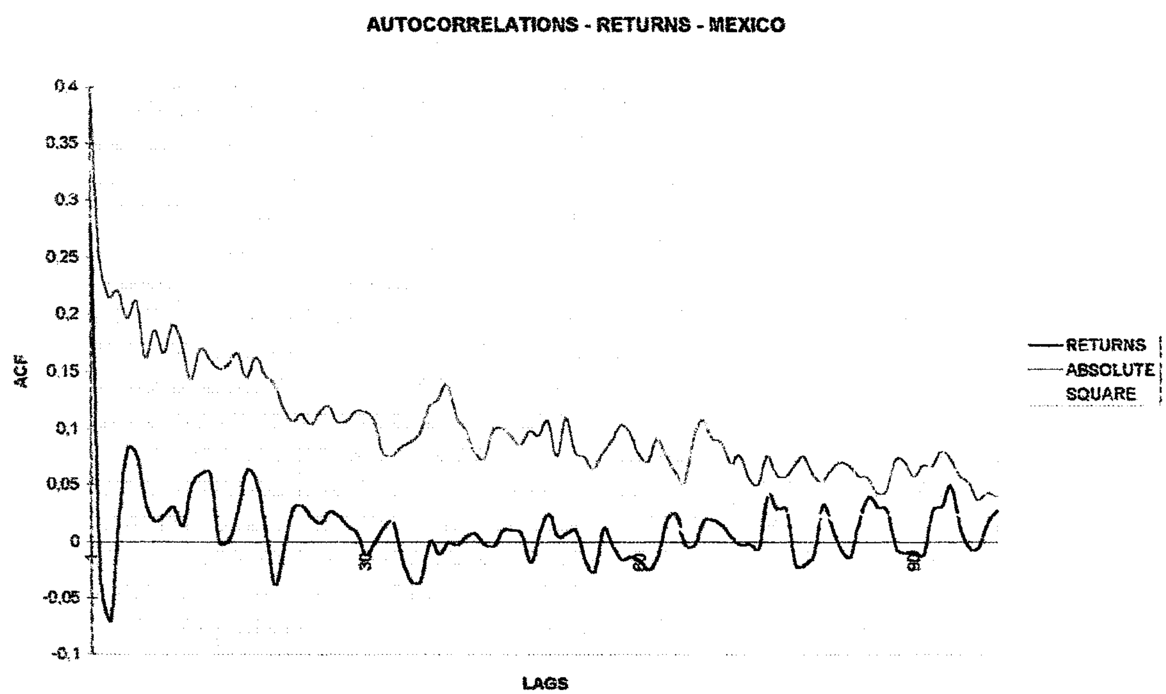
Graph 3.17



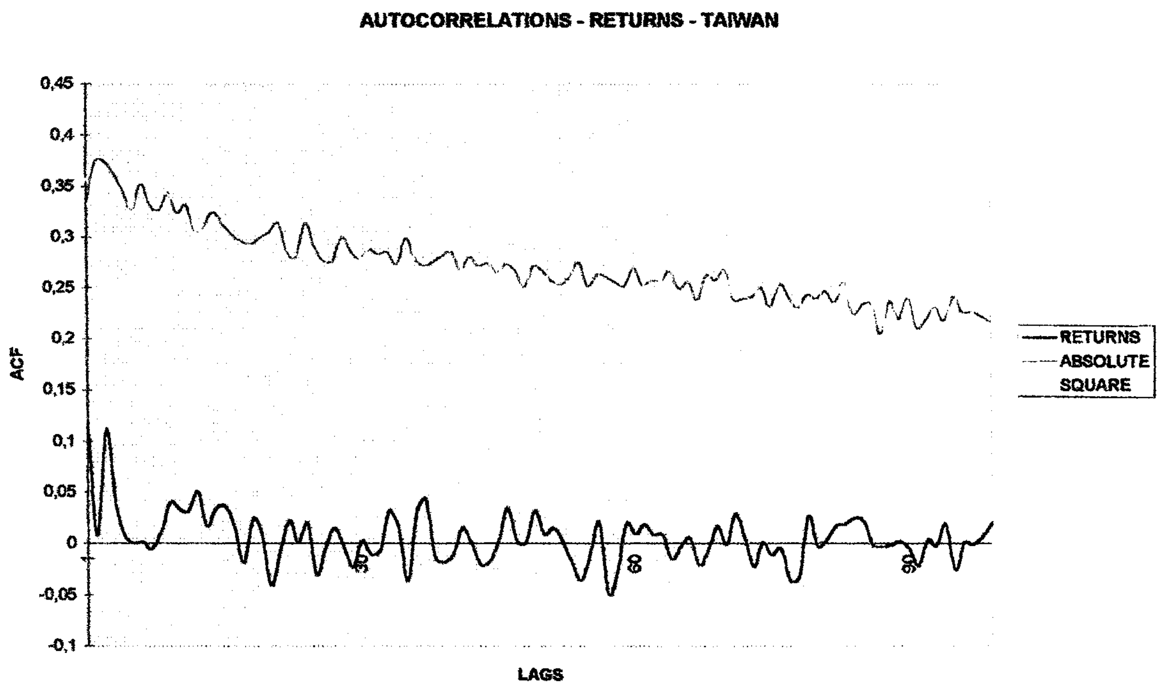
Graph 3.18



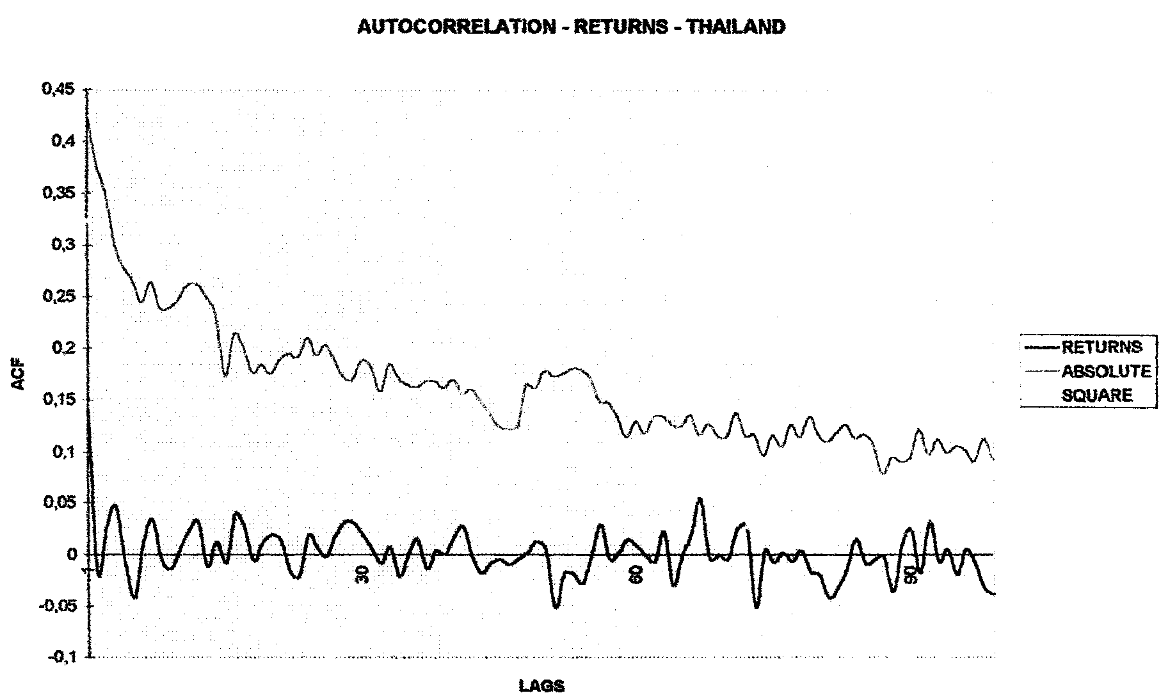
Graph 3.19



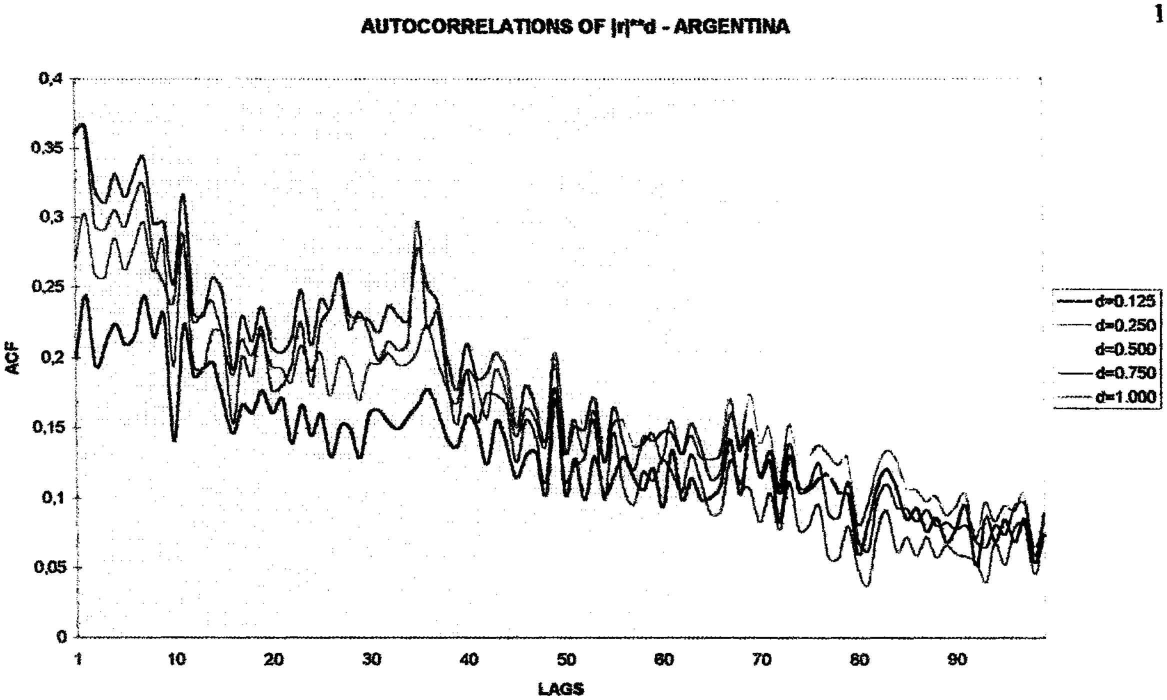
Graph 3.20



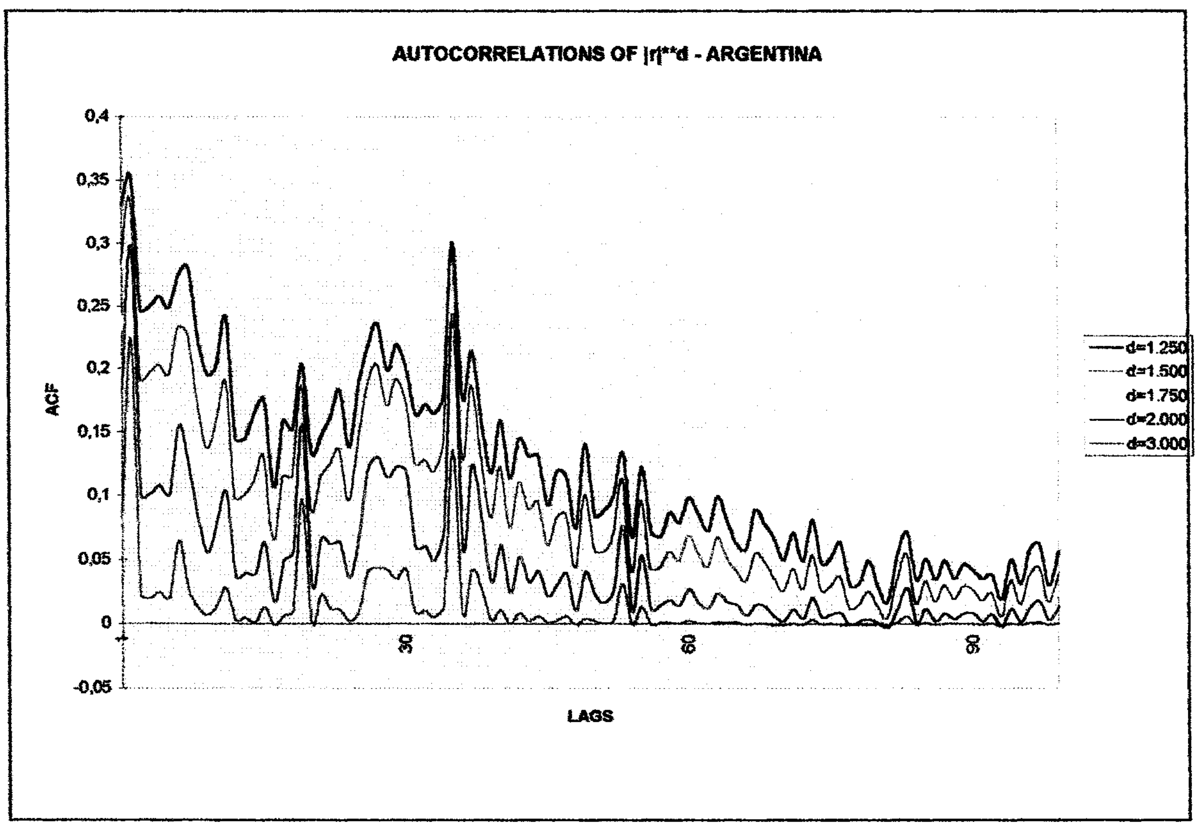
Graph 3.21



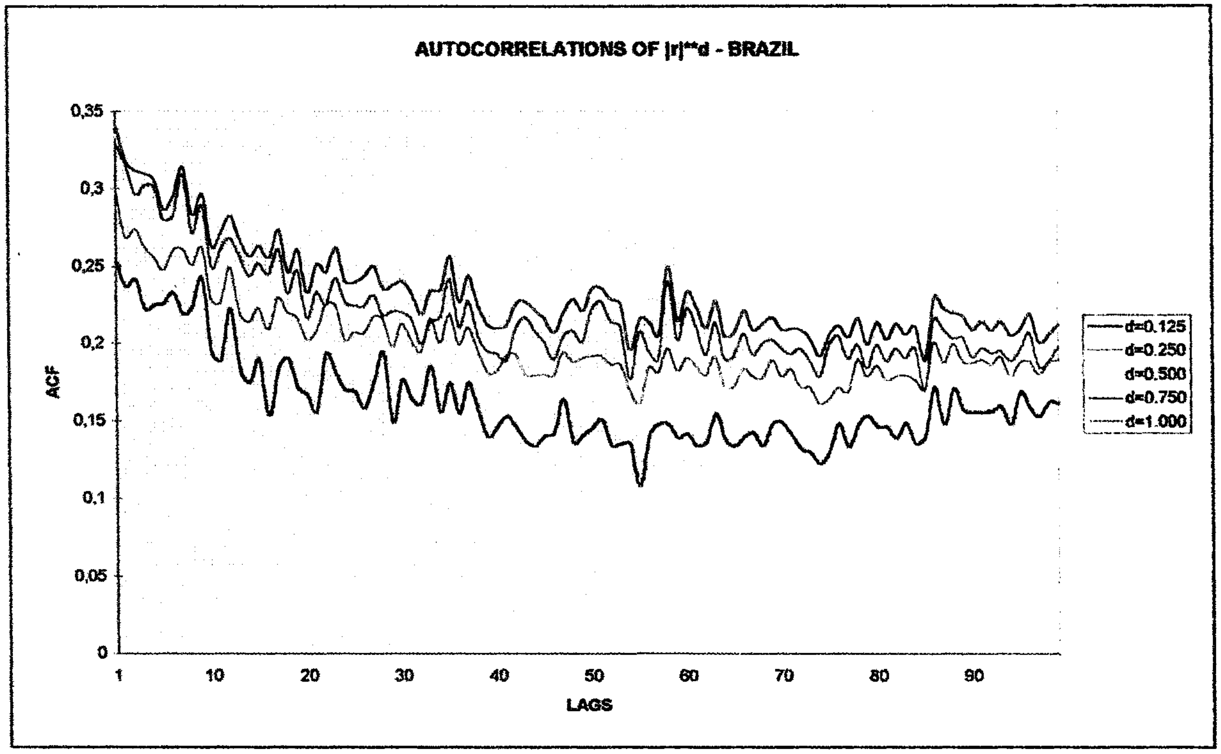
Graph 3.22a



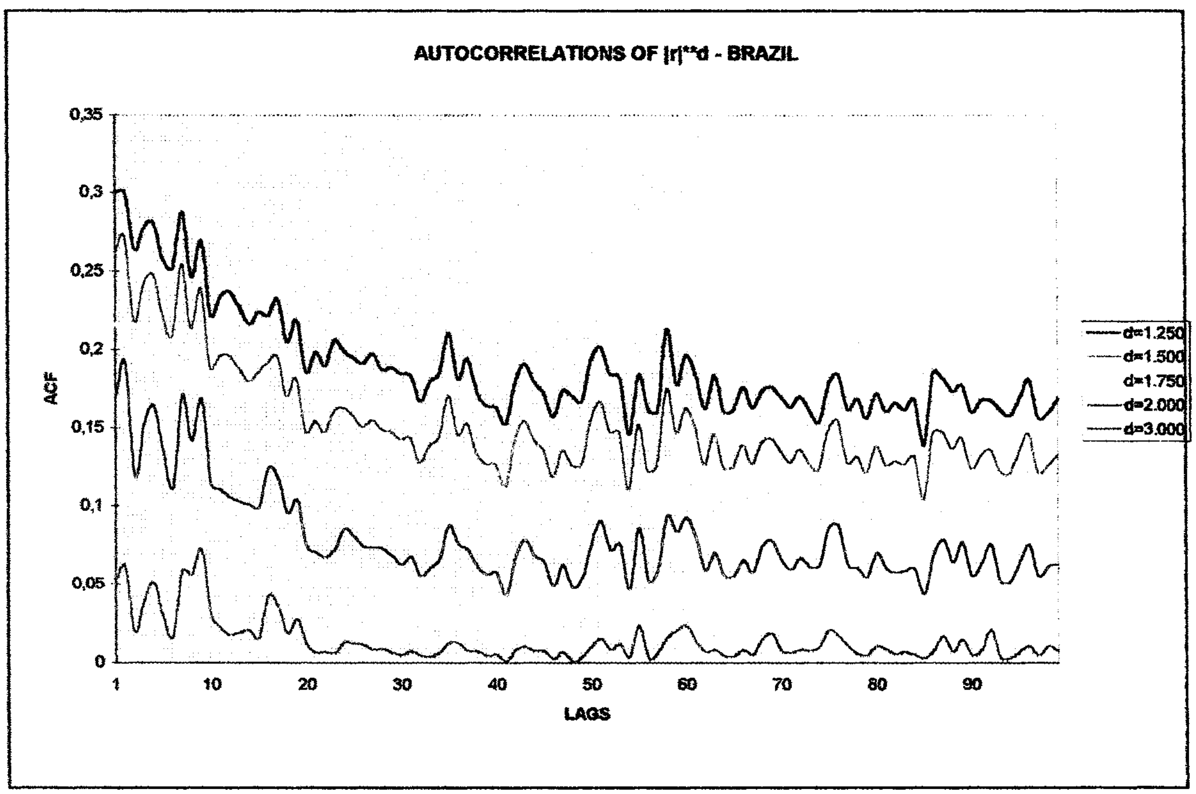
Graph 3.22b



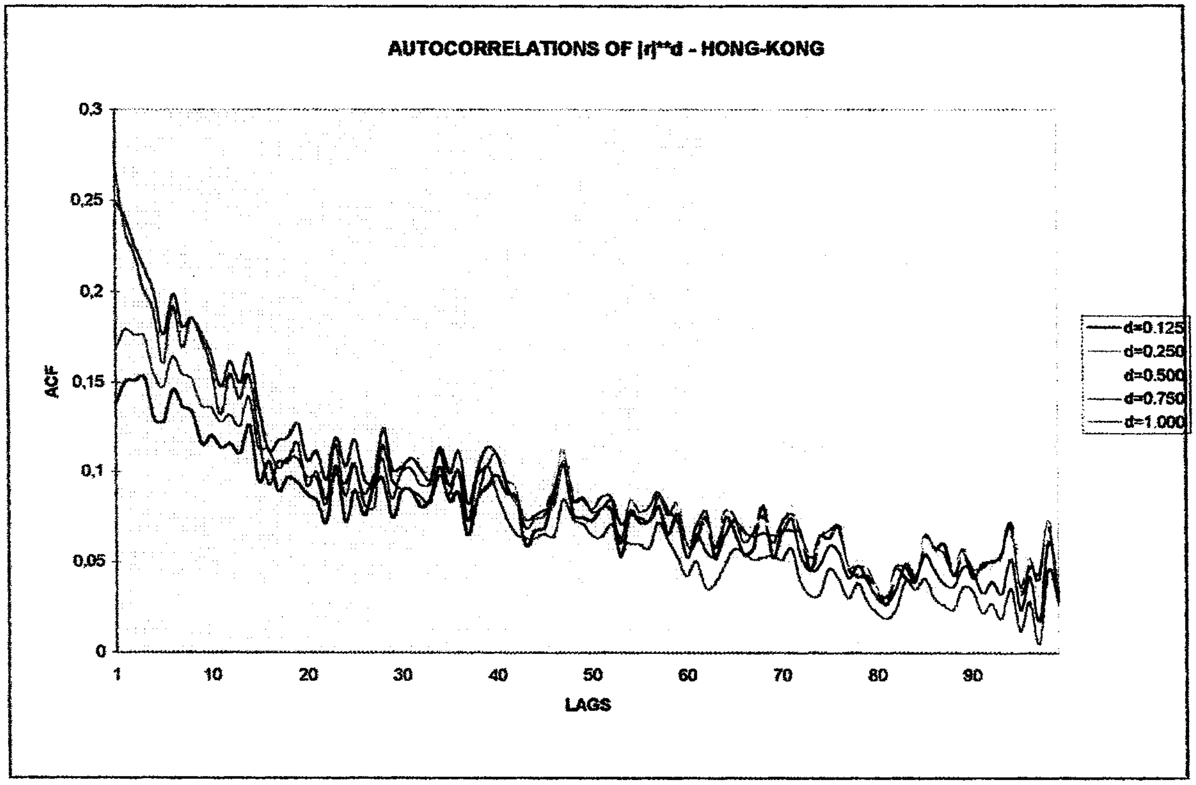
Graph 3.23a



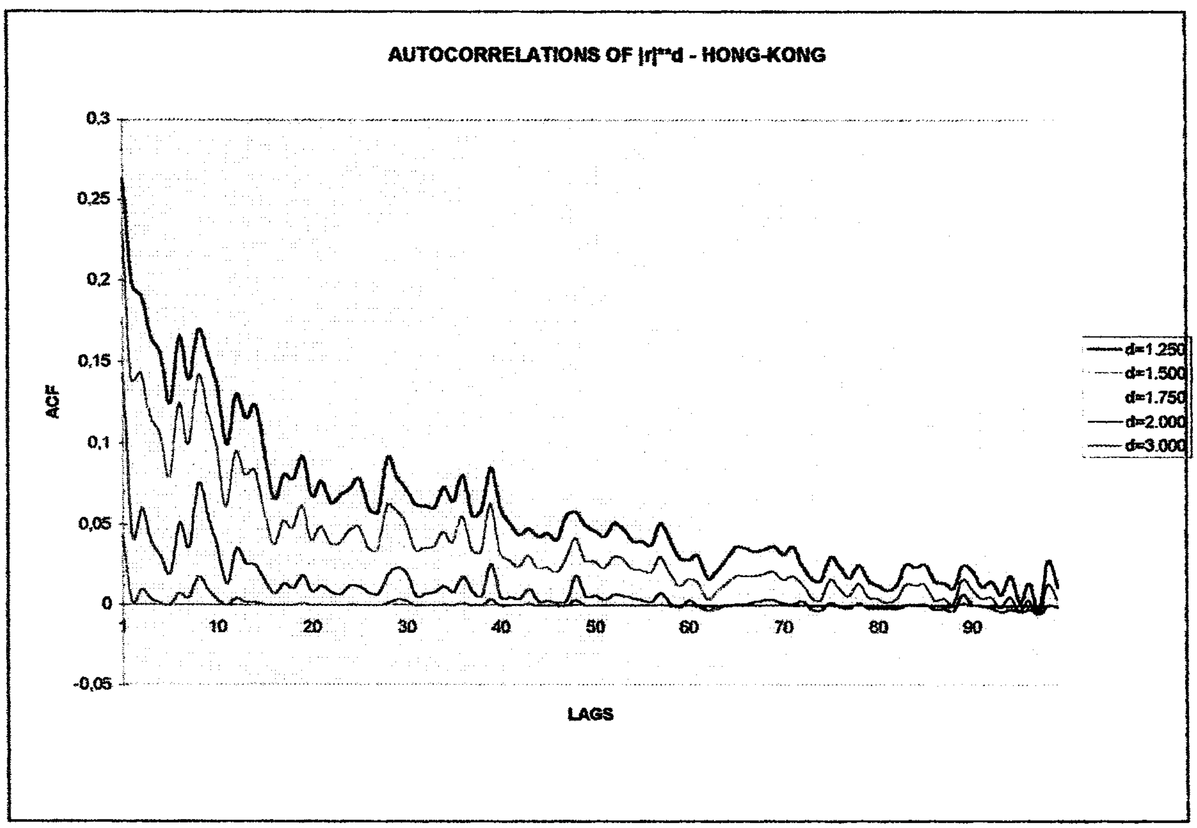
Graph 3.23b



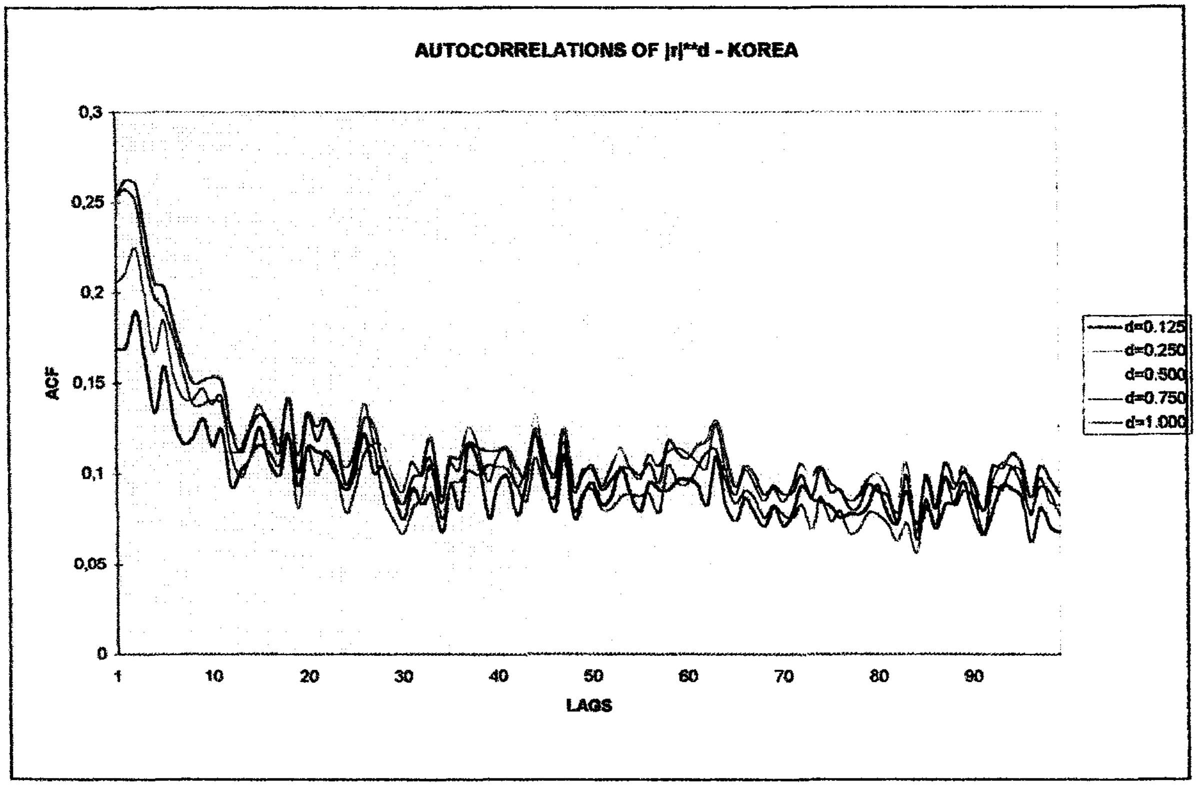
Graph 3.24a



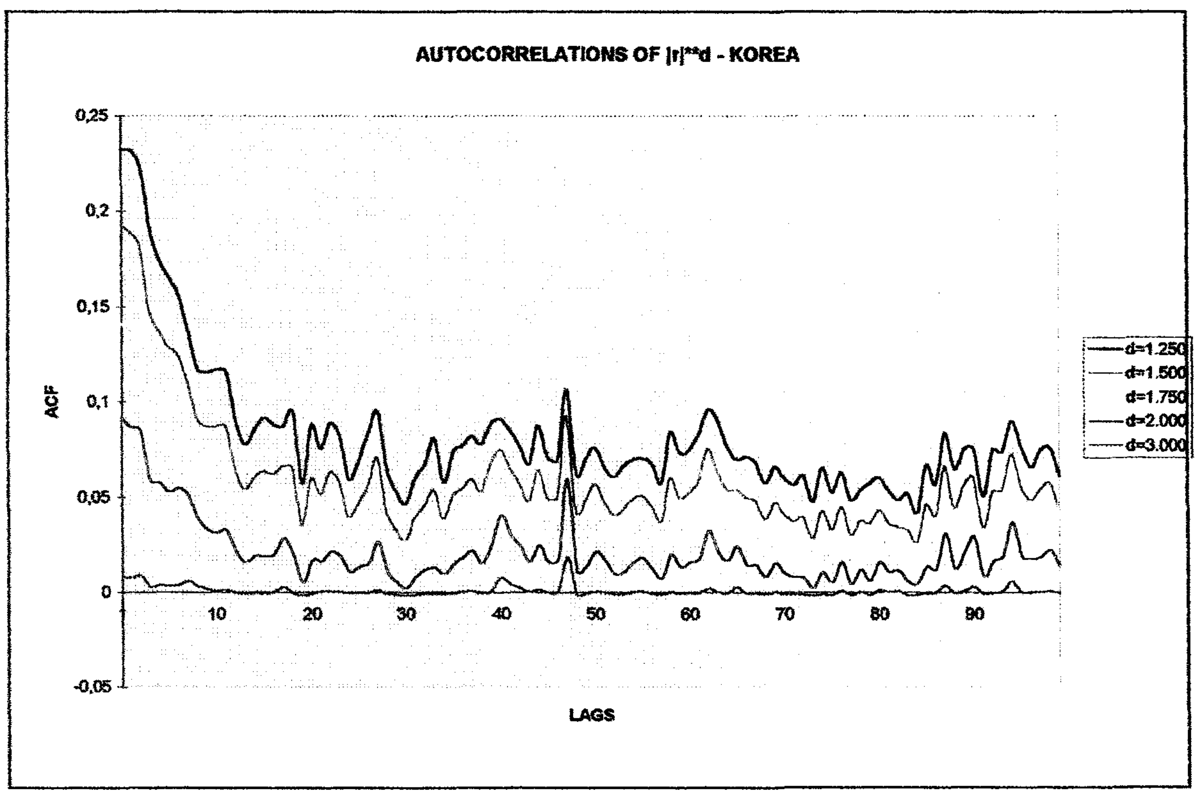
Graph 3.24b



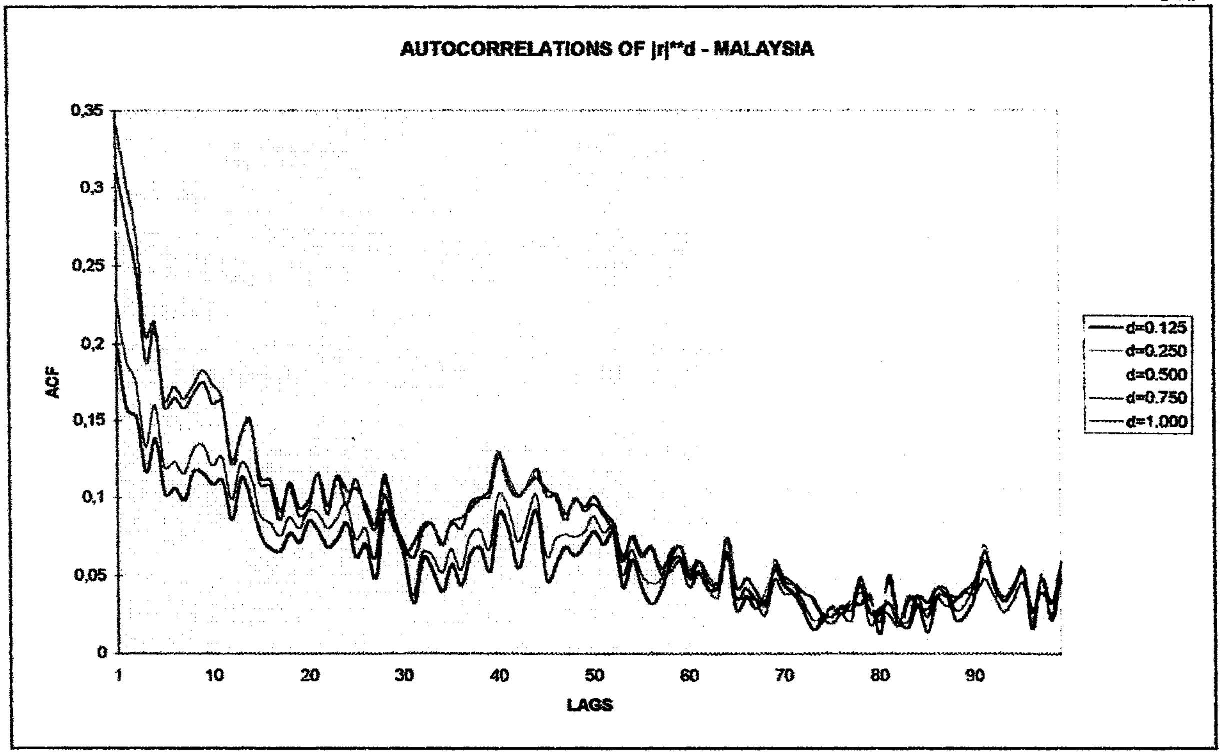
Graph 3.25a



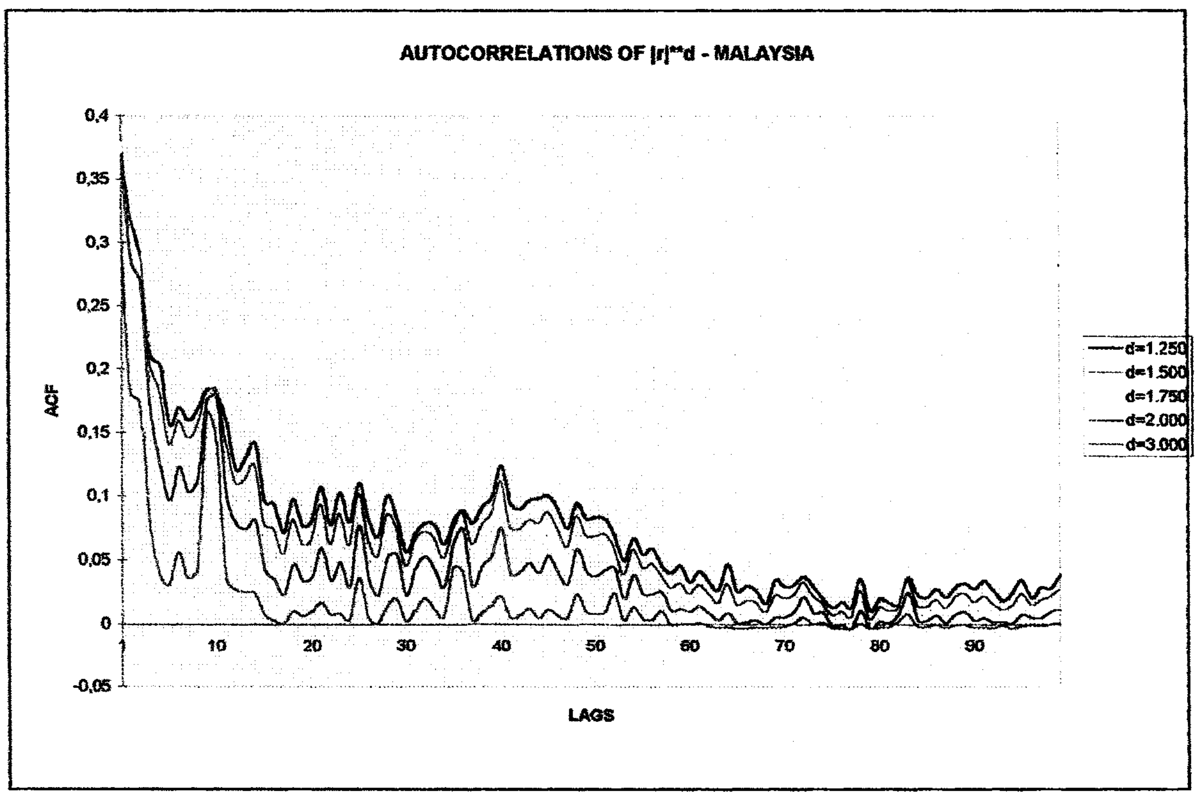
Graph 3.25b



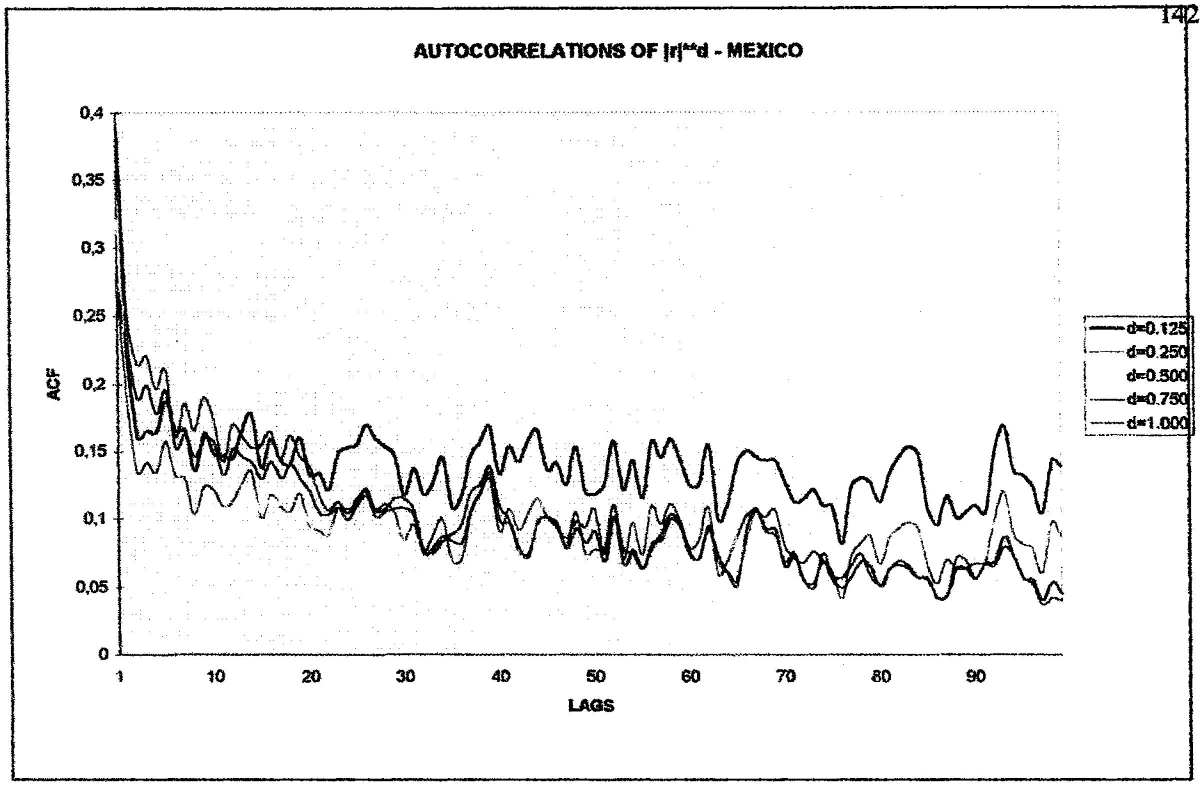
Graph 3.26a



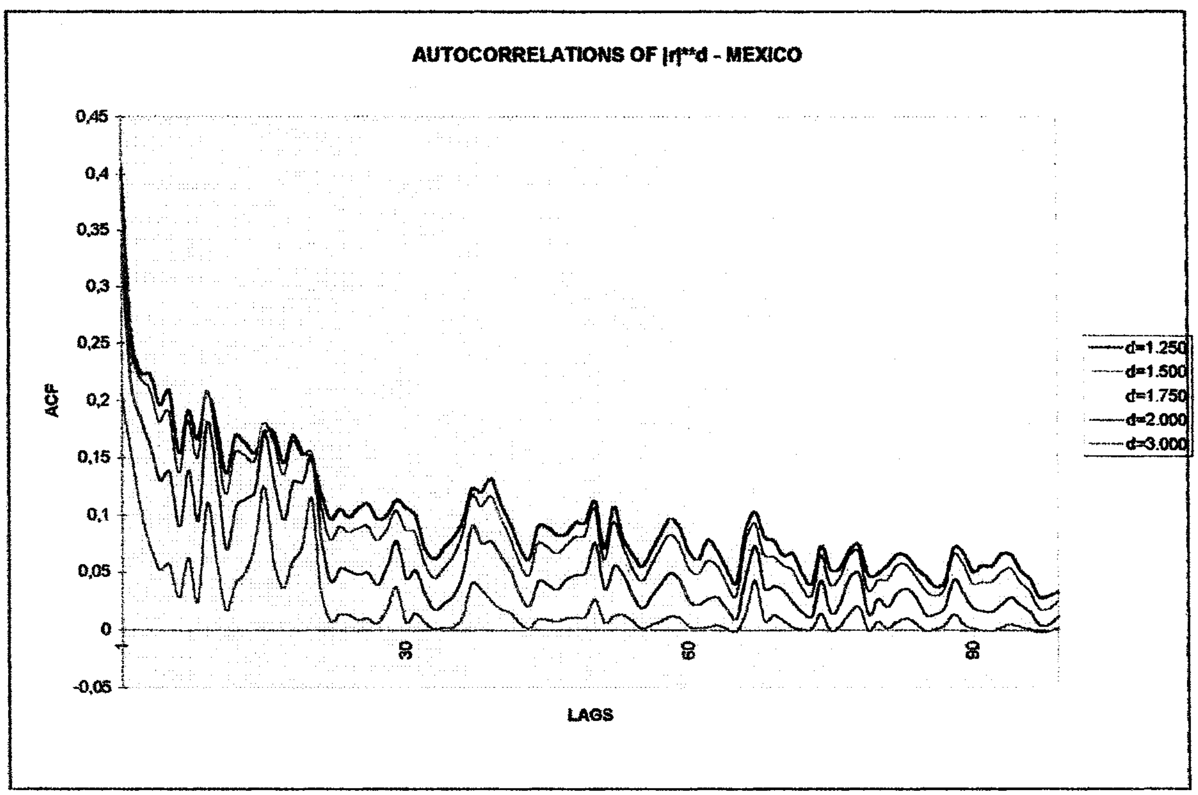
Graph 3.26b



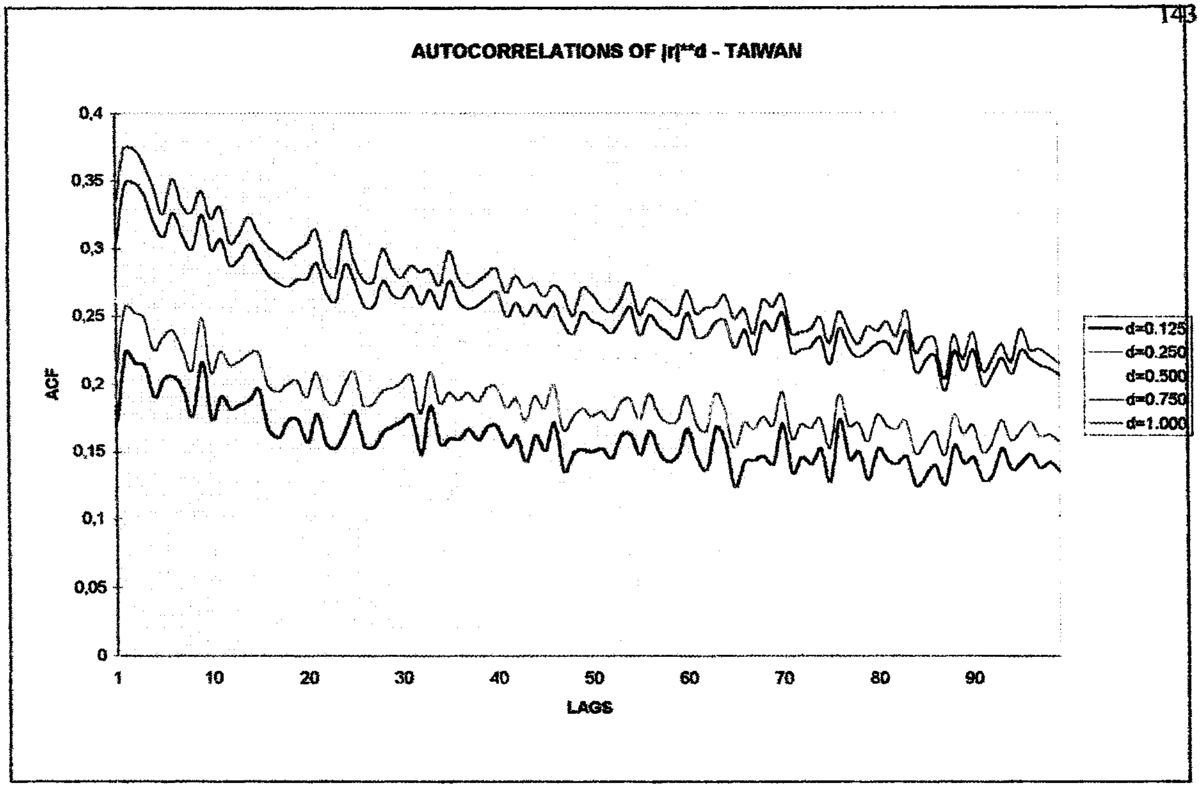
Graph 3.27a



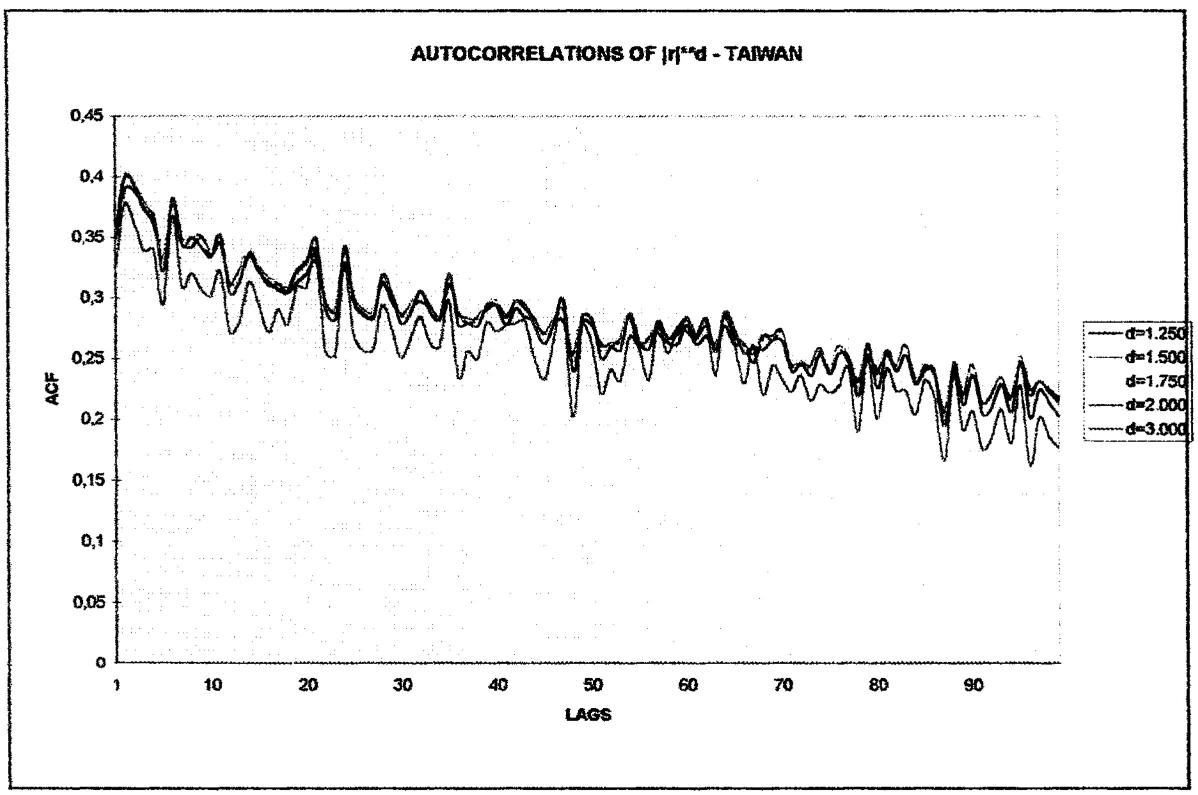
Graph 3.27b



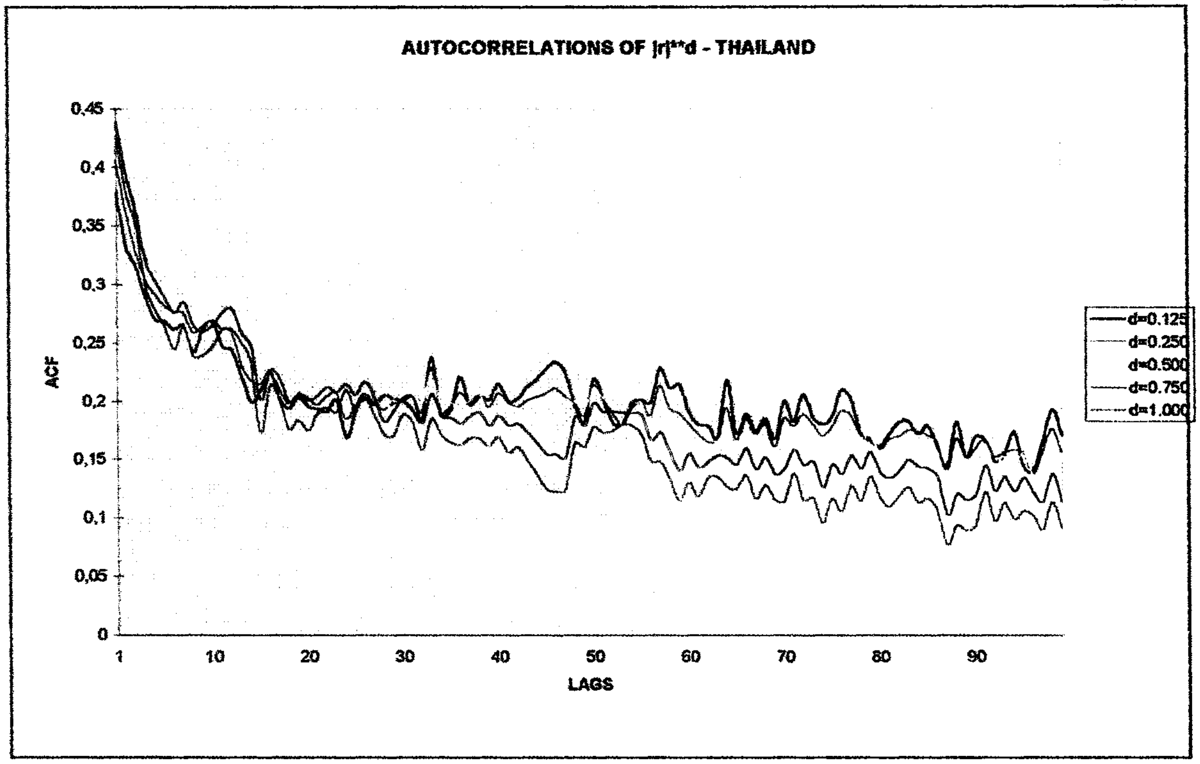
Graph 3.28a



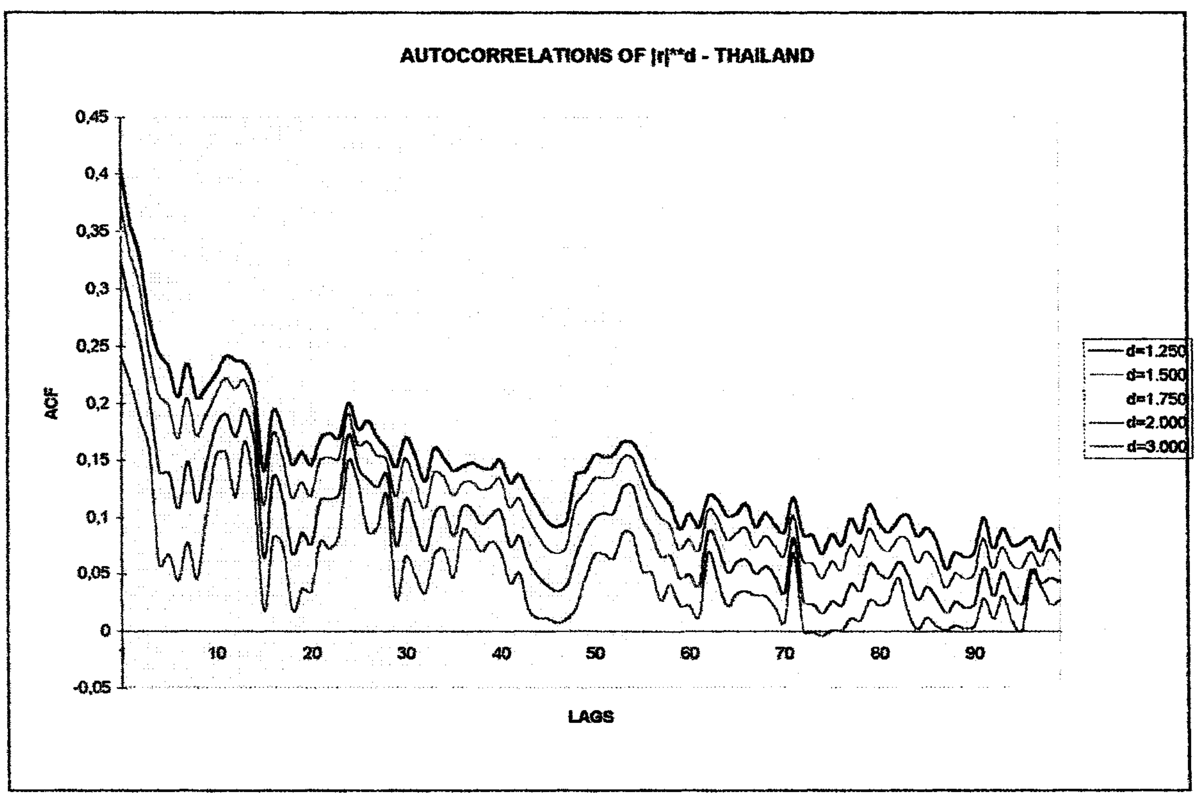
Graph 3.28b



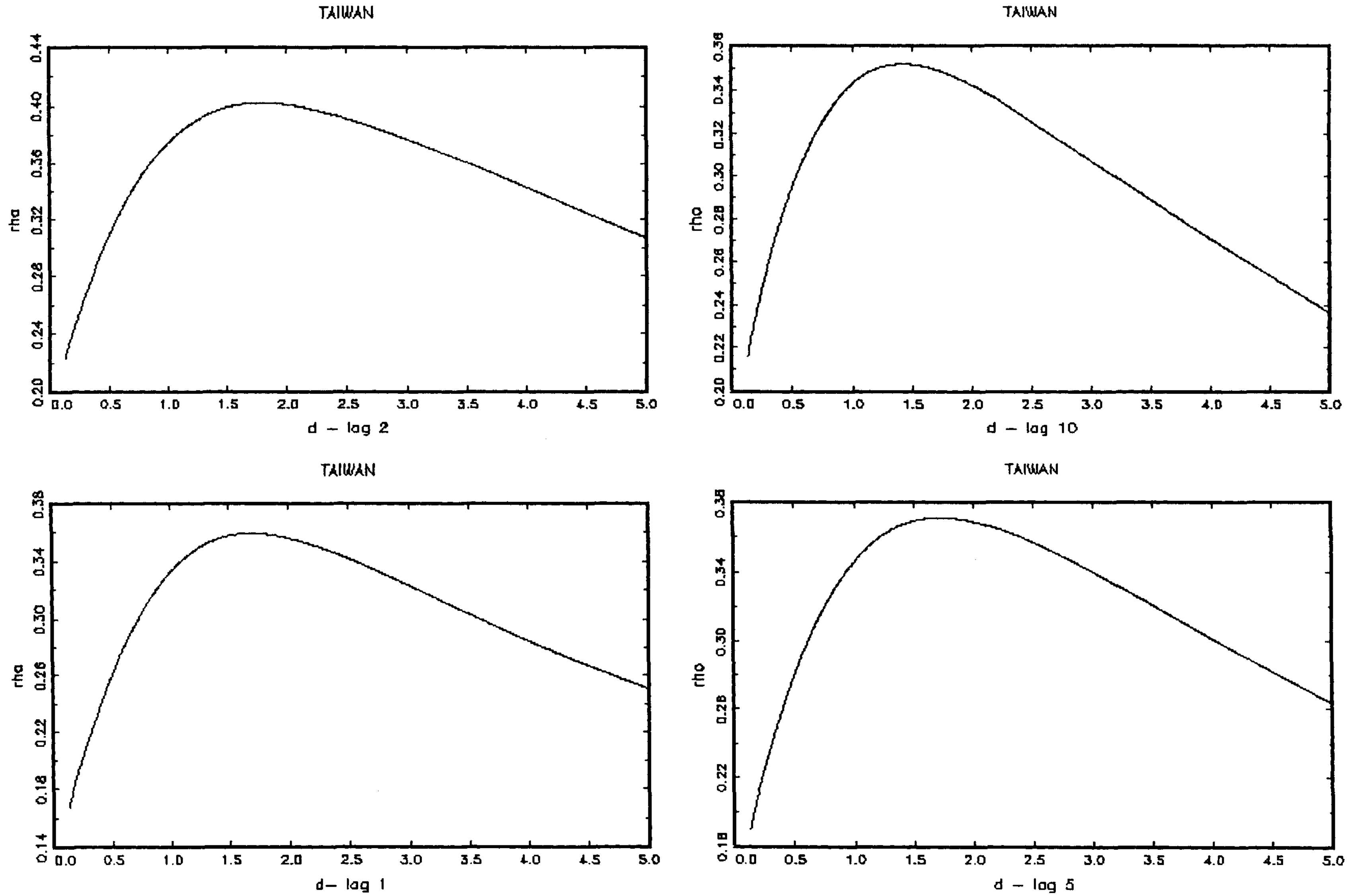
Graph 3.29a



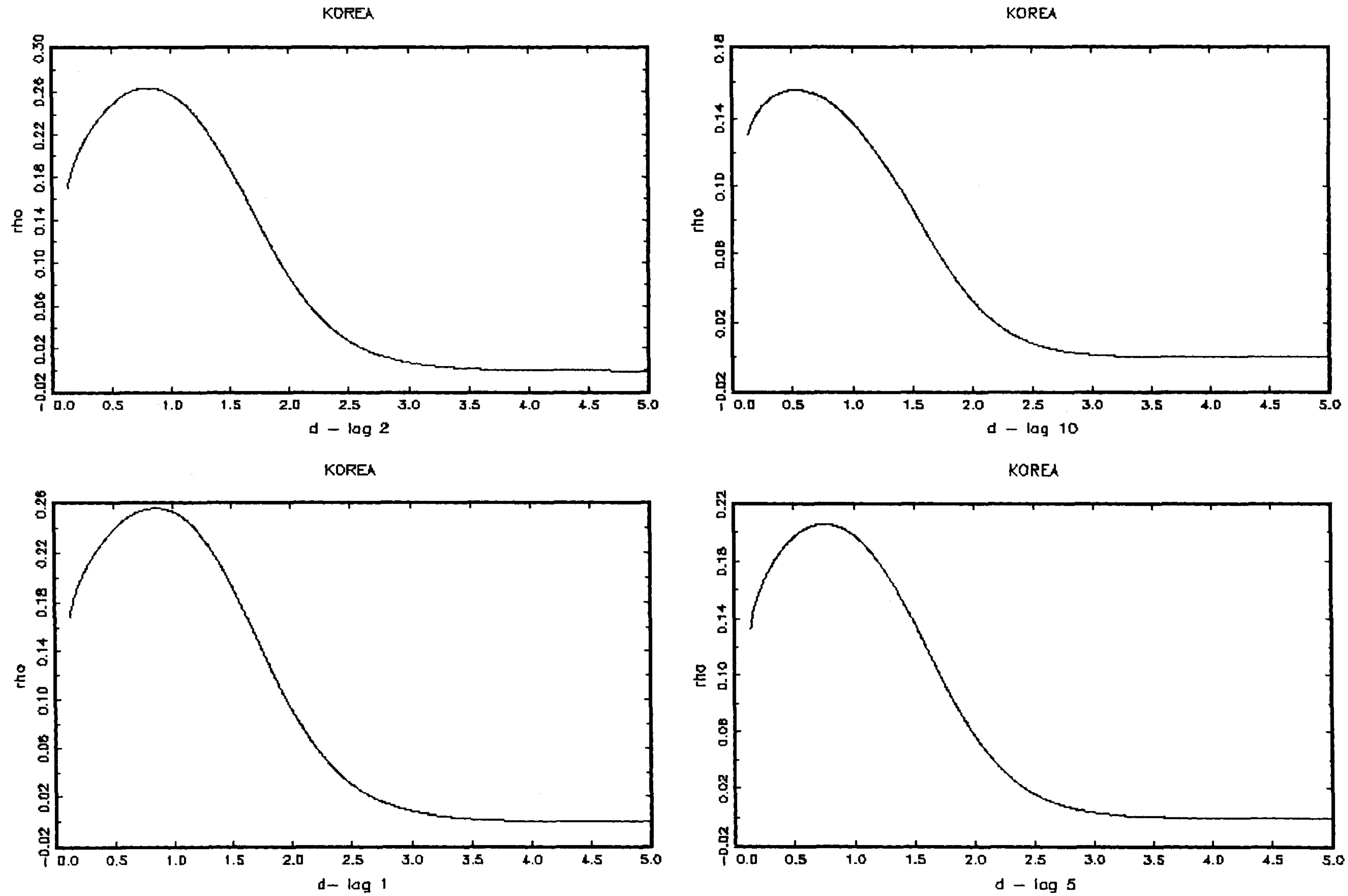
Graph 3.29b



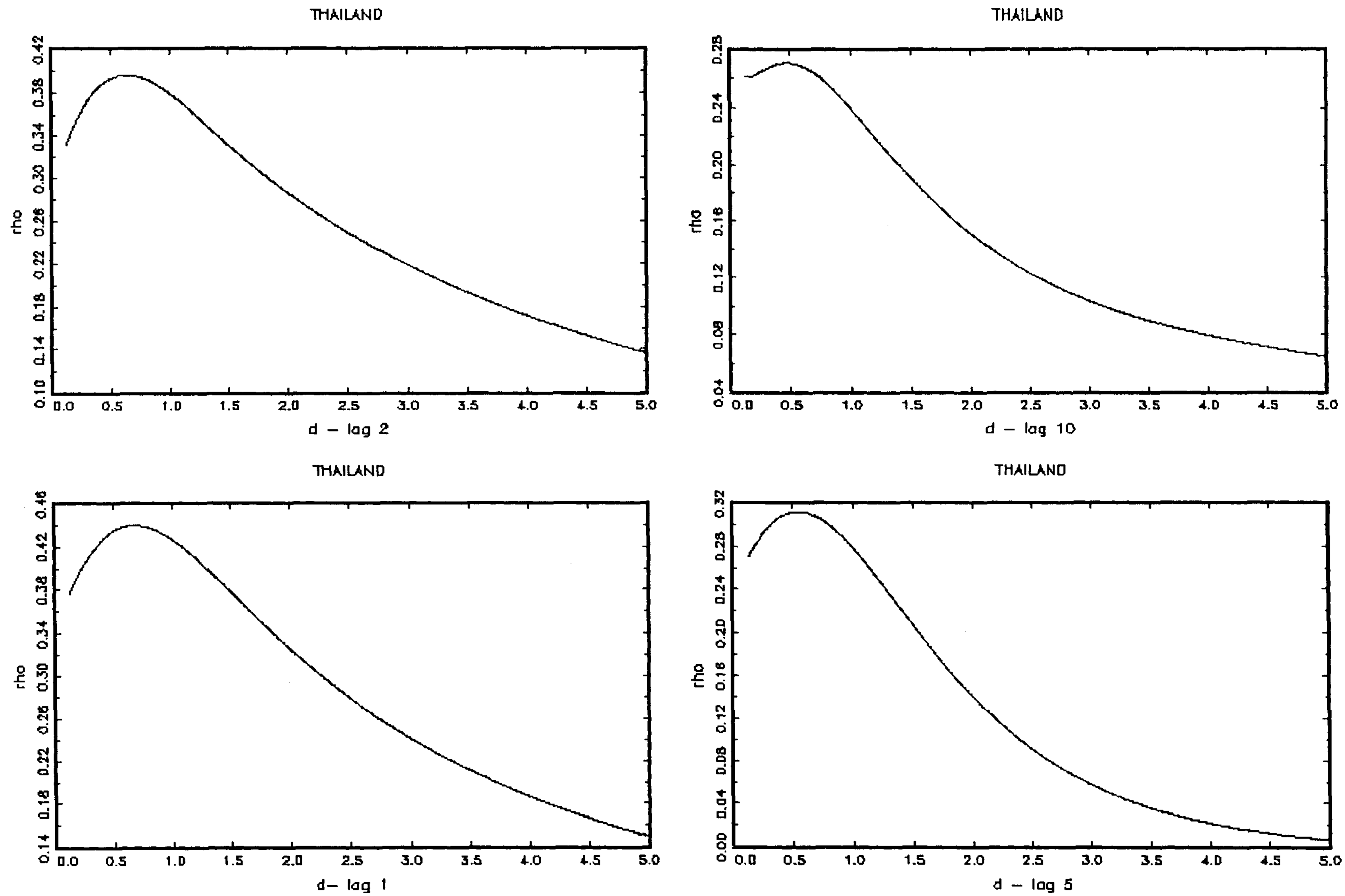
Graph 3.30 – Autocorrelations of |r|**d at lags



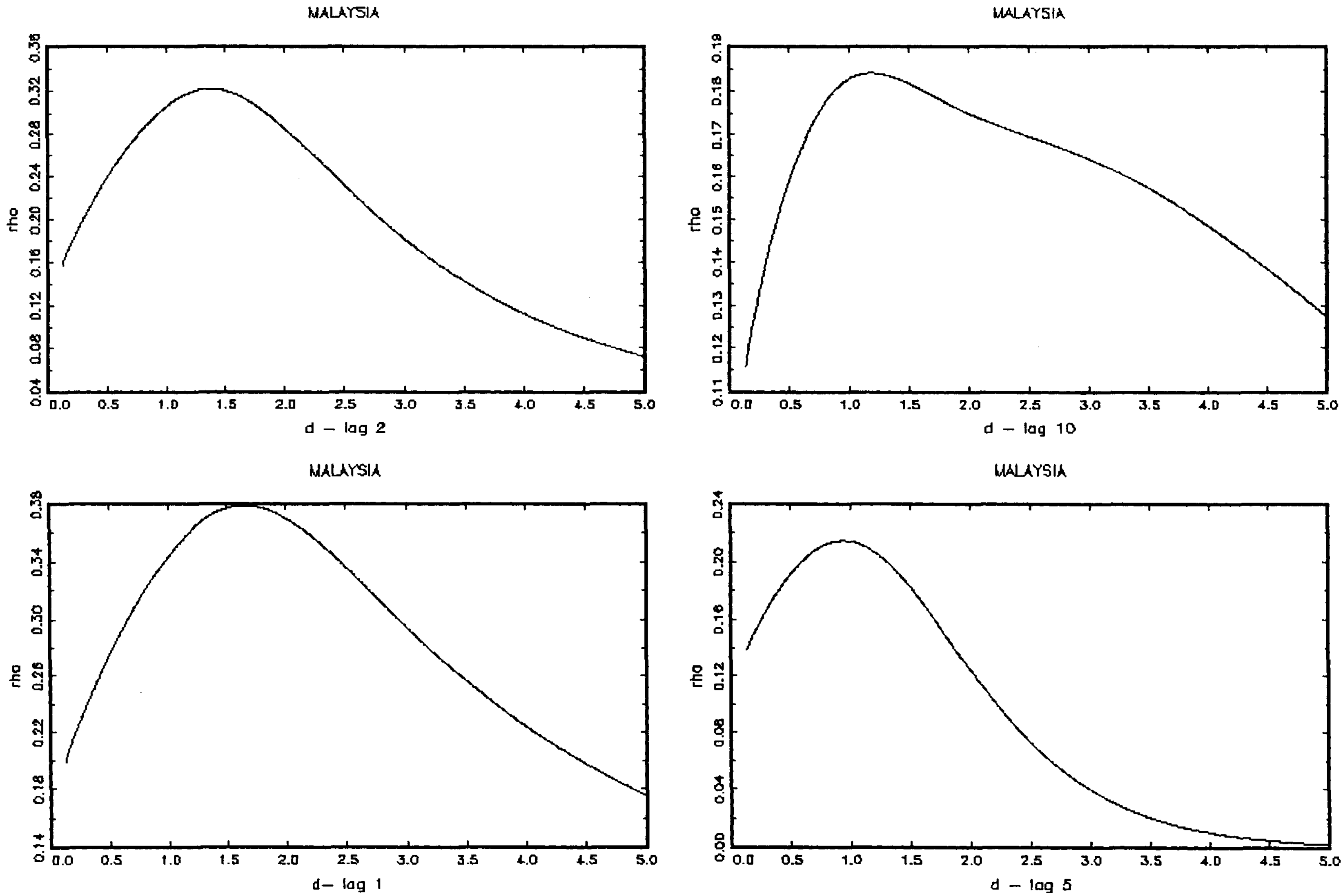
graph 3.31 – Autocorrelations of |r|**d at lags



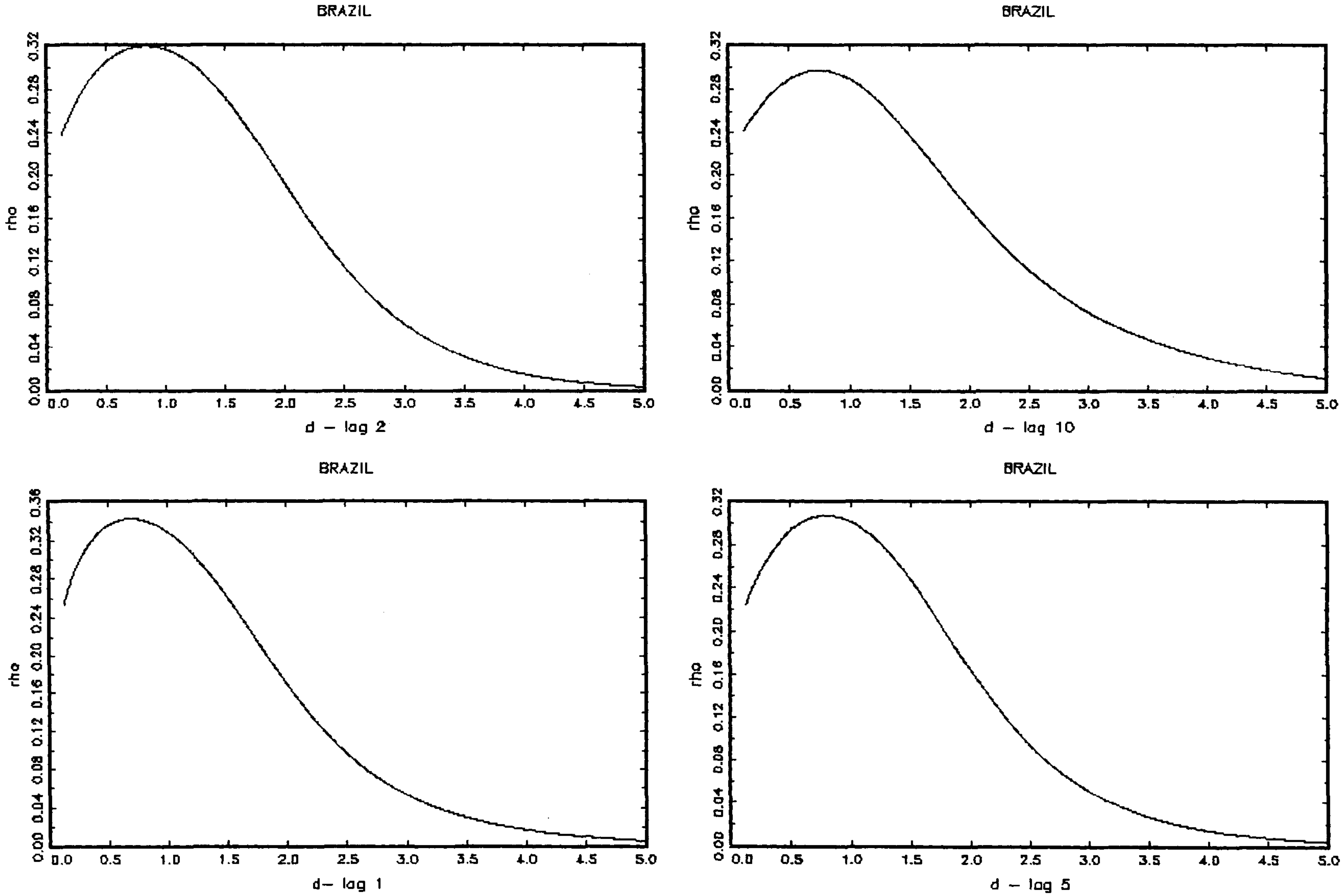
Graph 3.32 – Autocorrelations of |r|**d at lags



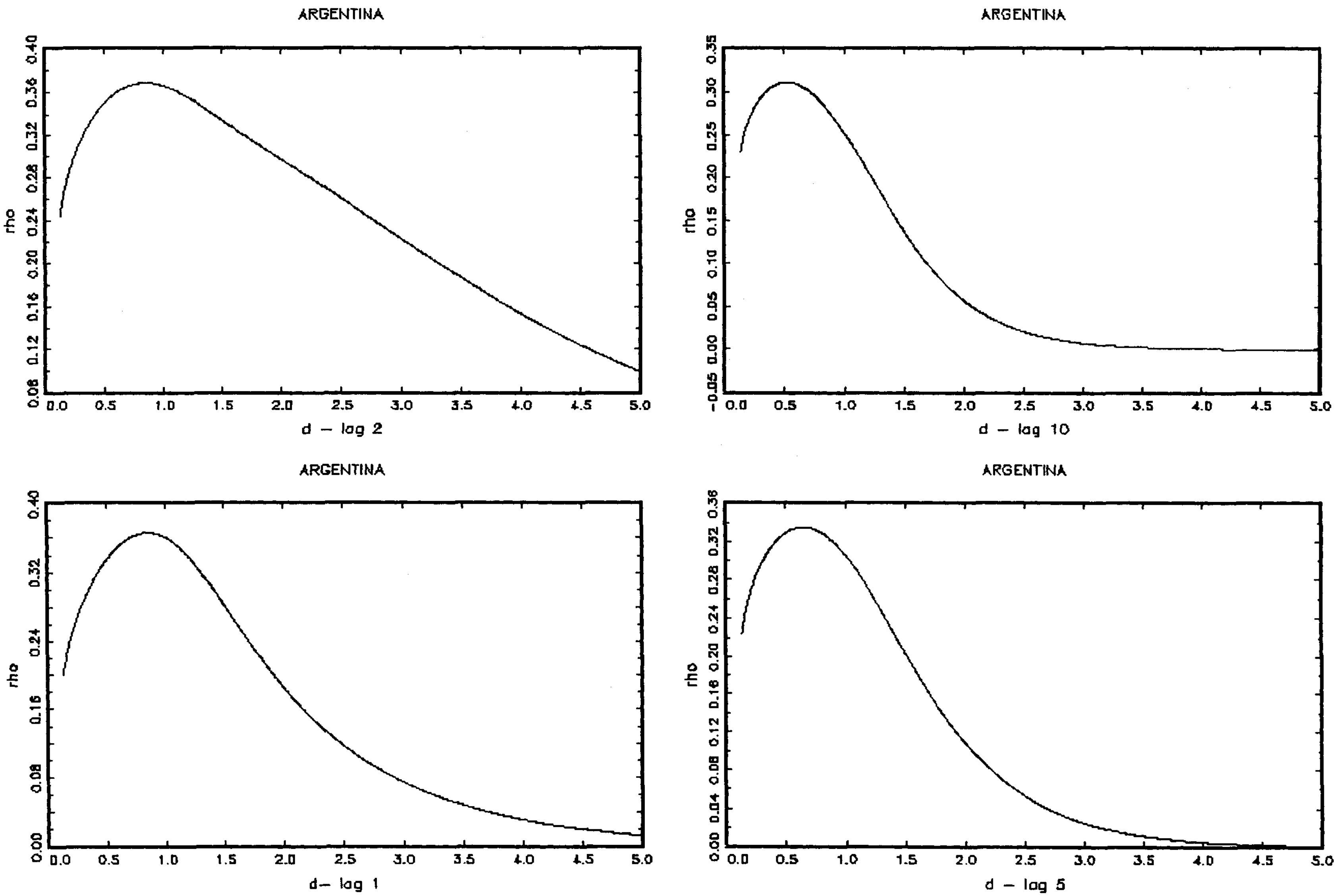
Graph 3.33 – Autocorrelations of lrl**d at lags



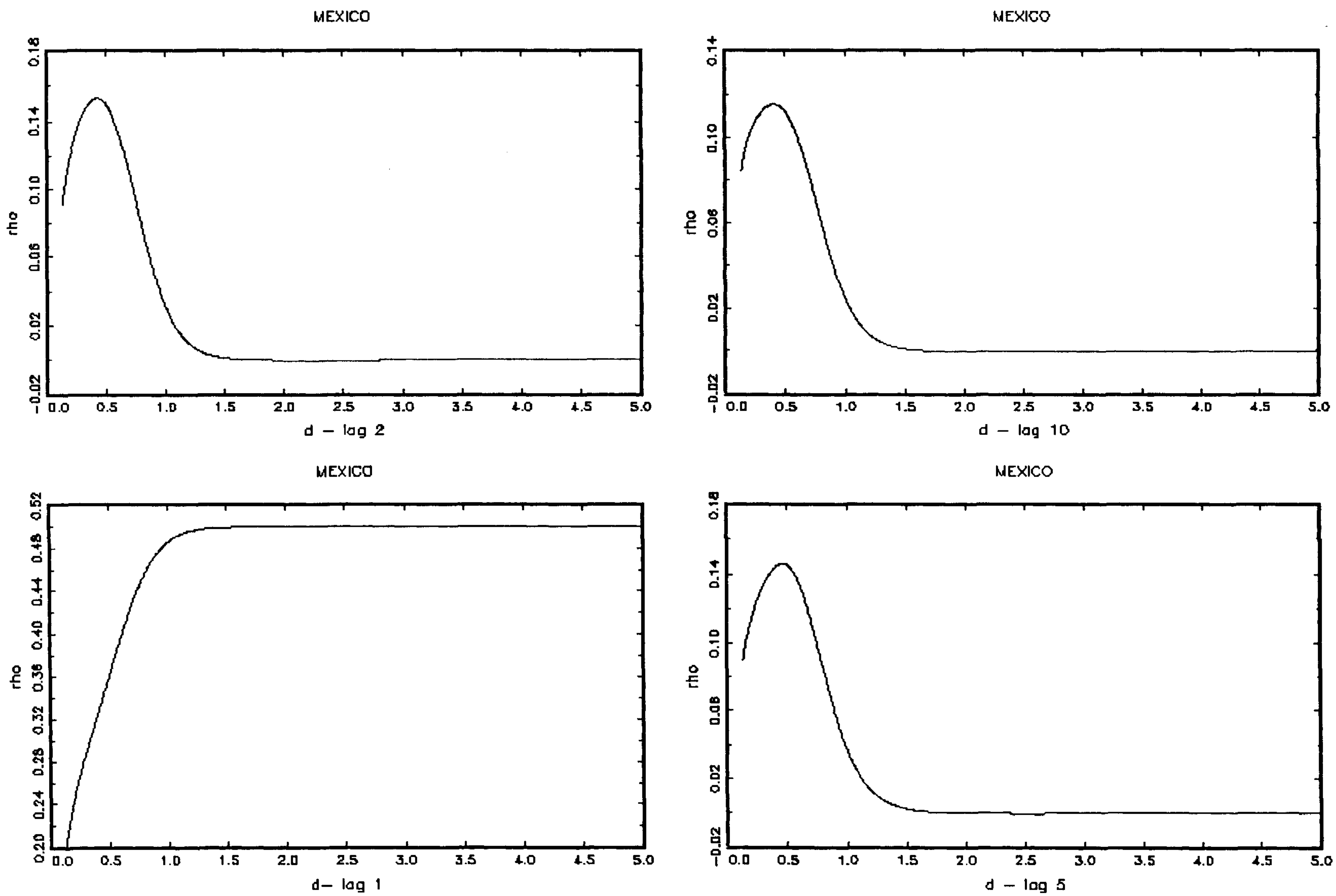
Graph 3.34 – Autocorrelations of lrl**d at lags



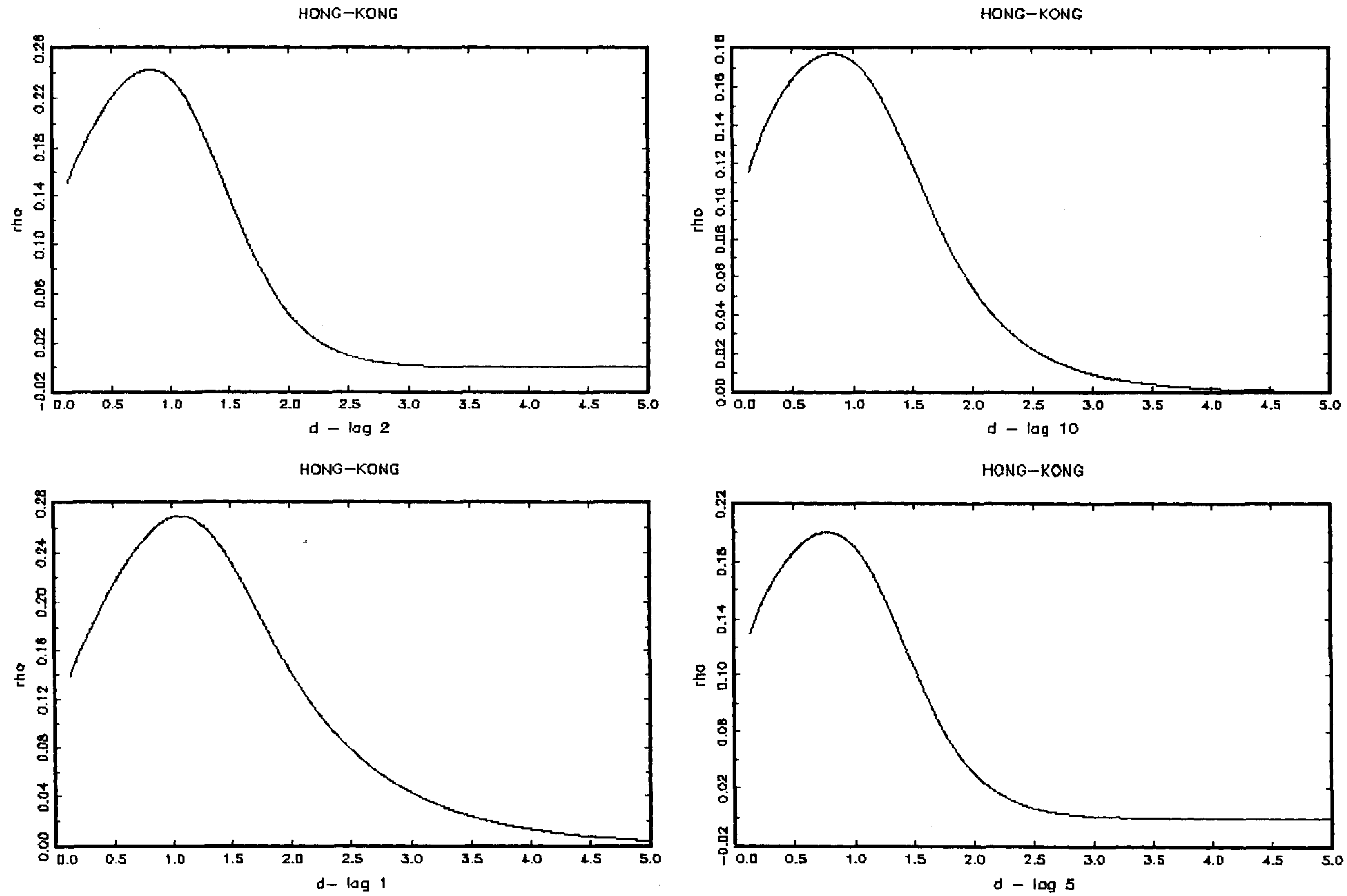
Graph 3.35 – Autocorrelations of |r|**d at lags



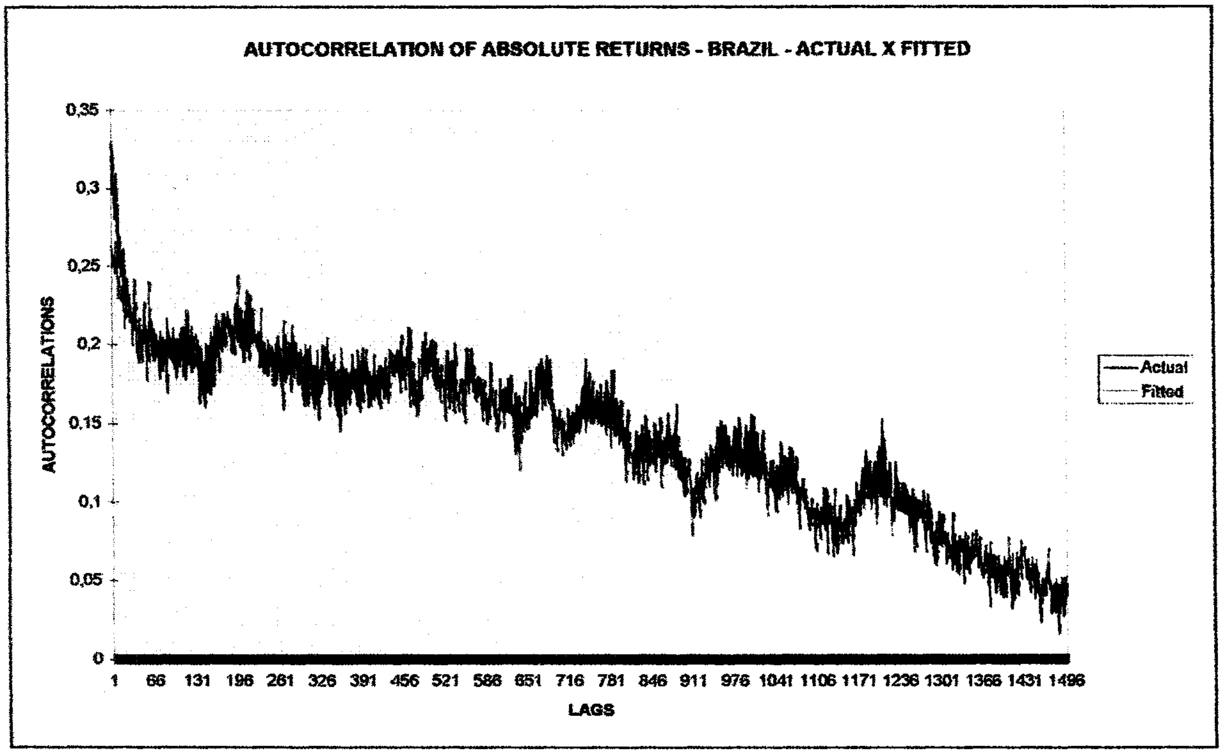
Graph 3.36 – Autocorrelations of |r|**d at lags



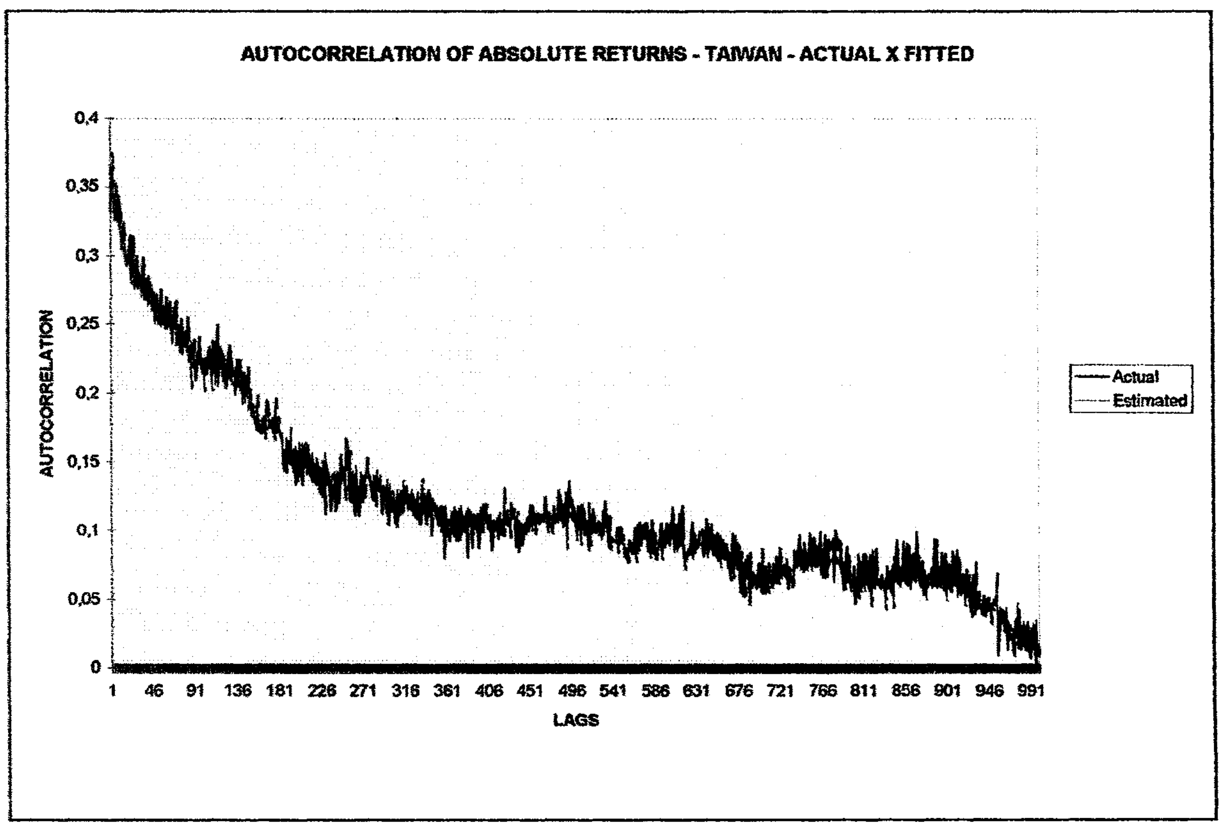
Graph 3.37 – Autocorrelations of Irl**d at lags



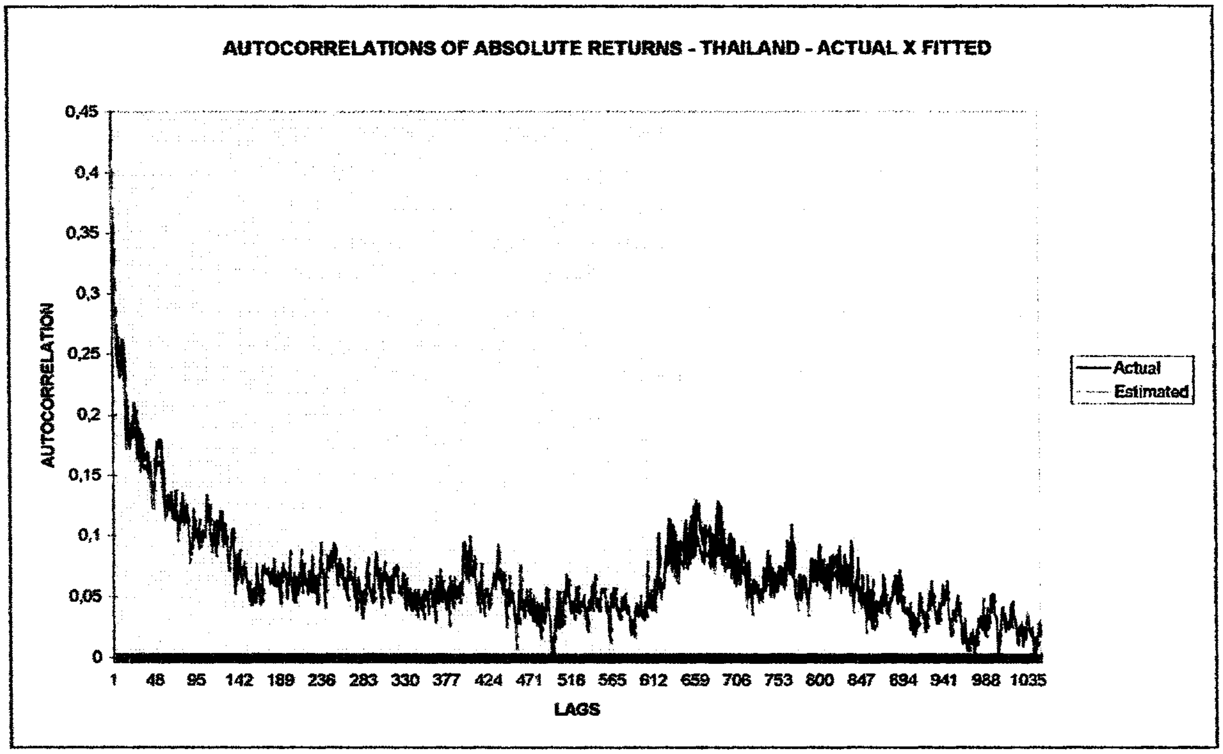
Graph 3.38



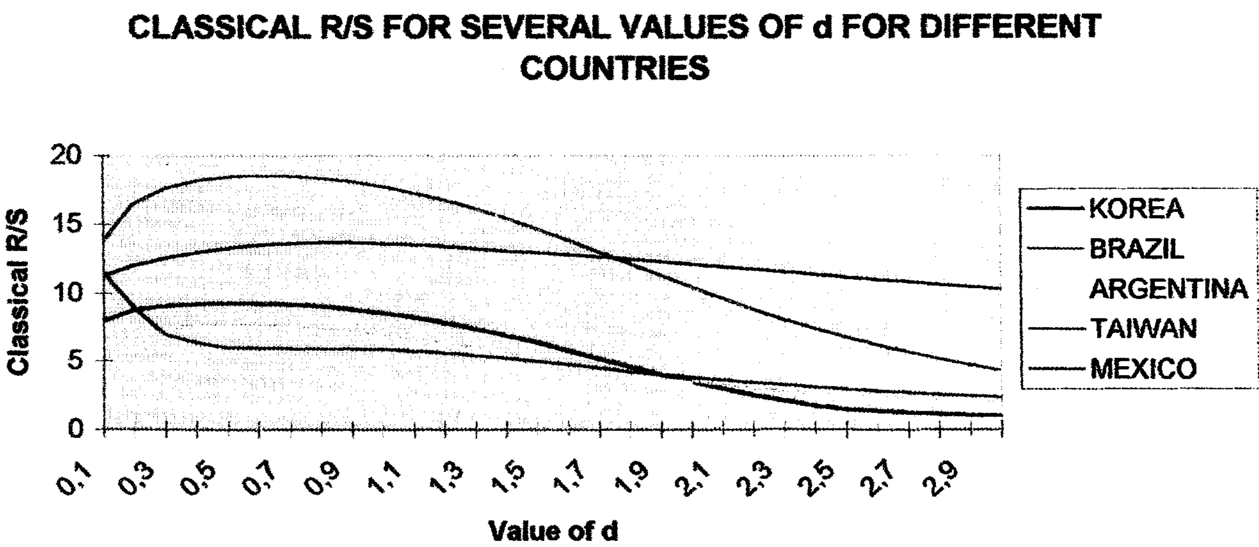
Graph 3.39



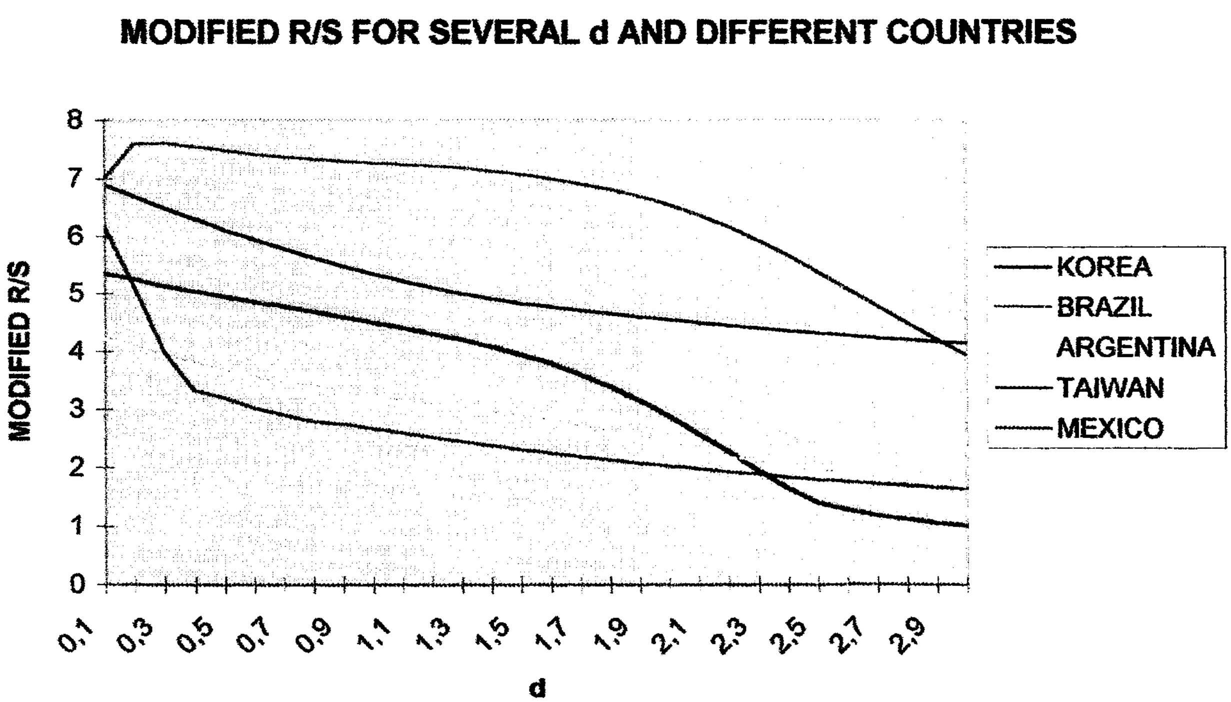
Graph 3.40


AUTOCORRELATIONS OF ABSOLUTE RETURNS - THAILAND - ACTUAL X FITTED
AUTOCORRELATION
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
1 48 95 142 189 236 283 330 377 424 471 518 565 612 659 706 753 800 847 894 941 988 1035
LAGS
Actual
Estimated

Graph 3.41



Graph 3.42



Chapter 4

Modelling Long Memory in the Volatility in the Emerging Markets

4.1 – Introduction

In this chapter we estimate the effects of long memory in the emerging markets economies. We begin by using the first development in this area, the ARCH model. We then proceed to a elaborate model designed to test for the existence of a unit root in the second moment, the FIGARCH model. We then present the conclusions we were able to reach using the data in our sample.

As we have seen in the previous chapter, there is much evidence that for the stock market series under study, the returns are correlated over time. Indeed, we can conclude that large changes tend to follow large changes, though the sign is unpredictable. The modified Box-Pierce test statistic on standardised absolute residuals and squared residuals for up to hundredth order serial correlation, seen below in table 4.1.[1]

Table 4.1

Country	Q(100) – Absolute Returns	Q(100) - Squared Returns
Taiwan	5542.66*	87.00
Thailand	409.47*	49.57
Korea	191.19*	0.61
Malaysia	648.52*	39.09
Hong-Kong	542.75*	40.46
Mexico	168.92*	0.10
Brazil	781.83*	50.11
Argentina	645.21*	45.12

(*) Significant at 95%

As we can see from the table above, the corresponding Q(100) is highly significant at any level of absolute returns, while the same is not true for squared returns. For a precise definition of the Box-Pierce test statistic and a discussion of its applicability in testing for absence of serial correlation see McLeod and Li (1983). Remember from the previous chapter that the autocorrelations and the partial autocorrelations die out fairly slowly. Ding, Granger and Engle (1993) find that there are substantially more correlations between absolute returns than the correlation between returns itself . This is also the case for the returns studied in this thesis. The

[1] The test statistic is a $\chi^2(30)$ = 50.892 at a 99% level and $\chi^2(30)$ = 53.67 at a 99,5% level. The critical values are $\chi^2(50)$ = 75.35 at a 99% level and $\chi^2(50)$ = 78.45 at a 99,5% level. Finally the test statistic is a $\chi^2(100)$ = 135.807 at a 99% level and $\chi^2(100)$ = 140.169 at a 99,5% level.

positive correlations for long lags indicate the existence of long memory in these financial return series. This correlation pattern suggests a predictable structure for volatility that can be exploited for forecasting. To statistically evaluate whether or not there might be a long run predictable structure in the volatility of daily returns, we could also have computed the Ljung and Box (1978) portmanteau statistic for the joint significance of the return autocorrelation for various lag lengths. Under the null hypothesis these statistics should have an asymptotic chi-square distribution. In the case of volatility autocorrelation the test statistics have very small p-values, so we can safely reject the null hypothesis. Table 4.2 shows the Ljung-Box test statistic[2] for no serial correlation up to the end of the 500th day. As suggested by these figures, there is no doubt that a significant serial correlation exists in the volatility of daily returns, even for quite long lags. These features of the data support the noted volatility clustering; however, up to the present, we have not been concerned with the long memory component in the volatility of daily returns. It is worthwhile to attempt to model this persistence so that we can use these forecasts to improve financial decisions.

Table 4.2

Country	Ljung-Box – Absolute Returns
Hong-Kong	935.66*
Brazil	17725.71*
Argentina	1220.75*
Mexico	3036.71*
Taiwan	75638.29*
Thailand	6362.57*
Korea	1089.93*
Malaysia	3043.28*

(*) Significant at 95%

4.2 - Review of some existing models and estimation results

In conventional econometric models, the variance of the disturbance term is assumed to be constant over time. However, the above empirical evidence suggests that this is not the case. In the previous chapters, we have seen that the financial time series in question exhibit periods of unusually large volatility followed by periods of relative tranquility. In such circumstances, the so-called assumption of constant variance (homoskedasticity) no longer seems appropriate. It is fairly simple to imagine situations where we might want to forecast the conditional variance of a series. As asset holders, for instance, we would be interested in forecasts of the rate of return and its

[2] The Ljung-Box Q-statistics is : $Q = T(T+2)\sum_{i=1}^{n} \rho(i) / (T-i)$

variance over the holding period. The unconditional variance (i.e. the long-run forecast of the variance) would be unimportant for anyone intending to buy an asset at t and sell at t+1.

The conditional variance appears to be the key variable. Let ε_t denote a discrete-time real-valued stochastic process. In order to forecast this variable, one possible approach would be to explicitly introduce an independent variable (with respect to ε_t) to help predict volatility. Consider a very simply case in which

$$y_t = \varepsilon_t z_{t-1}$$

where y_t = the variable of interest (returns, for example)

ε_t = a white-noise disturbance term with variance σ^2

z_{t-1} = an IID N(0,1) variable that can be observed at period t-1

If z_t is constant over time, y_t would be a white-noise with constant variance, the variance of y_t conditional on the observable value of z_{t-1} is then[3]

$$Var(y_t | z_{t-1}) = z^2_{t-1}\sigma^2_t$$

The conditional variance of y_t is then dependent on the observed value of z_{t-1}. Since we can observe z_{t-1} we can therefore form the variance of y_t conditionally on the realised value of z_{t-1}. Instead of using this *ad-hoc* variable choice z_{t-1}, the ARCH methodology proposes another way to proceed. If we define E_{t-1} to be the mathematical expectation conditional on the information set available at t-1, then, a discrete-time real-valued stochastic ARCH process $\{\varepsilon_t\}$ is such that $\varepsilon_t = z_t \sigma_t$ with $E_{t-1}(z^2_t) = 0$ and $E_{t-1}(z_t) = 1$ and σ_t is measurable with respect to the time t-1 information set. Note that the ε_t are, in the majority of cases, innovations for the conditional mean of a stochastic process such as $\{y_t\}$, that is, $\varepsilon_t \equiv y_t - E_{t-1}(y_t)$ such that the conditional variance is $\sigma^2_t = E_{t-1}(\varepsilon^2_t)$. We can therefore note that it is possible to model simultaneously the mean and the variance of a series of returns.

It is clear from above that for empirical application, the conditional variance function requires a specific parameterisation. Economic theory does not stipulate information on which

---

[3] $Var(aY_t) = a^2 Var(Y_t)$. In this case, $Var(y_t) = Var(\varepsilon_t z_{t-1}) = z^2_{t-1} Var(\varepsilon_t) = z^2_{t-1}\sigma^2$

expectations of volatility are formed. This means that various functions of past information on a process such as $\{y_t\}$ can be suggested.

4.2.1 The Autoregressive Conditional Heteroskedasticity (ARCH)

As outlined above, an ARCH process can be defined in a large variety of contexts. One possibility is to define it in terms of the distribution of errors of a dynamic linear regression model. The dependent variable is assumed to be generated by

$$y_t = x_t^{'}\xi + \varepsilon_t \quad (4.1)$$

where y_t the dependent variable and x_t is a k x 1 vector of exogenous variables, which may include lagged values of the dependent variable, and ξ is a k x 1 vector of regression parameters. The ARCH model characterises the distribution of the stochastic error ε_t conditional on the realised values of the set of variables ψ_{t-1} = { y_{t-1} , x_{t-1} , y_{t-2} , x_{t-2} , ...}. More specifically, Engle's (1982) original ARCH model assumes

$$\varepsilon_t | \psi_{t-1} \sim N(0, \sigma_t^2) \quad (4.2)$$

where

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_q \varepsilon_{t-q}^2 \quad (4.3)$$

with $\alpha_0 > 0$ and $\alpha_i \geq 0$, i = 1,.....,q , to ensure that the conditional variance is positive. We should, however, notice that as $\varepsilon_{t-i} = y_{t-i} - x_{t-i}^{'}\xi$, i = 1, 2, ... q, the conditional variance (σ_t^2) is clearly a function of the elements of ψ_{t-1}.

The distinguishing feature of the model (4.2) and (4.3) is not simply that the conditional variance (σ_t^2) is a function of the conditioning set ψ_{t-1}, but rather it is that the particular functional form is specified. Episodes of volatility are generally characterised by the clustering of large shocks to the dependent variable. The conditional variance function (4.3) is formulated to mimic this phenomenon. In the regression model, a large shock is represented by a large deviation of y_t from its conditional mean $x_t^{'}\xi$ or equivalently, a large positive or negative value of ε_t. In the ARCH regression model, the variance of the current error ε_t, conditional on the realised values of the lagged errors ε_{t-i}, i = 1 , ..., q, is an increasing function of the magnitude of the lagged errors, irrespective of their signs. Hence, large errors of either sign tend to be followed by a large error of either sign. And similarly, small errors of either sign *tend* to be followed by a small error of either sign. The order of the lag q determines the length of time for which a shock persists in conditioning the variance of subsequent errors. The larger the value of q, the longer the episodes of volatility will tend to be.[4]

A linear function of lagged squared errors, of course, is not the only conditional variance function that would produce clustering of large deviations. Actually any monotonic increasing function of the absolute values of the lagged errors would lead to this clustering. However, since variance is the expected measure of the squared deviations, it seems natural to use such a linear combination of lagged squared errors to lead with this trend in variance.[5]

For our present purposes, this means that the ARCH(q) parameterisation models the conditional variance as a weighted average of past squared forecast errors,

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \dots\dots\dots + \alpha_q\varepsilon_{t-q}^2 \qquad \text{or} \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 \qquad (4.4)$$

It follows that if $\alpha_0 = 0$ and $\alpha_i = 1/q$ then (4.4) is precisely the conditional variance of the sample variance of the most recent q returns. As Engle (1982) noted, the great advantage of the ARCH specification is that coefficients can be estimated from historical data and the resulting statistical model can be used to forecast future volatility and improve estimates.

Assuming the disturbances are normally distributed (or that the standardised residuals, $z_t = \varepsilon_t\sigma_t^{-1}$), the conditional log-likelihood function for the ARCH class of models can be written,

$$LogL(\theta;\varepsilon_1,\dots\dots,\varepsilon_T \,|\, \psi_0) = \sum_{t=1}^{T} -\frac{1}{2}\left[\log(2\pi) + \log(\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_t^2}\right] \qquad (4.5)$$

for which the initial conditions, ψ_0, used to start the recursions for the conditional mean and the variance function.[6] The method of estimation was based on the Berndt, Hall, Hall and Hausman

[4] Actually we could think as y_t as being the logarithm of the return of a given stock at t and $x_t'\xi$ as its conditional mean. In this context ε_t is an innovation for the conditional mean of the return.

[5] Consider the properties of $\{\varepsilon_t\}$ sequence. It is easy to show that this sequence has a mean of zero and is uncorrelated, that is, $E(\varepsilon_t \mid \psi_{t-1}) = 0$ and $Var(\varepsilon_t \mid \psi_{t-1}) = \sigma_t^2$.

[6] Actually to start up these recursions we need the pre-sample values for ε_0^2 and σ_0^2. A natural choice is given by the sample analogue $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2$

(BHHH) algorithm that is relatively simple although it sometimes requires more computation and iterations than necessary.[7]

The results found for the countries under study can be found in Table 4.3 below. We first estimated an ARCH(1) model and then an ARCH(4) model. The model estimated is then:

$$y_t = \mu + \varepsilon_t$$

$$\varepsilon_t = z_t \sigma_t \qquad z_t = \varepsilon_t \sigma_t^{-1} \text{ i.i.d. N(0,1)} \qquad (4.6)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots\ldots\ldots + \alpha_q \varepsilon_{t-q}^2$$

Following standard practice, we transformed the price index into a continuously compounded capital gains series, where the variable y_t is defined as the logarithm difference between the stock price at t and t – 1 for the stock price index in each of these countries. The parameter μ is the conditional mean of the returns and ε_t is an innovation for the conditional mean of the returns such that the conditional variance is $\sigma_t^2 = E_{t-1}(\varepsilon_t^2)$.

We can see all the parameters are very significant in the below models. In table 4.3 we have the results for the ARCH(1). In order to allow for longer effects of shocks we increased the order of the lag q. We know that the order of the lag determines the length of time for which a shock persists in conditioning the variance of subsequent error. As we are interested in examining the persistence of volatility it seems plausible to allow for a longer lag[8]. Indeed, a visual inspection of table 4.4 suggests that all coefficients are still highly significant for all of the countries under consideration. This means that estimating a model like ARCH(1) imposes restrictions that can lead to misspecification. In this case, as all the

[7] For a description of the BHHH algorithm see Greene (1993). It is a very effective class of algorithms that has been developed that eliminates second derivatives altogether and has excellent convergence properties, even for ill-behaved problems. Sometimes there are also called **quasi-Newton methods**.
[8] For some authors, it is useful to note that the ARCH(q) specification as (3.4) can be rewritten as an AR(q) process for ε_t^2, $[1-\alpha(L)]\varepsilon_t^2 = \alpha_0 + v_t$ in which L denotes the lag operator and $v_t \equiv [\varepsilon_t^2 - \sigma_t^2]$ is an unforecastable innovation to ε_t^2.

Table 4.3 - ARCH(1) Results*

Country	μ	α_0	α_1	Log L
Mexico	0.000438 (7.21)	0.0000378535 (58.10)	0.6321 (17.32)	17132.83
Taiwan	0.000521 (8.74)	0.000032892 (55.22)	0.4040 (18.34)	25901.40
Thailand	0.000741 (4.58)	0.0000130255 (99.11)	0.9332 (38.60)	23488.15
Korea	0.000855 (6.54)	0.0000172309 (99.85)	0.3187 (15.72)	25595.83
Malaysia	0.000121 (8.21)	0.0000248355 (69.18)	0.4200 (21.15)	22832.96
Brazil	0.000662 (7.45)	0.0000848185 (66.87)	0.5778 (35.02)	27995.33
Hong-Kong	0.000225 (4.54)	0.0000366787 (140.71)	0.3507 (25.80)	24088.75
Argentina	0.000748 (8.47)	0.0001355474 (62.71)	0.9471 (72.37)	6346.55

(*) All coefficients are significant at 95% and 99%.

Table 4.4 - ARCH(4) Results*

Country	μ	α_0	α_1	α_2	α_3	α_4	Log L
Mexico	0.000421 (6.21)	-0.0000013973 (-1134.61)	0.56 (10.15)	0.1836 (1153.2)	0.1625 (17.85)	0.3695 (1086.0)	16410.12
Taiwan	0.000221 (5.74)	0.000012914 (26.61)	0.1383 (10.31)	0.2030 (12.02)	0.2279 (12.87)	0.2026 (13.54)	26493.79
Thailand	0.000341 (6.45)	0.0000049434 (53.44)	0.5417 (28.95)	0.2890 (24.83)	0.1450 (11.06)	0.1891 (15.71)	24113.18
Korea	0.000345 (9.45)	0.0000075755 (28.88)	0.2434 (15.59)	0.2494 (17.57)	0.1551 (9.36)	0.2003 (13.93)	25862.57
Malaysia	0.000321 (9.12)	0.0000145581 (39.68)	0.2720 (14.93)	0.2112 (17.87)	0.1426 (12.90)	0.0595 (5.55)	23095.62
Brazil	0.000542 (8.32)	0.000020556 (30.54)	0.3263 (19.29)	0.2223 (16.43)	0.2977 (21.69)	0.2171 (14.91)	28868.16
Hong-Kong	0.000125 (5.54)	0.0000154753 (42.48)	0.25 (27.66)	0.2817 (19.21)	0.1810 (15.03)	0.1246 (9.88)	24495.37
Argentina	0.000923 (6.31)	0.00036583 (12.65)	0.3366 (13.37)	0.3293 (11.83)	0.1965 (7.20)	0.1690 (5.09)	6683.99

(*) All coefficients are significant at 95% and 99%.

coefficients are significant, there is a persistence that is described by this high order lag. We could probably continue to include more lags to take account this persistence in volatility. By using the estimated log-likelihood values, a nested test can easily be constructed to test ARCH(4) against ARCH(1). Let L_0 be the likelihood value under the null hypothesis that the true model is ARCH(1) (restricted model) and let L_1 be the likelihood under the alternative hypothesis that the true model is ARCH(4) (unrestricted). Then 2(log L_1 - log L_0) should have a chi-square distribution with 3 degrees of freedom when the null hypothesis is true. In our case, 2(log L_1 - log L_0) lies between

533,48 (Korea) and 3358,58 (Mexico), which is well beyond the critical value for any reasonable value. Hence, we can reject the hypothesis that the data are generated by an ARCH(1) in favour of ARCH(4).

4.2.2 Generalised Autoregressive Conditional Heteroskedasticity (GARCH)

There is a cost involved in estimating more parameters. In the first empirical applications of ARCH to the relationship between the level and the volatility of inflation, Engle (1982, 1983) found that a large lag q was required in the conditional variance function. This necessitated estimating a large number of parameters subject to inequality constraints. To reduce the computational burden, Bollerslev (1986) proposed an extension of the conditional variance function (4.4), which he termed generalised ARCH, or GARCH. This model has proven very useful in empirical work. GARCH was also proposed independently by Taylor (1986), who used a different acronym.

They suggested that the conditional variance be specified as

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots\ldots\ldots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \ldots\ldots\ldots + \beta_p \sigma_{t-p}^2 \qquad (4.7)$$

The main purpose of GARCH(p,q) is to capture the long memory property of the conditional variance process.[9] Taylor (1986) and Schwert (1989) proposed a similar model where the conditional standard deviation function instead of the conditional variance is defined as

$$\sigma_t = \alpha_0 + \alpha_1 |\varepsilon_{t-1}| + \ldots\ldots\ldots + \alpha_q |\varepsilon_{t-q}| + \beta_1 \sigma_{t-1} + \ldots\ldots\ldots + \beta_p \sigma_{t-p} \qquad (4.8)$$

At first glance, one might think that it would be better to use Taylor/Schwert's specification since this model is expressed in terms of absolute returns rather than squared returns.[10] However, as Ding, Granger and Engle (1993) have argued, this is not true when the model is highly non-linear one. Both models were estimated for all the countries under study. Following the example of other studies, such as Bollerslev et all (1986), we estimated a GARCH(1,1) process:

$$y_t = \mu + \theta_1 y_{t-1} + \varepsilon_t$$

$$\varepsilon_t = z_t \sigma_t \qquad z_t = \varepsilon_t \sigma_t^{-1} \text{ i.i.d. } N(0,1) \qquad (4.9)$$

[9] Actually there are some inequality restrictions to ensure that the conditional variance is strictly positive as $\alpha_0 > 0$, $\alpha_i \geq 0$ (for i = 1,....,q) and $\beta_i \geq 0$ (for i = 1,....,p).

[10] This is because, as noted before, there are a lot more of correlation among the absolute returns than the square of returns.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

where y_t stands for the stock index price for each market. The intention is that GARCH can parsimoniously represent a high order ARCH process, in the sense that there are fewer parameters to be estimated, while conserving the property of persistence in volatility[11].

The results for the GARCH(1,1) model are shown below in table 4.5. There are several interesting things to note. All of the coefficients concerning the GARCH parameters are highly significant at both 95% and 99% levels. This confirms what we already suspected, that is, there is indeed a persistence that can be described by an ARCH model with very high order lags. We probably could have included more lags to take on account this persistence in volatility. However, there is a cost involved in estimating more parameters. By estimating a GARCH(1,1)[12] we have comparatively fewer parameters to estimate while still managing to account for this persistence. We stopped here in estimating the GARCH model. Actually in applied work, it has been frequently demonstrated that the GARCH(1,1) process is able to represent a majority of financial time series. A data set which requires a model of order higher than GARCH(1,2) or GARCH(2,1) is very rare. For our present purposes we are satisfied by these results. The magnitudes of the parameters are also similar to those found in other studies of returns, such as Ding, Granger and Engle (1993) and McCurdy and Michaud (1996), among others. Ding, Granger and Engle reported results for the US market. It is clear that the coefficients for the US market are all significant. However, we can also see that the magnitudes of the coefficients are lower than those for emerging markets, with the exception of the β_1 term that relates the conditional variance to its lagged value. [13]

By using the log-likelihood values estimated, a nested test can easily be constructed to test GARCH(1,1) against ARCH(4). Let L_0 be the likelihood value under the null hypothesis that the true model is ARCH(1) (the restricted model) and let L_1 be the likelihood under the alternative

---

[11] It is easy to show that the GARCH(p,q) can be expressed as $\sigma_t^2 = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2$ with $\beta(L)$ and $\alpha(L)$ are polynomials in the lag operator. If the roots of 1 - $\beta(L)$ lie outside the unit circle, we can rewrite (4.7) as $\sigma_t^2 = \alpha_0^* + \sum_{i=1}^{\infty} \varepsilon_{t-i}^2$. Hence a convenient choice of parameters allows us to interpret the GARCH(p,q) as an infinite ARCH(q) with few parameters. Nelson and Cao (1992) have stated weaker conditions for the variance to be strictly positive. McCurdy and Michaud (1996) point out also that a GARCH(p,q) can also be written in an ARMA(m,p) representation, with m = max {p,q} as $\left[1-\alpha(L)-\beta(L)\right]\varepsilon_t^2 = \alpha_0 + \left[1-\beta(L)\right]v_t$, $\beta(L)$ is a polynomial in the lag operator of order p.

[12] There has also been many other extensions or alternative specifications of the very successful ARCH model. However, several authors have noted some shortcomings of the ARCH specification - particularly for forecasting (see Hamilton and Rusnel (1994)).

[13] For an extensive review of literature see Bollerslev , Chou and Kroner (1992)

that the true model is GARCH(1,1) (the unrestricted model). Then 2(log L_1 - log L_0) should have a chi-square distribution with 3 degrees of freedom when the null hypothesis is true. In this case, 2(log L_1 - log L_0) ranges from 136,78 (Argentina) to 2803.98 (Mexico), which is far beyond the critical value for any reasonable value. Hence we can reject the hypothesis that the data is generated by the ARCH(4) model in favour of the GARCH(1,1).

Table 4.5 - GARCH(1,1) Results

Country	μ	θ_1	α_0	α_1	β_1	Log L
Mexico	0.00115** (2.33)	0.32* (20.43)	0.00015* (31.44)	0.317* (23.24)	0.6012 (3.14)	17512.11
Taiwan	0.00128 (1.48)	0.43* (21.14)	0.00021* (25.42)	0.221* (12.25)	0.722* (142.21)	26659.53
Thailand	0.0214 (1.95)	0.46* (19.22)	0.00209* (20.54)	0.2214* (31.01)	0.714* (128.12)	24275.71
Korea	0.0221 (1.41)	0.47* (11.57)	0.000241* (20.24)	0.224* (20.14)	0.721* (102.46)	25994.68
Malaysia	0.0212 (1.55)	0.33* (30.30)	0.000311* (14.44)	0.2304* (18.14)	0.721* (54.20)	23178.63
Brazil	0.0314* (13.12)	0.52* (18.47)	0.00009* (31.41)	0.3617* (39.55)	0.6504* (165.12)	28998.03
Hong-Kong	0.000172* (3.41)	0.24* (30.42)	0.00001* (34.12)	0.2817* (18.41)	0.7001* (22.01)	24675.23
Argentina	0.0217* (3.95)	0.45* (47.15)	0.0005* (11.69)	0.2002* (12.12)	0.841* (65.34)	6742.38
USA	0.0000438* (7.20)	0.144* (18.40)	0.0000008* (12.50)	0.091* (50.7)	0.906* (43.40)	56822.0

(*) Coefficients significants at 95% and 99%. (**) Coefficients significant at 95%

We will now examine the Taylor/ Schwert GARCH(1,1) model, where

$$y_t = \mu + \theta_1 y_{t-1} + \varepsilon_t$$

$$\varepsilon_t = z_t \sigma_t \qquad z_t = \varepsilon_t \sigma_t^{-1} \text{ i.i.d. N(0,1)} \qquad (4.10)$$

$$\sigma_t = \alpha_0 + \alpha_1 |\varepsilon_{t-1}| + \beta_1 \sigma_{t-1}$$

Our results for the Taylor/Schwert model can be found in table 4.6 below. The parameters that describe the conditional variance are all highly significant at both 95% and 99% levels. Nevertheless, the log-likelihood value for Bollerslev's GARCH is significantly higher than that of Taylor/Schwert model. The same result was found in Ding, Granger and Engle (1993). The results for US markets shown in Table 4.6 come from this study. The magnitudes of all coefficients are lower than those found for emerging markets except for the β_1 term that relates the conditional variance to its lagged value. Comparing tables 4.5 and 4.6 we can see that the magnitudes of coefficients vary considerably from one situation from another. For some countries like Thailand and USA this variation is not very large, but in other cases such as Mexico, Taiwan and Argentina there is more variation. We can see that the log likelihood value for GARCH is larger than that of the Taylor/Schwert model for all countries except for Brazil, a finding that is consistent with Ding, Granger and Engle's results for the United States.

Table 4.6 - Taylor/Schwert Results

Country	μ	θ_1	α_0	α_1	β_1	Log L
Mexico	0.00102* (2.37)	0.32* (18.31)	0.00477* (58.32)	0.2158* (24.11)	0.7242* (101.21)	16255.21
Taiwan	0.00110 (1.41)	0.33* (26.42)	0.00748* (152.14)	0.2150* (20.12)	0.7100* (103.12)	26451.12
Thailand	0.00218 (1.34)	0.40* (28.14)	0.000441* (56.44)	0.2565* (42.22)	0.7102* (170.24)	24250.10
Korea	0.000200 (1.21)	0.37* (9.42)	0.000514* (75.10)	0.2414* (20.56)	0.7411* (135.14)	25920.14
Malaysia	0.00201 (1.44)	0.33* (14.21)	0.00408* (71.01)	0.2809* (20.10)	0.7114* (78.01)	23123.12
Brazil	0.00330* (12.12)	0.42* (19.32)	0.000657* (78.15)	0.2144* (39.99)	0.8102* (214.1)	29120.01
Hong-Kong	0.000195* (4.62)	0.39* (27.54)	0.00806* (74.54)	0.2511* (74.32)	0.7145* (280.20)	24554.54
Argentina	0.000417* (2.54)	0.41* (39.21)	0.00847* (49.21)	0.3142* (15.41)	0.64521 (45.21)	6677.27
USA	0.0004* (7.00)	0.139* (19.60)	0.000096* (12.60)	0.104* (67.00)	0.913* (517.00)	56776.00

(*) Coefficients significant at 95% and 99%.

One interesting fact that is apparent from a visual inspection of tables 4.5 and 4.6 is that the sums of α_1 and β_1 in equation (4.9) are very close to one. These range from 0.9182 to 1.02, and that a conventional Likelihood Test (LR) could be then performed to test the hypothesis that $\alpha_1 + \beta_1 = 1$. [14] For our present purposes, we have denoted the null hypothesis as the one associated with the GARCH model, and the alternative is that the model is IGARCH, where we impose the restriction $\alpha_1 + \beta_1 = 1$. Table 4.8 shows the maximum log likelihood value for each of the cases, the unrestricted (L_1 = GARCH) and the restricted (L_0 = IGARCH), in order to apply the

[14] If the restriction, $c(\theta) = 0$ is valid, imposing it should not lead to a large reduction in the log-likelihood function. Therefore, we can base the test on the difference log L - log L_R, where L is the value of the likelihood function at the unconstrained value of θ and L_R is the value of the likelihood function at the restricted value. The LR test, the Wald test and the Lagrange Multiplier test (LM) are a trio of testing procedures that can be applied in the context of maximum likelihood estimation. Asymptotically they are equivalent and we have then chosen to calculate the LR test, what should not be a problem as our samples are very big for each country.

LR test. In this case 2(log L_1 - log L_0) has a χ^2 distribution with 2 degrees of freedom when the null hypothesis is true.

In table 4.8 we see that GARCH (1,1) is indeed rejected in favour of IGARCH (1,1) for most of the countries in our sample. The exceptions are Korea and Argentina. This reinforces the effect of persistence

Table 4.7 - IGARCH Results

Country	μ	θ_1	α_0	α^{**}_1	Log L
Mexico	0.00105* (2.17)	0.39* (12.32)	0.00014* (541.23)	0.98* (472.05)	16601.56
Taiwan	0.00100 (1.04)	0.30* (20.12)	0.000019* (24.12)	0.2414* (31.99)	26646.10
Thailand	0.00218 (1.75)	0.65* (19.12)	0.000021* (16.44)	0.2401* (40.11)	24262.49
Korea	0.000175 (0.48)	0.47* (9.15)	0.0000110* (54.67)	0.2401* (65.12)	25994.12
Malaysia	0.00100 (1.31)	0.43* (27.30)	0.0000147* (20.55)	0.2470* (17.10)	23178.63
Brazil	0.00305* (13.20)	0.42* (29.40)	0.0000221* (31.61)	0.2104* (28.44)	28944.10
Hong-Kong	0.000187* (4.41)	0.41* (12.48)	0.0000140* (39.64)	0.2018* (34.58)	22501.37
Argentina	0.0114* (3.94)	0.51* (29.11)	0.0000154* (21.34)	0.2014* (19.32)	6741.75
USA	0.000457* (7.20)	0.179* (15.19)	0.0000034* (4.24)	0.3650* (7.15)	67212.00

(*) Coefficients significant at 95% and 99%. $\alpha_1^{**} = \alpha_1 + \beta_1$

in variance that we suspected when we dropped the ARCH specification in favour of GARCH. Engle and Bollerslev (1986) were the first to consider GARCH processes with $\alpha_1 + \beta_1 = 1$ as a distinct class of models, which they termed Integrated GARCH (IGARCH). They pointed out the similarity between IGARCH processes and processes that are integrated in the mean. In an IGARCH model a current shock persists indefinitely in conditioning the future variances. That is, from a forecasting perspective the difference between the covariance-stationary GARCH and the IGARCH model provides a natural analogue to the difference between I(0) and I(1) processes for the conditional mean. The IGARCH model is important because there is a remarkable empirical regularity found in applied work that these estimated two coefficients of a GARCH conditional variance sum close to one. We do show the results obtained by testing for a IGARCH model in Table 4.7. We can see that the coefficients reinforce what we have found previously. This result is also consistent with a variety of studies of stock markets in developed economies. In table 4.7 we report findings provided by the study of Bollerslev and Mikkelsen (1996). There we can see that all coefficients are significant and that, except for the term containing the sum of $\alpha_1 + \beta_1 = 1$, the coefficients for the emerging economies are larger than those found for developed economies.

Table 4.8 - Log Likelihood for both the IGARCH and GARCH models

Country	IGARCH (log L_0)	GARCH (log L_1)	2(log L_1 - log L_0)
Mexico	16601.56	17512.11	1821.10*
Taiwan	26646.10	26659.53	26.86*
Thailand	24262.49	24275.51	23.04*
Korea	25991.01	25994.12	6.22
Malaysia	23151.17	23178.63	54.92*
Brazil	28944.10	28998.03	107.86*
Hong-Kong	22501.37	24675.23	4347.72*
Argentina	6741.75	6742.38	1.26

(*) Reject the Null

For example, Baillie and Bollerslev (1989) estimated GARCH(1, 1) models for six U.S. exchange rates and found $\alpha_1 + \beta_1$ ranging between 0.94 and 0.99 for the six series.[15] The consistency of this finding lead Lamoureux and Lastrapes (1990b) to argue that large persistence may actually represent misspecification of the variance result from structural change in the unconditional variance of the process, represented by changes in α_0 in (4.9). A discrete change in the unconditional variance of a process produces clusters of large and small deviations that may show up as persistence in a fitted ARCH model. To illustrate this possibility, Lamoureux and Lastrapes used 17 years of daily returns data for stocks of 30 randomly selected companies and estimated GARCH(1,1) models holding α_0 constant and allowing α_0 to change discretely over sub-periods of the sample. For the restricted model, in which α_0 is constant, the average estimate of $\alpha_1 + \beta_1$ for 30 companies was 0.978, while for the unrestricted model, in which α_0 is allowed to change, the average estimate fell to 0.817. Lamoureux and Lastrapes a present Monte Carlo evidence that demonstrated that the MLE of $\alpha_1 + \beta_1$ has a large positive bias when changes in the unconditional variance are ignored. Taylor (1986) estimated GARCH(1,1) models for 40 different financial time series. The results show that, for all but six of the 40 series, the estimated value of $\alpha_1 + \beta_1$ is greater than or equal to 0.97. In Ding, Granger and Engle (1993), the estimated value of $\alpha_1 + \beta_1$ for daily S&P 500 returns is equal to 0.91.This regularity is widely considered to be a characteristic of volatility persistence, which we can see in table 4.5 above.

[15] Bollerslev and Engle (1989) considered multivariate IGARCH processes and defined a concept of co-integration in variance that they termed co- persistence. A set of univariate IGARCH processes are co-persistent if there exists a linear combination of the processes which is not integrated in variance. Nelson (1990) has cautioned that drawing an analogy with processes that are integrated in the mean, however may be somewhat misleading. Nelson (1990a) demonstrated that although IGARCH models are not weakly stationary, because they have infinite variances, they can be strongly stationary. Processes that are integrated in the mean are not stationary in any sense.

It appears that the series under analysis does indeed show long-run persistence in variance. Ding, Granger and Engle (1993) also argue that there is a stylised fact that should be taken on account. The asymmetric response of volatility to positive and negative 'shocks' is an established fact in the finance literature as the leverage effect of the stock market returns (see Black (1976), for example). This suggests that stock returns are negatively correlated with changes in return volatility, i.e. that volatility tends to rise in response to 'bad news' (lower returns than expected) and to fall in response to 'good news' (higher returns than expected). Nelson (1989) was the first to formally model this potential asymmetry. Several other empirical studies have shown that it is crucial to include the asymmetric term in financial time series model.[16] Ding, Granger and Engle (1993) also found that the empirical autocorrelation of alternative definitions of volatility $| y_t |^{\delta}$, were strongest for the power δ $1 < \delta < 1.25$, for most series we have seen this in and the previous chapter. This fact motivated them to propose a Box-Cox transformation so that the coefficient δ could be estimated rather than imposed as unity.

Therefore Ding, Granger and Engle (1993) proposed the asymmetric Power ARCH (named A-PARCH) with [17]the following specification:

$$\sigma_t^{\delta} = \alpha_0 + \sum_{i=1}^{q} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^{\delta} + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^{\delta} \quad (4.11)$$

There is another interesting property with the A-PARCH. This specification nests at least seven other members of the ARCH family. This new model is estimated for the countries under study and we used MLE and the BHHH algorithm. The model estimated is then:

$$y_t = \mu + \theta_1 y_{t-1} + \varepsilon_t$$

$$\varepsilon_t = z_t \sigma_t \qquad z_t = \varepsilon_t \sigma_t^{-1} \text{ i.i.d. N(0,1)} \qquad (4.12)$$

$$\sigma_t^{\delta} = \alpha_0 + \alpha_1 (|\varepsilon_{t-1}| - \gamma_1 \varepsilon_{t-1})^{\delta} + \beta_1 \sigma_{t-1}^{\delta}$$

The results are shown in table 4.9. Several observations[18] can be drawn from these results. First, the estimated δ is different from 1 with the Taylor/Schwert model and different from 2 with GARCH for all countries. However, for the United States this coefficient is much larger than those found for the emerging economies. The same observation is valid for the β_1 term that relates the conditional variance to its lagged value, which is higher for the US data. The t-statistic

[16] See Nelson (1989), Glosten, Jaganathan and Runkle (1989) and Engle and Ng (1992).
[17] Another possibility would be to introduce the E-GARCH (q,p) model as proposed by Nelson (1991).
[18] If $\delta = 2$, $\gamma_i = 0$,$\forall$ q and $\beta_j = 0$, $\forall$ p we have the ARCH model. If $\delta = 2$, $\gamma_i = 0$.$\forall$ q we have the GARCH model. If $\delta = 1$, $\gamma_i = 0$,$\forall$ q , we have the Taylor/Schwert model and so on. We should emphasise that other possible specifications to take on account this asymmetry is the Exponential GARCH by Nelson (1991),

for the asymmetric term (α_1) is highly significant, implying that the leverage effect does exist in all the countries. This result seems to indicate that asymmetries tend to be different between US markets and emerging markets, which seems plausible.

Actually by using the estimated log-likelihood values, a nested test can easily be performed against either Bollerslev's GARCH or the Taylor/Schwert model. Let L_0 be the log-likelihood value under the null hypothesis that the true model is GARCH or Taylor/Schwert and let L_1 be the log-likelihood value under the alternative that the true model is A-PARCH. Then, 2 (log L_1 - log L_0) should have a $\chi^2(2)$ distribution when the null hypothesis is true. For all the countries the calculated statistic test is beyond the critical value at any reasonable value. Hence we can reject the hypothesis the data is generated by either GARCH or Taylor/Schwert in favour of the more flexible A-PARCH specification. This result was reported in Ding, Granger and Engle (1993). Hence, we can reject the hypothesis that the data are generated by GARCH or Taylor/Schwert when we consider the APARCH as the true model. This implies that a more flexible structure should be used. We have also determined that the sum of the coefficients α_1 and β_1 is close to one. Even when we estimate the IGARCH model we get good results that are consistent with other findings for the developed economies. This leads us to a further investigation of this persistence in volatility as we begin to deal with a new variety of model, the so-called Fractionally Integrated GARCH (FIGARCH).

Table 4. 9 - A-PARCH Results

Country	μ	θ_1	α_0	α_1	γ_1	δ	β_1	Log L
Mexico	0.00108* (2.27)	0.69* (21.32)	0.000004 (1.44)	3.83* (9.76)	0.59* (52.58)	2.55* (17.84)	-0.00006 (-0.03)	17719.05
Taiwan	0.00108 (1.08)	0.73* (25.37)	0.073* (334.04)	0.15 (143.23)	0.14 (3.90)	0.39 (716.61)	2.44 (1495.6)	26369.79
Thailand	-0.00218 (-1.82)	0.80* (29.02)	0.000079* (2.73)	0.87* (21.40)	-0.107* (-8.23)	1.67* (25.85)	0.592* (3.17)	23514.22
Korea	0.000218 (0.26)	0.77* (10.58)	0.0527* (24.13)	0.183* (25.69)	0.09* (4.10)	0.426* (59.59)	3.24* (25.03)	25893.39
Malaysia	0.00184 (1.35)	0.83* (37.30)	0.058* (29.24)	0.23* (32.91)	0.00008 (0.004)	0.434* (66.18)	2.54* (34.53)	23087.23
Brazil	0.00303* (14.21)	0.82* (28.47)	0.152* (58.09)	0.161* (44.14)	0.072* (3.73)	0.267* (72.62)	1.1809* (83.96)	28705.41
Hong-Kong	0.000174* (4.56)	0.69* (32.48)	0.090* (39.32)	0.198* (75.03)	0.377* (21.36)	0.349* (67.37)	1.833* (41.50)	24451.85
Argentina	0.0117* (3.89)	0.81* (49.19)	0.099* (28.30)	0.239* (18.47)	0.121* (3.17)	0.375* (59.16)	1.502* (78.74)	6634.15
USA	0.000021* (3.20)	0.14* (19.00)	0.000014* (4.50)	0.083* (32.40)	-0.373* (-20.70)	1.43* (32.40)	0.920* (474.00)	56974.00

(*) Coefficients significant at 95% and 99%.

4.3 Persistence in Volatility – FIGARCH Models

It is now well known that using an ARMA-type specification such as GARCH to model the conditional variance of very high frequency returns commonly results in estimates of ($\alpha_1 + \beta_1$) that are close to unity in the AR polynomial. This suggests, as mentioned earlier, very high or infinite persistence for the effect of squared innovations to returns on conditional variance. In the covariance stationary GARCH, the effect of past squared innovations to returns, on the current conditional variance of returns, decay exponentially as the lag length increases. On the other hand, the integrated GARCH (IGARCH) process exhibits infinite persistence. That is, the effects of squared innovations on volatility never die out. As Baillie, Bollerslev and Mikkelsen (1996) point out, this knife-edge distinction between exponential decay and infinite persistence may be too restrictive and we should be willing to allow for some more flexibility.

Several recent studies[19] have reported the existence of long memory in the autocorrelations of some power of absolute returns. They have found that even if the GARCH specification is able to explain the short-run pattern of volatility, it fails to match the long-run volatility persistence. Therefore, motivated by this evidence of a long memory component in volatility, Baillie, Bollerslev and Mikkelsen (1996) proposed the fractionally integrated GARCH model of FIGARCH.

As was already mentioned, an extensive list of works concerning fractionally integrated processes has been done for the conditional mean.[20]

The Baillie, Bollerslev and Mikkelsen (1996) fractionally integrated GARCH model FIGARCH (p,d,q), allows for a fractional unit root, I(d), in the conditional variance process as it is the case for the mean in an ARFIMA (p,d,q) process. As we have seen above, in the GARCH (p,q) model, the conditional variance is parameterised as a distributed lag of past squared innovations,

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \dots\dots\dots + \alpha_q\varepsilon_{t-q}^2 + \beta_1\sigma_{t-1}^2 + \dots\dots\dots + \beta_p\sigma_{t-p}^2 = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i\sigma_{t-i}^2 \quad \text{or ,}$$

$$\alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2 \quad (4.13)$$

Rearranging the terms in (4.13) we could have written this as:

---

[19] Baillie, Bollerslev and Mikkelsen (1996), Dacorogna et al. (1993) and Harvey (1993) among others for exchange rates; Bollerslev and Mikkelsen (1996), Ding, Granger and Engle (1993) for S&P 500 equity returns and McCurdy and Michaud (1996) for the NYSE returns.

[20] A good survey of these works can be found in Baillie (1996).

$$\left[1-\alpha(L)-\beta(L)\right]\varepsilon_t^2 = \alpha_0 + \left[1-\beta(L)\right]v_t \qquad (4.14)$$

Alternatively, we could have factorised (4.13) the polynomial, $1 - \alpha(L) - \beta(L) \equiv (1-L)\phi(L)$[21].

$$\phi(L)(1-L)\varepsilon_t^2 = \alpha_0 + \left[1-\beta(L)\right]v_t \qquad (4.15)$$

The FIGARCH model is then obtained by replacing the difference operator in (3.14) with the fractional difference operator $(1-L)^d$. That is

$$\phi(L)(1-L)^d\varepsilon_t^2 = \alpha_0 + \left[1-\beta(L)\right]v_t \qquad (4.16)$$

where the roots of $\phi(z) = 0$ lie outside the unit circle. The fractional differencing operator has an infinite power series expansion. So, the FIGARCH(p,d,q) model nests the covariance-stationary GARCH(p,q) model for d = 0 and the IGARCH(p,q) model for d = 1[22]. Allowing for values of d in the interval between zero and unity gives added flexibility that may be interesting when modelling long-term dependence in the conditional variance. Sowell (1992) showed that an ARFIMA model of the conditional mean with parameter d would be able to capture the long memory component. This would be case for the FIGARCH. The d parameter would capture the long memory component of volatility while $\phi(L)$ and $\beta(L)$ would model the short-term structure. This specification could, then, prevent the mispescification of long memory components that could result from imposing less flexible GARCH/IGARCH structures on volatility data. As argued by Baillie, Bollerslev and Mikkelsen (1996) we might expect the fractionally integrated, I(d), specification to capture the volatility dependence in returns better than the knife-edge alternatives of either an I(0) or an I(1) specification.[23]

[21] In the case that the autoregressive polynomial has a unit root and $\phi(z)$ has all the roots outside the unit circle.

[22] This power series can be written as a binomial expansion which can be expressed as the Maclaurin series expansion $(1-L)^d = 1 - d\sum_{k=1}^{\infty}\Gamma(k-d)\Gamma(1-d)^{-1}\Gamma(k+1)^{-1}L^k \equiv 1-\delta_d(L)$ in which $\Gamma(.)$ denotes the gamma function. Also, by definition $(1-L)^0 \equiv 1$.

[23] While a shock to the optimal forecast of the future conditional variance decays at an exponential rate for the covariance-stationary GARCH (p,q) model, and remains and remains important for forecasts of all horizons for the IGARCH (p,q) model, in the FIGARCH (p,d,q) model the effect of a shock to the forecast of the future conditional variance will die out at a hyperbolic rate. The fractional differencing parameter is therefore identified by the rate of decay of a shock to the conditional variance.

In order to better understand the statistical properties and the estimation strategy used below, it is convenient to rewrite the FIGARCH (p,d,q) model in (4.15) in terms of the observationally equivalent infinite ARCH representation,[24]

$$\sigma_t^2 = [1-\beta(1)]^{-1}\alpha_0 + \{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\}\varepsilon_t^2 \equiv [1-\beta(1)]^{-1}\alpha_0 + \lambda(L)\varepsilon_t^2 \quad (4.16)$$

As we have already seen, the most common approach used for testing ARCH/GARCH models relies on the maximisation of a conditional likelihood function. In particular, assuming that the one-step ahead prediction errors are conditionally normally distributed, the likelihood function for the sample $\{y_1, y_2, \ldots, y_T\}$ equals (4.5) or ,

$$LogL(\theta; \varepsilon_1, \ldots\ldots, \varepsilon_T \big| \psi_0) = \sum_{t=1}^{T} -\frac{1}{2}\left[\log(2\pi) + \log(\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_{2t}}\right]$$

where the initial conditions, ψ_0 , are used to start-up the recursions for the conditional mean and variance functions. However, it has been argued that in many applications with high-frequency data, the assumptions of conditionally normally standardised innovations, $z_t = \varepsilon_t \sigma_t^{-1}$, is violated.

Following Weiss (1986) and Bollerslev and Wooldridge (1992) there is an alternative procedure regarding the normal Quasi-Maximum Likelihood Estimates (QMLE). The FIGARCH process is clearly not stationary as the IGARCH one. QMLE appear as an alternative to estimating the parameters of (4.15) or equivalently (4.16). The maximisation of QMLE realises $T^{1/2}$ consistent estimates of the parameters, then,

$$T^{1/2}(\hat{\theta} - \theta_0) \Rightarrow N\{0. A(\theta_0)^{-1} B(\theta_0) A(\theta_0)^{-1}\} \quad (4.17)$$

where A(.) and B(.) denote the Hessian and outer product of the gradient, respectively and θ_0 are the parameter values. Simulation evidence based in Bollerslev and Baillie (1996) indicates that the limiting distribution theory works well for sample sizes of 1500 and 3000 elements.[25]

[24] For the FIGARCH(p,d,q) model be a well-defined one and the conditional variance positive, all the coefficients in the infinite ARCH representation must be nonnegative on the basis of Nelson and Cao(1991).
For example for the FIGARCH (1,d,1) model estimated below, see Bollerslev and Mikkelsen (1996).

[25] Let's say that θ is the vector of estimates resulting from (4.16). An asymptotic robust covariance matrix for the parameter estimates is consistently estimated by $A(\theta)^{-1}B(\theta)A(\theta)^{-1}$, where A(θ) and B(θ) denote the Hessian and the outer product of the gradients, respectively, estimated at θ.

Of course, for the FIGARCH (p,d,q) model with d > 0.5 , the population variance does not exist. However, subject to some regularity conditions, conditioning on the pre-sample values will not affect the asymptotic distributions of the resulting estimators and test statistics. In most practical applications with high frequency financial data the standardised innovations $z_t = \varepsilon_t \sigma_t^{-1}$ are leptokurtic and not i.i.d. normal through time. In these situations, the robust Quasi MLE (QMLE) procedures discussed by Weiss (1986) and Bollerslev and Wooldridge (1992) may be invoked to permit valid asymptotic inference.

Unfortunately, the consistency and asymptotic normality of the QMLE based ARCH estimators and test statistics have only been formally established for the IGARCH (1,1) case to date. In particular, following Lee and Hansen (1994), and assuming that, if

1) z_t is stationary and ergodic
2) z_t^2 is non-degenerate
3) $E_{t-1}(z_t) \leq \kappa < \infty$ almost surely, and
4) $\sup_t E_{t-1}[\log\beta_1 + \alpha_1 z_t^2) < \infty$ almost surely

It is possible to show that the quasi-likelihood function and the corresponding score vector and Hessian are all strictly stationary and ergodic. Therefore, by a central limit theorem it follows that the QMLE obtained is both consistent and asymptotically normally distributed. While this result applies to IGARCH (1,1), this case extends directly to the FIGARCH(1,d,0) model through a dominance type argument. Proving the consistency and asymptotic normality of the estimators for the general FIGARCH (p,d,q) model remains an important avenue for future research.

Given the infinite series involved in computing $(1\text{-}L)^d$, initial conditions for fractionally integrated parameterisations are much more demanding than in ARCH/GARCH implementations. We have also used the unconditional sample variance for the pre-sample values of ε_t^2 and σ_t^2. The infinite power series expansion was truncated at 3000 lags rather than the 1000 lags used by Baillie et al. (1996). Since the fractional differencing operator is designed to capture the long-memory component of volatility, truncating it at too low a lag could destroy important long-run dependencies.

Using returns for eight of the countries in our sample, we estimated several models assuming the conditional variance to be FIGARCH (p,d,q). Since all of components of the index do not trade at the same time there is a lack of synchronisation that will generate serial correlation (see Scholes and Willian (1977) and Lo and MacKinlay (1990)). Therefore, it has been argued that an alternative model would parameterise the conditional mean function as an MA(1) rather than an AR(1). For this reason, we will consider both parameterisations as there is no conclusion about the superiority of one or the other.[26] The FIGARCH(1,d,1) model is, then:

$$y_t = \mu + \theta_1 y_{t-1} + \varepsilon_t$$

$$\varepsilon_t = z_t \sigma_t \qquad z_t = \varepsilon_t \sigma_t^{-1} \text{ i.i.d. N(0,1)} \qquad (4.18)$$

$$\sigma_t^2 = \left[1-\beta_1\right]^{-1}\alpha_0 + \left\{1-(1-\beta_1 L)^{-1}(1-\phi_1 L)(1-L)^d\right\}\varepsilon_t^2$$

The results can be seen in table 4.10 below for all countries. Several interesting facts can be seen from a visual inspection of this table. First of all, the parameters describing the conditional mean are all positive. Second, the parameters describing long memory in volatility (d) are also extremely significant

Table 4.10 - FIGARCH(1,d,1) Results

Country	μ	θ_1	α_0	β_1	ϕ_1	d	Log L
Mexico	0.00108* (5.87)	0.89* (54.32)	-0.585 (-0.85)	1.01* (74.4)	-0.70* (-2.07)	0.184 (0.477)	17671.09
Taiwan	0.0028* (2.08)	0.77* (25.37)	0.162 (1.17)	1.00 (539.15)	-0.479* (-4.10)	0.402* (3.33)	26712.21
Thailand	0.00418 (1.82)	0.81* (29.02)	0.0832 (0.83)	2.70* (2.85)	-0.008 (-1.82)	0.405* (2.94)	24351.23
Korea	0.0148 (1.50)	0.69* (30.58)	-0.035 (-0.890)	1.100* (8.41)	-0.049 (-0.65)	0.950* (84.9)	26001.25
Malaysia	0.0184 (1.35)	0.79* (35.30)	5.402* (3.27)	1.090* (4.76)	-0.0001 (-0.008)	0.554* (8.86)	23201.36
Brazil	0.0303 (1.21)	0.85* (20.47)	1.640* (1.160)	1.951* (3.06)	-0.0016 (-0.607)	0.644* (7.84)	29010.61
Hong-Kong	0.0174 (1.56)	0.79* (42.48)	3.780* (2.75)	1.360* (5.24)	0.001 (0.892)	0.680* (11.40)	24700.10
Argentina	0.0117 (1.89)	0.59* (49.19)	2.099* (2.30)	1.239* (8.47)	-0.001 (-0.17)	0.675* (9.16)	6751.21
USA	0.00048 (6.96)	0.182* (15.17)	0.0000013* (3.26)	0.657* (11.53)	0.387 (7.44)	0.447* (6.30)	67147.50

(*) Coefficients significant at 95%. The numbers in parenthesis refer to the Robust t-statistic derived from the QMLE. We have non-convergence of all estimates. T-stat are very high and reflect the non-convergence. As d is higher than 0.5 for a lot of countries, this is an indication that the variance does exist and by so will affect the asymptotic distribution of test statistics

[26] Bollerslev and Mikkelsen (1996) used the AR-FIGARCH parametrization while McCurdy and Michaud (1996) used the MA-FIGARCH parametrization.

. The values found for d range from 0.184 (Mexico) to 0.95 (Korea) and are fairly concentrated at 0.40. This seems to confirm the findings of other studies, such as Bollerslev and Mikkelsen (1996) and McCurdy and Michaud (1996). who found values of 0.447 and 0.471 respectively. Table 4.10 reports the findings for the US market using the S&P 500 as described in Bollerslev and Mikkelsen (1996). We can see that all coefficients are also significant and tend to be higher for the emerging economies. An exception is the ϕ_l term. Also, the fractional parameter is in the middle of the range of values found. It is worth a while that d higher than 0.5 indicates that the variance does not exits and so we should be sceptical in the interpretation of the t-stats (for instance, see Korea).

It is striking that all values for d are bigger than 0 and less than 1. This means that neither the GARCH nor IGARCH models are the correct specification for the conditional variance. Thus imposing either structure would produce specification error. Baillie, Bollerslev and Mikkelsen (1996) report the effects of estimating stable GARCH processes where the true data generating process is FIGARCH. The sum of the estimated GARCH (1,1) parameters is always close to one (as above) which implies integrated GARCH (IGARCH) behaviour and suggests that the apparent widespread IGARCH property so often found in high frequency studies of financial data may well be spurious. The IGARCH process is indeed poor at distinguish between integrated versus the long memory formulations of conditional variance. We also note that, as the parameter ϕ_1 is not significant for the majority of countries, we could re-estimate the model allowing for an FIGARCH (1,d,0).

We are not interested in modelling the conditional mean itself. Given, that the data used represent an index return at daily intervals, it is possible to have 'stale' returns, since the components of the index are not all traded at the same time. This lack of synchronised trading times has been pointed out by certain authors who argue that it will generate serial correlation. Actually, this would be a reason for using an MA specification for returns. [27] The exact structure of this autocorrelation will depend on the specific features of the non-synchronicity. In order to take account of such serial dependence, we extended the AR specification and parameterised the mean for all countries as an unrestricted AR(3) model. More complicated parametric formulations of the conditional mean, where the degree of serial correlation depends on the level of volatility, have been investigated by LeBaron (1992) and Bollerslev, Engle and Nelson (1994). According to Bollerslev and Mikkelsen (1996), the degree of predictability in the mean is marginal and of minor consequence for the conditional variances. We must also convey that the more conventional

[27] Scholes and Williams (1977) and Lo and MacKinlay (1990) argue that discontinuous trading in the stocks that make up the index may result in significant serial dependence in the index returns. Also the ARCH-M/GARCH-M has been proposed to deal with this kind of problem.

ARCH/GARCH type models that imply either exponential or infinite persistence may be overly restrictive. Thus, another reason for allowing a more rich specification is appreciated. However, no other more complicated structures will be pursued here.[28] The model, then, is the following:

$$y_t = \mu + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \theta_3 y_{t-3} + \varepsilon_t$$

$$\varepsilon_t = z_t \sigma_t \qquad z_t = \varepsilon_t \sigma_t^{-1} \text{ i.i.d. N}(0,1) \qquad (4.19)$$

$$\sigma_t^2 = \left[1-\beta_1\right]^{-1}\alpha_0 + \left\{1-(1-\beta_1 L)^{-1}(1-L)^d\right\}\varepsilon_t^2$$

The results for the AR(3)-FIGARCH(1,d,0) are depicted in table 4.11. Again the estimates for the fractional differencing parameter, d, is strikingly significant. Judged at standard significance levels, the estimated d is statistically different from zero or one. The parameter for the autoregressive part of the conditional mean is significant and conforms to other findings in the literature. Again the evidence supports the idea of specification error. The range for d lies between 0.284 (Mexico) and 1.15 (Hong-Kong) and is centred around 0.40 as well. For the US market we obtain the same results. All coefficients are significants. The coefficients for the emerging economies are generally larger than those found for the US market except for the β_1 and the fractional parameter. Again, we cannot reject the hypothesis of d different either from 0 (stable GARCH) and 1 (IGARCH). The largest value for d is for Korea (0.826) and the lowest for Mexico (0.284). The estimate for the USA is 0.447, around the mean value found for the overall data.

Our readings from the last tables (4.10 and 4.11) suggest that a GARCH, ARCH or IGARCH specification for the conditional variance may be inappropriate. In some cases the GARCH estimates may spuriously indicate an IGARCH process. Baillie, Bollerslev and Mikkelsen (1996) point out that inference with an IGARCH parameterisation will have less power to discriminate between integrated versus fractionally integrated data generating processes than will inference using the fractionally integrated parameterisation. This conforms the evidence found for the conditional mean. Therefore, we conclude that it is desirable to increase the flexibility of the conditional variance specification in the direction of the FIGARCH model.

[28] Actually it seems that imposing an AR(1) specification for the conditional mean could have imposed too much restrictions, so we allow for some more flexibility.

Table 4.11 -AR(3)-FIGARCH(1,d,0) Results

Country	μ	θ_1	θ_2	θ_3	α_0	β_1	d	Log L
Mexico	0.0008* (5.27)	0.193* (5.30)	-0.034* (5.32)	0.035* (5.33)	0.0001 (0.50)	0.646* (6.39)	0.284* (2.84)	17681.24
Taiwan	0.0010 (4.25)	0.291* (6.25)	0.051* (6.41)	0.062* (6.37)	0.050* (58.0)	0.266* (6.39)	0.346* (9.50)	26725.12
Thailand	-0.00218 (-4.82)	0.151* (9.02)	0.060* (9.32)	0.057* (9.45)	0.067* (3.96)	0.013 (0.17)	0.400* (6.31)	24352.12
Korea	0.00318 (3.21)	0.198* (4.52)	0.041* (4.89)	0.055* (5.00)	0.051 (0.45)	0.689 (0.857)	0.826 (0.956)	26011.26
Malaysia	0.00184 (4.35)	0.141* (3.30)	0.057* (3.31)	0.047* (3.35)	0.143* (4.29)	0.235* (2.155)	0.463* (3.85)	23204.33
Brazil	0.00303* (14.21)	0.278* (2.47)	0.061* (2.49)	0.055* (2.51)	0.138* (4.07)	0.256* (4.02)	0.423* (7.95)	29031.21
Hong-Kong	0.0074* (3.56)	0.220* (6.48)	0.031* (6.51)	0.061* (6.51)	0.039* (2.35)	0.864* (15.7)	1.15* (11.38)	24701.10
Argentina	0.0117* (3.89)	0.257* (8.19)	0.044* (8.49)	0.071* (8.01)	0.065 (1.33)	0.435* (4.08)	0.598* (8.33)	6753.41
USA	0.00048* (6.96)	0.182* (15.17)	-0.061* (-5.08)	0.026* (2.17)	0.0000013 (3.26)	0.365* (7.02)	0.447* (6.30)	67043.21

(*) Coefficients significant at 95%. The numbers in parenthesis refer to the Robust t-statistic derived from the QMLE

We confirm the stylised fact that for high frequency, the IGARCH model describes financial returns as the parameters for the GARCH (1,1) sum up to close one. Estimates of these persistence coefficients ($\alpha_1 + \beta_1$) are such that an integrated process may not be rejected by the data. On the other hand, the fractionally integrated GARCH shows that an integrated hypothesis would be rejected in favour of versions with a fractional unit root. Baillie et al (1996) and McCurdy and Michaud (1996) have found similar results. This finding supports the hypothesis that the IGARCH or infinite persistence case can be a specification that is not flexible enough to fit the data.

We recall, however, that for financial returns there is an asymmetric movement of volatility in response to the sign of innovations to returns. The asymmetric response of volatility to positive and negative 'shocks' is known in the finance literature as the leverage effect of stock market returns. This means that stock returns are negatively correlated with changes in return volatility, i.e. volatility tends to rise in response to 'bad news' (less return than expected) and to fall in response to 'good news' (returns higher than expected). As a result, McCurdy and Michaud (1996) proposed a Fractionally Integrated Asymmetric Power ARCH (FIAPARCH) specification. Another possibility would be the Fractionally Integrated Exponential GARCH (FIEGARCH) by Bollerslev and Mikkelsen (1996), which will not be pursued in this thesis.

Rewriting the A-PARCH specification (4.11) in an analogous form, we have:

$$\phi(L)(1-L)^d\left[|\varepsilon_t|-\gamma(L)\varepsilon_t\right]^\delta = \alpha_0+\left[1-\beta(L)\right]v_t \qquad (4.20)$$

in which $\phi(L) \equiv [1-\alpha(L)-\beta(L)]\ [1-L]^{-1}$ but , however, now $v_t \equiv (|\varepsilon_t| - \gamma\varepsilon_t)^\delta - \sigma_t^\delta$. The fractionally integrated form of the APARCH, which Mccurdy and Michaud label FIAPARCH (p,d,q) is then, given by:

$$\phi(L)(1-L)^d\left[|\varepsilon_t|-\gamma(L)\varepsilon_t\right]^\delta = \alpha_0+\left[1-\beta(L)\right]v_t \qquad (4.21)$$

In order to write the conditional variance, we can rearrange (4.21) finding:

$$\sigma_t^\delta = \left[1-\beta(1)\right]^{-1}\alpha_0+\left\{1-(1-\beta(L))^{-1}\phi(L)(1-L)^d\right\}\left(|\varepsilon_t|-\gamma(L)\varepsilon_t\right)^\delta \qquad (4.22)$$

Thus the FIAPARCH(p,d,q) nests the APARCH model (d=0), as well as any one of the other seven models listed before nested by the APARCH. This allows testing against more restrictive specifications using likelihood ratio (LR) tests. The FIAPARCH specification should also have advantages over the FIGARCH model in that it allows for asymmetry and also for the power exponent to vary to help match the temporal pattern of volatility found in previous studies. In the FIAPARCH model, we had the same problem as the FIGARCH, that is, given the infinite series involved in computing $(1-L)^d$, initial conditions for the fractionally integrated parameterisation are much more demanding. In particular, we had to use the unconditional sample standard deviation for the pre-sample values of ε_t and σ_t required in the terms $(|\varepsilon_t| - \gamma\varepsilon_t)^\delta$ and σ_t^δ . As already mentioned above, we truncate the infinite lag at 3000 rather than the 1000 lags used in Baillie et al. (1996).

Using the returns associated value-weighted index for each of the eight countries in our sample, we introduce the FIAPARCH model in our analysis. Our main concern is not estimating the conditional mean per se. However, the lack of synchronisation of trading times is expected to generate serial correlation. Another way to proceed involves modelling the conditional mean function as an MA(1) process rather than an AR(1). We follow McCurdy and Michaud (1996)[29] in using this specification. Therefore, the conditional mean function for each is parameterised as

[29] We must keep in mind that modelling the returns as an AR(1), that is, $y_t = \mu + \theta_1 y_{t-1} + \varepsilon_t$ allows us to find a Moving average representation of this AR(1). Using the lag operator L, we can rewrite this as $(1-\theta_1 L)y_t = \mu + \varepsilon_t$. Observing some restriction in the parameters θ_1 we can invert the polynomial in the lag operator leading to a convergent

$y_t = \omega \quad + \varepsilon_t + \psi_1 \varepsilon_{t-1}$ (4.23)

We then present the results of our estimations for the MA(1)-FIAPARCH (1,d,1) in table 4.12. Following the same steps used previously we are lead to conclude that this latter specification produces better results, which argues in favour of a long-memory component in volatility. In this case, any of the previous specifications (GARCH, ARCH, APARCH etc.) would lead to specification errors. All parameters have been estimated using approximate Maximum likelihood (QMLE) as above. We then present the following model:

$$y_t = \omega \quad + \varepsilon_t + \psi_1 \varepsilon_{t-1}$$

$$\varepsilon_t = z_t \sigma_t \qquad z_t = \varepsilon_t \sigma_t^{-1} \text{ i.i.d. N(0,1)} \qquad (4.24)$$

$$\sigma_t^\delta = \left[1 - \beta_1\right]^{-1} \alpha_0 + \left\{1 - (1 - \beta_1 L)^{-1}(1 - \phi_1 L)(1 - L)^d\right\}\left(\left|\varepsilon_t\right| - \gamma_1 L \varepsilon_t\right)^\delta$$

The results are depicted in the table 4.12 below. The results for the US market are from McCurdy and Michaud (1996). As we can see, most of the coefficients are significant. The d (fractional) and δ (asymmetry) parameters are positive for all countries in the sample. This highlights the facts that, not only are there are long memory effects but it seems there are also asymmetry effects that should be taken account of. On the basis of likelihood ratio (LR) tests, the fractionally integrated models provide a statistically significant improvement over the non-integrated specification. This is indeed another fact that points to the existence of long memory effects. The LR test statistic for FIGARCH (1,d,1) versus GARCH(1,1) is 54.6 and for FIAPARCH(1,d,1) versus APARCH (1,1) is 121.3. Under the null hypothesis that the standardised residuals have a conditional normal distribution, these test statistics have an asymptotic chi-square distribution with 1 degree of freedom (the restriction that d = 0). We can reject the non-integrated models in favour of the more flexible integrated versions, a conclusion found to be true above when the conditional mean of the returns is assumed to be AR(1) or AR(3).

---

series. If so, we can, then, find a moving average representation, $y_t = \frac{\mu}{(1-\theta_1 L)} +_t \frac{\varepsilon_t}{(1-\theta_1 L)}$. Thinking of the term in the denominator as the sum of an infinite geometric progression with first term ε_t and ratio equal to θ_1 we would have then: $y_t = \frac{\mu}{1-\theta_1} + \sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i}$ where the δ_i parameters are functions of the original parameter. This means that y_t can be represented by an MA(∞) process.

One can also reject the more restrictive FIGARCH (1,d,1) model in favour of the FIAPARCH (1,d,1) parameterisation. Therefore, the FIAPARCH(1,d,1) specification rejects all the alternatives and consequently all the existing models nested by the APARCH(1,1) structure.

Table 4.12 -MA(1)- FIAPARCH(1,d,1) Results

Country	ω	ψ_1	α_0	β_1	ϕ_1	d	γ_1	δ	Log L
Mexico	0.009* (4.31)	0.2471* (6.55)	0.0043 (1.54)	0.3711* (1.84)	0.2535* (9.54)	0.5412* (5.42)	0.2514* (4.22)	1.5798* (11.97)	17811.01
Taiwan	0.002* (3.08)	0.6807* (5.37)	0.0656* (3.04)	0.4432 (1.23)	0.0870* (3.90)	0.4289* (4.67)	0.0436* (5.45)	2.0451* (16.61)	26369.79
Thailand	0.003* (5.32)	0.4614* (8.44)	0.0051* (6.12)	0.4214* (3.21)	0.2478* (6.32)	0.3641* (4.99)	0.0179* (4.15)	1.5874* (21.32)	23612.34
Korea	0.005* (2.60)	0.4224* (5.58)	0.1427* (2.13)	0.4833* (2.69)	0.1158* (7.10)	0.4988* (8.12)	0.1436* (4.55)	0.7436* (9.59)	25893.39
Malaysia	0.002* (3.35)	0.3280* (4.35)	0.0068 (1.24)	0.1971* (3.91)	0.0203 (0.004)	0.3615* (5.55)	0.0865 (1.23)	2.2871* (6.18)	23087.23
Brazil	0.005* (4.21)	0.6584* (2.47)	0.0049* (5.09)	0.3690* (4.14)	0.0564* (13.73)	0.4639* (8.90)	0.0348* (9.12)	2.3062* (7.62)	28705.41
Hong-Kong	0.011* (4.63)	0.4684* (6.48)	0.0027* (3.32)	0.4693* (7.03)	0.1670* (21.36)	0.4878* (9.21)	0.1241* (4.32)	2.1308* (6.37)	24451.85
Argentina	0.019* (5.71)	0.6234* (4.32)	0.0016 (1.44)	0.6336* (7.34)	0.1461* (5.81)	0.6598* (3.45)	0.1361* (5.54)	2.2298* (17.84)	17719.05
USA	0.049* (5.71)	0.2060* (4.32)	0.0030 (1.44)	0.5730* (7.34)	0.377* (5.81)	0.2880* (3.45)	0.6100* (5.54)	1.7340* (17.84)	1036.36

(*) Coefficients significant at 95%. The numbers in parenthesis refer to the Robust t-statistic derived from the QMLE

Finally we could ask if the estimation methods and models used exhaust all possibilities. The answer is no. As an example, an alternative for the modelling persistence in variance is the stochastic volatility process developed by Breidt , Crato and de Lima (1993) and Harvey (1993). Their model is

$$y_t = z_t \sigma_t \qquad z_t \text{ is i.i.d. N(0,1)}$$

$$\sigma_t^2 = \sigma^2 \exp(h_t) \qquad (4.25)$$

In previous work on stochastic volatility models, it is commonly assumed that h_t is an AR(1) process, which implies an ARMA(1,1) representation for $\log(y_t^2)$. If is assumed that h_t is the fractional white noise process,

$(1-L)^d h_t = \varepsilon_t$ (4.26)

where ε_t is i.i.d $N(0,\sigma^2)$, then (3.25) and (3.26) generate a long memory stochastic volatility process. Estimation of regular stochastic volatility models has generally been through the state space representation and used Quasi-Maximum Likelihood Estimation (QMLE) via the Kalman filter. Since a state space representation does not exist for log memory processes, estimation of the long memory stochastic volatility process is difficult. Breidt, Crato and Lima (1993) use frequency domain approximate MLE to estimate and ARFIMA (0,d,1) model for $\log(y_t^2)$, while Harvey (1993) uses the GPH estimator to obtain an estimation of d in a fractional white noise model for $\log(y_t^2)$. The comparison of long memory ARCH and stochastic volatility models remains an interesting area for future research.

4.4 Conclusion

We have estimated a great variety of models in this thesis. Our goal was to investigate and verify whether or not an indication of the existence of long memory in volatility would persist when we used data collected from emerging markets. We have found several useful and interesting results.

First of all, we have found the expected result that ARCH/GARCH does not adequately represent the volatility present in the stock markets of emerging economies. This finding is consistent with what we found for more developed economies, and in particular the US market. We began to observe that in spite of the coefficients' significance, they tend to be systematically higher in emerging economies than in more developed economies. This may indicate that the effects are larger in emerging economies.

Secondly, we have looked at the existence of asymmetries. The findings of Ding, Granger and Engle (1993) and their A-PARCH model do apply for emerging economies: the asymmetry also appears to be higher in emerging economies as described by the magnitude of this parameter. This result may also be consistent with the fact that these markets subject to greater movements caused by intervention in the economy.

Third, in Chapter 3 we have found persistence in volatility using the autocorrelation analysis. The specification proposed by Engle and Bollerslev (1986) with their IGARCH model allows the shock to exist forever. We have seen that for all economies including the US, we cannot reject the null hypothesis that the sum of the coefficients is equal to one. This specification, however, is much too restrictive, and other possibilities must be explored. The natural progression is to use the FIGARCH approach.

Fourth, with respect to the results of the FIGARCH specification. we have found that the coefficients related to the existence of long memory are significant. It is striking that d (the fractional parameter) is different from 0 and 1. This is consistent for all economies, emerging or developed. We have also found, however, that the coefficients for emerging economies are higher than the ones found for the US. Having said this, we must stress that the samples differ in not only coverage period but also the amount of data used.

Even when using the FIGARCH specification, we should allow for asymmetries to be present. It appears that the FI-PARCH specification produces the best results. Again, all coefficients are found to be significant, and the coefficients are higher for emerging economies.

Chapter 5

Switching Regimes in Volatility – Testing the Change in Regimes for Volatility in the Emerging Markets

5.1 Introduction

In the past fifteen years, a great deal of attention has been paid to the properties of the second moments of financial data series. It is important to account for the conditional variance of such series not only for inference purposes but also for empirical implementation. Several models have been developed to capture the observed clustering of volatility found in these data, beginning with the ARCH model in 1982. Since then, numerous extensions of this model have been proposed in order to capture other phenomena such as asymmetry, kinks and discontinuities.

In two recent papers, Granger and Ding (1995) consider long return series containing first differences of log prices, or price indices. They establish a set of temporal and distributional properties for such series, suggesting that returns are well characterized by a double exponential distribution that displays persistence and long memory properties. Rydén et al. (1997) show that a mixture of normal variables with zero mean can generate a series with most of the properties that Granger and Ding (op. cit) have identified. In this case, they show that the temporal higher-order dependence observed in return series can be described by a hidden Markov model. They estimate this model for ten sub-series of the well-known S&P 500 series, including about 17,000 daily observations. This study reproduces the stylised facts of Granger and Ding quite well, but the parameter estimates vary considerably from one sub-series to the next.

In some statistical applications, the observed time series is an aggregation of many individual time series. The question arises as to whether aggregation has any relevant effect on the dependence structure. To clarify this point, suppose the individual time series $X_t(j)$ (j = 1,2,3,...) are summarised in a single aggregate time series as follows:

$$X_t = \sum_{j=1}^{\infty} X_t(j)$$

Assuming that the individual time series are all stationary with short memory, is it possible for the aggregate series X_t to exhibit long-memory? This question was addressed by

Granger (1980). He shows that aggregation of short-memory processes can indeed produce the appearance of long memory. The main point is that one cannot confirm the presence of long memory in individual series by observing long-range dependence in aggregated series since long memory can be artificially introduced by aggregation. To find the source of long memory, the behaviour of individual series needs to be examined. Hamilton (1994) treats certain Markov processes known as mixture distributions as an instructive case. In previous chapters we have explored some possible explanations of long-memory. However, as noted above, it is possible for aggregation to produce series with long-memory. Therefore, it is worthwhile considering other potential sources of long-memory, for example changes of regime. We would also like to examine other explanations for the presence for long-memory.

It is possible to reconcile the idea of mixture distributions with the occurrence of two different regimes. Let the regime that a given process is in at date t be indexed by an unobserved random variable s_t, where there are N possible regimes (s_t = 1, 2 , 3,..., N). When the process is in regime 1, the observed variable vt is presumed to have been drawn from a $N(\mu_1,\sigma_1^2)$ distribution. If the process is in regime 2, then y_t is drawn from a $N(\mu_2,\sigma_2^2)$ distribution, and so on. Hence, the joint density of yt conditional on the random variable s_t taking on the value j is

$$f(y_t/s_t = j;\theta) = \frac{1}{\sqrt{2\pi\sigma_j}} \exp\left\{\frac{-(y_t-\mu_j)^2}{2\sigma_j^2}\right\} \quad (5.1)$$

for j = 1, 2, 3,.....N. Here θ is a vector of population parameters that includes μ_1, μ_1, ...μ_N and σ_1^2 , σ_2^2 , ..., σ_N^2 .

The unobserved regime {s_t} is presumed to have been generated by some probability distribution, for which the unconditional probability that s_t takes on the value j is denoted π_j :

$$P\{s_t = j;\theta\} = \pi_j \quad for \;\; j = 1,2,....., N \quad (5.2)$$

The probabilities π_1 , π_2 , ...π_N are also included in θ. That is, θ is given by:

$$\theta \equiv (\mu_1,..., \mu_N, \sigma_1^2,....., \sigma_N^2, \pi_1,....., \pi_N)'$$

Recalling that for any events A and B, the conditional probability of A given B is defined as

$$P\{A/B\} = \frac{P\left\{A \quad and\; B\right\}}{P\{B\}}$$

assuming that the probability that the event B occurs is not zero. This expression implies that the joint probability of A and B occurring together can be calculated as

$$P\left\{A \quad and\; B\right\} = P\{A/B\}.P\{B\}$$

For example, if we are interested in the probability of the joint event that $s_t = j$ and that y_t falls within some interval [c.d], this could be found by integrating

$$P\{y_t, s_t = j;\theta\} = f\{y_t / s_t = j;\theta\}.P\{s_t = j;\theta\} \quad (5.3)$$

over all values of y_t between c and d. Expression (5.3) will be called the joint density-distribution function of y_t and s_t. From(5.1) and (5.2), this function is given by :

$$p(y_t, s_t = j;\theta) = \frac{\pi_j}{\sqrt{2\pi\sigma_j}} \exp\left\{\frac{-(y_t - \mu_j)^2}{2\sigma_j^2}\right\} \quad (5.4)$$

The unconditional density of y_t can be found by summing (5.4) over all possible values for j:

$$f(y_t;\theta) = \sum_{j=1}^{N} p(y_t, s_t = j;\theta) =$$

$$= \frac{\pi_1}{\sqrt{2\pi\sigma_1}} \exp\left\{\frac{-(y_t - \mu_1)^2}{2\sigma_1^2}\right\} \quad +$$

$$+ \frac{\pi_2}{\sqrt{2\pi\sigma_{j2}}} \exp\left\{\frac{-(y_t - \mu_2)^2}{2\sigma_2^2}\right\} \quad + \ldots\ldots\ldots$$

$$+ \frac{\pi_N}{\sqrt{2\pi\sigma_N}} \exp\left\{\frac{-(y_t - \mu_N)^2}{2\sigma_N^2}\right\} \quad (5.5)$$

Since the regime s_t is unobserved, expression (5.5) is the relevant density describing the actually observed data y_t. If the regime variable s_t is i.i.d. across different dates t, then the log likelihood for the observed data can be calculated from (5.5) as

$$L(\theta) = \sum_{t=1}^{N} \log f(y_t;\theta) \qquad (5.6)$$

The maximum likelihood estimate of θ can be obtained by maximizing (5.6) subject to the constraints that $\pi_1 + \pi_2 + \dots\dots + \pi_N = 1$ and $\pi_j \geq 0$ for j = 1, 2,...., N. Functions of the form of (5.5) can be used to represent a broad class of different densities. For example, if N = 2, then the unconditional density for the observed variable f (y_t ; θ) is the sum of these two magnitudes.

Once one has obtained estimates of θ, it is possible to make an inference about which regime was more likely to have been responsible for producing the date t observation of y_t. Again, from the definition of a conditional probability. it follows that

$$P\{y_t, s_t = j;\theta\} = \frac{p(y_t, s_t = j;\theta)}{f(y_t;\theta)} = \frac{\pi_j . f(y_t/s_t = j;\theta)}{f(y_t;\theta)} \qquad (5.7)$$

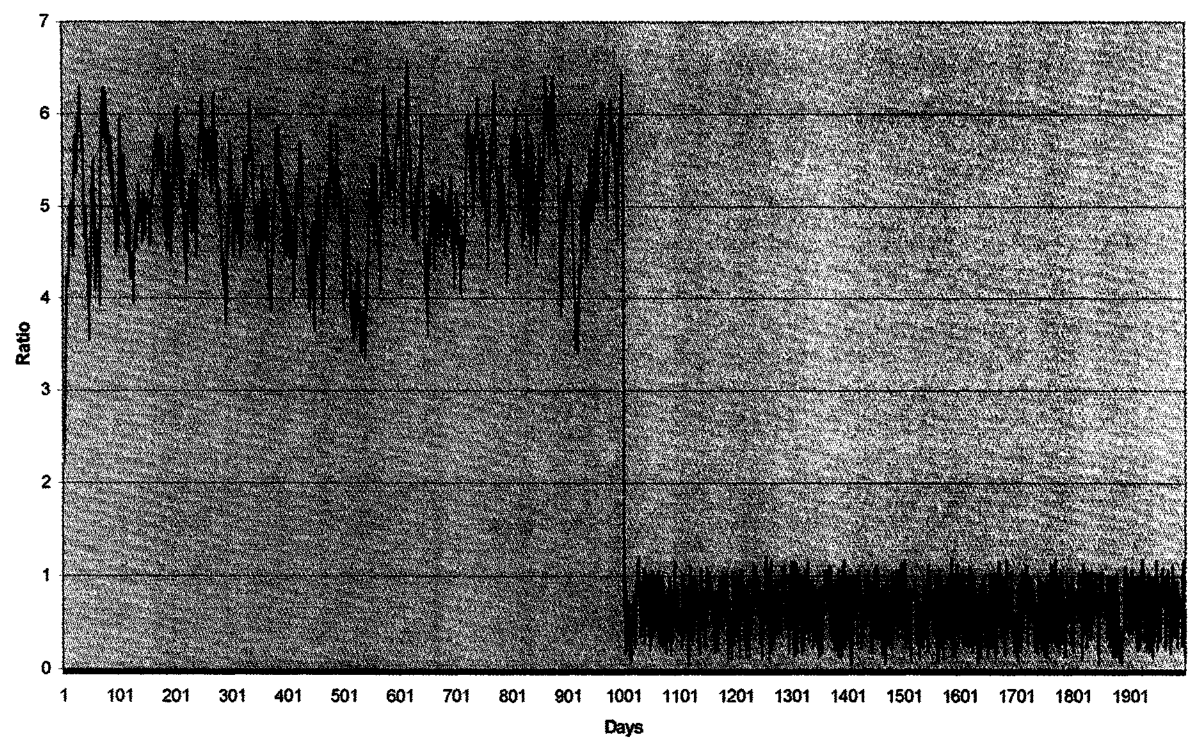
Given knowledge of the population parameters θ, it is possible to use (5.1) together with (5.5) to calculate the magnitude of (5.7) for each observation y_t in the sample. This number gives us the probability, given the observed data, that the unobserved regime responsible for observation t was regime j. For example let's imagine a density of mixture of two gaussian distributions with y_t / s_t = 1 ~ N(0,1) , y_t / s_t = 2 ~ N(4,1) and P[s_t = 1] = 0.8. If an observation y_t were equal to zero, one could be virtually certain that the observation had come from a N(0,1) distribution rather than a N(4,1) distribution.

Many variables undergo episodes in which the behaviour of the series seems to have changed quite dramatically. We have found this pattern for almost all series in our present study. Similar dramatic breaks will be seen if one follows any macroeconomic or financial time series for a sufficiently long period. Such apparent changes in the time series process can result from events such as wars, financial crises or significant changes in government policies. In the recent devaluation of the Mexican peso, there was quite a dramatic change in many of the underlying series for the Mexican economy. In Graph 5.1 we have plotted daily data for the log of the ratio of the peso value of US$ denominated bank accounts to the peso denominated bank accounts in Mexico from 1978-1992. The data show a sharp decrease in this ratio due to the devaluation of the Mexican peso. The Mexican government adopted various measures in 1982 to try to discourage the use of dollar-denominated accounts, and the effects are dramatically seen in a plot of the series.

Suppose that we are trying to better understand the data in graph 5.1. One simple explanation would be that the constant term for the autoregression changed at date 1000 (the point at which the devaluation happened). But if the process changed in the past, it could also change in the future. Moreover, the change in regime surely should not be regarded as the outcome of a deterministic event, but rather should be viewed as a random variable. These observations suggests that we might consider the process to be somehow influenced by an unobserved random variable s_t^*, which stands for the state or regime that the process was in date at date t. Note that one of the variables may be the volatility. This is the idea behind Hamilton and Susnel (1994) who proposed a Switching ARCH Model (SWARCH).

Graph 5.1

Log of the ratio of the peso value of dollar denominated bank accounts in Mexico to the peso-denominated bank accounts in Mexico, daily, 1978-1992



Source: Rogers (1992) in Hamilton (1994)

As we have seen in the preceding chapters, persistence and long memory in the volatility of asset returns appears to be very high. This observation has led to a consideration of a wide variety of models, from IGARCH to FI-GARCH. Volatility clustering is a well-documented feature of financial rates of return: price changes that are large in magnitude tend to occur in bunches rather than with equal spacing. A natural question is how long financial markets will remain volatile.

Two stylised facts that conventional volatility models, notably GARCH, have difficulty explaining are (1) that conditional volatility can increase substantially in a short amount of time at the onset of a turbulent period, and (2) that the rate of mean-reversion in stock-market volatility appears to vary positively and nonlinearly with the level of volatility. Hamilton and Susnel (1994) highlighted the forecasting difficulties of conventional GARCH models by showing that they can provide worse multi-period volatility forecasts than constant variance models. It would be possible to address this question by not allowing the conditional variance to respond proportionately to large and small shocks. In this way, the conditional variance could be prevented from increasing to a level at which volatility forecasts would be undesirably high. One drawback of this approach is that in such a model could understate the true variance by not responding sufficiently to large shocks and thereby never be pressed to display much mean reversion. Such "threshold" models do not address the two stylised facts listed above, i.e. sharp upward jumps in volatility, followed by fairly rapid reversion to near-normal levels.

Lamoreux and Lastrapes (1990) observed that structural breaks in the variance could account for high persistence and long memory in the estimated variance. This would be an alternative way to exploit long memory in volatility. It is then desirable to properly address the two stylised facts from within the class of ARCH/GARCH models with Markov-switching parameters. Markov-switching parameters ought to enable the volatility to experience discrete shifts and changes in the persistence parameters. A number of researchers have suggested that the poor forecasting performance and spuriously high persistence of ARCH models might be related to structural changes in the ARCH process[1]. Perron (1989) argued that changes in regime may give the false impression of unit roots in characterising of the level of a series. Cai (1994) stresses that volatility in Treasury bill yields are much less persistent when one models changes in parameters using a Markov-switching process. We will pursue this idea to verify whether this methodology can account for long memory in volatility.

The advantage of switching regime models relative to standard GARCH formulations is that they adapt more quickly to periods of high or low volatility. GARCH models are too persistent to capture a sudden increase in volatility. Similarly, shocks to the conditional variance die out too slowly to capture certain historical episodes.

[1] Diebold (1986) and Lamourereux and Lastrapes (1990) argued that the high estimated for the persistence parameter may reflect structural changes that occurred during the sample in the the variance process. This is related to Perron's (1989) observation that changes in regime may give the spurious impression of unit roots when dealing with the level of a series. Hamilton and Susnel (1994) found that ARCH/GARCH models have bigger MSE Loss than simple models that consider the constant unconditional sample variance to forecast variance. They also found that ARCH/GARCH models can lead to nonegligible consequences a full year later (high persistence).

It has been suggested that the conditional variance of stock returns may exhibit sudden jumps. A paper by Perron (1999) proposes a non-paramètric procedure to detect discontinuities in higher conditional moments, particularly the conditional variance. This procedure enables us to estimate the number and location of jumps. We use this procedure to detect several jumps in the conditional variance of daily returns.

In this chapter we exploit further the two ideas highlighted above. In section 5.2 we describe the methodology used by Perron (1999) and apply it to the data used in our thesis. In section 5.3 we present the methodology proposed by Hamilton and Susnel (1994) for the Switching ARCH (SWARCH) and Dueker (1997) in the context of a Switching GARCH (SW-GARCH) and we estimate these models for the countries in our study. Section 5.4 concludes.

5.2 - Jumps in the Volatility of Stock Returns

A large number of models using non-parametric estimation of the conditional variance have been developed as alternatives to parametric models. These estimators assume that the conditional variance process is smooth, so they will be inconsistent at any point of discontinuity. The first to propose such a methodology was Lamoureux and Lastrapes (1990). Hamilton and Susnel (1994) and Cai (1994) provided simple parametric forms by adding a Markov chain to an ARCH model. This was recently extended by Dueker (1997) by allowing for GARCH processes.

In view of the variety of available parametric models to describe time-varying volatility and long-memory. the non-parametric approach seems a good way to proceed. It has been established that misspecification of the continuous component of the conditional variance leads to erroneous inference in the presence of jumps. The test relies on Muller (1992) that derived it by using one-sided window in estimating the conditional variance. At points where a jump occurs, the left-hand side and right-hand side estimates converge respectively to the left and right limits at that point. The difference between these estimates provides the basis for determining the detection of the jump.

Let us assume that y_t is a random variable with zero mean. The conditional moments y_t of can be estimated using the following functional representation suggested by Perron (1999)

$$G(y_t) = m(X_t) + u_t = g(X_t) + S(Z_t) + u(X_t) \qquad (5.1)$$

where G(.) and g(.) are continuous functions, $X_t = (X_{1,t}\ , \ ......,\ X_{p+1,t})'$ is a p + 1 vector whose last element is time defined as a fraction of the sample $X_{p+1} = t/T$, T is the sample size, $S(X_t)$ is a step function with finite jumps whose argument is a member of X_t, and $u(X_t)$ is a stochastic disturbance term with $E(u(X_t) / X_t) = 0$. For simplicity, we can assume that the dependence of S and Z and u on X_t will be suppressed and written as S_t and u_t. The natural candidates for inclusion in X_t are lagged values of the dependent variable y_{t-1} , ..., y_{t-p}.

There are triggering changes in S_t that are supposed to be explained by the kth member of X, which we can denote Z. We can use "time" to identify the occurrence of jumps. Actually, any variable that causes the jumps must be consistent with the findings obtained by the assumption that "time" causes them.

Discontinuities can be detected by looking at the differences between kernel estimates with one-sided windows along the values of Z_t:

$$\hat{m}^{+}(z) = \frac{\sum_{t=1}^{T} K^{-}\left(\frac{\tilde{Z}_t - z}{b}\right) G(y_t)}{\sum_{t=1}^{T} K^{+}\left(\frac{\tilde{Z}_t - z}{b}\right)} = \frac{\hat{P}^{+}(z)}{\hat{f}(z)} \qquad (5.2)$$

$$\hat{m}^{-}(z) = \frac{\sum_{t=1}^{T} K^{-}\left(\frac{\tilde{Z}_t - z}{b}\right) G(y_t)}{\sum_{t=1}^{T} K^{-}\left(\frac{\tilde{Z}_t - z}{b}\right)} = \frac{\hat{P}^{-}(z)}{\hat{f}(z)} \qquad (5.3)$$

These estimators are referred to as the right and left side estimators, respectively, and they are the usual Nadaraya-Watson[2] kernels. At points of continuity of S(Z), all three estimators will

[2] The Nadaraya-Watson kernal estimator of m(x) is

converge to the same value, m(z). At points where S(Z) is not continuous, all three estimators will converge to different values .

Using the same idea, it is possible to decompose the conditional variance as:

$h_t = var(y_t / F_{t-1}) = E[y_t^2 / F_{t-1}] - \{ E[y_t / F_{t-1}]\}^2$ (5.4)

where F_{t-1} is the sigma-field parameter generated by past information and estimate each term by left-sided and right-sided kernel estimates. The appropriate estimator is:

$$\hat{h}^{\pm}(z) = \frac{\sum_{t=1}^{T} K^{\pm}\left(\frac{\tilde{Z}_t - z}{b}\right) y_t^2}{\sum_{t=1}^{T} K^{\pm}\left(\frac{\tilde{Z}_t - z}{b}\right)} - \left\{ \frac{\sum_{t=1}^{T} K^{\pm}\left(\frac{\tilde{Z}_t - z}{b}\right) y_t}{\sum_{t=1}^{T} K^{\pm}\left(\frac{\tilde{Z}_t - z}{b}\right) y_t} \right\}^2 = \quad (5.5)$$

$$= \frac{\hat{P}_2^{\pm}(z)}{\hat{f}(z)} - \left[\frac{\hat{P}_1^{\pm}}{\hat{f}(z)}\right]^2 = \hat{m}_2^{\pm}(z) - \left[\hat{m}_1^{\pm}(z)\right]^2$$

The asymptotic distribution of (5.5) is found in Perron (1999) and the behaviour for each $\hat{m}_1^{\pm}(z)$ and $\hat{m}_2^{\pm}(z)$ can be found in Delgado and Hidalgo (1986). The framework proposed is applicable to multiple jumps and these can be estimated sequentially. Suppose there are M jumps.

Define the following process:

$$\hat{m}(x) = \frac{\sum_{t=1}^{T} K\left(\frac{X_t - x}{b}\right) G(y_t)}{\sum_{t=1}^{T} K\left(\frac{X_t - z}{b}\right)} = \frac{P(x)}{\hat{f}(x)}$$

, where K, is a kernel, b is the usual bandwidth parameter

$$\Delta(z) = \hat{h}^{+}(z) - \hat{h}^{-}(z) \quad (5.6)$$

Defining the first jumps as $\hat{z}_1 = \arg\max_{overZ} \left[\hat{\Delta}(z)\right]^2$. After the first jump is estimated by $\hat{z}_1$, the second point estimate is obtained in a similar fashion, and so on. Perron (1999) has shown that the rate of convergence is slower than the usual root-T obtained in parametric models. However, the absence or presence of other jumps does not change the behaviour of the break point estimates due to the local nature of kernel estimation. If we accept that the mean function does not have any discontinuity, the problem simplifies since the estimates of the mean cancel out from (5.6) $\hat{m}_1^{+} = \hat{m}_1^{-} = \hat{m}_1$. In this case, we can look for jumps in the conditional variance by looking for jumps in y_t^2.

5.2.1 - Estimation Results of Jumps in the Volatility of Stock Returns

In this section we present results from the application of Perron's (1999) methodology to stock returns in emerging markets. We use a common GARCH(1,1) model:

$$y_t = \mu + h_t^{\frac{1}{2}} \varepsilon_t$$

$$h_t = \omega_t (1 - \alpha - \beta) + \alpha (y_{t-1} - \mu)^2 + \beta h_{t-1}$$

The bandwidth is choses as $b_j = c\hat{\sigma}_j T^{\frac{-1}{p+5}}$, where c is the bandwidth constant, $\hat{\sigma}_j$ is the estimated standard deviation of variable j. T is the sample size, and p is the lag length. Three values of c are allowed: 0.8, 1 and 1.2. Moreover, there is a data-determined selection rule for minimizing a variation of the following cross-validation criterion:

$$CV(c) = \frac{1}{T}\sum_{t=1}^{T}(y_t^2 - \hat{m}_{2t}^{-})^2 + \frac{1}{T}\sum_{t=1}^{T}(y_t^2 - \hat{m}_{2t}^{-})^2$$

where $\hat{m}_{2t}^{+}$ (respectively $\hat{m}_{2t}^{-}$) is the right side (respectively left side) estimate of $E_{t-1}(y_t^2)$. The criterion uses only the fit of the second moment of y_t to choose the bandwidth. The bandwidth constant is allowed to vary between 0.8 and 1.2 with a step of 0.1. Finally, the one-sided kernel is $k_+(x) = x(3-x)e^{-x} \quad (x \geq 0)$, while the two-sided kernel is Gaussian. All tests are carried out at a 5% significance level with normality imposed and ten percent of the observations deleted at the beginning and end of the sample.

In this section we apply the jump detection procedure to the series of daily returns for each of the countries under study. The moments $E(\varepsilon_t^3)$ and $E(\varepsilon_t^4 - 1)$ are estimated by using a two-sided kernel over the data remaining after deleting twice the bandwidth on each side of the currently estimated jump. The use of the two-sided kernel provides more reliable estimates of these moments. We allow the constant to vary between 0.8 and 1.2 in increments of 0.1. The results are quite interesting in two respects. First, the procedure did not run out of observations at all lags, and some jumps dates seem to be recurring in different countries. Second, some countries show more jumps than others.

The results for the detection of jumps are reported in Tables 5.1 to 5.9 for 4 lag lengths considered with the bandwidth chosen by the cross validation criterion described above. For each of the lags considered, we present the date the jump occurred, the size of the jump and the associated p-value. In what follows, almost all p-values are effectively 0. We report not only the breaks that reject the null strongly but also the next borderline break and its p-value in each case, so that the reader may judge the gap between the included and the excluded dates. We can see that the sizes of shocks vary a great deal from one country to another, with Brazil showing the biggest ones. This might be explained by the succession of economic plans since 1986.

It is interesting to note that some markets, including Brazil's, show a large number of jumps regardless of the number of lags used. Indeed, Brazil has experienced many interventions and considerable disruptions associated with its economic plans. The period June-July 1975 was one of sudden changes in volatility. This coincided with an abrupt increase in the inflation rate, which led the Brazilian Central Bank and the Ministry of Finance to announce a series of measures. Also, the period of June-July 1984 coincided with the election of the first civilian government following years of military dictatorship. The years 1988, 1989, 1990 and 1991 were marked by very high inflation (almost hyperinflation) and increases in the volatility of markets. These results appear to be consistent with those found by Perron (1999) for S&P returns, in the sense that he found a number of similar jumps in the U.S. economy.

Taiwan is also a country that has experience a number of large jumps in market volatility. For Hong-Kong, August 1980 was the most important period of turbulence. For Mexico, we only found evidence for one lag with three jumps. One of the jumps occurred before the debt crisis, and the other two were related to political crises. For Korea, 1980 and 1988 were the years where volatility increased. For Thailand, 1979-1980 was a period of instability, and we can identify various jumps in this period. Even though Brazil and Taiwan seem to have had more unstable periods than the other countries under consideration, the size of the jumps were bigger in Brazil than in Taiwan, but smaller than those found by Perron (1999) (see table 5.8). He found a jump in late 86 that preceded the October 1987 crash by a few months. This makes sense intuitively, as does the August 1990 jump that coincided with Iraq's invasion of Kuwait.

Table 5.1 - Results from estimation of jumps in conditional variance - Brazil

LAG 0 (c = 0.8)			LAG 1 (c = 1.2)			LAG 2 (c = 0.8)			LAG 3 (c = 0.9)		
Date	Size	p-value	Date	Size	p-value	Date	Size	p-value	Date	Size	p-value
0.7407 07.02.88	-0.456	0.00	0.3082 13.05.76	-693.0	0.00	0.6214 11.11.84	39228	0.00	0.8200 02.04.90	19421	0.00
0.8499 23.01.91	-0.277	0.00	0.6188 16.10.84	105.8	0.00	0.2778 17.07.75	-10450	0.00	0.6069 20.06.84	11248	0.00
0.6080 01.07.84	0.686	0.00	0.8474 12.12.90	-5.152	1.00	0.8063 17.11.89	1913	0.00	0.1533 29.02.72	-762.7	0.00
0.2688 19.04.75	0.0137	0.00	-	-	-	0.4455 02.02.80	254.9	0.00	0.4100 19.12.81	-13.62	0.1069
0.4496 21.05.82	-0.0390	0.083	-	-	-	0.00014 02.05.68	0.000	1.00	-	-	-

Table 5.2 - Results from estimation of jumps in conditional variance – Hong-Kong

LAG 0 (c = 1.2)			LAG 1 (c = 0.8)			LAG 2 (c = 1.10)			LAG 3 (c = 1.10)		
Date	Size	p-value	Date	Size	p-value	Date	Size	p-value	Date	Size	p-value
0.5547 17.06.86	0.022	0.00	0.2728 21.08.80	197.0	0.00	0.8200 09.12.91	3030	0.00	0.2712 09.08.80	-2213	0.00
0.6791 11.01.89	-0.011	0.00	0.4376 15.01.84	-219.6	0.00	0.1758 20.08.78	-2403	0.00	0.6598 18.08.88	559.8	0.00
0.4297 12.10.83	-0.031	0.743	0.6921 18.04.89	20.71	0.00	0.5932 15.03.87	-381.1	0.717	0.00018 19.05.75	0.000	1.00
-	-	-	0.8153 18.06.90	106.7	1.00	-	-	-	-	-	-

Table 5.3 - Results from estimation of jumps in conditional variance - Mexico

LAG 0			LAG 1 (c = 1.0)			LAG 2			LAG 3		
Date	Size	p-value	Date	Size	p-value	Date	Size	p-value	Date	Size	p-value
0.5458 25.03.88	0.195	0.073	0.5004 31.07.86	4755	0.00	0.5150 12.12.86	-237.4	1.00	0.5069 07.07.86	481.1	1.00
-	-	-	0.7545 11.02.90	345.1	0.00	-	-	-	-	-	-
-	-	-	0.1829 06.02.81	7.57	0.00	-	-	-	-	-	-
-	-	-	0.3417 12.03.82	0.658	1.00	-	-	-	-	-	-

Table 5.4 - Results from estimation of jumps in conditional variance - Korea

LAG 0 (c = 0,8)			LAG 1 (c = 1.10)			LAG 2 (c = 1.10)			LAG 3 (c = 1.10)		
Date	Size	p-value	Date	Size	p-value	Date	Size	p-value	Date	Size	p-value
0.7286 15.04.90	-0.0051	0.271	0.3519 11.02.84	-98.09	0.00	0.5854 31.10.88	36688	0.00	0.8352 17.11.93	249.9	0.00
-	-	-	0.5561 28.03.88	1.85	0.00	0.1876 17.10.80	-59.20	0.00	0.2667 23.05.82	-244.8	0.00
-	-	-	0.1574 08.03.80	-13.42	0.00	0.833 15.09.93	1.999	1.00	0.5929 25.12.88	86.8	0.00
-	-	-	0.777	-0.120	0.051	-	-	-	0.0002 15.04.75	0.00	1.00

Table 5.5 - Results from estimation of jumps in conditional variance - Malaysia

LAG 0 (c = 0.8)			LAG 1 (c = 1.1)			LAG 2 (c = 1.2)			LAG 3 (c = 0.8)		
Date	Size	p-value	Date	Size	p-value	Date	Size	p-value	Date	Size	p-value
0.6659 27.07.87	-0.010	0.00	0.3280 11.03.81	-42.77	0.00	0.7805 24.09.89	362.90	0.00	0.7654 12.06.89	2184	0.00
0.3548 19.12.81	0.011	0.00	0.7376 02.12.88	3.72	0.00	0.2870 01.06.80	-437.64	0.00	0.2975 13.08.80	-1683	0.00
0.2703 01.03.80	-0.021	0.067	0.5705 08.10.85	2.59	0.00	0.5457 26.05.86	3.429	1.00	0.6038 25.05.86	869.9	0.00
-	-	-	0.1629 05.03.78	0.802	1.00	-	-	-	0.0002 12.02.75	0.00	1.00

Table 5.6 - Results from estimation of jumps in conditional variance - Thailand

LAG 0 (c = 1.2)			LAG 1 (c = 1.20)			LAG 2 (c = 0.9)			LAG 3 (c = 1.10)		
Date	Size	p-value	Date	Size	p-value	Date	Size	p-value	Date	Size	p-value
0.7522 26.12.88	-0.012	0.00	0.2381 06.06.79	-531.4	0.00	0.6405 29.11.86	5338	0.00	0.2375 02.06.79	-792	0.00
0.5945 20.01.86	-9.06	0.00	0.6380 12.11.86	206.6	0.00	0.2880 10.05.80	-150.6	0.00	0.5920 15.12.85	901	1.00
0.2367 27.05.79	0.0001	0.00	0.8423 25.04.93	3.1488	1.00	0.4846 10.04.83	3.31	1.00	-	-	-
0.4024 30.06.81	0.0001	0.00	-	-	-	-	-	-	-	-	-
0.002 30.03.75	0.0000	1.00	-	-	-	-	-	-	-	-	-

Table 5.7 - Results from estimation of jumps in conditional variance - Taiwan

LAG 0 (c = 0.8)			LAG 1 (c = 0.9)			LAG 2 (c = 0.8)			LAG 3 (c = 1.0)		
Date	Size	p-value	Date	Size	p-value	Date	Size	p-value	Date	Size	p-value
0.8220 01.03.93	0.0139	0.00	0.4602 03.03.85	-595.6	0.00	0.6370 15.11.89	1531	0.00	0.7897 14.06.92	3601	0.00
0.7402 11.05.91	0.0447	0.00	0.2507 16.07.80	-104.8	0.00	0.4600 02.03.85	-965.7	0.00	0.1765 25.11.78	-9.324	0.00
0.6145 31.07.88	-0.1589	0.00	0.6199 12.09.88	178.9	0.00	0.1948 22.04.79	-805.2	0.00	0.4188 02.06.84	41.38	0.193
0.4840 11.09.85	0.0206	0.00	0.7724 26.01.92	4.95	0.00	0.8490 05.10.93	326.9	0.00	-	-	-
0.2526 31.07.80	-0.0163	0.00	0.00017 15.04.78	0.00	1.00	0.00017 15.04.78	0.00	1.00	-	-	-
0.3746 12.04.83	-0.0101	0.00	-	-	-	-	-	-	-	-	-
0.1707	0.010	1.00	-	-	-	-	-	-	-	-	-

Table 5.8 - Results from estimation of jumps in conditional variance - Argentina

LAG 0 (c = 1.2)			LAG 1 (c = 1.2)			LAG 2 (c = 1.2)			LAG 3 (c = 1.2)		
Date	Size	p-value	Date	Size	p-value	Date	Size	p-value	Date	Size	p-value
0.1505 16.12.86	-0.5606	0.000	0.2567 07.07.86	-755	0.00	0.7657 27.11.87	17349	0.00	0.7668 03.08.90	424.6	0.00
0.7082 07.03.89	0.0024	0.000	0.7410 23.08.90	290	0.00	0.2533 15.11.91	64.99	0.00	0.1947 01.03.87	-27.38	0.4792
0.3067 17.06.83	0.6001	0.000	0.4601 12.05.85	-0.139	1.00	0.4968 12.03.84	0.0848	1.00	-	-	-
0.4687 17.06.86	0.0825	0.1225	-	-	-	-	-	-	-	-	-

Table 5.9 - Results from estimation of jumps in conditional variance - USA

LAG 0 (c = 1.2)			LAG 1 (c = 0.9)			LAG 2 (c = 1.2)			LAG 3 (c = 1.1)		
Date	Size ($x10^7$)	p-value	Date	Size ($x10^4$)	p-value	Date	Size ($x10^4$)	p-value	Date	Size ($x10^3$)	p-value
0.4099 16.12.86	6.878	0.001	0.3829 07.07.86	32.8	0.00	0.4652 27.11.87	129.8	0.00	0.6229 03.08.90	25.15	0.00
0.5399 07.03.89	-0.452	0.001	0.6262 23.08.90	-20.7	0.00	0.6988 15.11.91	-0.038	0.00	-	-	-
0.2038 17.06.83	0.686	0.54	0.7720 16.02.93	4.1	-	-	-	-	-	-	-

The results from our examination of daily stock returns in emerging markets show that sudden changes in volatility have occurred for all countries in our sample, without exception. The findings that the volatility of financial markets cannot be described as a smooth function of lagged returns casts doubt on the practice of fitting GARCH-type models. It remains to be seen how much is lost by neglecting to model this feature of the data. It may very well be that GARCH models can still provide a good first approximation to the behaviour of the data.

Tables 5.1 to 5.8 show that it is possible to statistically detect changes in stock market return volatility processes. This finding poses a number of interesting questions. For example, how common are high volatility states? Do periods of high volatility coincide across countries? Is it possible to statistically identify groups of countries that jointly experience high volatility? A visual inspection of tables 5.4 (Korea) and 5.6 (Thailand) show that jumps occurred on December 25, 1988 in Taiwan and December 26, 1988 in Thailand. Also, comparing table 5.5 (Malaysia) to the others, we see that a jump happened in Malaysia on December 2, 1988. These results provide some preliminary evidence of (roughly) coincident volatility switches among countries.

5.3 - Switching Regimes in Volatility Models

5.3.1 - Markov Chains

In the previous section, we identified a number of discrete changes in market volatility that occurred in our sample of countries. We will now investigate the matter more fully. A natural extension of our analysis would be to attempt to deal with discrete changes in volatility.

How should we model changes in the process followed by a particular time series, such as the one shown by Figure 5.1? For the data plotted in this figure, a simple approach would be to assume that the constant term in the autoregression changed around observation 1000. For data prior to observation 1000, we could use a model such as:

$$y_t - \mu_1 = \phi(y_{t-1} - \mu_1) + \varepsilon_t \qquad (5.7)$$

while data after observation 1000 could then be described by

$$y_t - \mu_2 = \phi(y_{t-1} - \mu_2) + \varepsilon_t \qquad (5.8)$$

where $\mu_2 < \mu_1$. This specification would appear to provide a good description of the data in Figure 5.1. On the other hand, how can we use it for forecasting if the series keeps changing? If the process changed in the past it could certainly change again in the future, which would have to be taken into account in forecasting. The change in regime should not be regarded as the outcome of a perfectly deterministic event but as a random variable. A complete time series model would include a description of the probability law governing the change from μ_1 to μ_2. These observations suggest that we can consider the process to be influenced by an unobserved random variable s_t^*, which is commonly called the state or regime the process was in at date t. If $s_t^* = 1$, then the process is in regime 1, while $s_t^* = 2$, means that the process is in regime 2., so that we can rewrite (5.7) and (5.8) as:

$y_t - \mu_{s_t^*} = \phi(y_{t-1} - \mu_{s_{t-1}^*}) + \varepsilon_t$ (5.9) where $\mu_{s_t^*}$ indicates μ_1 when $s_t^* = 1$ and indicates μ_2 when $s_t^* = 2$.

We then need a description of the time series process for the unobserved variable s_t^*. Since s_t^* only takes on discrete values (in this case either 1 or 2), this model will differ slightly from those with continuous-valued random variables seen in the previous section.

Financial markets sometimes appear quite calm and at other times highly volatile. Describing how this volatility changes over time is important. Hamilton and Susnel (1994) argue that a promising alternative is to allow for the possibility of sudden. discrete changes in the parameter values of an ARCH(q) process, as in the Markov-switching model. The simplest time series model for a discrete-valued random variable is a Markov-chain.

We introduce switching volatility to capture possible exogenous shocks. An exogenous shock could be discrete if it resulted from a change in the underlying regime, or continuous if it took the form of a "market-news" event. Models of switching regimes follow different dynamics in each regime. Imagine the most basic model with a constant, where both the constant and the volatility of the process suffer from switching changes in regime as a function of a first order Markov Chain. The model is then defined as:

$$y_t = \mu_{st} + u_t \quad ,where \quad u_t \sim N(0, \sigma_{st}) \qquad (5.10)$$

Observing (5.10)[3], both the mean and the variance in this model are influenced by the variable s_t. This random variable is called a state variable. Let y_t denote the daily stock return in percent terms. For example, y_t = -2,0 would mean that stock prices fell 2% on day t. If $s_t = 1$, the prevailing regime would be the first, if $s_t = 2$, the regime would be the second, and so on. This state variable s_t follows a first order Markov process and can only assume the integer values {1,2,3,.....,N}. The probability that s_t equals some particular value j depends on the past only through the most recent value of s_t in the previous state: s_{t-1}:

$$P\{s_t = j / s_{t-1} = i,\ \ s_{t-2} = y, \ldots\ldots s_{t-k} = r\} = P\{s_t = j / s_{t-1} = i\} = p_{ij} \qquad (5.11)$$

Such a process is called a N-state Markov chain with transition probabilities $\{p_{ij}\}_{i,j=1,2,\ldots,N}$. The transition probability p_{ij} gives the probability that state i will be followed by state j. Note that we must have

$$p_{i1} + p_{i2} + \ldots\ldots + p_{iN} = 1 \qquad (5.12)$$

Given that we are in state 1, there will be certain probabilities of evolving to k different states of the nature. The economic *rationale*[4] is that the economy may be in different economic states. For instance, a stock market may have three regimes: one where stocks rise in a consistent fashion, with high returns and low volatility; another where stocks move neither up nor down, but where there are low returns and higher variance; and one last regime where there is a crisis, with negative returns and very high volatility.

It is often convenient to collect the transition probabilities associated with each regime in a (N x N) matrix P known as the transition matrix:

$$P = \begin{bmatrix} F & p_{12} & \ldots\ldots p_{1n} \\ F_{2} & p_{22} & \ldots\ldots p_{N2} \\ \ldots & \ldots & \ldots \\ F_{2} & p_{31} & \ldots\ldots p_{N1} \end{bmatrix}$$

---

[3] In the models we investigate, the process u_t that is described either by the ARCH or the GARCH process is the residual from a first order autoregression for stock returns $y_t = c + \theta_0 y_{t-1} + \xi_t$.

[4] A better rationale is that the k-discrete state model may serve as an approximation to a richer underlying model, being better than a one regime model, without being literally true.

The row j, column i element of P is the transition probability p_{ij}. For example, the element in row 2, column 1 gives the probability that state 1 will be followed by state 2. In the event that N = 2, we have the following Markov chain:

$$\begin{bmatrix} p_{11} & 1-p_{22} \\ 1-p_{11} & p_{22} \end{bmatrix}$$

Suppose that $p_{11} = 1$, so that the matrix P is upper triangular. Then once the process enters into state 1 there is no possibility of going back to state 2. In such a situation we would call state 1 an absorbing state and the Markov chain reducible[5]. A Markov chain that is not reducible is said to be irreducible. For example, take a two-state chain where $p_{11} < 1$ and $p_{22} < 1$

An important concept related to Markov chains is ergodic probability. This is the probability associated with each steady state. In the specific case of a two-state Markov chain, the eigenvalues of the transition matrix P satisfy[6]

$$|P - \lambda I_N| = 0$$

$$\det \begin{bmatrix} p_{11}-\lambda & 1-p_{22} \\ 1-p_{11} & p_{22}-\lambda \end{bmatrix} = 0$$

so we can write

$$= (p_{11} - \lambda)(p_{22} - \lambda) - (1- p_{11})(1- p_{22}) =$$
$$= - p_{11}\lambda - p_{22}\lambda - \lambda^2 - 1 + p_{22} - p_{11} \quad =$$
$$= (\lambda - 1)(\lambda - p_{11} - p_{22} + 1)$$

Thus, the eigenvalues for a two-state chain are given by $\lambda_1 = 1$ and $\lambda_2 = - 1 + p_{11} + p_{22}$ [7]. The eigenvectors associated with λ_1, for the two-state chains turn out to be

---

[5] More generally, an N-state Markov chain is said to be reducible if there is a way to label the states, that is a way to choose to call state 1, which to call state 2, and so on. This characteristics is not very common in financial series, for instance, in the previous example, if the market is rallying, there is no reason for it to continue always rising.

[6] Consider an N-state irreducible Markov chain with transition matrix P. Suppose that one of the eigenvalues of P is unity and that all other eigenvalues of P are inside the unit root interval. Then the Markov chain is said to ergodic. More generally the eigenvalues of the transition matrix P for any N-state Markov chain are found from the solutions to $| P - \lambda I_N | = 0$. Hamilton (1994) shows that the vector of ergodic probabilities can also be viewed as the uncoditional probability of each of the N-different sates. Also an ergodic Markov chain is a covariance-stationary process.

[7] The second eigenvalue will be inside the unit root as long as $0 < p_{11} + p_{22}$. Thus we would have a two-state markov chain ergodic provided that $p_{11} < 1$ and $p_{22} < 1$ and $0 < p_{11} + p_{22}$.

$$\begin{bmatrix} (1-p_{22})/(2-p_{11}-p_{22}) \\ \\ (1-p_{22})/(2-p_{11}-p_{22}) \end{bmatrix}$$

From here we can obtain the ergodic probabilities (or unconditional probabilities) for thetwo-state model. Thus, the unconditional probability that the process will be in regime 1 at any given time is

$$P\{s_t = 1\} = \frac{1-p_{22}}{2-p_{11}-p_{22}}$$

and the unconditional probability that the process will be in regime 2 at any given time is just 1 minus the magnitude above, that is:

$$P\{s_t = 2\} = \frac{1-p_{11}}{2-p_{11}-p_{22}}$$

In switching regime models it is useful to know the average duration of each regime given that we are in a specific state. If we define D as the duration in a determined state, we can write the following from the definition given above:

D = 1, if S_t = j and S_{t+1} = j ; P{D = 1} = $(1 - p_{11})$
D = 2, if S_t = S_{t+1} = j and S_{t-2} = j ; P{D = 2} = $p_{jj}\,(1 - p_{jj})$
D = 3, if S_t = S_{t+1} = S_{t+2} = j and Ss_{t+3} = j ; P{D = 3} = $p_{jj}{}^{2}(1 - p_{jj})$ and so on
....

The expected duration of state j can de derived by:

$$E(D) = \sum_{j=1}^{\infty} jP[D = j]$$
$$= P[S_{t+1} \neq j\, S_t = j] + 2P[S_{t+1} = j, S_{t+2} \neq j | S_t = j] + ...$$
$$= (1-p_{jj}) + 2p_{jj}(1-p_{jj}) + 3p_{jj}^2(1-p_{jj}) + ...$$
$$= \frac{1}{1-p_{jj}} \quad (5.13)$$

Equation (5.13) is very useful for collecting information on switching regimes. For example, when we estimate a model for changes in GDP we can observe the average duration of the cycle of growth and recession.

5.3.2. - SWARCH and SWGARCH

Many researchers have suggested that the poor forecasting performance and spuriously high persistence of ARCH models may be related to structural changes in the ARCH process, which produce so-called long-memory effects in volatility. This is related to Perron's (1989) observation that changes in regime may give the spurious impression of unit roots in characterisations of the level of a series. Cai (op. cit.) in particular noted that volatility in Treasury bill yields appears to be much less persistent when one models changes in parameters through a Markov-switching process. It seems promising to investigate whether a similar result might characterise stock returns.

A natural way to control for structural breaks is to use models of Markovian switching states for the variance. Several econometric models were developed in recent years to account for changes in states through a first order Markov chain.

A common empirical finding on volatility is the high persistence of the conditional variance. For example, Brenner et al. (1996) estimate a persistence parameter of 0.92 using weekly three-month U.S T-Bill data. Ball and Torous (1995) report a persistence parameter of 0.9152 using monthly one-month U.S. T-Bill data. Anderson and Lund (1997) report a persistence parameter of 0.98 for weekly data. Such long memory in volatility, as well as high persistence in interest rates or stock prices. might lead to biased forecasts.

High persistence, or long memory. in the conditional variance implies that shocks to the conditional variance do not die out quickly. That is, current information has a significant effect on the conditional variance for future horizons. Engle and Bollerslev (1986) proposed the IGARCH model to address this issue. Under an IGARCH model, shocks to the conditional variance never disappear.

There is some evidence that such long memory could be related to structural changes in the variance process that occurred during the sample period. Lamoreux and Lastrapes (1990) find that a single-regime GARCH specification leads to spurious high persistence in the presence of structural breaks. By allowing for the possibility of regime switching. the high persistence and long memory observed in single regime models no longer appears to be valid. Similar results have been documented by Hamilton and Susnel (1994), Cai (1994) and So, Lam and Li (1998). Ball and Torous (1995) consider regime switching in the variance but do not model persistence in

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 \quad (5.15)$$

variance. When low volatility can switch randomly between different regimes, and when each regime is associated with its own mean and variance parameter, we may find high persistence when we average data from different regimes. In our presentation we follow Hamilton and Susnel (1994), but we extend their model to allow for switching GARCH based on Duecker (1997) in order to deal with long memory effects in volatility.

Hamilton (1989) suggested the following regime-switching model for the conditional mean

$$y_t = \mu_{s_t} + \tilde{y}_t \quad (5.14)$$

Here μ_{s_t} denotes the parameter μ_1 when the process is in the regime represented by s_t = 1, while μ_{s_t} indicates μ_2 when s_t = 2. and so on. The variable $\tilde{y}_t$ is assumed to follow a zero-mean qth-order autoregression:

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \phi_2 \tilde{y}_{t-2} + \ldots\ldots + \phi_{2q} \tilde{y}_{t-q} + \varepsilon_t$$

The idea behind this specification is that ocasional abrupt shifts in the average level of y_t can be captured by the values of μ_{s_t}.

A natural extension of this approach to the conditional variance would be to model the residual u_t as Hamilton and Susnel (1994) did. They present a model of change in state for an ARCH, where we could also include the leverage effect along the lines of Glosten et al. (1989)[8]. This model is defined below, where u_t follows a standard ARCH (q) process and v_t is an i.i.d. sequence with zero mean and unit variance.

$$\xi_t = \sqrt{g_{s_t}}\, u_t \quad (5.16)$$

$$u_t = \sqrt{h_t}\, v_t \quad (5.17)$$

[8] Glosten's formulation takes in account leverage effects. A stock price decrease tends to increase subsequent volatility by more than would a stock price increase of the same magnitude. This is the so-called leverage effect. This specification follow Black (1976) and Nelson (1991).

Alternatively we could have Glosten's formulation that implies rewriting (5.18) as

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 - \delta d_{t-1} u_{t-1}^2 \quad (5.18')$$

where S_t assigns values 1,2,.....,k, and follows a S_t Markov Chain of first order. When k = 1, we have the following probabilities of transition: P{ S_t = 1 | S_{t-1} =1} = p; P{ S_t = 2 | S_{t-1} =1} = 1 - p; P{ S_t = 2 | S_{t-1} =2} = q and P{ S_t = 1 | S_{t-1} =2} = 1-q. However, g_{St} , is a function of S_t standardised for the first state (i.e. g_1 = 1). We then find that the conditional variance follows an ARCH(q) process, whereas u_t is multiplied by the constant $\sqrt{g_1}$ when it is state 1, by $\sqrt{g_2}$ when it is state 2, and so on. The main point is to model the changes in state as changes in the scale of the conditional variance process. In (5.18') d_{t-1} = 1 if $u_t > 0$ and d_{t-1} = 0 if $u_t < 0$.

In the absence of leverage effects (δ = 0), we will say that ξ_t follows a K-state q^{th}-order Markov-switching ARCH process, denoted SWARCH(K,q). In the presence of leverage effects, we will call it a SWARCH-L(K,q) specification. We investigate both Gaussian ($v_t \sim N(0,1)$) and Student's t (v_t distributed t with v degrees of freedom and unit variance) versions of the model.

These models all have the same problem common to all ARCH variants, namely the necessity of incorporating a large number of lags in order to describe the features of the model vis-à-vis GARCH (see Bollerslev (1986)).[9]

On the other hand, Duecker(1997) extends the model for four different specifications of GARCH. This procedure facilitates the estimation of the likelihood of the process without incurring in big losses. As such, we will be estimating equations (5.16), (5.17), (5.18) and[10]

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \quad (5.19)$$

The model comprising (5.16), (5.17) and (5.18) is called SWARCH(k,p) where k designates the number of states. If we use (5.18') we would have the SWARCH(k,p)-L model. In the same vein, the model comprising (5.16). (5.17) and (5.19) is called SWGARCH(k,p,q) where k designates the number of states. If we use (5.19') we would have the SWGARCH(k,p,q)-L model.

[9] This common empirical finding might be the result of estimation biases near the boundary in the QML-GARCH.

[10] We may also consider a slight different version of (5.9) called SWGARCH(k,p,q)-L that takes in account the leverage effect so common in the stock market $h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} - \delta d_{t-1} u_{t-1}^2 \quad (5.19')$

The estimation of the models SWARCH(k,q) and SWGARCH(k,p,q) is similar the estimation of GARCH models reported in the literature and is usually based on maximum likelihood estimation. For the normal distribution, the likelihood is given by[11]

$$L(\theta) = \sum_{t=1}^{T} \log f(y_t \mid s_t, s_{t-1}, \ldots, s_{t-pq}, Z_t; \theta)$$

$$f(y_t \mid s_t, s_{t-1}, \ldots, s_{t-k}, s_{t-pq}, Z_t) = \frac{1}{\sqrt{2\pi h_t(s_t, s_{t-1}, \ldots, s_{t-k}, s_{t-pq})}} \exp\left\{ \frac{-(y_t - Z_t'\beta)^2}{2h_t(s_t, s_{t-1}, \ldots, s_{t-k}, s_{t-pq})} \right\} \quad (5.20)$$

Or written differently

$$= -\left(\frac{T}{2}\right)\log(2\pi) - \left(\frac{1}{2}\right)\sum_{t=1}^{T} \log(h_t(s_t, s_{t-1}, \ldots, s_{t-pq}))$$

$$= -\left(\frac{1}{2}\right)\sum_{t=1}^{T} \frac{(y_t - Z_t'\beta)^2}{h_t(s_t, s_{t-1}, \ldots, s_{t-pq})} \qquad (5.21)$$

where pq = max{p,q}. Note that the conditional variance defined above is a function of all states.

We have also used a Student's t distribution for the estimation of the maximum-likelihood[12]. When the inference is based on information observed through date t it is called the 'filtered probability'. Alternatively, the full sample of observations can be used to construct the 'smoothed probability'.[13] In what follows we will report only the smoothed probability, as is usually done in the literature. When inference about a particular state of the process is based on the information through date t, not all information is used for the filtered probability.

---

[11] The normal distribution applies to the standardised innovations $z_t = \varepsilon_t/\sigma_t$ in the standard GARCH model.

[12] In this case v is supposed to be drawn from a t-Student distribution with v degrees of freedom normalised in order to have unit variance. In this case we would have (5.20) as

$$f(y_t \mid s_t, s_{t-1}, \ldots, s_{t-k}, s_{t-pq}, Z_t) = \frac{\Gamma(v-1/2)}{\Gamma(v/2)\sqrt{\pi}\sqrt{v-2}\; h_t(s_t, s_{t-1}, \ldots s_{t-pq})} \exp\left\{ 1 + \frac{(y_t - Z_t'\beta)^2}{(v-2)h_t(s_t, s_{t-1}, \ldots, s_{t-k}, s_{t-pq})} \right\}^{-(v+1)/2}$$

[13] The filtered probability $p(S_t, S_{t-1} / y_t, y_{t-1}, \ldots, y_{t-pq})$ stands for the conditional probability that the state in date t is S_t and that in date t-1 was S_{t-1}. These probabilities are conditional on the values of y observed through t (does not use all information). The smoothed probability $p(S_T / y_T, y_{T-1}, \ldots, y_{T-pq})$ are inferences about state t based on the data available through some future date T (end of the sample) using all data. The smoothed probability is an ex-post inference made about state t.

We fitted a variety of different SWARCH specifications to daily stock returns data for selected emerging markets. We estimated models with q = 2 (ARCH) terms and K = 3 states, with Normal and Student's t innovations and without the leverage parameter ξ. For the SWGARCH specification we have chosen p = q = 1 (ARCH and GARCH effects) and K = 2 states. For each model, the negative log-likelihood was minimized numerically using the optimisation program OPTIMUM in the GAUSS programming language.

Tables 5.10 to 5.25 show the results of the SWARCH and SWGARCH estimations.[14] A majority of the parameters are significant, including the parameters for the leverage effect.[15] The significance of these models indicates that they are superior. We based the choice of the models shown here on Akaike's Information Criterion (AIC) and Schwarz's Information Criterion (BIC). We report these in Tables 5.10 to 5.25, below. Our findings are similar to others in the literature. Standard errors are shown in parentheses. The tables also report the model selection statistics proposed by AIC and BIC. Based on these statistics, it is clear that the models accounting for leverage effects and using the Student's t distribution perform better.

In estimating these models we obtained two probabilities associated with different states of nature, that is, the smoothed probability and the filtered probability. To estimate the filtered probability we used information up to period t, whereas the whole series was used to estimate the smoothed probability. With this in mind, we see that the filtered probability is much more erratic than the smoothed one since it uses less information. We have estimated models with two and three regimes. For the SWARCH model, we used up to four lags but retained the specification SWARCH(3,2). As has already been stated, the GARCH specification can represent an ARCH structure with many lags in a parsimonious way. However, the estimation of SWGARCH models does pose a large computational burden since the rate of convergence is slower and we have to impose some constraints to make estimation tractable. We have worked with SWGARCH(2,1).

In all of the SWARCH/SWGARCH estimations we have allowed the distribution to be normal or Student's t and have also included the leverage effect. We can observe some interesting results in the tables below. With regard to the SWARCH models estimated for all countries, the scale parameter (g_2) that describes the moderate volatility was roughly between 2 and 4, independent of the alternative specifications used. This means that the moderate volatility is twice to four times as high as the low volatility. This is consistent with Hamilton and Susnel (1994), who found (g_2) to be 4.5. It is also interesting to note the scale parameter that describes

[14] We estimated only models with two and three states. For models SWARCH we estimated models with up to four lags. For models SWGARCH we have estimated only SWGARCH(1,1) models, and changing the distribution and using or not the leverage effect. This is explained by the model choice selection criteria on one hand and on the other hand the burdensome of estimating higher complex models. However, these specifications are the more common and found in the vast majority of studies.

[15] These models are highly non-linear and some times we had problems in getting standard errors as the inverse of the hessian sometimes was not positive definite. Multicolinearity could have led to this.

the high volatility (g_3). The estimates vary between 13.12 and 19.3 for all but two countries. Again this is consistent with Hamilton and Susnel (1994), who found (g_3) to be 13.8.

In what follows, we can make the overall remark that the Asian countries tend to produce smaller estimates of the parameters, while the Latin American countries tend to produce larger ones. This could be an indication of the higher volatility of Latin American markets relative to their Asian counterparts. For example, the constant term in the autoregression (c) is estimated between 0.02 and 0.06 with the exceptions of Thailand (-0.01) and Argentina (0.13). For the autoregressive parameter (θ_0), our estimates vary between 0.08 and 0.19 with the exceptions of Brazil (0.2615) and Mexico (0.2825).

When we turn to the parameters of the conditional variance (h_t), we find estimates for the ARCH parameters (α_1 and α_2) that follow the same pattern described above. It is indeed interesting that, for both parameters, the countries of Latin America are concentrated around the upper limit and the Asian countries around the lower limit.

We can still apply these general comments when we examine asymmetries (ξ) by estimating a SWARCH-L model except for the parameters α_0 (Hong-Kong and Malaysia) and α_1 (Taiwan). However, the general conclusion that Latin American countries show bigger effects remains valid. These results may be due to the greater degree of intervention in the Latin American economies compared with similar Asian countries, especially during the years in our sample.

With respect to the SWGARCH specifications, we must stress that the estimation of these models is more difficult, since there are not only more parameters to be estimated but also certain conditions that must for estimation to be feasible. For instance, we may have problems if $\alpha_1 + \beta_1 = 1$. In certain cases we were not able to obtain standard errors because convergence was not assured, so we used the eigenvalues as estimates to them.

Generally speaking the results follow what we have stated before. However, in the case of the SWGARCH estimation we can see the parameters varying a great deal across countries, much more so than those associated with the SWARCH estimates. Few studies have estimated SWGARCH effects across countries, so a comparison is not straightforward. However, Duecker (1997) used a SWGARCH model to assess volatility in the American market and found a lot of variation between different specifications. This reinforces the general idea that SWGARCH allows the parameters to vary a lot with compared to SWARCH.

Table 5.10 - SWARCH

Argentina										
Model		g_2	g_3	C	θ_0	α_0	α_1	α_2	ξ	v
SWARCH(3,2)		5.623	45.27	0.1339	0.08310	2.3575	0.06896	0.1133	-	-
AIC = 4.95	BIC = 4.99	(1.52)	(7.54)	(0.013)	(0.021)	(0.521)	(0.0154)	(0.021)	-	-
SWARCH(3,2)-L		5.698	45.97	0.1296	0.0821	2.31843	0.04031	0.11770	0.0691	-
AIC = 4.95	BIC = 4.99	(1.31)	(8.43)	(0.021)	(0.012)	(0.311)	(0.0127)	(0.065)	(0.0111)	-
SWARCH(3,2)-L –t		1.936	13.84	0.1370	0.0788	1.14826	0.14195	0.1687	0.0387	7.626
AIC = 4.95	BIC = 4.99	(0.52)	(6.51)	(0.041)	(0.013)	(0.214)	(0.0101)	(0.031)	(0.0107)	(1.587)

Table 5.11 - SWARCH

Brazil										
Model		g_2	g_3	c	θ_0	α_0	α_1	α_2	ξ	v
SWARCH(3,2)		2.642	13.40	0.1819	0.2615	0.8837	0.1385	0.1636	-	-
AIC = 4.41	BIC = 4.42	(0.33)	(4.51)	(0.052)	(0.043)	(0.204)	(0.2101)	(0.331)	-	-
SWARCH(3,2)-L		2.588	13.29	0.1731	0.2598	0.8801	0.09490	0.1603	0.10478	-
AIC = 4.41	BIC = 4.42	(0.42)	(7.42)	(0.062)	(0.053)	(0.214)	(0.0214)	(0.431)	(0.0307)	-
SWARCH(3,2)-L –t		4.351	13.15	0.1559	0.2649	0.9680	0.09744	0.1366	0.1266	9.7398
AIC = 4.39	BIC = 4.40	(1.12)	(5.32)	(0.070)	(0.035)	(0.297)	(0.0111)	(0.337)	(0.0127)	(2.274)

Table 5.12 – SWARCH

Mexico										
Model		g_2	g_3	c	θ_0	α_0	α_1	α_2	ξ	v
SWARCH(3,2)		0.249	117.0	0.0604	0.2825	0.8868	0.2351	0.1452	-	-
AIC =2.19	BIC = 2.21	(0.06)	(11.5)	(0.020)	(0.098)	(0.214)	(0.101)	(0.0251)		
SWARCH(3,2)-L		0.252	114.3	0.0570	0.2821	0.8578	0.1821	0.1434	0.1053	-
AIC =2.19	BIC = 2.21	(0.07)	(12.1)	(0.031)	(0.089)	(0.210)	(0.100)	(0.0312)	(0.0311)	
SWARCH(3,2)-L –t		2.806	23.04	0.0526	0.2924	0.2515	0.2172	0.1056	0.0704	6.028
AIC =2.15	BIC = 2.17	(0.92)	(6.15)	(0.091)	(0.092)	(0.251)	(0.141)	(0.041)	(0.0321)	(1.693)

Table 5.13 – SWARCH

Model		Thailand								
		g_2	g_3	c	θ_0	α_0	α_1	α_2	ξ	v
SWARCH(3,2)		2.166	18.92	-0.0115	0.0981	0.09925	0.3591	0.2886	-	-
AIC = 2.46	BIC = 2.48	(0.45)	(4.31)	(0.004)	(0.011)	(0.0128)	(0.0714)	(0.0651)	-	-
SWARCH(3,2)-L		2.166	18.92	-0.0115	0.0981	0.09925	0.3591	0.2886	5.6x10^{-11}	-
AIC = 2.46	BIC = 2.48	(0.50)	(4.27)	(0.004)	(0.013)	(0.0124)	(0.0704)	(0.0621)	(1.6x10^{-11})	-
SWARCH(3,2)-L –t		3.844	21.82	-0.0035	0.1033	0.1879	0.4428	0.1662	2.6x10^{-10}	3.120
AIC = 2.40	BIC = 2.42	(0.52)	(6.51)	(0.041)	(0.013)	(0.214)	(0.0101)	(0.031)	(1.1x10^{-11})	(1.151)

Table 5.14 – SWARCH

Model		Taiwan								
		g_2	g_3	c	θ_0	α_0	α_1	α_2	ξ	v
SWARCH(3,2)		3.655	13.12	0.0664	0.0864	0.7829	0.02455	0.0920	-	-
AIC = 3.47	BIC = 3.49	(0.45)	(2.54)	(0.0157)	(0.0241)	(0.221)	(0.0051)	(0.0147)	-	-
SWARCH(3,2)-L		3.670	13.29	0.0614	0.0864	0.7846	2.8x10^{-09}	0.0940	0.0795	-
AIC = 3.47	BIC = 3.49	(0.47)	(2.65)	(0.0141)	(0.0240)	(0.224)	(1.0x10^{-09})	(0.0150)	(0.0125)	-
SWARCH(3,2)-L –t		3.938	14.96	0.0599	0.0817	0.7944	1.4x10^{-08}	0.0972	0.0806	16.90
AIC = 3.42	BIC = 3.46	(0.53)	(2.71)	(0.0122)	(0.0222)	(0.235)	(1.0x10^{-07})	(0.0174)	(0.0121)	(3.54)

Table 5.15 – SWARCH

Model		Korea								
		g_2	g_3	c	θ_0	α_0	α_1	α_2	ξ	v
SWARCH(3,2)		3.780	19.30	0.0336	0.1361	0.1933	0.0122	0.0681	-	-
AIC = 2.70	BIC = 2.72	(0.62)	(2.91)	(0.010)	(0.0412)	(0.021)	(0.003)	(0.0014)	-	-
SWARCH(3,2)-L		3.800	19.39	0.0332	0.1357	0.1927	0.0019	0.0696	0.025	-
AIC = 2.70	BIC = 2.72	(0.66)	(3.04)	(0.012)	(0.0410)	(0.028)	(0.004)	(0.0012)	(0.0012)	-
SWARCH(3,2)-L –t		3.408	10.35	0.0215	0.1430	0.1778	0.0791	0.1366	0.0315	8.20
AIC = 2.69	BIC = 2.71	(0.61)	(2.01)	(0.011)	(0.0302)	(0.018)	(0.0101)	(0.0121)	(0.0021)	(1.05)

Table 5.16 – SWARCH

Hong-Kong										
Model		g_2	g_3	c	θ_0	α_0	α_1	α_2	ξ	v
SWARCH(3,2)		2.848	18.53	0.1162	0.1001	0.8721	0.0241	0.1016	-	-
AIC = 3.46	BIC = 3.48	(0.45)	(4.01)	(0.011)	(0.0203)	(0.210)	(0.005)	(0.012)	-	-
SWARCH(3,2)-L		2.948	19.48	0.1063	0.1101	0.6715	0.0291	0.1146	0.1028	-
AIC = 3.46	BIC = 3.48	(0.47)	(4.07)	(0.013)	(0.0214)	(0.125)	(0.004)	(0.021)	(0.0021)	-
SWARCH(3,2)-L –t		2.724	10.90	0.1009	0.1028	0.7102	0.0291	0.1146	0.1028	9.79
AIC = 3.45	BIC = 3.48	(0.46)	(4.04)	(0.012)	(0.0210)	(0.128)	(0.004)	(0.021)	(0.0021)	(2.54)

Table 5.17 – SWARCH

Malaysia										
Model		g_2	g_3	c	θ_0	α_0	α_1	α_2	ξ	v
SWARCH(3,2)		2.731	16.59	0.0211	0.1928	0.3338	0.12289	0.0973	-	-
AIC = 3.17	BIC = 3.19	(0.36)	(4.12)	(0.006)	(0.0374)	(0.047)	(0.042)	(0.012)	-	-
SWARCH(3,2)-L		2.461	8.96	0.0140	0.1804	0.4574	0.1845	0.1324	$3.0\text{x}10^{-09}$	-
AIC = 3.17	BIC = 3.19	(0.28)	(3.14)	(0.004)	(0.0381)	(0.041)	(0.041)	(0.045)	$(1.0\text{x}10^{-09})$	-
SWARCH(3,2)-L –t		2.461	8.95	0.0140	0.1804	0.4575	0.1845	0.1324	$2.9\text{x}10^{-09}$	5.46
AIC = 2.40	BIC = 2.42	(0.27)	(3.12)	(0.004)	(0.0380)	(0.040)	(0.041)	(0.052)	$(1.0\text{x}10^{-09})$	(1.25)

Table 5.18 – SWGARCH

Argentina								
Model		g_2	c	α_0	α_1	β_1	ξ	v
SWGARCH(2,1,1)		0.0116	0.1889	0.00533	0.00955	0.9232	-	-
AIC = -0.26	BIC = -0.25	(0.005)	(0.061)	(0.0012)	(0.0012)	(0.245)	-	-
SWGARCH(2,1,1)-L		0.0166	0.1262	$3.18\text{x}10^{-13}$	$7.6\text{x}10^{-13}$	0.8981	2.014	-
AIC = -0.27	BIC = -0.24	(0.004)	(0.041)	$(1.2\text{x}10^{-13})$	$(3\text{x}10^{-13})$	(0.245)	(0.821)	-
SWGARCH(2,1,1)-L		317.33	0.0353	0.03766	0.03259	0.8951	0.8442	2
AIC = -0.28	BIC = -0.26	(12.25)	(0.001)	(0.0005)	(0.0021)	(0.245)	(0.241)	(0.512)

Table 5.19 – SWGARCH

Brazil								
Model		g_2	c	α_0	α_1	β_1	ξ	v
SWGARCH(2,1,1)		0.0475	0.0739	0.1398	0.1124	0.9289	-	-
AIC = -0.33	BIC = -0.33	(0.0125)	(0.014)	(0.051)	(0.052)	(0.211)	-	-
SWGARCH(2,1,1)-L		0.00014	13.79	8500.3	4095.61	0.71282	82.97	-
AIC = -0.33	BIC = -0.32	(0.001)	(2.54)	(32.54)	(25.12)	(0.254)	(21.12)	-
SWGARCH(2,1,1)-L		8481.63	0.024	672.69	1593.56	0.7905	28.76	2
AIC = -0.35	BIC = -0.35	(12.32)	(0.001)	(45.21)	(124.2)	(0.321)	(12.1)	(0.254)

Table 5.20 – SWGARCH

México								
Model		g_2	c	α_0	α_1	β_1	ξ	v
SWGARCH(2,1,1)		37.73	0.070	0.0666	0.0753	0.0658	-	-
AIC = -0.80	BIC = -0.79	(5.24)	(0.021)	(0.0012)	(0.015)	(0.0125)	-	-
SWGARCH(2,1,1)-L		0.0002	0.0697	309.86	349.40	0.0717	21.92	-
AIC = -0.79	BIC = -0.78	(0.001)	(0.002)	(32.12)	(54.21)	(0.0123)	(5.41)	-
SWGARCH(2,1,1)-L		65.76	0.0755	0.2515	0.3191	0.7766	$1.8x10^{-06}$	2
AIC = -0.83	BIC = -0.81	(12.54)	(0.021)	(0.052)	(0.021)	(0.214)	$(1x10^{-05})$	(0.125)

Table 5.21 – SWGARCH

Thailand								
Model		g_2	c	α_0	α_1	β_1	ξ	v
SWGARCH(2,1,1)		0.0003	0.0084	6.662	305.68	0.2644	-	-
AIC = -0.88	BIC = -0.87	(0.0001)	(0.0032)	(2.12)	(24.51)	(0.110)	-	-
SWGARCH(2,1,1)-L		$1.2x10^{-05}$	0.0093	217.97	9590.44	0.2033	13.97	-
AIC = -0.88	BIC = -0.87	$(1x10^{-04})$	(0.0012)	(51.24)	(241.21)	(0.051)	(2.45)	-
SWGARCH(2,1,1)-L		$6.3x10^{13}$	-0.0063	$1.2x10^{09}$	$1.2x10^{-09}$	0.3019	$1.5x10^{12}$	2
AIC = -0.94	BIC = -0.93	$(2x10^{13})$	(0.0012)	$(1x10^{08})$	$(2x10^{-08})$	(0.112)	$(3x10^{11})$	(0.212)

Table 5.22 – SWGARCH

Taiwán								
Model		g_2	c	α_0	α_1	β_1	ξ	v
SWGARCH(2,1,1)		0.9226	0.0627	0.0030	0.00107	0.9716	-	-
AIC = -0.50	BIC = -0.49	(0.214)	(0.021)	(0.0011)	(0.0001)	(0.351)	-	-
SWGARCH(2,1,1)-L		0.8727	0.0561	0.1610	0.0021	$1.8\text{x}10^{-10}$	$2.3\text{x}10^{-10}$	-
AIC = -0.49	BIC = -0.48	(0.312)	(0.012)	(0.312)	(0.005)	($4\text{x}10^{-09}$)	($5\text{x}10^{-10}$)	-
SWGARCH(2,1,1)-L		191.78	-0.0243	12.69	6.684	0.4306	0.0888	2
AIC = -0.51	BIC = -0.50	(25.12)	(0.005)	(3.45)	(2.54)	(0.012)	(0.021)	(0.05)

Table 5.23 – SWGARCH

Korea								
Model		g_2	c	α_0	α_1	β_1	ξ	v
SWGARCH(2,1,1)		0.20	-0.0179	0.00666	0.02166	0.8406	-	-
AIC = -0.66	BIC = -0.65	(0.01)	(0.002)	(0.0012)	(0.002)	(0.21)	-	-
SWGARCH(2,1,1)-L		1	-0.0181	0.00666	0.02166	0.8406	$1.6\text{x}10^{-10}$	-
AIC = -0.66	BIC = -0.65	(0.02)	(0.004)	(0.0002)	(0.001)	(0.232)	($6\text{x}10^{-09}$)	-
SWGARCH(2,1,1)-L		2232.6	-0.0228	52.12	117.34	0.8404	4.094	2
AIC = -0.69	BIC = -0.68	(123.1)	(0.005)	(12.4)	(25.4)	(0.212)	(1.542)	(0.121)

Table 5.24 – SWGARCH

Hong-Kong								
Model		g_2	c	α_0	α_1	β_1	ξ	v
SWGARCH(2,1,1)		0.781	-0.0101	0.0174	0.0129	0.844	-	-
AIC = -0.49	BIC = -0.48	(0.212)	(0.001)	(0.002)	(0.004)	(0.251)	-	-
SWGARCH(2,1,1)-L		3.551	-0.0086	0.0169	0.0114	0.836	$9.1\text{x}10^{-10}$	-
AIC = -0.49	BIC = -0.48	(1.23)	(0.0007)	(0.003)	(0.0012)	(0.14)	($3\text{x}10^{-10}$)	-
SWGARCH(2,1,1)-L		6.67	-0.0312	11.766	7.200	0.822	$7.4\text{x}10^{-08}$	2
AIC = -0.52	BIC = -0.51	(1.25)	(0.012)	(3.54)	(2.12)	(0.212)	($4\text{x}10^{-07}$)	(0.257)

Table 5.25 – SWGARCH

Malaysia								
Model		g_2	c	α_0	α_1	β_1	ξ	v
SWGARCH(2,1,1)		1.200	-0.056	0.0171	0.0073	0.768	-	-
AIC = -0.58	BIC = -0.57	(0.012)	(0.001)	(0.002)	(0.0002)	(0.254)	-	-
SWGARCH(2,1,1)-L		1.786	-0.06	0.0178	0.00412	0.6410	0.011	-
AIC = -0.58	BIC = -0.57	(0.232)	(0.01)	(0.002)	(0.0002)	(0.231)	(0.01)	-
SWGARCH(2,1,1)-L		15.94	-0.076	21.469	3.560	0.4013	15.94	2
AIC = -0.62	BIC = -0.60	(4.12)	(0.002)	(5.41)	(1.21)	(0.145)	(5.21)	(0.365)

An inspection of the tables above reveals some interesting results. With regard to the SWARCH models estimated for all countries, the scale parameter (g_2) that describes the moderate volatility is roughly between 2 and 4, except for Argentina (5.623) and Mexico (0.249), regardless of the alternative specifications used. The same observation is valid for the scale parameter for high volatility (g_3). The estimates vary between 13.12 and 19.3 except for Mexico (117) and Argentina (45.27). For the constant term in the autoregression (c) the estimates range from 0.02 to 0.06 with the exceptions of Thailand (-0.01) and Argentina (0.13). For the autoregressive parameter (θ_0) estimates vary from 0.08 to 0.19 with the exceptions of Brazil (0.2615) and Mexico (0.2825).

For the parameters of the conditional variance (h_t) most estimates were between 0.09 and 0.88. Again, the exception was a Latin American country: Argentina (2.35). The ARCH parameters (α_1 and α_2) follow the same pattern described above. For we have α_1 estimates from 0.01 and 0.13. And for α_2 the results indicate the parameter to vary between 0.09 and 0.16. It is indeed interesting that for both parameter, the countries of Latin America are concentrated around the upper limit and the Asian countries are around the lower limit except for Thailand (α_1 found to be 0.28).

These remarks are still valid when we allow for asymmetries (ξ) estimating a SWARCH-L model except for the parameters α_0 (Hong-Kong and Malaysia) and α_1 (Taiwan). But the general finding that Latin American countries show bigger effects remains valid.

The results are slightly different when we allow the distribution to be Student's t by incorporating v (degrees of freedom). In this case the scale parameter (g_2) that describes the moderate volatility ranges from 2 to 4 and (g_3), the parameter describing the high volatility ranges from 10 to 15 with Mexico presenting the higher estimate (23.04). For the constant term in the autoregression (c) we have the estimates ranging from 0.05 to 0.15 with exceptions for Thailand (-0.0035)) and Korea (0.0215). For the autoregressive parameter (θ_0) the estimates vary from 0.08 to 0.19 with exceptions for Argentina (0.03259), Brazil (0.2649) and Mexico (0.2924).

Turning to the parameters of the conditional variance (h_t), we found estimates between 0.09 and 0.88 for the constant (α_0). Again the exceptions are countries form Latin America: Argentina shows a big variation from 2.35 (SWARCH) to 1.14 (SWARCH-L-t) and Mexico varies from 0.89 (SWARCH estimates) to 0.25 (SWARCH-L-t estimates). This sharp change may be due to the adoption of a market practice that prevents stocks from rallying or dropping suddenly: the circuit breaker. This instrument was only introduced in Brazil and Argentina in the late nineties. It can create large difference in asymmetric effects when comparing Asian and Latin American countries. Another possibility, which we already examined in chapter three, is that the distribution of returns for Latin American countries departs much more from the Normal distribution than does the Asian countries' distribution. This explains why we have different estimates when we change from Gaussian to Student-t errors.

Estimates of the ARCH parameter (α_1) range from 0.02 and 0.21, the exception being Taiwan, which shows a dramatic change from 0.02 (SWARCH) to 1.4×10^{-8}. This result is more difficult to explain, since this only happens for one parameter and one country. For α_2, the results indicate that the parameter varies between 0.09 and 0.16, which were the results reported above for the SWARCH case. For the asymmetric parameter (ξ), bigger effects are found for Brazil, Argentina and Mexico (0.12, 0.04 and 0.07, respectively) than for the Asian countries. For the parameter related to the Student's t distribution (v) we have estimates ranging from 3 to 10, with Taiwan's estimate of 16 being an exception. Again, the abnormality of the estimate of (α_1) may be related to this. A general conclusion is that the effects found for the Latin American countries are larger than those found for the Asian countries. It is worth noting that nearly all of the coefficients are significant at a 95% level.

We also estimated SWGARCH models and found the scale parameter (g_2) that describes the low volatility to be roughly between 0.01 and 1.2. The exceptions are Thailand (0.0003) and Mexico (37). However these estimates are highly sensitive to the specification used. In the case of the SWGARCH-L model with Gaussian errors and leverage effect, estimates range from 0.87 to 3.5 except for Thailand (0.000012). The most striking results are due to the SWGARCH-L-t specification with Student's t innovations and leverage effect. In this case, the estimates are between 6.67 and 65.76. Brazil and Argentina show rather high estimates of 8481 and 317, respectively. On the other hand, Malaysia shows a small value of 1.5. Again we still have the Latin American countries showing bigger estimates than the Asian countries.

For the constant term in the autoregression (c) we have the estimates ranging from 0.06 to 0.18 with exceptions for Korea (-0.0179) and Argentina (0.18). For the SWGARCH-L model we have an interval of –0.06 to 0.12, with Brazil showing the biggest estimate (13.79) and Hong-Kong the lowest (-0.0086). For the SWGARCH-L-t, we have estimates between –0.0226 and 0.0755, with Mexico at the upper limit of the interval and Thailand at the lower bound (-0.0063).

We now turn to the parameters of the conditional variance (h_t). Beginning with the ARCH component α_0, we found estimates of between 0.0171 and 6.662 for the SWGARCH specification. For the other specifications we found estimates to vary between 0.0169 and 217.87, a wider range than in the case of SWGARCH-L, with Brazil showing the biggest value (8500). When we dealt with SWGARCH-L-t, we found estimated ranging from 0.25 to 21.469. Yet again, the biggest value is for a Latin American country, Brazil (672.69), and the lowest value is for an Asian country, Thailand (1.225×10^{-9}). It is indeed interesting that for all specifications, the countries of Latin America are concentrated around the upper limit and the Asian countries around the lower limit. With respect to α_1 . when we deal with the SWGARCH specification, the estimates are comprised between 0.00107 and 0.1124. When we adopt the SWGARCH-L specification, the amplitudes increase substantially to between 0.00114 and 349. For the SWGARCH-L-t the interval is 0.03257 to 7.2. In this case, Brazil has the biggest value (1593.56) while Thailand has the lowest (7.6×10^{-13}).

Turning to the GARCH parameter. β_1 ,the overall results follow the pattern described above, but in this case the variability of the estimates is lower when we switch from one model to another. For the SWGARCH specification. we have estimates ranging from 0.76 to 0.92. For the SWGARCH-L specification, estimates are between 0.07 and 0.84 with Argentina having an estimate of 0.8987 and Taiwan 1.8×10^{-10}. In the SWGARCH-L-t specification, the estimates are within the interval 0.30 and 0.84, Argentina being at the upper limit and Thailand at the lower bound.

These remarks are still valid when we allow for asymmetries (ξ) in estimating a SWARCH-L model. For the SWGARCH-L, the estimates lie between 2 and 21, with Brazil having a figure of 82 and Hong-Kong showing 9.1×10^{-10}. For the SWGARCH-L-t, the interval is 0.84 to 15.94. Again Brazil has an estimate of 28.76, while Hong-Kong has 1.8×10^{-6}.

Finally, when we look at the parameters for the Student's t distribution incorporating v (degrees of freedom), we find that it is around 2 irrespective of the country.

The general conclusion is that the Latin American countries do present bigger estimates than the Asian countries, an indication that the effects of volatility tend to be higher in these countries. We also found this result in the previous chapter. A nice conclusion is that using SWARCH/SWGARCH models reduces the ARCH/GARCH effects. This reinforces the idea that ARCH/GARCH tends to produce too much persistence and long memory.[16] Diebold (1986) remarks that this high persistence (or long memory) may be due by the non-observation of the change in the non-conditional volatility in the series under analysis. Lamureux and Lastrapes (1990) verify, through simulations of GARCH models, that the persistence is much higher when there are structural breaks in the unconditional variance. That is, we have periods in where the series shows higher volatility than it does in other periods where we do not control for this. This may result in an estimated persistence that is higher than the actual persistence. Hamilton and Susnel (1994) address this problem with models such as the GARCH-L and the high degree of persistence they imply for stock volatility

An interesting observation is that, except for Malaysia and Thailand, the estimated transition probabilities describe each state as being highly persistent with durations ranging from 84 days to 256 days. These results are similar to the ones reported by Hamilton and Susnel (1994). Also, the variance of each state shows that some countries have a more volatile profile than others. For example, for Mexico the high-volatility state is between 37 and 117 times as great as the low-volatility state, whereas the spread is lower for other countries. For Thailand the variance in the high-volatility state is almost twice as large as the low-volatility variance. The same observations apply to Korea, Thailand and Taiwan.

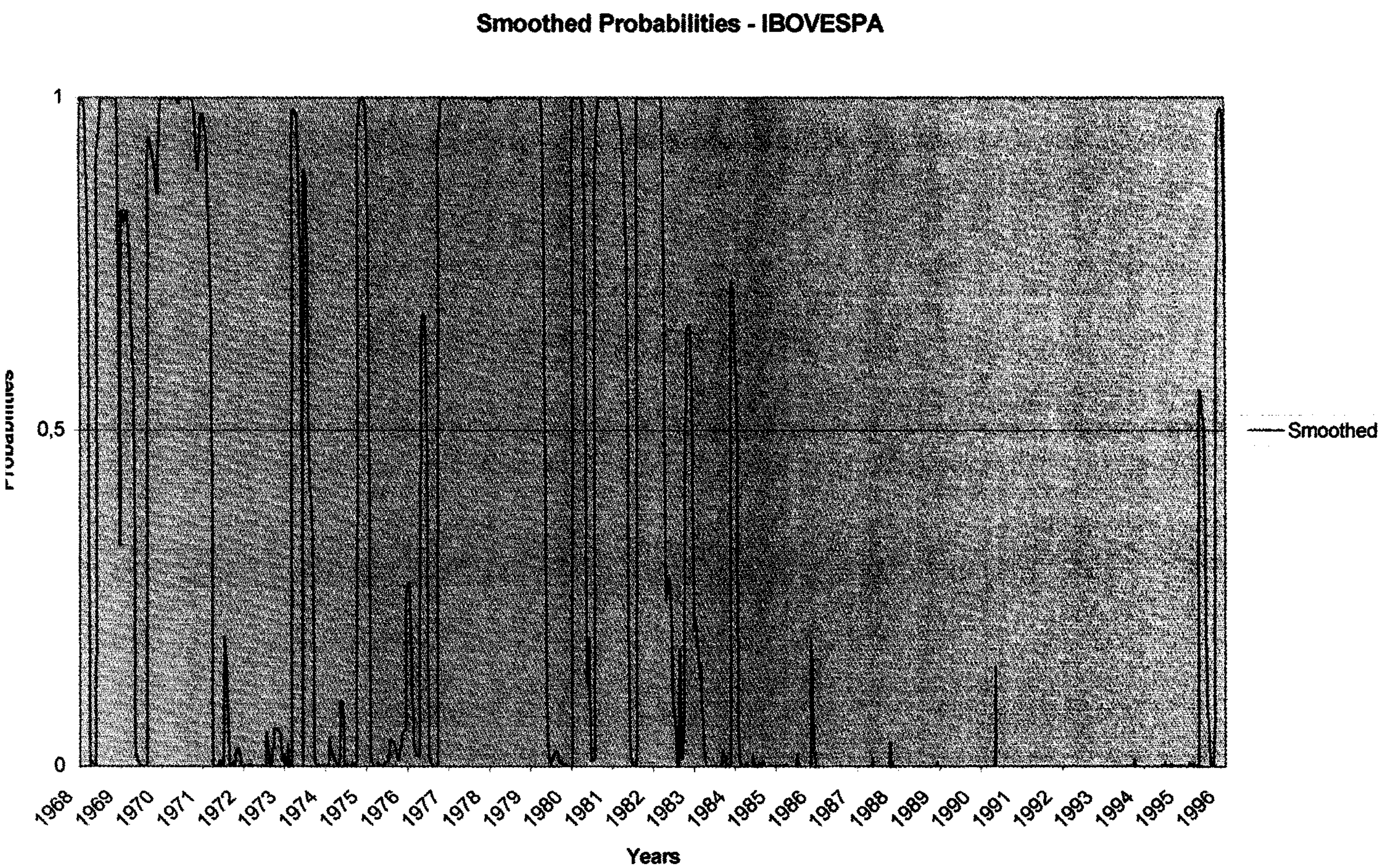
Again our finding lead us to the conclusion that there are different degrees of persistence in volatility among emerging markets, with a general tendency for Latin American countries to be more volatile than Asian countries. Again the explanation seems straightforward. During the period under consideration each of the Latin American countries were subjected to several economic plans. For example, Brazil had 10 different plans since 1967. The degree of intervention is also much lower in Asian countries.

In graph 5.2, we see the smoothed probability of Brazil being in state S_t. Initially the series is in regime one (low volatility). This period of growth is known as the "Miracle" in the Brazilian economy. However, approximately two years later the regime of high volatility begins to

[16] One could argue that the fact that SWARCH/SWGARGH produce less persistence doesn't imply that ARCH/GARCH has too much persistence per se, although this is one possible interpretation. Another possibility is that it's the SWARCH/SWGARCH that are bad because they falsely attribute genuine persistence to regime switches.

be dominant (1971). At this stage there is a crash in the stock market of Rio de Janeiro that affects the market in São Paulo. We can also see the effect of the 1973 oil shock, which pushes the Brazilian economy into a recession that lasts until 1974. In 1985 the high volatility regime becomes important again due to the first civilian government after 21 years of dictatorship. This event is followed by a succession of economic plans: the Cruzado Plan (1986), the Bresser Plan (1987) and the Verão Plan (1989). In 1990 a new government is elected which produces two more economic plans (Collor I and Collor II). After President Collor resigns in 1992 there is a period of volatility that lasts until 1996. In June 1994 the Real Plan was adopted. This plan was successful in bringing the inflation rate to a very low level. We can see the estimated transition matrix for the model SWGARCH(2,1,1)-t with leverage effect below.

Graph 5.2 – Probability that market was in regime one for each indicated day (Smoothed)



We can compute, based on the transition probability, the average duration of each regime. That

$$P = \begin{bmatrix} 0.9931 & 0.00688 \\ 0.01636 & 0.98364 \end{bmatrix}$$

is, we calculate the expected duration of each regime given that we are in a specific regime. As stated previously, the transition probabilities p_{kk} stand for the probability associated with each state. For example, p_{kk} shows the probability of state $S_t = k$, given that $S_{t-1} = k$. For Brazil p_{11} p_{22} are very high 0.9931 and 0.98364 respectively. This means that the probability that state 1 is followed by state 1 is high (0.9931). The probability that state 2 is followed by state 2 is also high (0.983364), indicating a high persistence in states that is compatible with the series for Brazil and with the findings of earlier chapters. On the other hand, the probability that[17] state 1 is followed by state 2 and the probability that state 2 is followed by state 1 is very small 0.01636 and 0.00688. We have information about duration dependence in the sense that, if we are in state S_t, we can say something about the "exit probability," which is related to the transition probability. This confirms the existence of long memory in volatility in the sense that a state tends to persist until there is a change in regime, which may not happen very frequently.

We then find that the average durations of low and high volatility periods are given by $\frac{1}{1-0.9931} = 144.93 \quad days$ and $\frac{1}{1-0.98364} = 61.12 \quad days$, which is clearly high for both states. Indeed, from a visual inspection it appears that the periods of low volatility last longer than the high volatility periods. From other side we can compute the ergodic probabilities discussed above as $\frac{1-0.98364}{2-0.9931-0.98364} = 0.70335$ and $\frac{1-0.9931}{2-0.9931-0.98364} = 0.29665$, so that there is a larger probability that the economy is in a period of a low volatility than high volatility. This could be the result of numerous crises in the international environment that affected the Brazilian economy. It is interesting to note that this result is very similar to that found by Almeida and Pereira (1997) for the Brazilian economy. The results for the rest of the countries in our sample can be found in the appendix to this chapter. The reason for this is that an explanation of these different regimes requires a great deal of specific knowledge of the individual economies. However, we can clearly see the conditional variance switching from high volatility to low volatility period (or vice-versa). As has been stated previously, g represents a scale effect related to volatility. That is, it represents the number of times the volatility in the high volatility state is larger than the one in the low volatility state. For Brazil, we have g assuming very high values. In the SWGARCH-L model g is found to be 8000, which is as large as the volatility in the low volatility state. This indicates that there is a sharp increase in volatility during crises, a finding that is consistent with some of the experiences we have faced during crashes and other turbulent periods in the Brazilian economy.

Below we show all of the transition matrices for the countries in our sample. As discussed above, the transition probabilities p_{kk} stand for the probability associated with each state. For

---

[17] Generally speaking p_{ij} is defined as the probability of S_t assuming a value j given that the state prevailing in t-1 was i.

instance p_{kk} shows the probability of state S_t = k, given that S_{t-1} = k. In general the transition probabilities are very high, and this seems to apply to all of the countries.

For Mexico, p_{11} and p_{22} are 0.9769 and 0.9908, respectively. This means that the probability that state 1 is followed by state 1 is 0.9769, and the probability that state 2 is followed by state 2 is also high 0.9908, indicating high persistence in states that is compatible with the series for Mexico and our findings in the other chapters. On the other hand, the probability that[18] state 1 is followed by state 2 is low, 0.0231, and the probability that state 2 is followed by state 1 is very small, 0.0092, for Mexico. Again, long memory in volatility is present.

The numbers appear to be consistent across countries. The only one that deserves attention is Thailand, where p_{11} and p_{22} are 0.9696 and 0.5743, respectively. Here the probability that state 1 is followed by state 1 is high (0.9696), but the probability that state 2 is followed by state 2 is only moderate (0.5743), indicating greater persistence in state 1 than state 2. On the other hand, the probability that[19] state 1 is followed by state 2 is small, 0.0304, and the probability that state 2 is followed by state 1 is moderate 0.4257.

We now have information about duration dependence in the sense that if we are in state S_t we can say something about the "exit probability" that is related to the transition probability. That is, we can compute the average duration of each state along lines sketched above. The average durations of a periods of low and high volatility are $\frac{1}{1-p_{11}}$ and $\frac{1}{1-p_{22}}$, respectively. We have computed these average durations for each country in our sample (see table 5.25). Indeed, from a visual inspection it seems that the periods of low volatility last longer than periods of high volatility, and these periods tend to have a high persistence (i.e. longer memory), except that Thailand shows a small duration for each of the periods.

From the other side we can compute the ergodic probabilities discussed above as $\frac{1-p_{11}}{2-p_{22}-p_{11}}$ and $\frac{1-p_{22}}{2-p_{22}-p_{11}}$. These probabilities are also shown in table 5.25 below, so that there is a larger probability of the economy being in a period of low volatility than one of high volatility. We can see from the table that the probability of being in state 1 is greater than that of being in state 2. This might be explained by the fact that these economies have been subjected to a high degree of intervention. However, these interventions are not so frequent so that they make the periods of low volatility more likely.

[18] Generally speaking p_{ij} is defined as the probability of S_t assuming a value j given that the state prevailing in t-1 was i.

[19] Generally speaking p_{ij} is defined as the probability of S_t assuming a value j given that the state prevailing in t-1 was i.

Figure 5.2- Transition Matrices for the Countries under Study

$$\begin{bmatrix} 0.9769 & 0.0092 \\ 0.0231 & 0.9908 \end{bmatrix}$$

Mexico

$$\begin{bmatrix} 0.9923 & 0.0053 \\ 0.0071 & 0.9947 \end{bmatrix}$$

Argentina

$$\begin{bmatrix} 0.9696 & 0.4257 \\ 0.0304 & 0.5743 \end{bmatrix}$$

Thailand

$$\begin{bmatrix} 0.9961 & 0.0119 \\ 0.0039 & 0.9881 \end{bmatrix}$$

Taiwan

$$\begin{bmatrix} 0.9899 & 0.0358 \\ 0.0101 & 0.9642 \end{bmatrix}$$

Korea

$$\begin{bmatrix} 0.9925 & 0.022 \\ 0.0025 & 0.9780 \end{bmatrix}$$

Hong-Kong

$$\begin{bmatrix} 0.9776 & 0.0390 \\ 0.0224 & 0.9690 \end{bmatrix}$$

Malaysia

Table 5.25 – Duration of Both Regimes

	Duration of the First Regime	Duration of the Second Regime
Mexico	43.29	108.70
Argentina	129.87	188.68
Thailand	32.29	2.35
Taiwan	256.41	84.03
Korea	99.01	27.93
Hong-Kong	133.33	45.45
Malaysia	44.64	25.64

Note: All test statistics are significant at 1% level

Table 5.26 – Ergodic Probabilities for Each Regime

	First Regime	Second Regime
Mexico	0.2848	0.7152
Argentina	0.4077	0.5923
Thailand	0.4334	0.7532
Taiwan	0.7532	0.2468
Korea	0.7799	0.2201
Hong-Kong	0.7458	0.2542
Malaysia	0.6352	0.3648

5. 4 - Conclusion

In this chapter we have considered different approaches to modelling the phenomena of persistence and long memory. The search for new models that produce these features must deal with two stylised facts that conventional volatility models have difficulty reconciling. The first is that conditional volatility can increase substantially in a short amount of time with the advent of a turbulent period. The second is that the rate of mean-reversion in stock-market volatility appears to vary positively and nonlinearly with the level of volatility. For example, this can happen when governments intervene in markets.

Several authors have brought this subject to our attention. Hamilton and Susnel (1994), for instance, highlighted the difficulty of forecasting with conventional GARCH models by showing that they can provide worse volatility forecasts than constant variance models on the basis of the MSE loss criterion. One possible way to deal with this question would be to not allow the conditional variance to respond proportionately to "large" and "small" shocks. If this approach were taken, the conditional variance could be restrained from increasing to a level at which volatility forecasts would be undesirably high. One problem with this strategy is that such a model might understate the true variance by not responding sufficiently to large shocks, thereby not showing much mean reversion. Such "threshold" models do not treat the two stylised facts listed above, namely sharp upward jumps in volatility followed by fairly rapid reversion to near-normal levels. Thus, we should look for a different way to deal with this problem.

Lamoreux and Lastrapes (1990) suggest that structural breaks in the variance could account for high persistence and long memory in the estimated variance. Along these lines, they argue in the same direction as Perron (1989) in dealing with the mean of a series. This alternative could exploit persistence and long memory in volatility. A natural way to do this that would also address the two stylised facts above would be to allow for Markov-switching parameters. This approach was suggested by Hamilton and Susnel (1994) and has been pursued in this chapter.

Markov-switching parameters allow the volatility to experience discrete changes in persistence, leading to jumps. Many researchers have suggested that the poor forecasting performance and long memory properties of ARCH models might be related to structural changes in the ARCH process. As discussed previously, Perron (1989) argues that changes in regime may give the spurious impression of unit roots.

Based on this view, we have used Perron's (1999) methodology in order to measure the occurrence of jumps in volatility that could suggest structural breaks. Our results indicate that, for the stock returns under study, we cannot reject the hypothesis that many of these jumps occur in the data. One interesting result is that some countries do show more jumps than others. This is the case with Brazil, a country that experienced significant interventions during the period covered by the data. However, the overall results do show the occurrence of jumps. In this sense, we revert to the fact that changes in regime may give the spurious impression of unit roots in characterising the volatility of a series. This is consistent with the work of Hamilton and Susnel (1994), which suggested "spuriously high persistence" in volatility.

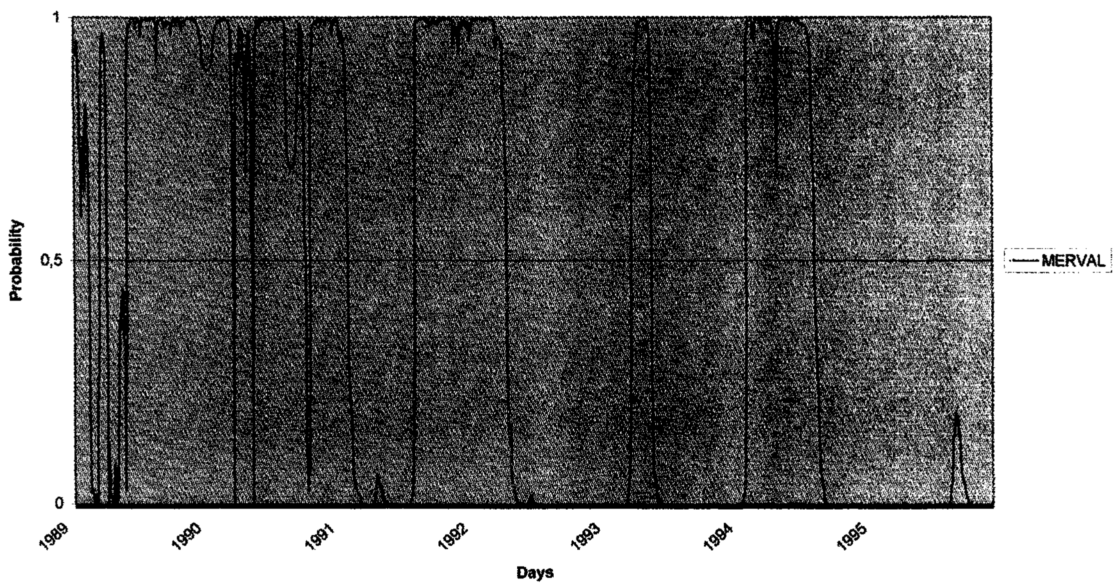
This chapter introduced Markov-Switching ARCH and GARCH to describe the volatility of stock prices. SWARCH and SWGARCH can offer alternatives for estimating the persistence and long memory effects in volatility. Our estimates attribute most of the persistence and long memory in stock price volatility to the persistence of low and high volatility regimes, which typically last for several months. The high-volatility regime is to some degree associated with economic recessions. The analysis also seems to confirm the finding of researchers that stock price decreases lead to bigger increases in volatility than would price increases of similar magnitude, and that Student's t innovations are much better for describing fundamental innovations on the basis of the AIC and BIC criteria.

We have also found Latin American markets to be more persistent and to show greater long memory effects than Asian markets. This finding is intuitive given the degree of intervention in these markets vís-à-vís the Asian economies. It is also consistent across different model specifications. A nice conclusion is that using SWARCH/SWGARCH models causes the ARCH/GARCH effects to be reduced. This reinforces the idea that ARCH/GARCH tends to produce too much persistence.

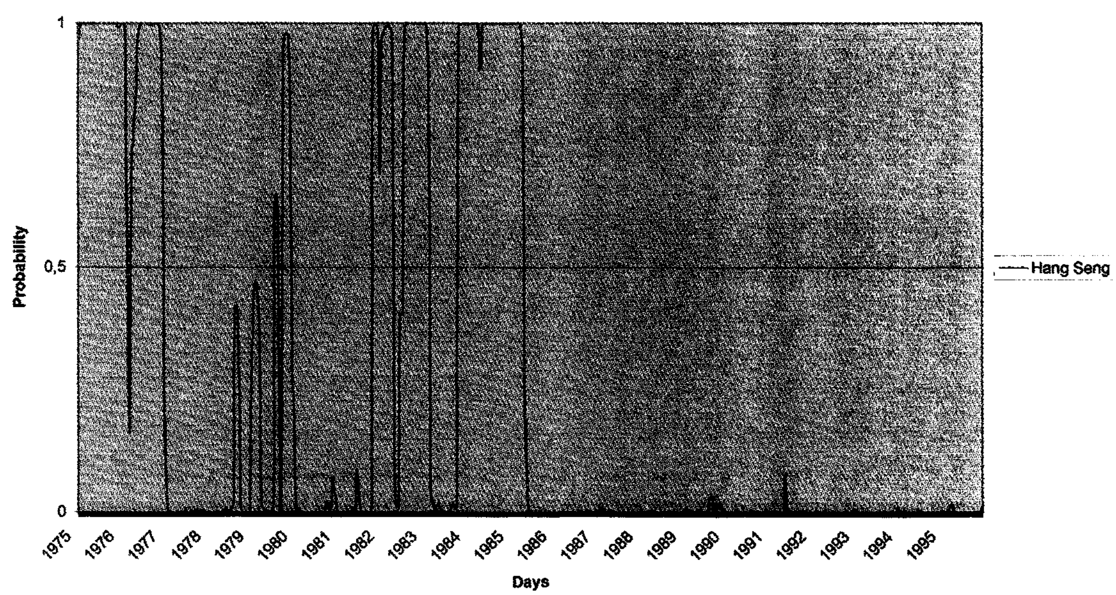
We could pursue our analysis further by examining whether periods of high volatility spillover across countries. If such evidence were found, it could support the idea of contagion models. The extent of changes in volatility has played an important role in discussions of whether emerging economies have indeed been subject to "contagion". This could be done with a simple analysis of the behaviour of correlation coefficients. But however, it could lead to a misleading picture of contagion if the country in question experiences changes in volatility regimes.

The fact that high volatility states roughly coincide across countries is indeed suggestive, but does not constitute statistical evidence in favour of the "volatility contagion" hypothesis. This might be confirmed using a bivariate switching volatility model, and would be an interesting avenue for future research.

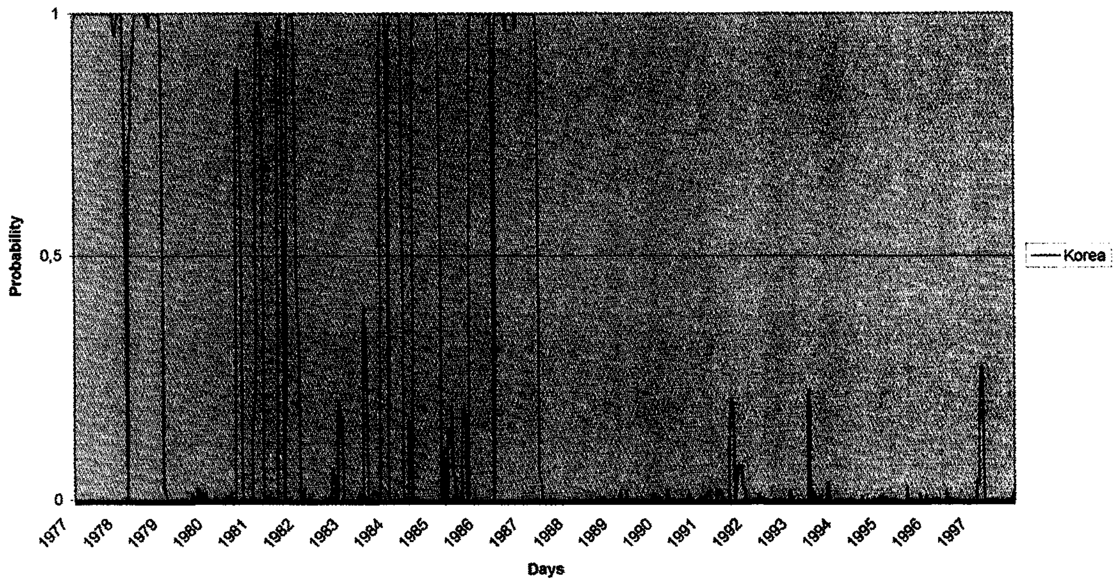
Graph 5.3 - Probability that market was in regime one for each indicated day -Smoothed Probabilities for Argentina



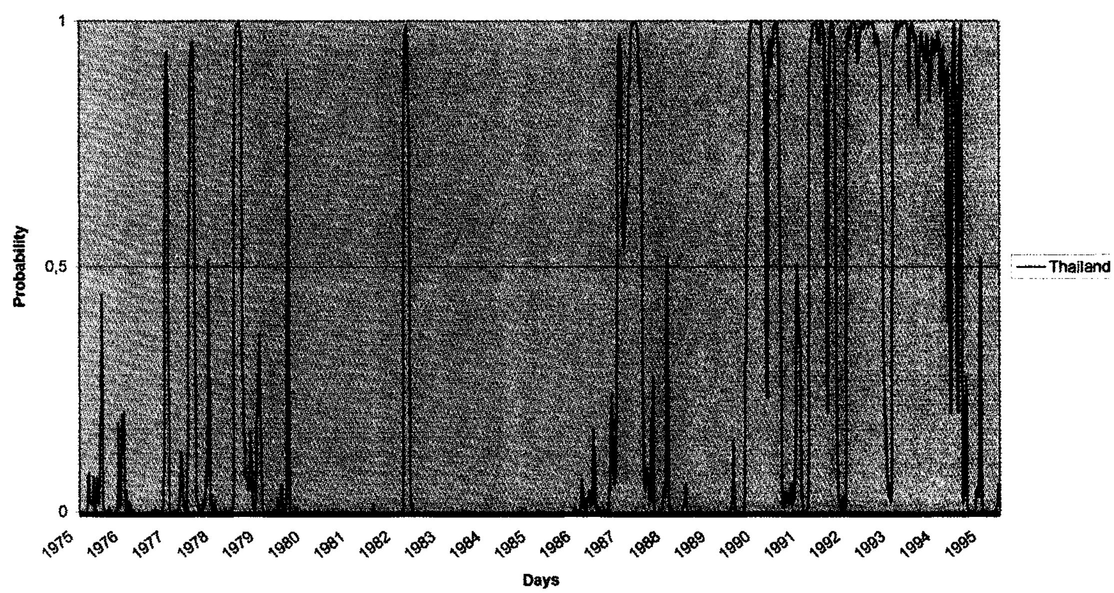
Graph 5.4 - Probability that market was in regime one for each indicated day- Smoothed Probabilities for Hong-Kong



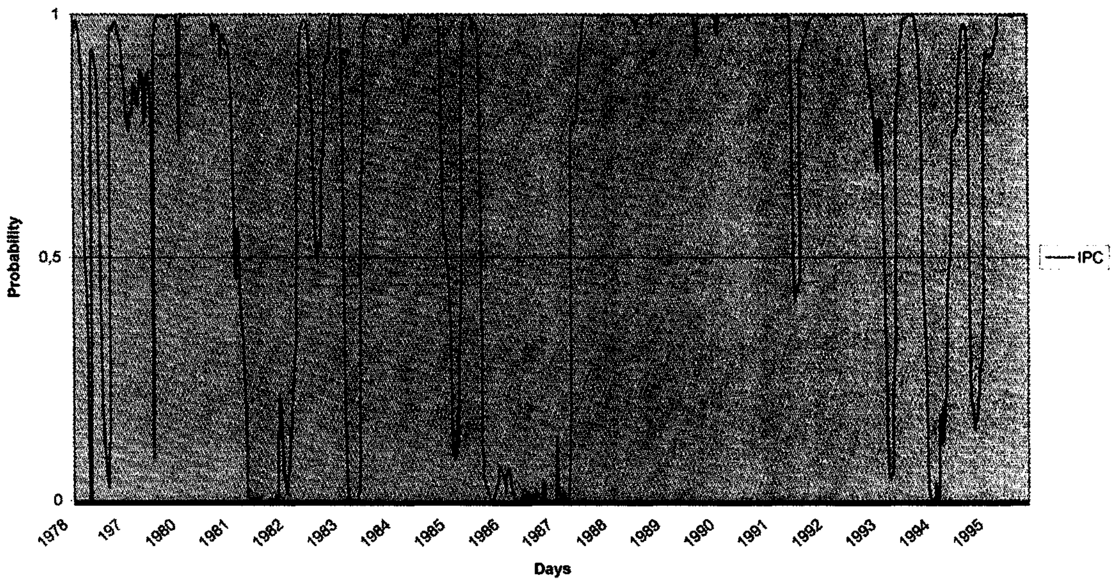
Graph 5.5 - Probability that market was in regime one for each indicated day - Smoothed Probability for Korea



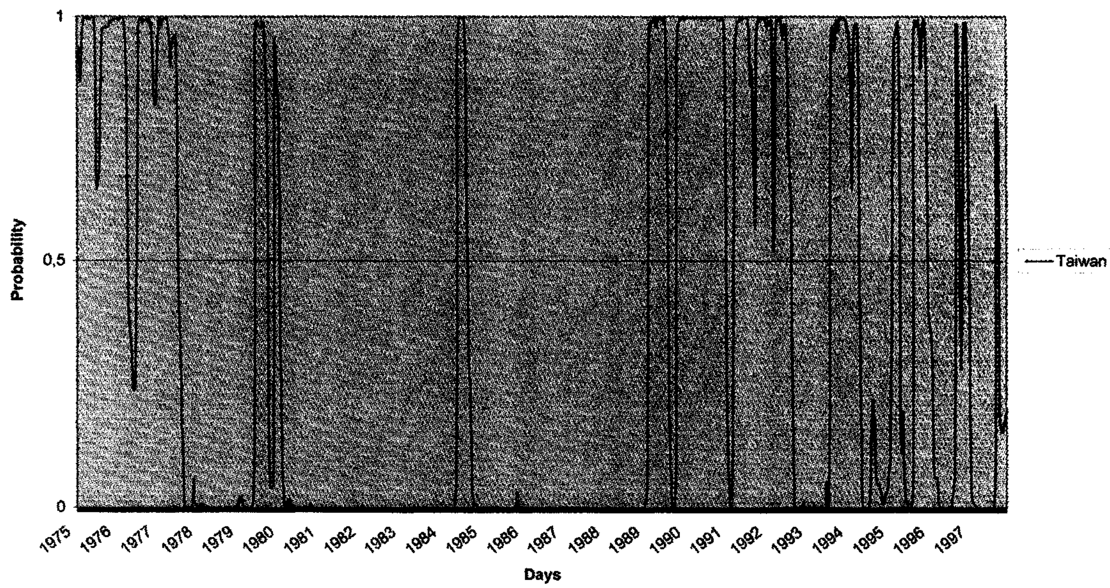
Graph 5.6 - Probability that market was in regime one for each indicated day -Smothed probability for Malaysia



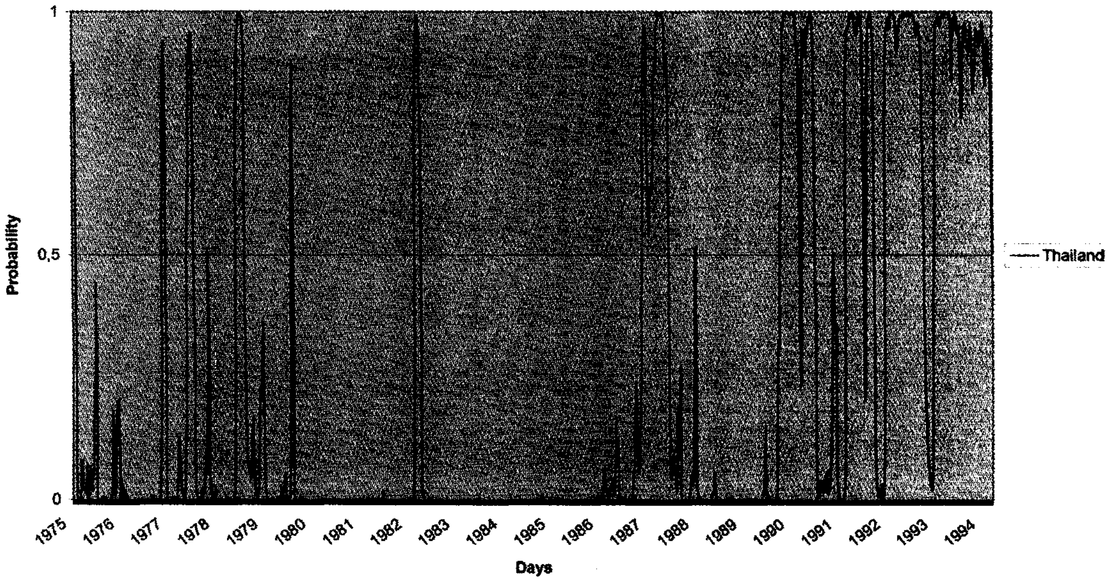
Graph 5.7- Probability that market was in regime one for each indicated day - Smoothed Probability for Mexico



Grpah 5.8 - Probability that market was in regime one for each indicated day - Smoothed probabilty for Taiwan



Graph 5.9 - Probability that market was in regime one for each indicated day - Smoothed Probability for Thailand



Chapter 6

Conclusion

6.1 Background

In developed countries during the late 1980s, a huge amount of financial capital was available through pension and investment funds that could be drawn to developing countries, providing that they liberalised their markets externally and developed their markets internally. Liberalisation was imperative: foreign bank loans, which dominated inward capital flows to developing countries in the 1970s, were decreasing in the aftermath of the Latin American debt crisis.

In less developed countries, the speed and extent of stock market developments in the last twenty years has been unprecedented. and has led to fundamental shifts in the capital flows received from developed countries. In leading developing countries during the 1970s and 1980s, the capitalisation ratio (market capitalisation as a proportion of GDP). a key indicator of stock market development, rose at an unprecedented rate. In the course of twenty years, the ratio climbed from 10% to over 70% in countries like Chile and Taiwan. In contrast, it probably took the US over 80 years to achieve a similar ratio. In terms of comparative market capitalisation, many emerging markets have now surpassed the average-medium-sized European stock market. In addition, the numbers of new listings and investors in these markets have soared. The total value of shares traded on LDC markets rose over twelve-fold during 1986-95, an increase from just over 2% to nearly 9% of the total world value.

The development of these markets was aided by external financial liberalisation and by the consequent influx of foreign portfolio capital flows. As seen in Table 6.1, the scale and composition of capital flows to developing countries have undergone major changes. the salient one being the huge increase in private finance. The average net private capital flow to developing countries was $15.1 billion in 1983-88, whereas the figure surged to $107.6 billion in 1989-95, and reached $200.7 billion in 1996. While net direct investment has remained an important component of these flows, the surge in net portfolio investment is mostly responsible for the increased flows, which rose $3.4 billion to $44 billion.

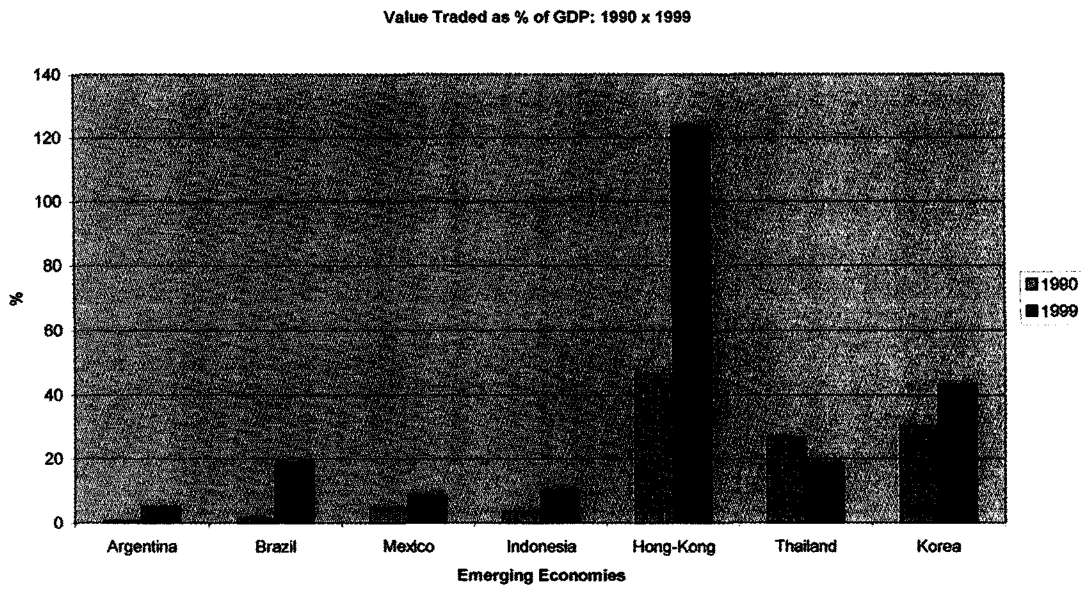
Table 6.1

COUNTRIES	1983-88	1989-95	1996
Net Private Flows	15.1	107.6	200.7
Net Direct Investment	10.4	41.8	90.7
Net Portfolio Investments	3.4	44.0	44.6
Other Net Investments	1.3	22.1	64.9

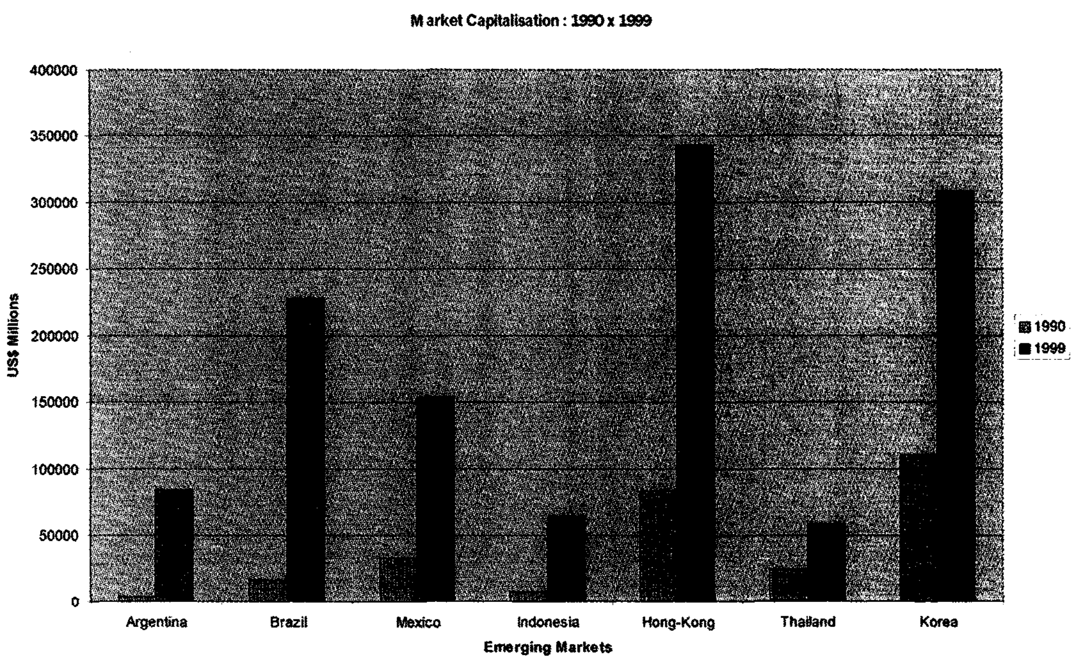
Developing countries currently account for a disproportionate share of global issuance. In 1994, developing countries accounted for only 12.6% of total world market capitalisation, but a full 37% of global equity issues. Portfolio capital inflows are invariably short-term and speculative. They are often not related to economic fundamentals, but rather to whims and fads prevalent in international financial markets. Thus, rapid reversals can lead to crises, collapses in economic growth and increased volatility. The volatility of share prices tends to be higher in emerging markets than in developed economies. Some authors argue that the increase in volatility is largely the result of "excessive" capital mobility. Accordingly, the imposition of controls on capital flows should help countries reduce externally induced financial instability. Some authors argue that changes in volatility have played an important role in discussions of whether or not emerging markets have indeed been subject to "contagion". Volatility is a crucial variable present in most financial markets. For instance, it plays a central role in many areas of finance and foreign exchange. It is important for government as well as the private sector to act in order to contain volatility, which may in turn cause disturbances in the economy.

From an empirical standpoint, it is very important to carefully model any temporal variation in the volatility process to understand the way it affects the economy. Interest in the study of volatility has increased since the beginning of the eighties. Models were first applied to developed economies mainly because the data was readily available: however, some countries are making efforts to change and improve their economies with the goal of raising their performance to that of more advanced nations. These economies are called emerging markets. Currently, some investors favour emerging market stocks and bonds since they have the potential for high returns in a relatively short period of time. There is a great deal of risk involved in these investments because by definition, emerging markets are in a state of transition, and thus subject to unexpected political and economic upheavals. The values of their stocks, bonds and currency can change dramatically and unpredictably. Irrespective of the measure chosen (see graph 6.1 and graph 6.2), emerging markets have shown a huge increase in their transactions and so, it is important to analyse these markets more precisely and this was the major aim of this thesis.

Graph 6.1



Graph 6.2



6.2 Long Memory Volatility in Emerging Countries

This thesis examines the discussion about volatility, and expands it to the so-called emerging markets that currently occupy an important place in the menu of choices for foreign investors. The growth in the attention dedicated to volatility, however, has not been symmetrical: emerging markets take an unimportant place in the vast majority of studies on this theme. Our investigation takes some of the findings for more established markets, and applies them instead to emerging markets to determine whether or not the findings hold true. Taking the more developed markets as a benchmark. we further explore the differences between both markets. After considering a non-technical discussion about volatility in Chapter One and characterising volatility in the US, we conclude that there is no clear trend for volatility. The results reveal only that movements in volatility are more frequent than they were in the past. After extending this discussion to encompass the existence of emerging markets, we show that changes in volatility are indeed much greater in emerging markets than in the US. In other words, if the findings for the US stock markets (Dow Jones) show significant results, we can expect more pronounced movements in emerging markets.

To narrow the focus of this thesis, only the more capitalised markets were chosen since these have a stronger tradition of computing stock market indices for a longer period of time. For this reason, we have not chosen the ex-socialist countries, but instead have worked with Latin American countries (Argentina, Brazil and Mexico) as well as countries from Southeast Asia (Taiwan, South Korea, Thailand, Indonesia and Hong-Kong). These countries are more representative of the emerging economies and for this reason they are studied.

Chapter one puts in perspective the question of volatility and discuss the different definitions of volatility and establishes a comparison among the US stock market and the emerging economies. It can be seen that movements in the emerging economies tend to be more pronounced than the movements observed for developed economies such as the U.S. economy. The higher volatility is related to the flow of capitals and liberalisation in the emerging economies. But however a more technical approach is needed in order to study the differences in terms of modelling and treatment of volatility.

There is a great deal of interest in the subject of volatility in financial markets, as evidenced by the enormous quantity of papers and different specifications described in the literature review in Chapter two. The ARCH model and its various extensions have proven to be very effective tools, and as a result, the literature on ARCH has expanded dramatically since the seminal paper by Engle (1982). There has been much speculation about whether or not volatility has changed in past years primarily because recent episodes like the Asian Crisis and the decay of the NASDAQ market in the US are linked to the new economy. Derivatives, which have been greatly developed in all markets, have contributed to the increase of liquidity as well as the change in volatility.

Financial integration increases greatly as lower information costs and technology provide more opportunities for investors. One effect of this integration is to create a link between all markets, which in turn makes it difficult to talk anymore about isolated effects. It is particularly interesting, however, that the effects tend to last much longer than previous movements noticed in capital markets, i.e. long memory in volatility. From a technical perspective, the traditional ARCH methodology does not satisfactorily account for the changes in volatility that remain for long periods after the first movement in stock prices is detected. The general objective of this thesis was to establish a comparison between developed and emerging markets, then conclude that while studies of established markets confirm the existence of long memory in volatility, the results are much more pronounced in emerging markets.

Chapter three clearly shows that regardless of the way we choose to characterise long memory in volatility, there is indeed a long memory effect in these economies. We have seen the Hurst exponent is higher than 0.5 (persistence) for all economies. We have also seen that the autocorrelation analysis shows the correlation among returns to be significantly higher than zero after long lags. In fact, not only is there substantially more correlation between absolute returns than returns themselves, but the power transformation of the absolute returns also has quite a high autocorrelation for long lags. This finding is consistent with others in the literature, which shows that emerging markets mimic the pattern established for other developed markets.

In Chapter four, we estimate the long-memory effects in the volatility of the retuns for the emerging economies. After estimating these effects in emerging markets, we have seen that the ARCH/GARCH specification imposes too much restriction on the parameters of the variance function. Even the cutting-edge IGARCH specification proposed by Engle and Bolerslev (1986) proves to be unnecessary. What emerges from our analysis is that the long memory volatility models are extremely significant and useful to estimate persistence in variance. The FIGARCH and FI-APARCH specification produces estimates significant for all countries studied. We have found that the evidence in emerging economies follows that found in developed economies. As we have already stressed, the movements are stronger in emerging economies, which indicates that investors should approach these stock markets carefully.

Finally, Chapter five discusses the occurrence of structural breaks in the series of volatility. This offers an alternative to the parametric methods of the FI-GARCH estimation when dealing with high persistence in the series we have studied. With this intention in mind, we have researched the number of jumps occurring in selected emerging markets. The results indicate that a certain number of jumps occur during the periods in our data. Some of these series, like Brazil, show that more jumps indicate more volatility. Based on the estimation of the number of jumps, we have chosen to model volatility using a Markov chain and switching regimes modelling. It is a useful description of the process. The results also prove that we cannot disregard this methodology when describing the volatility in emerging markets.

We have produced some evidence of the volatility in emerging markets by pursuing different modelling strategies as well as tests that could lead us to draw conclusions about high persistence in volatility. We have found that regardless of the way we choose to measure volatility, there seems to be no doubt about the so-called long-memory property. Although much work can be done on this topic, the scope of this thesis must remain focused. We have produced some evidence on a subject that has attracted the attention of academics and non-academics alike, and expanded knowledge in areas we judged important. Some developments are necessarily beyond the scope of this thesis, but remain a fertile ground for further study.

Bibliography

Aggarwal, R., C. Inclan, R. Leal, 1995, Linkages and Sudden Changes in Variance in Emerging Markets, Journal of European Financial Management, forthcoming.

Aggarwal, R., R. Leal, 1995, Linkages and Volatility in Emerging Capital Markets, in G. Tsetsekos and M. Papaioannou, (editors), Portfolio Management and Hedging Strategies for Emerging Capital Markets, Irwin, forthcoming.

Aggarwal, R. and Leal, R., 1996, Integration and Annomalies in the emerging markets of Asia and Latin America. In Research in International Business and Finance, eds. J. Doukas and L. Lang, Supplmt 1, pp. 75-99, Jai Press, Greenwhich, Connecticut.

Aggarwal, R. and Tandom, K., 1994, Anomalies or illusions ? Evidence from stock markets in eighteen countries. Journal of International Money and Finance, 13, 83-106.

Almeida, G. Rodrigues E. and Lemgruber, E., 1993, O efeito mensal no mercado brasileiro de ações. In Anais do 17º Encontro da Anpad – Finanças, pp. 111-121, ANPAD, Salvador.

Almeida, Nuno M. C. C. and Pedro L. Valls Pereira, 1999, SWGARCH models applied to the IBOVESPA index. Annals of the XXI Brazilian Econometric Society Meeting.

Anderson, T., J. Lund, 1997, Estimating Continuous Time Stochastic Volatility Models of the Short-Term Interest Rates, Journal of Econometrics, 77, 343-77.

Bachelier, L., 1900, Thèorie de la Spéculation. Annales de l'Ecole Normale Supérieure, 17: 21-86.

Ball, C. , Tourous, W.N., 1995, Regime Shifts in Short Term Riskless Interest Rates, Working Paper 15-95, The Anderson School at UCLA.

Bailey, W. , J. Jagtiani, 1994, Foreign Ownership Restrictions and Stock Prices in the Thai Capital Market, Journal of Financial Economics, 36, 57-87.

Bailey, W. , R. Stulz, 1990, Benefits of International Diversification: the Case of Pacific Basin Stock Markets, Journal of Portfolio Management, 16, 57-61.

Baillie, Richard T. , Tim Bollerslev. 1989. The message in Daily Exchange rates: A Conditional Variance Tale. Journal of Business and Economic Statistics ,7, 297-305.

Baillie, Richard T., Tim Bollerslev. and Hans Mikkelsen, 1996, Fractionally integrated generalized autoregressive conditional heteroskedasticity, Journal of Econometrics. forthcoming.

Baillie, Richard, Ching-Fan Chung, and Margie A. Tieslau. 1995, Analysing inflation by the fractionally integrated ARFIMA-GARCH model. Journal of Applied Econometrics, 11, 23-40.

Barry, C. and M. Rodriguez, "Risk, Return and Performance of Latin America's Equity Markets, 1975-1995, Latin American Business Review, 1(1), 1998, p.51-76.

Bekaert, G. , C. Harvey, 1995, Time-Varying World Market Integration, Journal of Finance, forthcoming.

Bekaert, G. 1992, The time-variation of expected returns and volatility in foreign exchange markets. Mimeo, Northwestern University.

Berndt, E.K. Hall, B.H Hall, and J, A. Hausman. 1974, Estimation and Inference in Nonlinear Models. Annals of Economic and Social Measurement. No.3,pp. 653-65.

Bera, Anil , M. Higgins. 1993, ARCH models: Properties, Estimation and Testing. Journal of Economic Surveys, 7, p.p.100-137.

Black, F. And M. Scholes, 1973, The pricing of options and corporate liabilities, Journal of Political Economy, 81 :637-659.

Bollerslev, Tim ., Hans Mikkelsen, 1996. Modelling and pricing long-memory in Stock mrket volatility. Journal of Econometrics, 73, 151-184.

Bollerslev, T., R. Chou, and K. Kroner, 1992, ARCH Modelling in Finance: A Review of The Theory and Empirical Evidence, Journal of Econometrics, 52, 5-59.
Bollerslev, Tim. 1986 , Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, 31. 307-27.

Bollerslev, Tim. 1988 , On the Correlation Structure for the GARCH Process. Journal of Time Series Analysis, 9, 121-132.

Bollerslev, T., Engle ,R. F. and Wooldridge. J. M. 1988, A capital asset pricing model with time-varying covariances. Journal of Political Economy, 96: 116-131.

Bollerslev, Tim. 1987, Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. Review of Economics and Staristics, 69: 542-47.

Bollerslev, Tim. 1988, On the Correlation Structure for the Generalized Autoregressive Conditional Heterokedastic Process. Journal of Time Series Analysis, 9: 121-31.

Bollerslev, T., Ray Y. Chou, and Kenneth F. Kroner. 1992, ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. Journal of Econometrics, 52: 5-59.

Bollerslev, T., M. Wooldridge. 1992, Quasi-Maximum Likelihood Estimation and Inference in models with Time Varying Covariances. Econometric Reviews, 11: 143-72

Bollerslev, T. , R. J. Hodrick. 1992, Financial market efficiency tests. Kellogg Graduate School of Management Working Paper, n. 132.

Bollerslev, T., Engle, R. F. 1993, Common persistence in conditional variances. Econometrica, 61: 67-186.

Booth, G.G., F.R. Kaen, and P.E. Koveos. 1982, R/S Analysis of foreign exchange rates under two international money regimes, Journal of Monetary Economics, 10 :407-417

Breidt,F.J., N.Crato and P.J.F. de Lima. 1993, Modelling long-memory stochatsic volatility, Working Paper (John Hopkins University, Baltimore, MD).

Brenner, R.J., R. Harjes aand K. Kroner, 1996, Another look at models of short term interest rates, Journal of Quantitative Analysis, 31, 85-107.

Brockwell, P. J. . R. A. Davis, 1991, Time Series: Theory and Methods. Second Edition, New York:Springer-Verlag.

Cai, J, 1994, A Markov Model os Switching Regime ARCH, Journal of Business and Economic Statistics, 12, 309-316.

Campbell, J.Y. , Y. Hamao, 1992, Predictable Stock Returns in the United States and Japan: A Study of Long-Term Capital Market Integration, Journal of Finance, 47. 43-69.

Chan, K.C., B.E. Gup. e M. Pan, 1992, An Empirical Analysis of Stock Prices in Major Asian Markets and the United States, Financial Review, 27, 289-307.

Cheung, C.S. e J. Lee. 1993, Integration vs. Segmentation in the Korean Stock Market, Journal of Business Finance and Accounting. 20, 267-273.

Cheung, Y.L. , Y.K. Ho. 1991, The intertemporal stability of the relationship between the Asian emerging markets and the developed equity markets, Journal of Business Finance and Accounting, 18, 235-254.

Cheung, Y.L., 1993, A Note on the Instability of the Intertemporal Relationships Between the Asian-Pacific Equity Markets and the Developed Markets: a Non-Parametric Approach, Journal of Business Finance and Accounting, 20. 229-236.

Claessens, Stijn. Moon-Whoan Rhee. 1994. The Effect of Equity Barriers to Equity Investment in Developing Countries, in Jeffrey A. Frenkel, editor, The Internationalisation of Equity Markets, The University of Chicago Press, pp. 231-271.

Corhay, A., A. T. Rad. and J.P. Urbain, 1994, Co-movements and dynamic interrelationships in the pacific basin stock markets. in Bos, T. and T.Fetherston, editors, Research in International Business and Finance, 18, 93-105.

Costa Jr. ,N. ,1990, Sazonalidades do Ibovespa. Revista de Administração de Empresas, 30, 79-84.

Costa Jr, N. and Lemgruber, E. ,1993, O efeito fim-de-semana durante períodos de abertura e de fechamento das bolsas de valores. In Anais do 17º Encontro da Anpad – Finanças, Salvador, 103-110.

Dacorogna, M.M..U.A. Muller, R.J. Nagler , R.B. Olsen and O.V. Pictet, 1993, A geographical model for the daily and weekly seasonal volatility in the foreign exchange market, Journal of International Money and Finance, 12. 413-438.

Dalton, J. M., 1988, How The Stock Market Works, 2nd edition, pages 175-200,Institute of Finance, New York.

DeFusco. R., J. Geppert, and G. Tsetsekos, 1994, Emerging Capital Markets: Results from Cointegration Tests , working paper, University of California.

Delgado, M.A., J. Hidalgo, 1996, Nonparametric Inference on Structural Break Points, mimeo.

Dick van Dijk, Philip Hans Franses and André Lucas, 1998, Testing for ARCH in the presence of additive outliers, Journal of Applied Econometrics, 14, 539-62.

Dickey, D. A. , W. A. Fuller. 1979, Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association, 74: 427-31.

Diebold, F. X. 1988, Empirical Modelling of Exchange rate Dynamics. Springer Verlag, New York.

Diebold, Francis X. , Mark Nerlove. 1989, The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model. Journal of Applied Econometrics. 4: 1-21.

Diebold, F. X. , G. D. Rudebusch. 1991, On the power of Dickey Fuller tests against fractional alternatives. Economics Letter , 35: 155-60.

Diebold, F. X. , G. D. Rudebusch. 1989a. Long memory and persistence in aggregate output. Journal of Monetary Economics, 24: 189-209.

Diebold, F. X. , G. D. Rudebusch. 1989b. Is consumption too smooth? Long memory and the Deaton paradox. Review of Economics and Statistics, 73.

Divecha, A.B., J. Drach, and D. Stefek, 1992, Emerging Markets: A Quantitative Perspective, Journal of Portfolio Management, Fall, pp. 41-50.

Ding, Zhuanxin, C. W. J. Granger and R F. Engle. 1993, A long memory property of stock market returns and a new model. Journal of Empirical Finance, 1: 83-106.

Ding, Zhuanxin, C. W. J. Granger, 1994, Modeling Volatility of Speculative Returns: a New Approach. Mimeo. University of California Los Angeles.

Drost, F. C. , T. E. Nijman, 1992, Temporal aggregation of GARCH processes. Econometrica.

Duffee, R. F, 1992. Reexamining the relationship between stock returns and stock return volatility. Federal Reserve Board Finance and Economics Discussion Series, n.191.

Duecker, M.J. (1997). Markov switching in GARCH process and mean-reverting stock-market volatility. Journal of Business and Economic Statistics. 15, 26-34.

Engle, Robert F. 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50: 987-1007.

Engle, Robert F. 1983, Estimates of the variance of US inflation based upon the ARCH model. Journal of Money Credit and Banking, 15: 2286-301.

Engle, Robert F., C. W. J. Granger. 1987a. Co-Integration and Error Correction: Representation, Estimation and Testing. Econometrica, 55: 251-76.

Engle, Robert F. , Tim Bollerslev. 1986, Modelling the Persistence of Conditional Variances. Econometric Reviews, 5: 1-50.

Engle, Robert F. , Victor K. Ng. 1991, Measuring and Testing the Impact of News on Volatility. Mimeo.University of California San Diego.

Engle, R. F. , D. Kraft, 1983, Multiperiod forecast error variances of inflation estimated from ARCH models . In A. Zellner (editor.), Applied Time Series Analysis of Economic Data, Bureau of the Census, Washington, D.C.

Engle, R. F. , Lilien, D. and Robins, R.P. 1985, Estimating time-varying risk premia in the term structure: The ARCH-M model. UCSD, Discussion Paper, n.85-17.

Errunza, V., 1994, Emerging Markets: some new concepts, Journal of Portfolio Management, 20, 82-87.

Errunza, V., E. Losq, and P. Padmanabhan, 1992, Test of Integration, Mild-Segmentation and Segmentation Hypothesis , Journal of Banking and Finance, 16, 949-972.

Errunza, V., K. Hogan Jr., O. Kini, and P. Padmanabhan, 1994, Conditional Heteroskedasticity and Global Stock Return Distributions, Financial Review, v. 29, n. 3, p. 293-317.

Eun, Cheol S. , Sangdal Shim, 1989, International Transmission of Stock Market Movements, Journal of Financial and Quantitative Analysis, 24, 241-256.

Ferson, Wayne E. 1989, Changes in Expected Security Returns, Risk, and the Level of Interest Rates. Journal of Finance, 44: 1191 -1218.

Fiorentini, Gabriele, Giorgio Calzolari and Lorenzo Panattoni. 1995, Analytic derivatives and the computation of GARCH estimates. Journal of Applied Econometrics, 11. 399-417.

Fornari, Fabio, Antonio Mele, 1996, Sign and volatility-switching ARCH models : theory and applications to international and applications to international stock markets, Journal of Applied Econometrics, 12. 49-65.

Gallant, A. Ronald, David A. Hsieh, and George Tauchen. 1989, On Fitting a Recalcitrant Series: The Pound Dollar Exchange Rate 1974-83.Mimeo. Duke University.

Geweke, John. 1986, Modeling the Persistence of Conditional Variances: A Comment. Econometric Reviews 5: 57-61.

Geweke, J. , S. Porter-Hudak. 1983, The estimation and application of long memory tirne series models.Journal of Time Series Analysis, 4: 220-38.

Giraitis, L., Piotr Kokoszka and Remigijus Leipus, 2000, Stationary Arch models : Dependence Structure and Central Limit Theorem , Econometric Theory, 16, 3-22.

Glosten, Lawrence R., Ravi Jagannathan. and David Runkle, 1989. Relationship between the Expected Value and the Volatility of the Nominal Excess Return on Stocks.Mimeo. Northwestern University.

Granger, C. , P. Newbold, 1977, Forecasting Economic Time Series, New York: Academic Press.

Granger, C. W. J. , R. Joyeux. 1980, An introduction to long memory time series models and fractional differencing. Journal of Time Series Analysis, 1: 15-29.

Granger, C. W. J. ,1980, Long Memory Relationships and the Aggegation of Dynamic Models. Journal of Econometrics, 14: 221-238

Granger, C. W. J. , Z. Ding. 1993, Varieties of Long Memory Models. UCSD Economics Department Working Paper.

Granger, C. W. .J. , Z. Ding, 1993, Some properties of absolute returns. an alternative measure of risk. UCSD Economics Department Working Paper.

Greene, M. and B. Fielitz, 1977, Long-Term dependence in common stock returns. Journal of Financial Economics, 4:571-581.

Guerreiro, J. , R. Leal, 1995, An Application of Multivariate Mutual Information with Emerging Stock Markets, Information Sciences, forthcoming.

Hamao, Y.R., R.W. Masulis, and V. Ng, 1990, Correlation in Price Changes Across International Stock Markets, Review of Financial Studies. 3, 281-307.

Hamilton, J.D. , 1994, Time Series Analysis. Princeton University Press.

Hamilton, J.D. , R. Susnel, 1994, Autoregressive conditional heteroskedasticity and change in regimes . Journal of Econometrics. 64, 307-333.

Hamilton, J.D. . Gang Lin, 1996, Stock market volatility and the business cycle, Journal of Econometrics, 11, 573-593.

Hansen, B.E., 1995, Regression with nonstationary stochastic volatility, Econometrica, 63, 1113-1132.

Hargis, K., 1994, Time-Varying Transmission of Prices and Volatility: Latin American Equity Markets: 1988-1994, University of Illinois at Urbana-Champaign Working Paper.

Harvey, A C.1993, Long Memory in stochastic volatility, Working Paper (London School of Economics, London)

Harvey, A C., E. Ruiz and N.G. Shephard, 1994, Multivariate stochastic variance models, Review of Economic Studies, v.63, 435-463.

Harvey, C. R., 1995a. Predictable Risk and Returns in Emerging Markets, Review of Financial Studies, forthcoming.

Harvey, C. R., 1995b. The Cross-Section of Volatility and Autocorrelation in Emerging Markets, Finanzmarkt und Portfolio Management, forthcoming.

Harvey, C. T., 1995c, The Risk Exposure of Emerging Markets, The World Bank Economic Review, 9:19-50.

He, C. and Timo Terasvirta, 1999, Fourth Moment Structure of the GARCH(p,q) process, Econometric Theory, 15, 824-846.

Hosking, J. R. M. ,1981, Fractional differencing. Biometrika, 68: 165-76.

Hosking, J. R. M. ,1982, Some models of persistence in Time Series. in O.D. Anderson (Editor), Time Series Analysis: Theory and Practice .1: 641-54, Amsterdam: North-Holland.

Hosking, J. R. M., 1984, Modelling Persistence in Hydrological Time Series Using Fractional Differencing. Water Resources Research, 20: 1898-908.

Hurst, H. ,1951. Long-Term Storage Capacity of Reservoirs. Transactions of the American Society of Civil Engineers,116: 770-99.

Hsieh, David A. ,1988, The statistical properties of daily foreign exchange rates:1974-1983. Journal of International Economics, 24: 129-145.

Hsieh, David A. .1989a, Modeling Heteroscedasticity in Daily Foreign-Exchange Rates. Journal of Business and Economic Statistics, 7: 307-17.

Hsieh, David A. ,1989b, Testing for nonlinear dependence in daily foreign exchange rates changes. Journal of Business, 62: 339-368.

International Finance Corporation, 1995, Emerging Stock Markets Factbook 1995 .

Johnson N. L. and S. Kotz, 1970, Continuous Univariare Disnibutions -2 ed., New York, John Wilev & Sons.

Kashyap, R. L. and Eom, K.-B. ,1988, Estimation in Long Memory Time Series Models. Journal of Time Series Analysis, 9: 35-42.

Klien, B. ,1977, The demand for quality-adjusted cash balances: price uncertainty in the US demand for money function. Journal of Political Economy, 85: 692-715.

Koedijk. K. G., Stork. P. A. and De Vries, 1990, Tail index of exchange returns. Journal of International Economics. 29 : 93-105.

Koedijk. K. G., Stork, P. A. and De Vries, 1992, Conditional Heteroskedasticity, realignements and the European monetary system, Mimeo, Katholieke Universiteit Leuven.

Kunsch, H. ,1986, Discrimination between Monotonic Trends and Long-Range Dependence. Journal of Applied Probability, 23: 1025-30.

Lamoureux, Christopher G., and William D. Lastrapes, 1993, Forecasting Stock Return Variance: Towards an Understanding of Stochastic Implied Volatilities. Review of Financial Studies, 5:293-326.

Lamoureux, Christopher G., and William D. Lastrapes, 1990, Heteroskedasticity in stock return data: volume versus GARCH effects. The Journal of Finance, 45: 221-229.

Lamoureux, Christopher G., and William D. Lastrapes, 1990, Persistence in Variance. Structural Change and the GARCH Model. The Journal of Business and Economic Statistics, 8: 225-234.

Leal, R. e M. Austin, 1996, Time Series Properties of Emerging Markets, in Advances in Pacific Basin Financial Markets Vol. II, T. Bos and T. Fetherston, editores, JAI Press, forthcoming.

Leal, R. and Sandoval, E., 1994, Anomalias nos mercados de ações de países de desenvolvimento. In Anais do 18º Encontro da Anpad – Finanças, pp. 213-230, Anpad, Curitiba.

Lemgruber, E.; Becker, J. and Chaves, T., 1988, O efeito do fim de semana no comportamento dos retornos diários de índices de ações. In Anais do 12º Encontro da Anpad, pp. 873-883.

Le Roy, S. F. and Porter, R. D.,1981, Tests Based On Implied Variance Bounds. Econometrica. 49, 555-74.

Leal, R. e E. Sandoval, 1994, Anomalias nos Mercados de Ações dos Países em Desenvolvimento. Anais do 18º Encontro Anual da ANPAD.

Liesenfeld. Roman, Robert C. Jung, 1999. Stochastic volatility models: conditional normality versus heavy-tailed distributions, Journal of Applied Econometrics, 15, 137-60.

Liu, Y. Angela, Ming-Shiun Pan, L. Paul Hsueh, 1994, Volatility Spillover Effects in the U.S. and Pacific-Basin Stock Markets, Working Paper, National Chung Cheng University.

Lo, W.C., H.G. Fung, S.K. Chen, e G.C. Lai, 1993, Dependency in Pacific-Basin Stock Returns, International Review of Financial Analysis, 2, 199-210.

Lo, A.W., 1991, Long-Term Memory in Stock Market Prices. Econometrica,59:1279-313.

Loudon, Geoffrey. Wing H. Watt and Pradeep K. Yadav, 1999, An Ampirical analysis of alternative parametric ARCH models, Journal of Applied Econometrics, 15, 117-36.

Lumsdaine, R.L.. 1991, Asymptotic propertie of the quasi-maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models, Mimeo, Princeton University.

Lumsdaine, R. L.., 1991, Finite sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: a Monte Carlo investigation. Mimeo, Princeton University.

Maddala, G.S., 1992, Introduction to Econometrics, MacMillan.

Mahieu, Ronald J., Peter C. Schotman, 1997, An Empirical application of stochastic volatility models, Journal of Applied Econometrics, 13, 333-60.

Mandelbrot, B., 1963a. The variation of certain speculative prices. Journal of Business, 36: 349-419.

Mandelbrot, B., 1963b, New methods in statistical economics. Journal of Political Economy, 71:421-440.

Mandelbrot, B., 1967, The variation of some other speculative prices. Journal of Business, 40:393-413.

Mandelbrot, B., 1969. Long-Run Linearity. Locally Gaussian Process,H-Spectra, and Infinite Variances. International Economic Review, 10: 82-111.

Mandelbrot, B., 1972, Statistical Methodology for Nonperiodic Cycles: From the Covariance to R/S Analysis. Annals of Economic and Social Measurement, 1, 259-90.

Mandelbrot, B. and Wallis, J.R.,1969, Some Long-Run Properties of Geophysical Records. Water Resources Research, 5: 321-40.

Mandelbrot, B. and M. Taqqu, 1979, Robust R/S Analysis of Long-Run Serial Correlation. Bulletin of the International Statistical Institute, 48, Book 2: 59-104.

McLeod, A.I. and Hipel,K.W., 1978, Preservation of the Rescaled Adjusted Range: A Reassessment of the Hurst Phenomenon. Water Resources Research. 14: 491-518.

Milhoj, Anders, 1985, The Moment Structure of ARCH Processes. Scandinavian Jornal of Statistics. 12:281-92.

Milhoj, Anders, 1987, A Multiplicative Parameterization of ARCH Models. Mimeo. Department of Statistics. University of Copenhagen.

Mills,T.C. ,1990, Time Series for economists.Third Edition,London:Cambrige University Press.

Mills,T.C., 1994, Applied Time Series. First Edition, London:Cambridge University Press.

Mullin, J.. 1993, Emerging Equity Markets in the Global Economy, Federal Reserve Bank of New York Quarterly Review, Summer, pp. 54-83.

Mullins, M., 1988, Meltdown Monday or Meltdown Money: Consequences of the Causes of a Stock Market Crash. Special Paper Number 11 from the LSE Financial Markets Group Special Paper Series.

Nelson, Daniel B., 1989, Mode;ing Stock Market Volatility Changes, Proceedings of the American Statistical Asoociation, Business and Economic Statistics Section, 93-98.

Nelson, Daniel B., 1990, Stationarity and Persistence in the GARCH(1, 1) Model. Econometric Theory, 6:318-34.

Nelson, Daniel B., 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach. Econometrica, 59:347-70.

Nelson, D. B. and Charles Q. Cao, 1992, Inequality Constraints in the Univariate GARCH Model. Journal of Business and Economic Statistics, 10:229-35.

Ng, V.K. and Pirrong, S.C., 1992, Disequilibrium adjustment, volatility, and price discovery in spot and future markets, Mimeo, University of Michigan.

Nicholls, D. e D. Tonuri, 1995, Modelling Stock Market Volatility in Australia, Journal of Business Finance and Accounting, 22, 377-396.

Pagan, A.R. and Wickens, M.R., 1989, A Survey of Some Recent Econometric Methods. Economic Journal. 99:962-1025.

Pantula, Sastry G., 1986, Modeling the Persistence of Conditional Variances: A Comment. Econometric Reviews, 5:71-74.

Park, J. e A. Fatemi, 1993, The Linkages Between the Equity Markets of Pacific-Basin Countries and Those of the U.S. and Japan: a Vector Autoregression Analysis, Global Finance Journal, 4, 49-64.

Parzen,E., 1982, ARARMA Models for Time Series Analysis and Forecasting. Journal of Forecasting,1: 67-82.

Perron, Benoit, 1999. Jumps in the volatility of financial markets, Cahier 0999, Université de Montréal.

Perron, P., S. Ng. 1996,. Useful modifications to some unit roots tests with dependent errors and their local asymptotic properties., Review of Economic Studies, 9, 63-71.

Reston, Steven L., K. Geert Rouwenhorst, 1994, Does industrial structure explain the benefits of international diversification? Journal of Financial Economics, 36, 3-27.

Robinson, P.M.,1991. Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression, Journal of Econometrics, 47, 67-84.

Robinson. P.M.,1995, Gaussian semiparametric estimation of long range dependence, Annals of Statistics, 23, 1048-72.

Robinson, P.M. . M. Henry, 1999, Long and Short memory conditional heteroskedasticity in estimating the memory parameter of levels, Econometric Theory, 15, 299-336.

Rogers, John H., 1992, The currency substitution hypothesis and Relative money demand in Mexico and Canada, Journal of Money Credit and Banking, 24:300-318.

Roll, Richard, 1992, Industrial Structure and the Comparative Behaviour of International Stock Market Indices, Journal of Finance, 47, p.3-42.

Ruiz, E., 1994, Quasi-maximum likelihood estimation of stochastic volatility models., Journal of Econometrics, 63. 289-306.

Rydén, Tobias, Timo Terasvita and Stefan Asbrink, 1997, Stylised facts of daily return series and the hidden Markov model, Journal of Applied Econometrics, 13, 217-44.

Schwert, Willian, 1990. Stock Market Volatility, Financial Analysts Journal, 46:23-34.

Sewell, S., S. Stansell, I. Lee, and M. Pan, 1993, Nonlinearities in Emerging Foreign Capital Markets, Journal of Business Finance and Accounting, 20, 237-248.

Shea, G.S., 1989. Uncertainty and implied variance bounds in long-memory models of the interest rate term structure, Mimeo, Department of Finance, Penn State University Park, Pa.

Shiller, R. J., 1990, Market Volatility - pages 1-4, 71-76, 197- 214, MIT Press, London.

Shiller,R.J. ,1981. Do Stock Prices Move Too Much To Be Justified by Subsequent Changes in Dividends? . The American Economic Review, 71, 421-435.

So, M., K. Lam, 1998. A Stochastic Volatility Model with Markov Switching, Journal of Business and Economic Statistics, 16, 244-53.

Sowell,F., 1987. Maximum-likelihood estimation of fractionally integrated time-series models. Research Paper n. 87-07, Duke University.

Sowell,F.. 1992a. Maximum likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models. Journal of Econometrics. 53:165-88.

Sowell, F., 1992b, Modelling Long-run behaviour with the Fractional ARIMA model. Journal of Monetary Economics, 29:277-302.

Speidell.L.S., R. Sappenfield, 1992. Global diversification in a shrinking world, Journal of Portfolio Management, Fall, 19, 57-67.

Taylor, S., 1986, Modeling Financial Time Series. New York, John Wiley & Sons.

Theodossiou, P.. G. Koutmos, and C. Negakis, 1993, The Intertemporal relation between the U.S. and Greek Stock Markets: A Conditional Variance Tale Analysis, The International Journal of Finance, 5, 492-508.

The Wall Street Journal, several editions.

Watanabe, Toshiashi, 1998, A non-linear filtering approach to stochastic volatility models with na application to daily stock returns, Journal of Applied Econometrics, 14, 101-21.

Weiss,A.A., 1986. Asymptotic Theory for ARCH models: estimation and testing. Econometric Theory, 2:107-131.

West, Kenneth D. Hali J, Edison, and Dongchul Cho., 1993, A Utility Based Comparison of Some Models of Foreign Exchange Volatility. Journal of International Economics.

The World Bank Review, 1995.

Woo, C., A. Lai, Y. Cheung, 1995, Specification Tests of A Market Model of Asia-Pacific Stock Returns: Thailand And Hong Kong, Journal of Business Finance and Accounting, 22, No. 3, pp. 363-375.

Wright, Jonathan H., 1999, Testing for a unit Root in the Volatility of Asset Returns , Journal of Applied Econometrics, 14, 309-318.