CITY UNIVERSITY

TRANSITION EQUITY MARKETS OF CENTRAL EUROPE: VOLATILITY, PREDICTABILITY, INTEGRATION

A THESIS SUBMITTED TO THE DEPARTMENT OF ECONOMICS IN CANDIDACY FOR THE DOCTOR OF PHILOSOPHY IN ECONOMICS

DEPARTMENT OF ECONOMICS

MARIA KASCH-HAROUTOUNIAN

SEPTEMBER 2000

CONTENTS*

CHAPTER 5

The International Capital Asset Pricing Model and the Integration of the Transition Equity Markets of the Central Europe into the International Capital Markets

CHAPTER 6

Conclusions

LIST OF TABLES

LIST OF FIGURES

* The version of the chapter 3 of this thesis is forthcoming in the *Applied Financial Economics*. The version of the chapter 4 is submitted to the *Emerging Markets Review*.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Professor Simon Price who has been an exceptional supervisor. His consistent support, guidance and encouragement, as well as ability to keep my spirits high throughout two and a half years of my PhD were crucial in helping me to complete my thesis.

My PhD research was funded by grants from the British Council, the BFWG Charitable Foundation and the Benlian Trust. This thesis would not have been possible without the financial support I received from these organisations.

Finally, I would like to thank my husband, Lorenz, for his inspiration, love and support.

DECLARATION

The author of this thesis grants power of discretion to the university librarian to allow this thesis to be copied in whole or in part without further reference to her. This permission covers only single copies made for study or research purposes subject to normal conditions of acknowledgement.

ABSTRACT

The objective of this thesis is to add evidence from the transition equity markets of Central Europe to the econometric modelling of financial time series by addressing the issues of volatility, predictability and international asset pricing in these markets.

In Chapter Two we start from an overview of the transition stock markets by presenting their historical background, basic regulations, statistics, and stock market indices.

Chapter Three focuses on the modelling of univariate and multivariate volatility in transition equity markets. Our sample has all the previously documented characteristics of the unconditional distribution of stock returns normally used to justify the use of the GARCH class of the models of conditional volatility. Strong GARCH effects are apparent in all series examined. The estimates of asymmetric models of conditional volatility show rather weak evidence of asymmetries in the markets. The results of the multivariate specifications of volatility have implication for understanding the pattern of information flow between the markets. The constant correlation specification indicates significant conditional correlation between three pairs of countries: Hungary and Poland, Hungary and Czech Republic, and Poland and Czech Republic. The BEKK model of multivariate volatility shows evidence of return volatility spillovers from Hungary to Poland, but no volatility spillover effects are found in the opposite direction.

Chapter Four examines the linear and nonlinear predictability of transition equity returns with simple technical trading rules. The application of the moving average trading rules to the data reveals that technical analysis helps to predict stock price changes. Firstly buy signals consistently generate higher returns than sell signals; secondly the returns following buy signals are less volatile than returns following sell signals. The application of the bootstrap methodology to check whether three popular null models of stock returns with linear conditional mean specification replicate the trading rule profits indicates that returns obtained from trading rules signals are not likely to be generated by these models. Comparison of the out-of-sample forecast performance of linear and nonlinear (feedforward networks) conditional mean estimators with past trading signals in the conditional mean equation indicates substantial forecast improvements of the feedforward network regression.

Chapter Five addresses the issue of integration of the transition equity markets into the global capital market by testing pricing restrictions of the international CAPM simultaneously for four national equity markets: two developed markets (U.S. and Germany) and two new transition markets (Hungary and Poland). Methodologically, we extend the BEKK multivariate GARCH specification to accommodate GARCH-M effects, and propose an alternative specification of the conditional CAPM, which allows return volatility transmissions between the markets in the system. The results reveal that the world price of covariance risk is positive and equal across the markets. This is consistent with the international CAPM and supports the hypothesis of integration of the transition markets into the global market. However, our further results indicate individual significance of the Hungarian idiosyncratic risk, pointing to some level of segmentation of the Hungarian market. Moreover, the introduction of world-wide information variables into the system reveals that some variation in the excess national returns is still predictable after accounting for the measure of market-wide risk.

CHAPTER 1

INTRODUCTION

1.1 Objectives and Structure of the Thesis

Financial econometrics has developed a range of models to account for empirical regularities in financial data. Most contributions to this literature deal with mature, large and liquid markets. However, the financial markets in the transition economies of Central and Eastern Europe have been studied only by a small number of authors.[1] The objective of the thesis is to add evidence from the transition equity markets of Central Europe to the econometric modelling of financial time series.

The analysis focuses on four markets of Central Europe: Hungary, Poland, Czech Republic and Slovakia. These countries were chosen because of the advancement of their capital markets due to a greater political stability and a rapid economic growth compared to other transition markets in the region. Chapter 2 of the thesis gives an overview of the transition equity markets by presenting their historical background, basic regulations and statistics, as well as stock market indices.

[1] Chobanov et al. (1996) estimate the distribution of the Bulgarian lev / US$ rate. Emerson et al. (1996) propose the state space model to study the evolving market efficiency for Bulgarian bank stocks. Erb, Harvey and Viskanta (1996) analyse returns and volatility for 135 markets, including Albania, Bulgaria, Czech Republic, Hungary, Poland and Russia. Meszaros (1996) studies the EMH on the Bratislava Stock Exchange. Zalewska-Mitura and Hall (1999) consider whether the Hungarian stock market becomes more weak-form efficient over time. Wajerowska (1999) examines validity of the optimal portfolio allocation model with price constraints on the example of the Warsaw Stock Exchange. Linne (1999) uses a Kalman filter approach to consider the integration of the Central and East European equity markets into the international capital markets. Charemza and Syczewska (1999) apply the Dickey-Fuller and KPSS tests to the East European time series.

The rest of the thesis is divided into three major parts, examining issues of volatility, predictability, and integration of the transition markets of the Central Europe into the international capital markets.[2]

Chapter 3 focuses on modelling of volatility of the transition equity markets. Although economists have long been interested in the analysis of behaviour under uncertainty, it was not until the beginning of the 1980s that econometricians have begun developing an analytical framework to deal with uncertainty. A central feature of this framework is the modelling of second and, possibly, higher moments as well. One of the prominent tools used to model the second moments is due to Engle (1982). Engle (1982) suggested that these unobservable second moments could be modelled by specifying a functional form for the conditional variance and modelling the first and second moments jointly, giving what is called in the literature the Autoregressive Conditional Heteroscedastic (ARCH) model. The linear ARCH model was later generalised by Bollerslev (1986), and called Generalised ARCH (GARCH). The ARCH models were first formal specifications that seemed to capture the stylised facts characterising financial data. Within the last two decades the GARCH literature has grown in spectacular fashion. The aim of this chapter is to add evidence from the transition equity markets of Central Europe to the univariate and multivariate GARCH modelling of conditional volatility of financial time series.

We start from examination of the statistical properties of the sample from transition markets, which reveals that it has all of the previously documented characteristics of the unconditional distribution of stock returns that are used to justify the GARCH specification of the time-varying volatility. In the univariate section of our empirical analysis we estimate and contrast several symmetric and asymmetric GARCH models of conditional volatility. The alternative asymmetric models allow examination of the different types of asymmetry in the impact of news on volatility of

[2] The estimation of all models employed in this thesis is carried out using RATS (Regression Analysis of Time Series) econometrics software.

the transition equity markets. In the multivariate section we consider the interdependence between conditional second moments of the return series, which helps understanding the pattern of information flow between the markets. We employ two alternative multivariate GARCH specifications of conditional volatility, which allow analysis of the conditional correlation and volatility spillover effects between the markets in the system.

Chapter 4 addresses the issue of return predictability in the transition markets by employing technical trading strategies for buying and selling in the market. Technical analysis is considered by many to be the original form of investment analysis. Over the years, investors and technical analysts have devised hundreds of technical market indicators in an effort to forecast the trend of the security market. Technical traders base their analysis on the premise that the patterns in stock prices are assumed to recur in the future, and thus, these patterns can be used for predictive purposes. Historically, technical analysis has never enjoyed the same degree of acceptance by the academic finance community that, for example, fundamental analysis has received. However, the recent studies on technical trading strategies[3] signal a growing interest in technical analysis among financial academics.

The chapter is divided in three major parts. In the first part, the application of the technical trading rules to the data from four transition equity markets reveals that technical analysis helps to predict stock price changes. In the second part, we employ the bootstrap methodology to check whether popular null models of stock returns are able to replicate trading rule profits obtained in the first part of the chapter. And finally, in the third part we compare the out-of-sample forecast performance of linear and nonlinear (feedforward networks) conditional mean estimators of stock returns with past trading signals in the mean equation.

[3] See e.g. Blume, Easley and O'Hara (1994), Brock, Lakonishok and LeBaron (1992), Brown and Jennings (1989), LeBaron (1996), Neftci (1991), Pau (1991), Taylor and Allen (1992), and Treynor and Ferguson (1985)

Chapter 5 uses models of international asset pricing to consider the issue of integration of the transition equity markets of Central Europe into the international capital markets. The question of why different countries' market indices command different expected returns lies at the foundation of international finance. In studying assets of one particular country, we would say that differing expected returns are due to differing risk exposures of these assets. In international markets, the answer is more difficult. Aside from the obvious complications arising from country-specific exchange rates, the risk is hard to quantify if a country is not fully integrated into the world capital markets. Markets are fully integrated if assets with the same risk have identical expected returns irrespective of the market. Risk refers to exposure to some common world factor. If a market is segmented from the rest of the world, its covariance with a common world factor may have little or no ability to explain its expected return.

The Capital Asset Pricing Model (CAPM), developed by Sharpe (1964), Lintner (1965), Mossin (1966), and Black (1972), was the first equilibrium asset pricing model, and it remains one of the foundations of financial economics. In our analysis we test an international version of the CAPM (ICAPM), which implies that if international markets are fully integrated, the world market risk is the only relevant pricing factor[4], and the price of this risk should be positive and equal across all markets. Methodologically our model extends the multivariate GARCH specification of Engle and Kroner (1995) ("BEKK") to accommodate GARCH-M effects, and allows volatility spillovers between the markets in the system.

To test the performance of the benchmark ICAPM we compare it to two alternative specifications. The first alternative allows for some level of market segmentation by introducing country-specific risk and a constant term (to reflect other forms of segmentation) into the model. The second

[4] Our approach assumes that investors do not cover their exposure to exchange rate risk or, equivalently, that the price of exchange risk is equal to zero.

alternative tests the hypothesis that some of the variation in the excess returns can still be explained by a number of information variables, even after accounting for the world market risk. All the models we specify are first estimated with the additional assumption that the price of risk is constant. Then this assumption is relaxed, and the measure of market risk is allowed to vary through time.

Finally, chapter 6 concludes the thesis by summarising the main results and identifying interesting issues for future econometric research.

CHAPTER 2

OVERVIEW OF THE TRANSITION MARKETS OF CENTRAL EUROPE

2.1 Introduction

At the end of the 1980s the post-communist countries of Central and Eastern Europe (CEE) started to transform their economies from centrally planned to free market systems. The stock markets of these countries are still very young and small compared to the advanced markets in industrial countries. The nine largest stock markets in CEE (Croatia, Czech Republic, Estonia, Hungary, Latvia, Poland, Slovakia, Slovenia and Russia), although having grown rapidly in recent years, had only a combined market capitalisation at the end of 1998 of some USD 70 bn. This was about 10 per cent of the capitalisation of the Frankfurt Stock Exchange.

In our study we focus on a sample of Central European stock markets, namely the Czech, Hungarian, Polish, and Slovak markets. Although, they share the common goal, a well-functioning free market economy, the paths chosen to obtain this differ from country to country. Different approaches have been necessary due to discrepancies in the infrastructure, different levels of economic development, and different historical and cultural traditions. For instance, privatisation, along with the creation of a capital market, was achieved through a mass (the Czech Republic) or gradual (Hungary, Poland) share distribution schemes. The choice of privatisation methods significantly determined the way in which financial structures were created. In the Czech Republic, where the transfer of ownership from the state to private citizens was accomplished mainly through voucher privatisation, shares of 1500 enterprises were offered in one public offering, with only minimum disclosure requirements and market regulations. The government's main contribution to the development of the capital market's infrastructure was to create a system where ownership of shares was recorded in a central, computerised share registry. The government has adopted a hands-off approach to regulation and has allowed the various markets and exchanges to compete with each other. Furthermore, the government has only ensured that prices of the large trades made outside the organised exchanges been publicly disclosed. In contrast, in Hungary and Poland the creation of a capital market and privatisation were introduced as an integrated process. Offering shares of medium and large-sized companies to the public through the stock market was the main method of privatisation. Regulation of public offering and trading on the secondary market closely imitated western standards. Transparency and prudent behaviour were the primary objectives in designing the privatisation programme and the secondary trading market.

The rest of this chapter presents a short historical background, basic regulations and statistics of the Central European stock markets.

2.2. The Hungarian stock market

2.2.1. Historical background of the Budapest Stock Exchange

Established under a decree by Emperor Franz Joseph I in the city of Pest in 1864, the Budapest Commodity and Stock Exchange began its really spectacular course of development after the political compromise of 1867. With the turn of the century approaching, Hungary took an active role in European economy, and the quotes of the Budapest Exchange were regularly published in Vienna, Frankfurt, London and Paris. The great economic crisis of the early 1930s also hit Hungary, and the Exchange was closed from summer 1931 to fall 1932. The BSE stopped operating in 1948.

Hungary was the first country of the post-Soviet block to start to reform its financial system. In 1986 a two-tier banking system was introduced, with the domestic commercial banking operations of the National Bank of Hungary and State Development Bank taken over by three new commercial banks.[5] These banks inherited the loans and deposits of their predecessors. The Foreign Trade Bank was also allowed to provide commercial banking services. In 1988-89, banks were granted additional rights to provide foreign exchange related services to their clients. With the creation of the two-tier banking system, the National Bank of Hungary gradually shifted from direct credit ceiling to the use of indirect monetary instruments. The next step was establishment of the law on commercial banks, which was adopted in January 1992, imposing the Basle-defined standard for capital adequacy on Hungarian banks. Simultaneously with changes in organisation of the financial structures the Budapest Stock Exchange re-opened its gates on June 21, 1990, with share trading for eight major companies.

[5] The commercial banks were still owned by the state and state enterprises.

The Exchange has come a long way in the 9 years since then only to develop into an internationally recognised market with advanced regulatory and technical background, and to operate as a major factor in the Hungarian economy and the focal point of the country's capital market. The decision made by the London Stock Exchange, to enter the BSE on its list of approved exchanges as the first such institution from Central and Eastern Europe, was regarded as a significant achievement.[6]

2.2.2. Trading on the BSE

The Budapest Stock Exchange is organised on the Anglo-Saxon model, hence all transactions are to be implemented on the Floor of the Stock Exchange during the opening hours by brokers authorised by the Members of the Stock Exchange. During the trading session the brokers may make bids and offers orally or in writing and may respond to the bids and offers of other brokers in accordance with the provision of the Rules of the BSE. If no trade is carried out for a given security on a given day, then the price of the last trade of the security is regarded as the closing price, which means that the closing price of non-traded asset remains unchanged as long as non-traded period lasts.

As mechanisms to stabilise the movement of prices, breaks (or suspensions) of trading were introduced. A break means a suspension of trading in a given share, and therefore no bids or offers can be made for it. The temporary suspension of the sale and purchase of a given security can be ordered by a Speaker if the change in the price of the of shares admitted to the Stock Exchange is greater than +/-10%, but less than +/-20%, compared with the opening price. The trading break may last 5 minutes. If the change in price compared with opening price is bigger than +/-20%, the break may last 5-10 minutes. Another suspension, ordered by the Chief Executive of the

[6] For the summary of market statistics of the Budapest Stock Exchange over the period 1990-1998 see Appendix 2.2, Table 2.5 at the end of this chapter.

Exchange, can stop trading in a certain type of security or all the securities on the Exchange, if the CEO deems that continued trading would endanger the well-being of investors or the operation of the Stock Exchange. This kind of suspension may last one day. The Stock Exchange Board or the Supervision Board may suspend trading in certain types of securities or in trading overall on the Stock Exchange, if the general financial, economic or political situation does not provide for the organised and transparent trading on the Stock Exchange. This kind of suspension should not exceed ten days. If the need for suspension exceeds this period, it has to be approved by the Ministry of Finance.

In the spring of 1997 the Annual General Meeting (AGM) of the BSE developed a new model for its operation. The major elements of the new model are the establishment of market segments for different instruments (sections), changes in composition of the exchange members entitled to trade on the stock exchange, as well as changes in the voting rights of participants in the general meeting and in the structure of the fees.

The AGM of the BSE set up three sections with the following product lines effective from 1 June 1997:

Corporate securities - shares, company bonds, mutual fund notes, compensation coupons and non-standardised options;

Government securities - government bonds, treasury bills;

Derivative section - standardised futures, and from autumn 1998, standardised options.

The new act has significantly changed the principle of regulation by opening certain markets of the stock exchange to market participants, primarily banks, which were kept at the arm's length since the foundation of the BSE in 1990. Until 1997, only stockbrokers operating in a company form (limited liability or share company) were allowed to be members of the BSE and, consequently, traders. The new act has opened up markets on government securities and derivative products for credit institutions. Moreover, private individuals can make transactions personally (on their account) in the derivative market.

The new act also changed voting rights granting members from 1 to 10 votes depending on their turnover per section. This means that each member may have from 1 to 30 votes (three sections in total) at the calendar year preceding the convocation of the general meeting.

The new regulation created two main types of fees in the stock exchange: entry (stock exchange/section) fees and turnover fees. The turnover fees, to be paid by both buyers and sellers, are different in the various sections but identical within one section.

Public offers of securities within Hungary may be made following a prospectus approved by the Hungarian Banking and Capital Market Supervision (HBCMS). There are two listing categories on the BSE, 'listed A' and 'listed B'. The principal requirements for a full listing on the BSE are:

Listed A category	Listed B category
Only full share series can be listed	Only full share series can be listed
Share series to be listed minimum HUF 5 billion market value	Share series to be listed minimum HUF 100 million market value
Public ownership proportion*: either minimum 25 per cent and HUF 2 billion market value or minimum HUF 5 billion market value	Public ownership proportion: either minimum 10 per cent and HUF 50 million market value or minimum HUF 200 million market value or minimum 100 shareholders
Minimum1000 owners	Minimum 36 owners
3 completed and audited financial years	1 completed and audited financial year
Series can only consist of registered shares	Series can consist of registered or bearer shares

* Public ownership proportion: the portion of the share series whose owners have less than a five per cent stake.

Once listed, a company must conform to certain continuous disclosure requirements.

2.2.3. The Budapest Stock Index (BUX)[7]

The Budapest Stock Index (BUX) replaced the unofficial Budapest Stock Exchange Index, which was used during the initial phases of economic transition. Currently, the BUX contains 20 stocks. To qualify for the index, a stock has to comply with three out of five requirements, including a certain minimum face value of a stock, a defined minimum price, a minimum number of transactions, and a cumulated minimum turnover of 10% of the registered capital during the six months preceding the revision of the index.

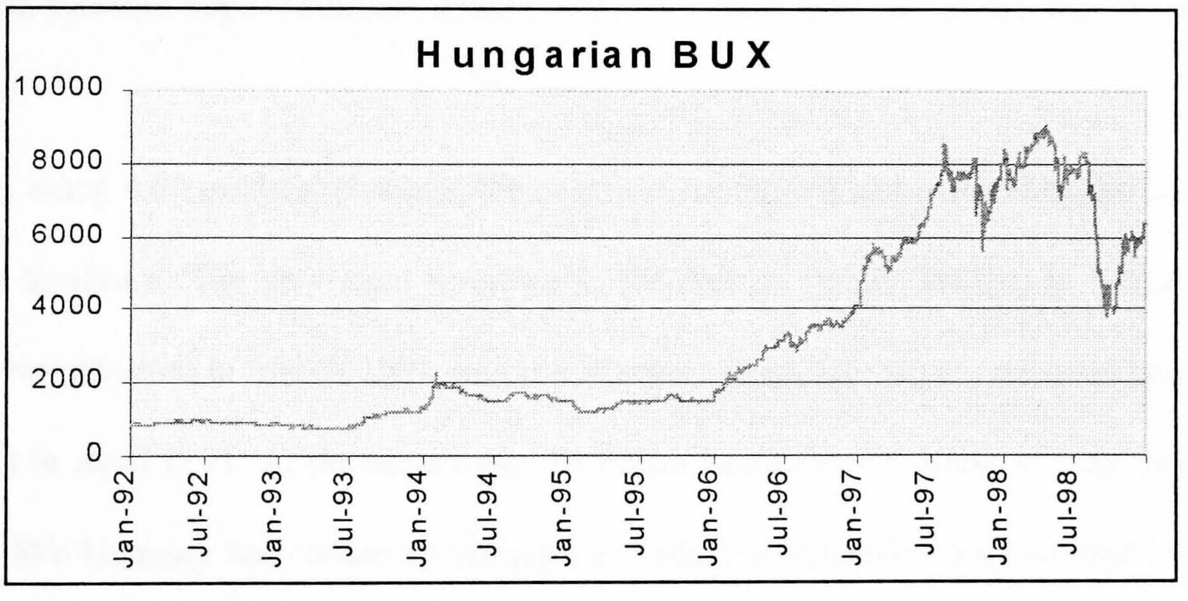
If more than 25 stocks meet at least three out of these five requirements at the time of the semi-annual updating of the index sample (on March 31 and September 30), the index stocks are selected according to their so-called global rank, which is derived from a weighted average across the five criteria.

The Figure 2.1 below presents historical development of the Hungarian BUX index in 1992 – 1998. The Hungarian market boomed in 1996 and 1997, experiencing some of the fastest rates of

[7] For the composition of the index see Appendix 2.1, Table 2.1 and Figure 2.5 at the end of the chapter.

growth among emerging markets. In April 1998 the index reached an all time high of 9,016.31, up 12.7 % from the year-end 1997. However, the BUX fell to 7,049.35 in the week after the Fidesz election victory. Instability in the emerging markets further pushed down the index to a low of 4,891.65 in early September. Foreigner investors account for the bulk of trading in the stock market (more than 50%, compared e.g. with 30% in Warsaw), which accounted for the greater volatility in the market in the wake of the Russian devaluation and default of August 1998. Foreigners may have sought to compensate for Russian losses by taking advantage of the liquidity of the Hungarian market.

Figure 2.1



2.3. The Polish Stock Market

2.3.1. Historical Background of the Warsaw Stock Exchange (WSE)

The first stock exchange in Warsaw was opened on May 12, 1817. Trading sessions were held between 12.00 and 13.00. In the nineteenth century mostly bonds and other debt instruments were traded on the Warsaw bourse. Trading in equities developed in the first half of twentieth century. Before the World War II seven stock exchanges operated in Poland. Warsaw accounted for more than 90% of the total trading. In 1938 there were 130 securities traded on the Warsaw Exchange: municipal, corporate and government bonds as well as shares. Due to the change of political and economic systems, capital markets could not be re-created after the World War II was over.

In 1989, along with political changes, the new non-communist government began creating a capital markets structure. The new legal framework, the Act on Public Trading in Securities and Trust Funds, was adopted in March 1991, and the Warsaw Stock Exchange was established by the State Treasury in April 1991. At the same time, the Polish Securities Commission was created.
Poland, like Hungary has chosen to undergo a gradual privatisation process, and in October 1990 the new post-communist government passed the first laws allowing the privatisation of the state-owned firms. Among 8,500 companies, reported to belong to the state, five firms (*Exbud, Krosno, Prochnik, Slaska Fabryka Kabli and Tonsil*) were privatised. They were also the first five companies listed at the first trading session of the Warsaw Stock Exchange. The new listings on the WSE have been the combination of initial public offerings by post-1989 private companies and privatisation floatations, first by the state property fund and more recently by the 15 mass-

privatisation National Investment Funds (NIFs). The NIFs' own shares began to trade on the WSE during 1997. Bank shares have been particularly popular with WSE investors.[8]

2.3.2. Trading on the Warsaw Stock Exchange

The Warsaw Stock Exchange is a 'Lyon type' capital market, which can be described as 'order-driven', centralised and paperless. This system is based on the establishment of a single price for each stock by comparing buy and sell orders submitted to the Stock exchange by the brokerage houses before session. This method of fixing the price is similar to the French 'par casier' or German 'Einheitkurs' quotation system. In July 1996, the trading system started to be reformed in order to allow trades to be executed by continuous trading and block trading systems, in addition to the established single-price auction system.

A. Single price-auction

The main feature of the single-price auction is that a single price per security emerges at each session as a result of the orders submitted. Transaction can be concluded even for one individual security. In their orders, clients of brokerage houses define the quantity and the price of securities. In his order investor can indicate the price limit or "at the market". The validity of an order can not exceed the end of the next month. The price on the equities market can only be higher (upper limitation) or lower (lower limitation) from the previous session's price by a maximum of 10%. Bonds maximum price change is 5 percentage points. Price for shares, subscription rights, allotment certificates and NIF certificates is given in Polish zlotys. For bonds, price is given in

[8] For the summary of market statistics of the Warsaw Stock Exchange over the period 1991-1998 see Appendix 2.2, Table 2.6 at the end of this chapter.

percentage of the nominal value. The settlement price for bonds is calculated by adding accrued interest to the market price.

During trading sessions, a key role is played by brokerage houses operating on the WSE as specialists. The issuer of the security with the consent of the Management Board nominates the specialist. The main role of the specialist, represented on the trading floor by specialist brokers, is to establish the session's price. After the opening of the trading session, the specialist broker receives a list of orders participating in the price setting process qualified for the session, verifies them and determines the price for a given security. In determining the session's price, the specialist is obliged to follow the rules listed below in the given order of importance:

- maximise turnover of that security;
- reach the smallest possible difference between demand and supply;
- minimise the difference in price between that of the current session and the previous one.

After the specialist has established the price all transactions in a given security are executed at that price. If it is possible that all market orders are executed at a given price, then this price is called an equilibrium price, and the market on which this price is established is called a balanced market. If the price, which maximises turnover and minimises the difference between demand and supply, exceeds the lower or upper price limit (previous trading session's price +/- 10%), it cannot be accepted by the exchange. In this case, the specialist establishes the session's price at the lowest or highest acceptable level. The price established in such a way is not the equilibrium price, and the market is called unbalanced.

In the case of an unbalanced market, the specialist is obliged to determine the degree of market imbalance between supply and demand at the given price. If the imbalance ratio of demand to supply (or supply to demand) is greater than 5:1, the imbalance is too high for transactions to be concluded. In the case of a predomination of buy orders, a non-transactional price is published on the stock exchange quotation list followed by the symbol *ok*. In the case of a predomination of sell orders, the price is followed by the symbol *os*.

The specialist can intervene to reduce the market imbalance. This means that during the session, after the price has been established, the specialist broker may buy or sell securities from his own inventory.

*B. Continuous Trading System*The system of continuous quotations used by the WSE was developed in 1992 and initially was used for Treasury bonds. On July 8, 1996 the Warsaw Stock Exchange introduced to continuous trading first five shares of listed companies selected by a ballot from among companies with highest liquidity. On August 12, 1996 National Investment Funds certificates and on July 1, 1997 shares of fifteen National Investment Funds started to be traded in continuous trading. In 1998 futures on WIG20 share index (January 16), warrants (March 9) and futures on USD exchange rate (September 25) were introduced to continuous trading. Unlike in the single price auction system, where a single share or certificate can be traded, in the continuous trading the transaction units consist of round lots. The size of a round lot is determined individually for each security.

The continuous trading session starts at 1.00 p.m. Before the session, between 8 a.m. and 1.00 p.m., the Exchange accepts only limit orders. Orders may carry an indication of time validity, although much shorter than in the single-price auction, since it cannot exceed the end of the week following the one in which the order was placed. At

1.00 p.m., the Exchange stops accepting orders until the opening price has been announced, or continuous trading has commenced. The opening price is determined based on the orders placed between 8 a.m. and 1.00 p.m., as well as orders which have not been executed during previous sessions and the validity of which had not expired. The mechanism for setting the opening price is the same as for single-price auction.

The opening price cannot be higher or lower from the reference price by more than 3 percentage points for bonds and 5% for shares and NIF certificates. The reference price is the price from the previous session's single-price auction, and if a security is not traded in the single-price auction, it is the price of last transaction from the previous continuous trading session. If the opening price cannot be established at

1.00 p.m., the trading session is commenced and the first transaction's price becomes the opening price. The session lasts until 4.00 p.m. The session of continuous trading starts with identifying the best bid and ask. The first criterion of selecting orders to be executed is the price limit, and with orders with the same price limit, the time of their placement. Market orders can be accepted only when limit order on the opposite side of the market is waiting in the system.

C. Block Trades

Large blocks of securities can be traded off-session as so-called block trades. Such a transaction takes place, when Exchange member presents to the Exchange buy and sell orders for the same number of securities at the same price. The value of securities in order to be accepted as a block trade must be at least equal to 300 000 PLN or at least 1% of securities introduced to trading on the exchange. For securities, which are admitted to trading but have not yet been introduced to trading, the block must be at least 5% of the number of securities admitted.

The price of securities in block trade may differ from the last session's price by no more than 15%, when the number of securities in the block is less than 5% of securities admitted to trading. When the block is larger than that, the price may differ from the last session's price by no more than 30%. The Exchange charges smaller fee for the block trades, than for transactions concluded for the same security during the session.

2.3.3. The Warsaw Stock Exchange Indices[9]

The Warsaw Stock Exchange Index WIG was the first index to be introduced after the reopening of the Warsaw Stock Exchange. It is calculated as a total return index for the main market once per trading day after each session. The weight of an individual stock according to its market capitalisation is limited to 10% of the index sample. Furthermore, a single sector may not account for more than 30% of the index. The index is regularly revised every three months, mainly to account for the introduction of new stocks.

In addition to the WIG, the Warsaw Stock Exchange also publishes an index for the parallel market (WIRR) and the Warsaw Stock Exchange Price Index (WIG20). The WIG20 measures the aggregate price change of twenty domestic stocks positioned best according to the number of ranking points obtained. The formula for determining the number of index points draws upon both the turnover and market capitalisation during the month preceding the quarterly revision of the index sample in a proportion of 60:40. In addition, to qualify as a new stock in the index sample, a

[9] For the composition of the WIG 20 see Appendix 2.1, Table 2.2 and Figure 2.6 at the end of the chapter.

stock must have ranged among the twenty leaders in at least two of the last three months in terms of turnover.

In order to avoid a dominating influence of the largest and most actively traded companies on the index, the maximum number of ranking points attainable equals 150 times the number of ranking points obtained for a stock with average turnover and market capitalisation. Based on this modified record of ranking points, the number of shares admitted to enter the index portfolio is determined, thereby assigning a specific weight to each index stock.

The Figure 2.2 below presents historical development of the Polish WIG 20 index in 1992 – 1998. Similarly to Hungarian market, the Polish stock market had seen significant increase in the index value in 1996 and 1997. However, it was hit by instability in emerging markets in summer-autumn 1998.

Figure 2.2

2.4. The Czech Stock Market

2.4.1. Historical Background of the Prague Stock Exchange

While the Prague Stock Exchange has a very short contemporary history, it has a rich tradition extending back for over a century. On March 23, 1871, an Exchange for securities and commodities trading opened its doors in Prague for the first time. It was, however, closed during World War I, and its operation recommenced on February 3, 1919. In the period between the two world wars, the Prague exchange recorded a great boom and rivalled virtually any other market on the continent. Its growth was interrupted by the onset of World War II, following which the exchange was closed and, during the communist era, constitutionally eliminated. It took over a half-century for the tradition of securities trading to resume in the Czech Republic.

In May 1991 the Preparatory Committee for the Establishment of a Stock Exchange was created. A new company was established by eight banks and, on August 24, 1992, this was transformed into an association and subsequently converted into a trading company pursuant to the newly established Stock Exchange Act. The Prague Stock Exchange, with 17 founding members, was incorporated as a joint stock company on November 24, 1992, upon its entry into the Trade Register. The first trading session was held on its floor on April 6, 1993.

The stock exchange's renewal is related to the voucher privatisation, which denationalised extensive assets. The two-stage voucher privatisation took place in the period 1992-1995. First, in 1993 almost 1,000 issues, and subsequently, in 1995, more than 700 were registered at the stock exchange. Marker capitalisation amounted to CZK500 billion in 1996.

It soon became apparent after registration the voucher privatisation's new titles that not all issues were sufficiently attractive. Gradually, the stock exchange split them into three markets, depending on the degree of their liquidity and the companies' market capitalisation. In 1997 1,301 non-liquid share issues were excluded from the Exchange's Free market.[10]

2.4.2. Trading on the Prague Stock Exchange

Since September 1, 1995, the Prague Stock Exchange (PSE) has provided trading in three markets: Main, Secondary and Free markets. The Main and Secondary markets have emerged from the original Listed market, and the Free market comprises the former Unlisted market. The whole process of the market segmentation was primarily motivated by the Exchange's efforts to clearly profile two basic groups of securities.

The prestigious markets - Main and Secondary - trade in top-quality securities. They have to provide regularly financial information and report corporate actions, which may have an impact on the official price of the security.

The conditions for admission of securities issued by joint stock companies (JSCs), investments funds and unit trusts to the Main market or Secondary market are presented below.

[10] For the summary of market statistics of the Prague Stock Exchange over the period 1993-1998 see Appendix 2.2, Table 2.7 at the end of the chapter.

A. Main market

Conditions	JSCs excluding funds	Investment funds and unit trust funds
Registered capital amount(for unit trusts – value of issued units)	according to the Commercial Code	minimum CZK 500 million
Value of the part of issue that was issued throughpublic offer	minimum CZK 200 million	–
Percentage share of the part of issue that was issued through public offer of the total value of the issue	minimum 20 %	–
Duration of the business activities	minimum 2 years	minimum 2 years
For shares and units each issue has to be liquid enough; the liquidity criteria are set by the Exchange Listing Committee, and its decision is published well before their effectiveness. For the year 1998 the average daily trade value must be higher than CZK 1 million.		

B. Secondary market

Conditions	JSCs excluding funds	Investment funds and unit trust funds
Registered capital amount(for unit trusts – value of issued units)	according to the Commercial Code	minimum CZK 250 million
Value of the part of issue that was issued throughpublic offer	minimum CZK 100 million	–
Percentage share of the part of issue that was issued through public offer of the total value of the issue	minimum 15 %	–
Duration of the businessactivities	minimum 2 years	minimum 2 years

The basic types of trades carried out on the PSE are:

* Trading at fixed price (fixing)* Continual trading at fixed price (supplementary orders)

* Continual trading at variable price (KOBOS)

* Trading under the Trades in Shares and Bonds Supporting System (SPAD)

* Purchase and sale of units of open unit trust funds

Trading at fixed price (fixing) is based on the processing of buying and selling orders accumulated to a single moment of time. Once the orders have been processed price for each issue is set in such a way, as to enable the achievement of the maximum possible trade volume of securities. The newly set price may only vary from preceding price by the allowable spread, which has been set at 5 %.

The purpose of the *continual trading at fixed price* is offsetting the excess occurred under continual trading at fixed price by inputting supplementary orders in the trading system. Only orders with limit price equal to the set auction price may be input.

Under *continual trading at variable price*, or so called *KOBOS*, trades are concluded on the basis of a continual inputting of buying and selling orders for securities. Depending on the immediate bid and offer, the price of securities is set continually, while the official price is always equal to the price of the last concluded trade. Principle of price and, subsequently, of time priority applies to order matching. Under *KOBOS* trading, it is possible to input orders that are valid for a period longer than one exchange day.

System Supporting the Market for Shares and Bonds (*SPAD*) is a trading segment based on exploitation of function performed by market makers maintaining continual quotation of bid and ask price for selected issues. Trading under *SPAD* is divided into two parts: open phase, the phase with obligatory quotation of prices by appointed market makers, and closed phase, the phase without obligatory quotation of prices by appointed market makers.

2.4.3. The Prague Stock Exchange Index (PX 50)[11]

The PX-50 is the leading representative of a family of indexes (so far 21 indices) introduced by the Prague Stock Exchange.[12] It consists of 50 issues, which are selected according to market

[11] For the composition of the index see Appendix 2.1, Table 2.3 and Figure 2.7 at the end of the chapter.

[12] Other 21 indices introduced by the Prague Stock Exchange include PX-Glob (encompasses all listed stocks and investment fund shares that have been traded at least once), PXL (was introduced to document the performance of stocks which are traded in the official market segment without being included in the PX-50, but was abolished in

capitalisation, liquidity and branch classification. In detail, the following rules for updating the PX-50 index base apply:

* Definition of the index sample

  - Market capitalisation (issue ranging among the 100 largest issues according to the fourth trading day preceding the index revision)

  - Trading volume (issue ranging among the 80 most actively traded issues during the preceding quarter)

  - Sectoral balance (individual issues are selected starting from those with the highest market capitalisation under the condition that the sectoral composition of the index sample reflects the sectoral balance of total market capitalisation)

* Criteria for non-admission

  - Stocks of bankrupt companies

  - Stocks of companies in which trading has been suspended for a long time

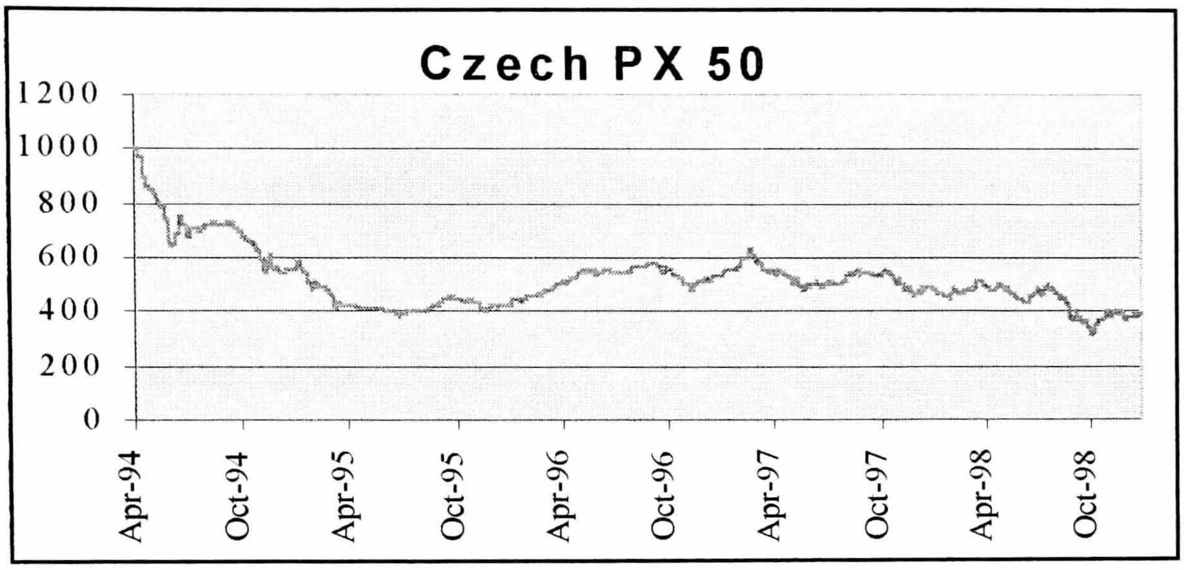
  - Stocks of investment funds

* Semi-annual revision of the index

The Figure 2.3 below presents historical development of the Czech PX 50 index in 1994 – 1998. The volatility of the Czech stock market in the second half of 1998 was not as high as in many other emerging markets.

September 1997, when issues with low liquidity were excluded from the Free market), and a set of 19 sectoral indexes.

Figure 2.3



2.5. The Slovak Stock Market

2.5.1. Historical background of the Bratislava Stock Exchange

The Bratislava Stock Exchange (BSSE) is the principal organiser of spot markets in Slovakia. It was founded in 1991, one year after the beginning of the transformation of the country's economic system. The most important point in the history of BSSE came in April 1993, when the two-year preparatory stage was accomplished, during which the basic rules were drawn up governing the activities of the Stock Exchange. Commercial activity of the Exchange commenced on 6 April 1993. On that day members of the BSSE executed the first trades with listed bonds. The turning point was launching the book-entered (dematerialised) share trading. The shares were products of the first wave of voucher privatisation. The first such deals were made by the members on 1 June 1993. Trading takes place via an electronic stock exchange trading system, without the traditional intermediary function of traders or specialists, by means of on-line communication between the Stock Exchange and the members. In July 1997 trading of foreign securities was launched on the

BSSE. The same year a new structure of BSSE markets was introduced. Trading in securities was divided into three main markets.[13]

2.5.2. Trading on the Bratislava Stock Exchange

The Stock Exchange organises trading on the following three markets: Market for Listed Securities, Market for Registered Securities, Free Market for Securities. The table below summarises basic criteria (minimum values) for listing on the Market for Listed Securities.

Amount of registered capital	Sk 500 million
Market capitalisation of the issue	Sk 100 million
Market capitalisation of the part of issue that is publicly held	Sk 100 million
Duration of issuer's business activities	3 years
Number of completed years for which information is provided in prospectus	3 years

In case an issue does not meet the basic criteria for listing, but the issuer is interested in trading his securities on the BSSE, he can register his issue on the Market for Registered Securities. All publicly tradable securities issued in accordance with valid legislation are admitted to the Free Market for Securities on the basis of an issuer's request.

Issuers whose securities are listed or registered on the BSSE are obliged to continuously inform the Stock Exchange about all important facts that could influence trading of their respective issues. Issuers of listed securities are obliged to submit their performance results to the Stock Exchange quarterly. Issuers of registered securities are obliged to do so semi-annually. The Stock Exchange

[13] For the summary of market statistics of the Bratislava Stock Exchange over the period 1993-1998 see Appendix 2.2. Table 2.8 at the end of the chapter.

provides these results for publication to domestic periodicals, as well as to information agencies operating on a worldwide basis in the shortest possible term.

The trading of securities on the BSSE runs daily by means of the Elektronický Burzový Operaèný Systém - EBOS (Electronic Stock Exchange Trading System). The system is based on the active entering of offers to buy and sell into a computer by members of the Exchange. Orders can be entered on the Stock Exchange floor or at any other location that has an on-line interconnection between the member's workstation and the BSSE's trading system. Using this system, members can execute *anonymous transactions*, *direct transactions* and *repo transactions*. In addition, they can declare and carry out the purchase of all shares of an issuer, the so-called "*takeover bid*".

Anonymous transaction An anonymous transaction on the EBOS is executed between an order to buy and an order to sell. In anonymous transaction, the member who is buying does not know the member who is selling and vice versa. Members can execute anonymous transactions in one of the three sub-systems of the EBOS:

1. Auction Trading (trading at one price) – under this method of trading Stock Exchange members deliver to the Stock Exchange orders to buy and orders to sell securities (stored in order files) via modem or on diskette. One single realisation price, at which all transactions are then executed, is calculated on the basis of an algorithm for every issue of securities that has been a subject of at least one offer. The calculated algorithm ensures a maximum amount of traded securities and a minimum leftover, that is the difference between the total purchase and sale.
2. Continuous Trading – under this method of trading Stock Exchange members place orders to buy and orders to sell securities through terminals. Received orders are processed according to the following rules:

- priority is given to the order with the best price;

- in case of orders with equal prices, priority is given to the order placed earliest.

When the price of the best order to buy is equal to or higher than the price of the best order to sell, the computer automatically concludes the transaction by matching these orders. The price of securities in one issue can vary in the course of continuous trading.

3. Block Trading – under this method of trading orders to buy and orders to sell securities are not automatically matched by the computer. Members submit orders to buy and sell - non-address orders - through terminals. These are sorted in the system according to price, volume of order, and the time when the order was made. Members can respond to the most advantageous non-address order by making an address order, which will conclude the transaction. The price of securities in one issue can vary in the course of block trading.

Direct Transactions

A direct transaction is concluded between the buying and selling members when both parties know each other. One party reports the transaction by entering it into the EBOS. The transaction can be concluded after confirmation by the counter-party.

Repo Transactions

Conclusion of repo transactions is similar to the procedure of reporting direct transactions. Each repo transaction consists of a transfer and a re-transfer of securities, by which (the re-transfer) the securities of the same type and number are transferred to the account of the original owner. The re-transfer is also called the closing of repo transaction. In case there are obstacles to the closing of a repo transaction, the participating parties may agree on a change of closing date of the repo transaction or termination of the repo transaction which they report to the BSSE.

Takeover Bid

According to the Securities Act, a legal or a physical entity is allowed to acquire more than 30% of publicly tradable shares of one issuer only through an offer to purchase all publicly tradable shares of that issuer, that is a takeover bid. After the takeover bid has been published, its time of validity (30 - 60 days), price, and minimum required amount of securities is set up in the EBOS. The takeover bid is successful only if the bid's underwriter has purchased a minimum required amount of securities through anonymous or direct transactions in the course of the bid's validity.

2.5.3. The Bratislava Stock Exchange Index (SAX)[14]

The BSSE's official share index is the Slovenský Akciový Index - SAX (Slovak Share Index). The SAX is a capital-weighted index, and reflects an overall change of assets connected with an investment in shares that are included in the index. This means that, besides fluctuations of prices, the index also includes dividend payments and revenues connected with changes of share capital amount (that is, the difference between current market price and the issue price of new shares).

The initial value of the SAX index - 100 points - refers to September 14, 1993. The index reflects development only on the BSSE, and has been based to date on the average daily prices stated in the price lists. A calculation based on closing prices will be applied in the future, although only an increased liquidity of the market will allow this step.

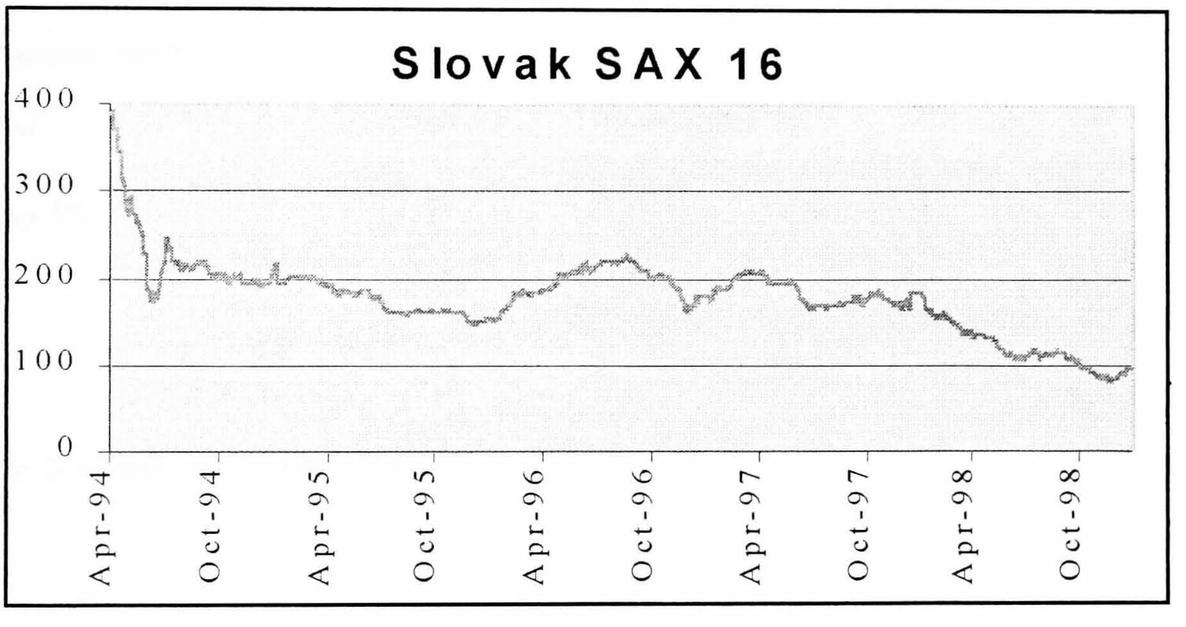
The SAX index formula is flexible and allows for the participation of various companies in the index, as well as changes in the number of companies, proportional to changes in their tradability or in the case of a new company entering the capital market. In the case of a change in the index's structure, the correction factors are set in such a way that the index with the new structure

[14] For the composition of the index see Appendix 2.1, Table 2.4 and Figure 2.8 at the end of the chapter.

continuously follows the development of the index with the previous structure. At present, the SAX index includes 16 most liquid Slovak joint-stock companies.

The Figure 4 below presents historical development of the Slovak SAX index in 1994–1998. Similar to other transition markets of the region, 1996 and partially 1997 were successful for the Slovak stock market. However, once again, it was hit by the emerging markets crisis in summer-autumn 1998.

Figure 2.4



APPENDIX 2.1

Composition of the Stock Exchange Indices (December 1998)

Table 2.1 Composition of the BUX

Stock	Weight(%)	Stock	Weight (%)
BorsodChem	4.31	NjlBI	1.41
Danubius	2.82	OTP	15.53
Egis	5.64	Pannonplast	2.24
POTEX	0.86	Pick, Szeged	2.57
Graboplast	2.85	Primagáz	1.07
Human Rt.	0.65	Rába	3.24
Inter-Európa, Bank	0.46	Ricbter	14.93
MATAV	15.44	Skála-Coop "T"	0.54
Mezógép Rt.	2.45	TVK	6.33
MOL	14.69	Zalakerámia	1.99

Figure 2.5 Sectoral Structure of the BUX

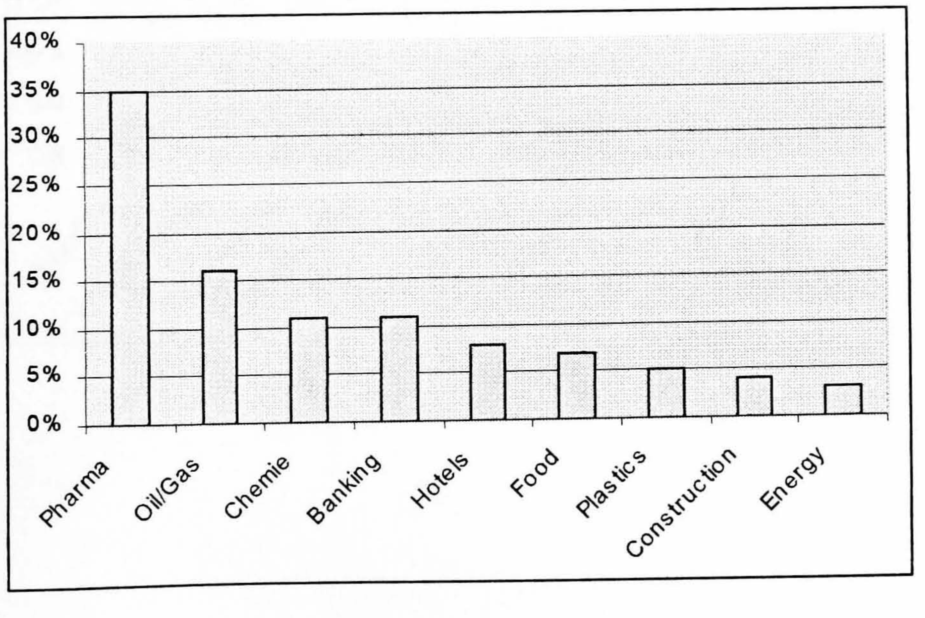


Table 2.2 Composition of the WIG 20

Stock	Weight (%)	Stock	Weight (%)
ELEKTRIM	11.7	**OKOCIM**	3.4
HANDLOWY	11.4	**STALEXPORT**	3.1
KGHM	9.9	**ZYWIEC**	3.1
BSK	8.1	**CELULOZA**	2.9
BPH	7.0	**OPTIMUS**	2.8
BRE	6.5	**DEBICA**	2.7
WBK	5.3	**POLIFARB-CW.**	2.7
BUDIMEX	4.5	**AGROS**	2.6
UNIVERSAL	4.5	**STOMIL**	2.4
MOSTOSTAL-EXP.	3.5	**GORAZDZE**	1.9

Figure 2.6 Sectoral Structure of WGI 20

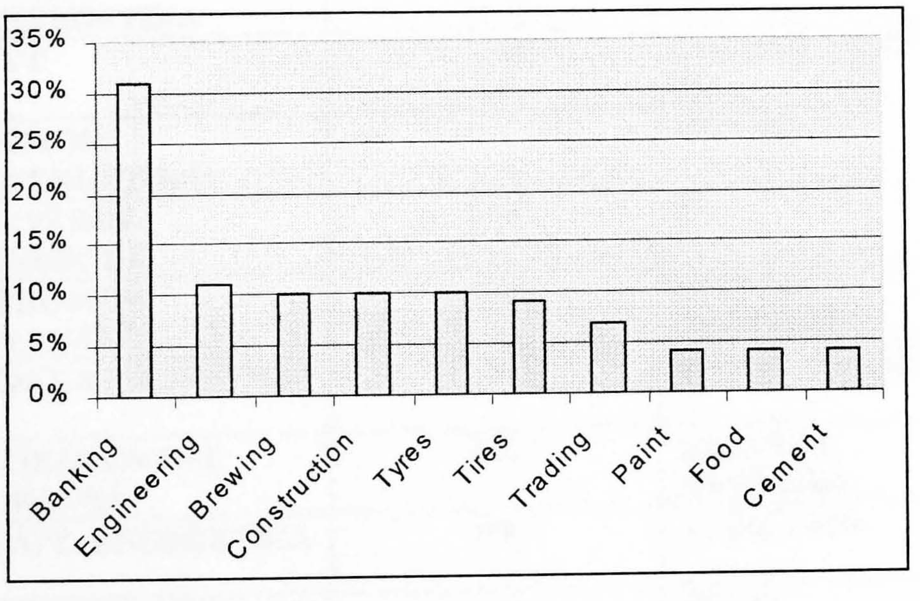


Table 2.3 Composition of the PX-50

Stock	Weight (%)	Stock	Weight (%)
SPT TELECOM	40,86	NOVÁ HU	0,62
È. RADIOKOMUNIK.	12.77	ŠKODA PLZEÒ	0,56
ÈEZ	10,19	SEMOR. PLYNÁRENSKA	0,49
IPB	4,07	INTERCONTINENT AL	0,46
KOMERÈNÍ BANKA	2.79	IPS PRAHA	0,36
UNIPETROL	2,61	KERAMIKA HOB	0,35
ÈESKÁ SPOØITELNA	2.46	BVV BRNO	0,34
SEVEROÈ. DOLY	2,12	PIVOV.RADEGAST	0,34
VERTEX	1,3	VÍTKOVICE	0,32
PRASKÁ TEPLÁREN.	1,26	MOSTECKÁ UH.SPOL.	0,29
ELEKTRÁRNY OPATOV.	1.23	METROSTAV	0,27
JIHOMOR. ENERGET.	1.18	PRASKÉ PIVOVARY	0,25
VÝCHÈ. ENERGETIKA	1,04	DEZA	0,25
PVT	1,01	CHLUMÈAN.KER.Z ÁV.	0,2
SEMOR. ENERGETIKA	1	SETUZA	0,17
SEVEROÈ. ENERGET.	0,97	SS	0,16
JIHOMOR. PLYNÁREN.	0,93	PLZEÒSKÝ PRAZDROJ	0,16
ÈESKÁ POJIŠOVNA	0,87	BIOCEL PASKOV	0,16
SOKOLOVSKÁ UHELNÁ	0,82	SKLÁRNY KAVALIER	0,15
ZÁPÈ. ENERGETIKA	0,8	ALIACHEM	0,13
MORSLEZS.TEPLÁR NY	0,68	PRVNÍ SEVZÁP.TEPL.	0,12
PRASKÁ ENERGETIKA	0,68	ÏAS	0,1
STØEDOÈ. ENERGET.	0,66	ZPS ZLÍN	0,07
JIHOÈ. ENERGETIKA	0,63	ÈESKÁ ZBROJOVKA	0,07
IVNOSTENSKÁ BANKA	0,63	VODNÍ STAVBY PRAHA	0,05

Figure 2.7 Sectorla Structure of PX 50

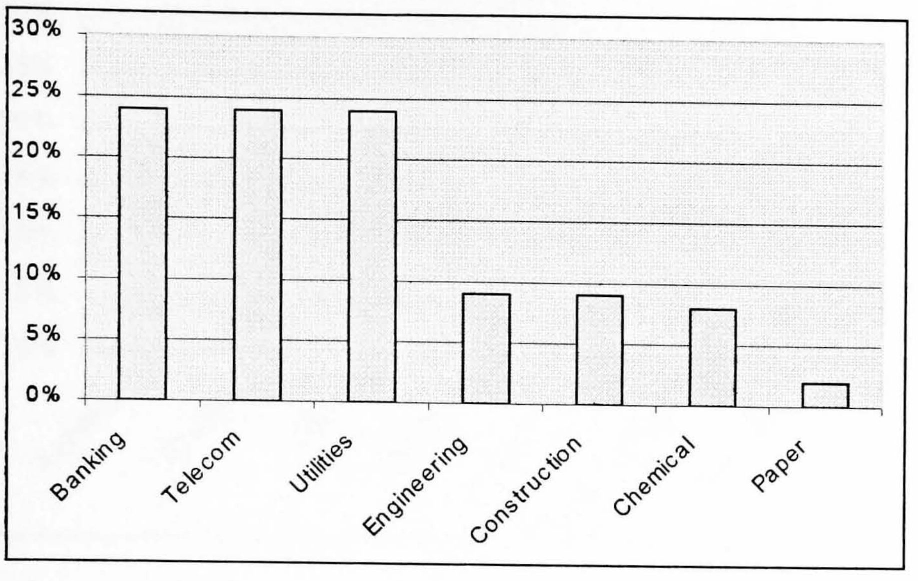
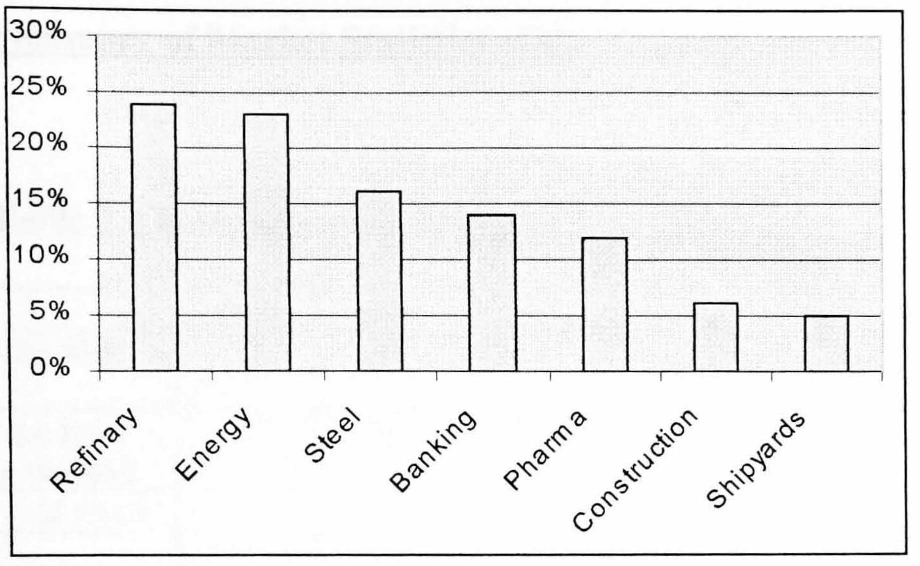


Table 2.4 Composition of the SAX

Stock	Weight (%)	Stock	Weight (%)
BIOTIKA	0.45%	**SLOVAKOFARMA**	21.92%
DOPRASTAV	3.01%	**SLOVENSKA POISOVOA**	3.57%
DROTOVOA	0.48%	**SLOVNAFT**	33.40%
FIGARO	1.56%	**VAHOSTAV**	0.62%
NAFTA	5.10%	**VSZ**	15.05%
PLASTOKA	0.75%	**VÚB**	7.65%
POVASKE STROJAME	2.00%	**ZAVODY SNP**	2.83%
SES TLMAEE	0.69%	**ZELEZIARNE PPDBREZOVA**	0.91%

Figure 2.8 Sectoral Structure of SAX



APPENDIX 2.2

Summary of Market Statistics of the Transition Stock Exchanges

Table 2.5 Hungarian market statistics (1990-1998)

	1990	1991	1992	1993	1994	1995	1996	1997	1998
Number of the BSE Members	42	48	48	47	51	56	57	63	
Number of sec-ties Admitted to the BSE	6	22	40	62	120	166	167	149	144
of which:									
Equities	6	20	23	28	40	42	45	49	55
Government bonds	-	1	4	16	25	34	36	28	30
corporate bonds	-	1	1	1	3	3	2	7	8
T-bills	-	-	10	10	31	50	50	44	42
Investment funds	-	-	1	6	20	36	33	20	8
Compensation notes	-	-	1	1	1	1	1	1	1
Securities admitted to the BSE at nom. value (US$mln)	99,3	440,4	2 261,5	4 387,3	7 736,0	8 731,4	12 008,7	12 123,3	14 023,2
of which:									
Equities	99,3	240,7	343,0	380,3	720,9	1 517,7	1 873,6	1 956,9	2 443,5
Government bonds	-	198,4	952,7	2 353,5	3 380,6	2 999,6	4 035,7	4 439,8	6 166,1
Corporate bonds	-	1,3	1,2	3,0	14,5	11,3	7,7	181,7	236,7
T-bills	-	-	706,2	763,7	2 318,2	2 964,8	4 912,4	4 725,7	4 490,0
Investment funds	-	-	20,2	89,4	215,0	308,0	264,4	132,8	38,2
Compensation notes	-	-	238,2	797,4	1 086,8	930,0	914,9	686,5	648,6
Equities cap-n of BSE in % of GDP					20,40	22,32	36,06	61,29	53,71
Capitalization on the BSE (US$mln)	266,9	708,8	2 404,4	4 538,3	7 984,6	8 757,0	15 658,7	25 136,3	25 395,7
of which:									
Equities	266,9	505,2	562,1	811,3	1 639,7	2 350,2	5 582,9	15 028,8	14 021,46
Cash turnover on market val. (US$mln)	96,8	131,5	426,3	2 017,5	1 961,3	2 021,8	7 501,8	36 040,1	64 234,5
of which:									
Equities	96,8	131,1	76,1	198,4	530,0	696,9	3 212,6	15 369,4	32 208,34
Government bonds	-	0,4	156,5	797,3	500,1	616,2	2 738,9	10 483,3	22 159,21
corporate bonds	-	0,0	0,0	0,0	1,1	0,4	3,7	209,3	166,25
T-bills	-	-	190,6	923,5	753,8	647,7	1 351,3	9 725,1	9 614,50
Investment funds	-	-	0,4	1,1	2,4	27,4	58,2	12,4	7,73
Compensation notes	-	-	2,7	97,2	173,9	33,2	137,1	240,6	87,12
Number of tran-ns	4 255	13 676	8 145	21 185	70 744	71 240	170 956	504 879	1 026 646
Average daily number of trans-ns	24	54	32	84	281	286	689	2 044	4 140

Source: Budapest Stock Exchange

Table 2.6 Polish market statistics (1991-1998)

	1991	1992	1993	1994	1995	1996	1997	1998
SHARES								
Single-Price Action								
Number of listed companies	9	16	21	36	53	66	96	117
Capitalisation (PLN mln)	161	351	5 803	7 149	10 902	23 036	38 107	68 082
Average P/E ratio	4,1	3,4	13,3	16,4	7,8	12,3	14,9	12,6
Dividend yield (%)	0,0	5,5	0,4	0,4	2,3	1,2	1,5	0,9
Number of trans-ns per sess-n	877	1 233	9 832	24 594	7 164	8 074	9 891	9 766
Total turnover (PLN mln)	30	230	7 750	22 640	12 200	25 611	36 038	39 170
Continuous Trading								
Number of transactions	-	-	-	-	-	31	460	1 578
Total turnover (PLN mln)	-	-	-	-	-	251	3 338	8 487
Block Trades								
Number of transactions	-	-	-	-	63	143	426	1 058
Total turnover (PLN mln)	-	-	3	-	395	1 575	4 332	7 925
BONDS								
Single-Price Action								
Number of listed bond issues		3	8	12	13	14	15	16
Interest paid per one three-year bond (PLN)	-	-	40,17	34,76	29,67	25,18	22,65	23,65
Total turnover (PLN mln)	-	7,1	115,4	322,6	419,3	1898,23	2781,57	3752,68
Number of trans-ns per sess-n	-	48	69	112	136	1 038	1 225	591
Continuous trading								
Number of transactions	-	67	447	491	1 115	1 917	1 961	2 540
Total turnover (PLN mln)	-	14,3	185,2	1826,1	3734,6	4007,36	2871,25	1196,61
Block Trades								
Number of transactions	-	-	-	-	660	480	422	327
Total turnover (PLN mln)	-	-	256,4	1151,3	15121,7	10313,5	7835,3	3631,77

Source: Warsaw Stock Exchange

Table 2.7 Czech market statistics (1993-1998)

	1993	1994	1995	1996	1997	1998
Number of members	62	71	101	106	88	66
Number of issues	982	1,055	1,764	1,750	412	402
Total of which: stocks + units	971	1,028	1,716	1,670	320	304
Total of which: bonds	11	27	48	80	92	98
Total trade value (CZK bln)	9.0	62.0	195.4	393.2	679.5	860.2
Average daily trade value (CZK bln)	0.220	0.385	0.835	1.579	2.718	3.427
Market cap.: stocks + units (CZK bln)	-	353.1	478.6	539.3	495.7	416.2
Market cap.: bonds (CZK bln)	-	50.0	87.8	136.9	174.0	198.1
Number of Trading sessions	41	161	234	249	250	251

Source: Prague Stock Exchange

Table 2.8 Slovak market statistics (1993-1998)

	1993	1994	1995	1996	1997	1998
Stock Exchange Floor						
Volume (Sk mln)	37.883	829.443	470.799	15,824.168	5,922.935	22,613.220
Number of securities (in units)	36,603	604,513	839,307	12,810,809	5,709,482	4,828,064
Number of Transactions	1.361	11,187	10,182	33,857	14,985	4,344
Direct Trades						
Volume (Sk mln)	128.437	5,454.056	39,597.936	98,292.194	158,140.924	276,467.690
Number of securities (in units)	79,260	7,987,351	44,319,363	118,424,574	134,510,913	79,779,326
Number of Transactions	21	1,780	18,573	147,053	142,812	55,526
Total						
Volume (Sk mln)	166.320	6,283.500	40,068.735	114,116.363	164,063. 859	299,080.910
Number of securities (in units)	115 863	8 591 864	45 158 670	131 235 383	140 220 395	84 607 390
Number of Transactions	1 382	12 967	28 755	180 910	157 797	59 870

Source: Bratislava Stock Exchange

CHAPTER 3

VOLATILITY IN THE TRANSITION MARKETS OF CENTRAL EUROPE[15]

3.1. Introduction

Until some twenty years ago, the focus of statistical analysis of time series centred on the conditional first moment. While it has been recognised for quite some time that the uncertainty of speculative prices, as measured by variances and covariances, is changing through time, it was not until the beginning of the 1980s that applied researchers started to explicitly model time variation in second and higher moments. Researchers used only informal procedures to take account for changing variances and covariances.[16] A breakthrough occurred when, in line with Box-Jenkins type models for conditional first moments, Engle (1982) put forward the Autoregressive Conditional Heteroskedastic (ARCH) class of models for conditional variances, which proved to be extremely useful for analysing economic time series. Since then an extensive literature has been developed for modelling higher order conditional moments. The ARCH model has applications to numerous and

[15] The version of this chapter is forthcoming in *Applied Financial Economics*.

[16] For example, Mandelbrot (1963a) used recursive estimates of the variance over time and Klien (1977) took five period-moving variance estimates about a 10-period moving-sample mean.

diverse areas.[17]

As emphasised by Pagan (1996) and Bollerslev et al. (1994), many financial time series have a number of characteristics in common. First, asset prices are generally nonstationary and often have a unit root, whereas returns are normally stationary. Second, return series usually show little autocorrelation, while serial independence between the squared values of the series is often rejected pointing towards the existence of nonlinear relationships between the subsequent observations. Volatility of the returns appears to be clustered. Returns go through periods of high and low variance. These facts point towards time-varying conditional variances. Third, beginning with the seminal works of Mandelbrot (1963a) and Fama (1965), evidence indicates that the empirical distribution of return series differs significantly from sampling independent observations from an identical Gaussian distribution. The series are characterised by leptokurtosis, which could be related to the time-variation in the conditional variance. Fourth, some series exhibit asymmetric behaviour in the conditional variance, related by a number of authors to leverage effects. Finally, empirical evidence indicates that volatilities of different securities have common persistent components, pointing towards the linkages between markets.

The ARCH models were first formal specifications that seemed to capture the stylised facts characterising financial series.[18] Most contributions to the ARCH literature deal with mature, large

[17] For example, it has been used in asset pricing to test the CAPM and APT; to develop volatility tests for market efficiency and to estimate the time-varying systematic risk in the context of the market model. 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. In macroeconomics, it has been successfully used to construct debt portfolios of developing countries, to measure inflationary uncertainty, to examine the relationship between the exchange rate uncertainty and trade, to study the effects of central bank interventions, and to characterise the relationship between the macroeconomy and the stock market.

[18] The history of ARCH is quite a short one but its roots go far in the past – possibly as far as Bachelier (1900), who was the first to conduct a rigorous study of the behaviour of speculative prices. There was then a period of long silence. 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). He was first to observe the 'stylised facts' characterising many economic and financial variables, such as unconditional distributions have thick tails, variances change over time, and large (small) changes tend to be followed by large (small) changes of either sign.

and liquid markets. The motivation for this chapter is to add evidence from the transition equity markets of Central Europe to the econometric modelling of univariate and multivariate volatility of financial time series. This will be accomplished by testing a range of models, which has proven to be able to account for empirical regularities characterising financial series, using the data from four transition equity markets.

3.2. ARCH Models of Conditional Volatility: A Review

The ARCH type models will be employed in different contexts for the parameterisation of the second conditional moments of the considered time series in this, as well next two chapters of our work. Therefore, the aim of this section is to provide a review of the ARCH models of volatility, which covers different aspects of the modelling of conditional second moments considered throughout the thesis.

Part 3.2.1 of this section provides the basic parameterisation of the ARCH type models. Part 3.2.2 presents unconditional moments of the ARCH process. After building the theoretical framework of the basic ARCH models, the emphasis of the part 3.2.3 are on more flexible ARCH specifications. The frequent inability of the conditionally normal ARCH models to pass the simple diagnostic tests has lead to the use of the conditional distributions more general than the normal distribution. These issues are discussed in part 3.2.4. In part 3.2.5 we present several multivariate specifications of the ARCH process, while part 3.2.6 covers the issues of estimation of the discussed models. Finally, Appendix 3.2 at the end of the chapter provides the models of the conditional volatility alternative to the ARCH specification.

3.2.1 The Definition of the ARCH Process

An ARCH model can be formulated for the original series or for an error term. In most financial applications, the error ε_t of a regression or time series model is conditionally heteroskedastic. In our work the ARCH process will be defined in terms of the distribution of the errors of a dynamic linear regression model. The dependent variable y_t is assumed to be generated by

$$y_t = x_t'\xi + \varepsilon_t, \quad t = 1, \ldots, \mathrm{T}, \tag{3.1}$$

where x_t is a $k \times 1$ vector of predetermined variables, which may include lagged values of the dependent variable, and ξ is $k \times 1$ 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 $\Psi_{t-1} = \{y_{t-1}, x_{t-1}, y_{t-2}, x_{t-2} \ldots\}$. The Engle's (1982) original ARCH model of order q can be written as

$$\varepsilon_t \mid \Psi_{t-1} \sim N(0, h_t) \tag{3.2}$$

where

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2, \tag{3.3}$$

with $\alpha_0 > 0$ and $\alpha_i \geq 0$, $i = 1, \ldots$, q, to ensure that the conditional variance is positive.

The particularity of the model (3.2) and (3.3) is not simply that conditional variance h_t is a function of the conditioning set $\Psi_{t-1} = \{y_{t-1}, x_{t-1}, y_{t-2}, x_{t-2} \ldots\}$ (since $\varepsilon_t = y_t - x_t'\xi$), but rather it is the particular functional form that is specified. Episodes of volatility are generally characterised as the clustering of large shocks to the dependent variable. The conditional variance function (3.3) is formulated to mimic this phenomenon. 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. As a result, the larger errors of either sign tend to be followed by a large error of either sign. 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.[19]

The generating equation for an ARCH process is

$$\varepsilon_t = \eta_t \sqrt{h_t} \,, \tag{3.4}$$

where $\eta_t \sim N(0,1)$ and h_t is given by (3.3). Since h_t is a function of the elements of Ψ_{t-1}, and is therefore fixed when conditioning on Ψ_{t-1}, it is clear that ε_t as given in (3.4) will be conditionally normal with $E(\varepsilon_t|\Psi_{t-1}) = \sqrt{h_t}E(\eta_t|\Psi_{t-1}) = 0$ and $Var(\varepsilon_t|\Psi_{t-1}) = h_t Var(\eta_t|\Psi_{t-1}) = h_t$. Hence, the process specified by (3.4) is identical to the ARCH process (3.2). The generating equation (3.4) reveals that ARCH rescales an underlying Gaussian innovation process η_t by multiplying it by conditional standard deviation, which is the function of the information set Ψ_{t-1}.

Engle (1982, 1983) in the first empirical applications of ARCH to the relationship between the level and volatility of inflation found that a large lag q was required in the conditional variance function (3.3). To reduce the computational burden, Engle parameterised the conditional variance as

$$h_t = \alpha_0 + \alpha_1 \sum_{i=1}^{q} w_i \varepsilon_{t-i}^2 \,, \tag{3.5}$$

[19] A linear function of lagged squared errors, of course, is not the only conditional variance function that will produce clustering of large and small deviations. Any monotonically increasing function of the absolute values of the lagged errors will lead to such clustering. However, since variance is the expected squared deviation, a linear combination of the squared errors is a natural measure of the recent trend in variance to translate to the current conditional variance h_t.

where the weights

$$w_i = \frac{(q+1)-i}{\frac{1}{2}q(1+q)}, \tag{3.6}$$

decline linearly and are constructed so that $\sum_{i=1}^{q} w_i = 1$. With this parameterisation, a large lag can be specified and yet only two parameters are required to be estimated in the conditional variance function. However, the important thing to note is that this formulation does put undue restrictions on the dynamics of the ARCH process.

Bollerslev (1986) and Taylor (1986) independently proposed a generalisation of the ARCH (q) model that allows for both parsimonious parameterisation and flexible lag structure by introducing autoregressive term into the conditional variance equation. Bollerslev (1986) termed the new model generalised ARCH (GARCH). According to GARCH the conditional variance is specified as:

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}, \tag{3.7}$$

Sufficient, but not necessary conditions for the nonnegativity of h_t are

$$\alpha_0 > 0,\ \ \alpha_i \geq 0 \text{ for } i = 1,\ldots, \text{q, and } \beta_i \geq 0 \text{ for } i = 1,\ldots, \text{p}. \tag{3.8}$$

To demonstrate the motivation of the GARCH process we express (3.7) as

$$h_t = \alpha_0 + \alpha(B)\varepsilon_t^2 + \beta(B)h_t,$$

where $\alpha(B) = \alpha_1 B + \ldots + \alpha_q B^q$ and $\beta(B) = \beta_1 B + \ldots + \beta_p B^p$ are polynomials in the backshift operator B. In the case where the roots of $1-\beta(Z)$ lie outside the unit circle, the GARCH (p,q) model can be written as an ARCH (∞) model

$$h_t = \frac{\alpha_0}{1-\beta(1)} + \frac{\alpha(B)}{1-\beta(B)}\varepsilon_t^2 = \alpha_0^* + \sum_{i=1}^{\infty}\delta_i \varepsilon_{t-1}^2 \tag{3.9}$$

where $\alpha_0^* = \frac{\alpha_0}{1-\beta(1)}$ and δ_i is the coefficient of B^i in the expansion of $\frac{\alpha(B)}{1-\beta(B)}$. Hence expression (3.9) reveals that a GARCH (p,q) process is an infinite order ARCH process with a rational lag structure imposed on the coefficients. As a result, GARCH parsimoniously represents a high order ARCH process.[20] 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 GARCH (2,1) is very rare.

3.2.2 Unconditional moments of ARCH process

The derivation of the unconditional moments of the ARCH process is possible through the use of the *law of iterated expectations*[21], which relates unconditional and conditional moments of the series. Since the ARCH model is specified in terms of its conditional moments, it provides a method for deriving unconditional moments.

Consider the unconditional mean of a GARCH (p,q) error ε_t with conditional variance (3.7). Applying the law of iterated expectations, $E(\varepsilon_t) = E[E(\varepsilon_t \mid \Psi_{t-1})]$. However, because the GARCH

[20] Nelson and Cao (1992) demonstrated that weaker sufficient conditions that those in (3.8) above can be found to ensure that the conditional variance of a GARCH (p,q) process is strictly positive. They pointed out that from the inverted representation of h_t in (3.9): $\alpha_0^* > 0$ and $\delta_i \geq 0$, $i = 1,\ldots,\infty$ are sufficient to ensure that the conditional variance is strictly positive. In terms of original parameters of the GARCH model, Nelson and Cao showed that restrictions above do not require all the inequalities in (3.8) to hold. For example in a GARCH (1,2) process, $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and $\beta_1\alpha_1 + \alpha_2 \geq 0$ are sufficient to guarantee that $h_t > 0$. Therefore, α_2 may be negative. They presented general results for GARCH (1, q) and GARCH (2, q), but suggest that derivation for GARCH processes with $p \geq 3$ is difficult.

[21] For definition of the law of iterated expectations, see e.g. Appendix B.10 in „Econometric Theory" by Davidson (2000).

model specifies that $E(\varepsilon_t \mid \Psi_{t-1}) = 0$ for all realisations of Ψ_{t-1}, it follows that $E(\varepsilon_t) = 0$. As a result, the GARCH process has mean zero.

To evaluate the unconditional variance of the GARCH process for simplicity let us consider GARCH (1,1) process. On the basis of the law of the iterated expectations we get:

$$
\begin{aligned}
E(\varepsilon_t^2) &= E[E(\varepsilon_t^2 \mid \Psi_{t-1})] = E(h_t) \\
&= \alpha_0 + \alpha_1 E(\varepsilon_{t-1}^2) + \beta_1 E(h_{t-1}) \\
&= \alpha_0 + (\alpha_1 + \beta_1) E(\varepsilon_{t-1}^2),
\end{aligned} \tag{3.10}
$$

which is a linear difference equation for the sequence of variances. Assuming the process began infinitely far in the past with a finite initial variance, the sequence of variances converges to the constant

$$\sigma^2 = E(\varepsilon_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \tag{3.11}$$

if $\alpha_1 + \beta_1 < 1$. For the general GARCH (p,q) process, Bollerslev (1986) gave the necessary and sufficient condition for the existence of the variance:

$$\alpha(1) + \beta(1) = \sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i < 1 \tag{3.12}$$

When this condition is satisfied, the variance is:

$$\sigma^2 = E(\varepsilon_t^2) = \frac{\alpha_0}{1 - \alpha(1) - \beta(1)} \tag{3.13}$$

Although the variance of ε_t conditional on Ψ_{t-1} changes with the elements of the information set, unconditionally the ARCH process is *homoskedastic*.

The nature of the unconditional density of an ARCH process can be analysed by the higher order moments. Due to the fact that ε_t is conditionally normal, for all odd integers *m*: $E(\varepsilon_t^m \mid \Psi_{t-1}) = 0$.

The skewness coefficient is immediately seen to be zero. Since ε_t is continuous, this implies that the unconditional distribution is symmetric.

A general expression for the fourth moment of GARCH (p,q) process is not available, but Engle (1982) gave it for ARCH (1) process, and Bollerslev (1986) generalised it for the GARCH (1,1) case. According to Engle's result for ARCH (1)

$$E(\varepsilon_t^4) = \frac{3\alpha_0^2}{(1-\alpha_1)^2} \frac{1-\alpha_1^2}{1-3\alpha_1^2} \tag{3.14}$$

and the kurtosis

$$Kurt(\varepsilon_t) = \frac{E(\varepsilon_t^4)}{(E[\varepsilon_t^2])^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} \tag{3.15}$$

exists if $3\alpha_1^2 < 1$.

The kurtosis of a GARCH (1,1) process

$$Kurt(\varepsilon_t) = \frac{E(\varepsilon_t^4)}{(E[\varepsilon_t^2])^2} = 3 + \frac{6\alpha_1^2}{1-\beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2} \tag{3.16}$$

exists if $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$.

Both (3.15) and (3.16) are clearly greater than 3, the kurtosis coefficient of the normal distribution. Therefore, the ARCH process has tails heavier than the normal distribution. Because of this leptokurtic unconditional distribution, ARCH models are particularly suited for financial series, which often exhibit such behaviour.

Although no known closed form for the conditional density function of an ARCH process exists, Nelson (1990b) demonstrated that, under suitable conditions, GARCH(1,1) process approaches a continuous time process whose stationary unconditional distribution is a Student's *t*. This result indicates why heavy-tailed distributions are so prevalent with high frequency financial data.

Another important characteristic of the model is that the parameterisation of the ARCH process does not *a priori* impose the existence of the unconditional moments. Very early in the history of financial econometrics Mandelbrot (1963a) raised the point that asset returns had second unconditional moments that may not exist. The fact that the ARCH model admits 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.

Above we considered the univariate distributions of a single ε_t. The moments of the joint distribution of the ε_t also reveal important properties of the ARCH process. The autocovariances of the GARCH (p,q) process are

$$\begin{aligned} E(\varepsilon_t \varepsilon_{t-k}) &= E[E(\varepsilon_t \varepsilon_{t-k} \mid \Psi_{t-1})] \\ &= E[\varepsilon_{t-k} E(\varepsilon_t \mid \Psi_{t-1})] \\ &= 0 \end{aligned} \tag{3.17}$$

Since the GARCH process is serially uncorrelated with constant mean zero, the process is weakly stationary if the variance exists, i.e. if (3.12) holds. A property of GARCH process, first demonstrated by Nelson (1990a) for GARCH (1,1), is that it may be strongly stationary without being weakly stationary. That the GARCH process may be strongly stationary without being weakly stationary stems from the fact that weak stationarity requires mean, variance and autocovariance be finite and time invariant. Strong stationarity requires that the distribution function of any set of ε_t is invariant under time translations. Finite moments are not required for strong stationarity. The results of Nelson (1990a) show that the unconditional variance may be infinite and yet the GARCH process may still be strongly stationary.

3.2.3 Extensions of the basic ARCH models

ARCH-M

In the ARCH in the mean, or ARCH-M model, introduced by Engle, Lilien and Robbins (1987), the conditional mean is an explicit function of the conditional variance of the process. Many theories in finance involve an explicit tradeoff between the risk and expected return. The ARCH-M model is ideally suited to handling such questions in the time-series context where the conditional variance may be time-varying. In the regression set-up an ARCH-M model is specified as

$$y_t = x_j^{'}\xi + \delta g(h_t) + \varepsilon_t, \tag{3.18}$$

where $\varepsilon_t \mid \Psi_{t-1} \sim N(0, h_t)$, and h_t is determined by an ARCH or GARCH process. To examine the properties of the ARCH-M model, we consider a simple version of

(3.18):

$$y_t = \delta h_t + \varepsilon_t$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \tag{3.19}$$

Then, we can write

$$y_t = \delta\alpha_0 + \delta\alpha_1\varepsilon_{t-1}^2 + \varepsilon_t. \tag{3.20}$$

From this expression, and using $E(\varepsilon_{t-1}^2) = \alpha_0 / 1 - \alpha_1$ (unconditional variance of ARCH process), it follows that

$$E(y_t) = \delta\alpha_0 / 1 - \alpha_1 , \tag{3.21}$$

which can be viewed as the unconditional expected return for holding risky asset.

Also, using result for $E(\varepsilon_t^4)$ in (3.14)

$$Var(y_t) = \frac{(\delta\alpha_1)^2 2\alpha_0^2}{(1-\alpha_1)^2(1-3\alpha_1^2)} + \frac{\alpha_0}{1-\alpha_1}. \tag{3.22}$$

In the absence of a risk premium $Var(y_t) = \frac{\alpha_0}{1-\alpha_1}$. Therefore, the first component of $Var(y_t)$ is due to the presence of a risk premium, which makes y_t more dispersed.

Finally, the ARCH-M effect makes y_t serially correlated (see Hong (1991)):

$$\rho_1 = Corr(y_t, y_{t-1}) = \frac{2\alpha_1^3\delta^2\alpha_0}{2\alpha_1^3\delta^2\alpha_0 + (1-\alpha_1)(1-3\alpha_1^2)}$$

$$\rho_k = Corr(y_t, y_{t-k}) = \alpha_1^{k-1}\rho_1, \qquad k = 2,3,\ldots \tag{3.23}$$

Bollerslev (1988) obtained similar results for GARCH process.

Models with nonlinear conditional variance

In the original ARCH model Engle (1982) assumed that the conditional variance function was linear in the squared errors and that the conditional distribution was normal. He acknowledged, however, that linearity and conditional normality assumptions might not be appropriate in particular applications. In this section we survey alternative nonlinear formulations of the conditional variance, and, in the section 3.2.4 we look at the alternative specifications of the conditional distribution, which have proven useful in applied research.

One of the problems encountered with the linear ARCH model was that the estimated coefficients α_i were sometimes found to be negative. Geweke (1986) and Milhoj (1987a) suggested the log ARCH model.

$$\ln(h_t) = \alpha_0 + \alpha_1 \ln(\varepsilon_{t-1}^2) + \ldots + \alpha_q \ln(\varepsilon_{t-q}^2) \tag{3.24}$$

Taking the exponential of both sides of (3.24), $h_t = e^{(\cdot)}$ is strictly positive, and therefore, no inequality restrictions are required for the α_is to ensure that the conditional variance is strictly positive.

To determine whether the linear model (3.3) or the logarithmic model (3.24) provided a better fit for actual data, Higgins and Bera (1992) proposed a non-linear ARCH (NARCH) model, which still requires non-negativity restrictions, but includes linear ARCH as a special case and log ARCH as a limiting case. Their specification of the conditional variance is:

$$h_t = [\phi_0(\sigma^2)^\delta + \phi_1(\varepsilon_{t-1}^2)^\delta + \ldots + \phi_q(\varepsilon_{t-q}^2)^\delta]^{1/\delta}, \tag{3.25}$$

where $\sigma^2 > 0, \phi_i \geq 0, \delta > 0$ and the ϕ_is are such that $\sum_{i=0}^{q} \phi_i = 1$. The motivation of the NARCH model can be seen by rearranging (3.25) to give:

$$\frac{h_t^\delta - 1}{\delta} = \phi_0 \frac{(\sigma^2)^\delta - 1}{\delta} + \phi_1 \frac{(\varepsilon_{t-1}^2)^\delta - 1}{\delta} + \ldots + \phi_q \frac{(\varepsilon_{t-q}^2)^\delta - 1}{\delta}, \tag{3.26}$$

from which it is evident that the NARCH model is a Box-Cox power transformation[22] of both sides of the linear ARCH model. It is apparent that when $\delta = 1$, (3.26) is equivalent to the linear ARCH model, and as $\delta \to 0$, (3.26) approaches the log ARCH model (3.24). Higgins and Bera (1992) estimated (3.25) with weekly exchange rates and found that δ was typically significantly less than 1 and much closer to zero, indicating that the data favoured the logarithmic rather than the linear ARCH model. Extensions of the above functional forms to the GARCH model are straightforward.

---

[22] Box and Cox (1964) suggested the following transformation

$$Y_t = \begin{cases} \dfrac{X_t^\theta - 1}{\theta} & \theta \neq 0 \\ \ln X_t & \theta = 0 \end{cases}, \text{ ....where } Y_t = \ln X_t \text{ if } \theta = 0, \text{ since } \lim \frac{x^\theta - 1}{\theta} \to \ln x \text{ when } \theta \to 0.$$

It should be noted that this transformation is only defined for all values of θ if $X_t > 0$ (see Schlittgen and Streitberg (1994)).

The limitation of the models described above is that the conditional variance function h_t is symmetric in the lagged ε_ts. Many financial series however are strongly asymmetric. Negative equity returns are followed by larger increases in volatility than equally large positive returns. Black (1976) interpreted this phenomenon as the leverage effect, according to which since a lower stock price reduces the value of equity to corporate debt, a sharp decline in stock prices increases corporate leverage and could thus increase the risk of holding stocks. Based on the leverage effects noted in Black (1976), Christi (1982) and French, Schwert, and Stambaugh (1987), Nelson (1991) proposed the exponential GARCH (EGARCH) model, which allows for asymmetry. To avoid non-negativity restrictions on parameters of the model, Nelson maintained the logarithmic specification (3.24) and proposed

$$\ln(h_t) = \alpha_0 + \sum_{i=1}^{q} \alpha_i g(\eta_{t-i}) + \sum_{i=1}^{p} \beta_i \ln(h_{t-i}), \tag{3.27}$$

where

$$g(\eta_t) = \theta\eta_t + \gamma[|\eta_t| - E|\eta_t|]. \tag{3.28}$$

The conditional variance in (3.27) and (3.28), is known as exponential GARCH (EGARCH). The properties of EGARCH model are determined by construction of the function $g(\cdot)$ in (3.28). The function $g(\cdot)$ is piecewise linear in η_t ($\eta_t = \varepsilon_t / \sqrt{h_t}$). It contains two parameters that determine what can be termed the 'size impact' and the 'sign impact' of news on volatility. If $\gamma > 0$, the model implies that a deviation of $|\eta_t|$ from its expected value causes the conditional variance to be larger than otherwise, producing the size impact of news on volatility (an effect similar to the idea behind the GARCH specification). The parameter θ allows the asymmetric impact of news on volatility. If $\theta = 0$ then a positive surprise, $\eta_t > 0$, has the same effect on volatility as a negative surprise,

$\eta_t < 0$, of the same magnitude. If $-\gamma < \theta < 0$, a positive surprise increases volatility less than a negative surprise. If $\theta < -\gamma < 0$, a positive surprise actually reduces volatility while a negative surprise increases volatility. A number of researchers have found evidence of asymmetry in stock price behaviour – negative surprises seem to increase volatility more than positive surprises.[23]

Nelson (1991) fitted the EGARCH model to the excess daily return on the CRSP value-weighted stock-market index. The estimate of θ was -0.118 with a standard error of 0.008, confirming a highly significant negative correlation between the excess return and subsequent volatility.

Building on the success of the EGARCH model to represent asymmetric responses in the conditional variance to positive and negative errors, a series of papers have proposed other ARCH models, which allow a very general shape in the conditional variance function. Although these models are parametric, and estimated by maximum likelihood, they are non-parametric in spirit because the shape of the conditional variance function is largely determined by the data themselves. Zakoian (1990) suggested a conditional standard deviation of the form

$$\sqrt{h_t} = \alpha_0 + \sum_{i=1}^{q} \alpha_i^+ \varepsilon_{t-i}^+ + \sum_{i=1}^{q} \alpha_i^- \varepsilon_{t-i}^- \ , \tag{3.29}$$

where $\varepsilon_t^+ = \max\{\varepsilon_t, 0\}$ and $\varepsilon_t^- = \min\{\varepsilon_t, 0\}$. The parameters are constrained by $\alpha_0 > 0$, $\alpha_i^+ \geq 0$ and $\alpha_i^- \geq 0$ for $i = 1,\ldots,q$ to ensure that the conditional standard deviation is positive. Zakoian referred to this formulation as a threshold ARCH (TARCH) model because the coefficient of ε_{t-i} changes when ε_{t-i} crosses the threshold of zero. When $\varepsilon_{t-i} > 0$, the conditional standard deviation is linear in ε_{t-i} with slope α_i^+, and when $\varepsilon_{t-i} < 0$, the conditional standard deviation is linear in ε_{t-i} with slope α_i^-. This allows for asymmetry in the conditional variance in the fashion of EGARCH.

[23] See, for example Pagan and Schwert (1990), Engle and Ng (1993), and the studies cited in Bollerslev, Chou, and Kroner (1992).

Glosten, Jarannathan and Runkle (1993) independently suggested a similar formulation, which allowed for asymmetric effects by including an indicative dummy taking the value of 1 if $\varepsilon_{t-i} < 0$ and zero otherwise. This model will be presented more in detail later on in our work (see section 3.3.2).

Gouroeroux and Monfort (1992) proposed that a step function over the support of the conditioning error vector $\underline{\varepsilon}_{t-1} = (\varepsilon_{t-1},\ldots,\varepsilon_{t-q})'$ could approximate a highly non-linear conditional variance function. Letting $A_1,\ldots,A_m$ be partition of the support of ε_t, Gourieroux and Monfort considered a conditional variance of the form

$$h_t = \alpha_0 + \sum_{i=1}^{m}\sum_{j=1}^{q}\alpha_{ij} 1_{A_i}(\varepsilon_{t-j}), \tag{3.30}$$

where $1_A(\varepsilon_t)$ is the indicator function of the set A, which takes the value 1 when $\varepsilon_t \in A$ and zero otherwise. The authors describe (3.30) as a qualitative TARCH (QTARCH) model because the conditional variance is determined by the region in R^q in which $\underline{\varepsilon}_{t-1}$ lies, rather than by the continuous values of the elements of $\underline{\varepsilon}_{t-1}$.

Engle and Ng (1993) provided a summary of asymmetric ARCH models and introduced several new models of their own. They concentrated on the GARCH (1,1) process and the functional relationship $h_t = h(\varepsilon_{t-1})$, which they term 'news impact curve'.[24] Three new asymmetric models they proposed are

$$h_t = a_0 + \beta_1 h_{t-1} + a_1(\varepsilon_{t-1} + a_2)^2 \tag{3.31}$$

[24] The differences between news impact curves of several asymmetric GARCH models of conditional volatility, employed for empirical analysis of the data from transition markets, will be analysed in the section 3.3.2 of this chapter.

$$h_t = a_0 + \beta_1 h_{t-1} + a_1(\varepsilon_{t-1} / \sqrt{h_{t-1}} + a_2)^2 \tag{3.32}$$

$$h_t = a_0 + \beta_1 h_{t-1} + a_1(\varepsilon_{t-1} + a_2\sqrt{h_{t-1}})^2 \tag{3.33}$$

Engle and Ng (1993) also proposed a very flexible functional form, which is similar to QTARCH model but is piecewise linear over the support ε_{t-1} rather a step function as in (3.30). They characterised this model as 'partially non-parametric' (PNP). They partitioned the support of ε_{t-1} into intervals, where the boundaries of the intervals are $\{\tau_{m^-},\ldots,\tau_{-1},0,\tau_1,\ldots,\tau_{m^+}\}$, and m^- is the number of intervals below zero and m^+ is the number of intervals above zero. Engle and Ng then specified

$$h_t = \alpha + \sum_{i=0}^{m^+} \theta_i P_{it}(\varepsilon_{t-1} - \tau_i) + \sum_{i=0}^{m^-} \delta_i N_{it}(\varepsilon_{t-1} - \tau_i) + \beta h_{t-1} \tag{3.34}$$

where the variables P_{it} and N_{it} are defined as

$$P_{it} = \begin{matrix} 1 & if \quad \varepsilon_{t-1} > \tau_i \\ 0 & otherwise \end{matrix}$$

and

$$N_{it} = \begin{matrix} 1 & if \quad \varepsilon_{t-1} < \tau_{-i} \\ 0 & otherwise \end{matrix}.$$

From (3.34), h_t will be linear with a different slope over each interval. For example, if ε_{t-1} is positive and lies in the interval (τ_i, τ_{i+1}), then the slope coefficient is $\theta_1 + \ldots + \theta_i$. Engle and Ng chose the τ_i s to be multiples of the unconditional standard deviation of the series.

Engle and Ng (1993) also conducted an experiment to compare the ability of asymmetric ARCH models to represent the conditional variance of stock returns. Using daily observations on the Japanese Topix stock index from January 1980 to September 1987 Engle and Ng fitted the asymmetric GARCH models discussed above. Although based on only one data set, the authors' results indicated that a parsimonious and highly parametric EGARCH could represent the volatility of stock returns remarkably well. Whether any inadequacies in the EGARCH functional form for representing the volatility of stock returns justifies a more flexible model like TARCH, QTARCH or PNP models, may largely depend on the peculiarities of the individual data set and the ultimate purpose of the empirical analysis.

In the context of estimating risk premia, Pagan and Hong (1991) suggested that no parametric functional form is sufficiently general to represent the diverse types of data that display conditional heteroskedasticity. Pagan and Hong used a non-parametric kernel estimator of the conditional variance and demonstrated that the non-parametric estimators give different conclusions about the effect of the risk premium on asset returns than do the standard parametric ARCH models. We will briefly discuss the non-parametric approach suggested by Pagan and Hong (1991) in section 3.2.6 on the estimation of the ARCH type models.

Integrated GARCH

Engle and Bollerslev (1986) were the first to consider GARCH process (3.7) with $\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i = 1$ as a distinct class of models, which they term integrated GARCH (IGARCH). They pointed out the similarity between IGARCH processes and the processes that are integrated in the mean. For a process that is integrated in mean (one that must be differenced to induce stationarity) a shock in the current period affects level of the series in 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 the applied work, is that the estimated coefficients of a GARCH conditional variance sum close to 1.

The consistent finding of very large persistence in variance in financial time series is puzzling because currently no theory predicts that this should be the case. Lamoureux and Lastrapes (1990b) argued that large persistence may actually represent misspecification of the variance and result from the structural change in the unconditional variance of the process (as represented by changes in α_0 in (3.7)). A discrete change in the unconditional variance of the process produces clustering of large and small deviations, which may show up as persistence in a fitted ARCH model. To illustrate this possibility, Lamoureux and Lastrapes used 17 years of daily returns on the 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 the 30 companies is 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 also 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.

3.2.4 Non-normal conditional distribution

As described above an attractive feature of the ARCH process is that, even though the conditional distribution of the error is normal, the unconditional distribution is non-normal with tails thicker than the normal distribution.

In spite of this property, empirical work with ARCH models indicated that the implied unconditional distributions of estimated ARCH models were not sufficiently leptokurtic to represent the distribution of returns. In the linear regression model with conditionally normal ARCH errors, suppose that $\hat{\varepsilon}_t$ and $\hat{h}_t$ are estimates of the error and conditional variance. Then the standardised residuals $\hat{\varepsilon}_t / \hat{h}_t^{1/2}$ should be approximately $N(0,1)$. However, many studies demonstrated that the sample kurtosis coefficient of the standardised residuals often exceeded 3. The frequent inability of the conditionally normal ARCH model to pass this simple diagnostic test has lead to the use of conditional distributions more general than the normal distribution.

Let $\eta_t = \hat{\varepsilon}_t / \hat{h}_t^{1/2}$ be the standardised error, and the conditional distribution of η_t be specified as

$$\eta_t \mid \Psi_{t-1} \sim f(\eta, \theta), \tag{3.35}$$

where θ is a low dimension parameter vector whose value determines the shape of the conditional distribution of η_t. In the conditionally normal ARCH model, θ is absent and $f(\eta)$ is simply $N(0,1)$ density. Bollerslev (1987) specified $f(\eta, \theta)$ as a conditional t distribution, where θ is the degrees of freedom of the distribution. The conditional t distribution allows for heavier tails than the normal distribution and, as the degrees of freedom go to infinity, includes the normal distribution as a limiting case. Using the daily rate of return in the spot market for the Deutschmark and the British pound from March 1980 to January 1985, Bollerslev estimated GARCH (1,1) models with conditional t distributions and rejected the hypothesis of conditional normality. The sample kurtosis coefficients of the standardised residuals were very close to the kurtosis coefficients of the t distribution evaluated at the estimated parameters. Engle and Bollerslev (1986), Baillie and Bollerslev (1989) and Hsieh (1989) also found that employing a conditional t distribution helped account for excess kurtosis in daily exchange rates.

Nelson (1991) employed a generalised error distribution (GED) with the EGARCH model. The GED encompasses distributions with tails both thicker and thinner than the normal, and includes the normal as a special case. For a stock price, Nelson found evidence of non-normality in the conditional distribution, but concluded that tails of the estimated GED were not sufficiently thick to account for a large number of outliers in the data. Lee and Tse (1991) suggested that the conditional distribution must be not only leptokurtic but also asymmetric. They argued that for the rates of return that cannot be negative, such as nominal interest rates, the conditional distribution should be skewed to the right. Using interest rates from the Singapore Asian dollar market, Lee and Tse estimated their model but failed to find any evidence of skewness.

Unfortunately, no single parametric specification of the conditional density (3.35), with parametric specification of the conditional variance function, appears to be suitable for all conditionally heteroskedastic data. Applications in which none of the above conditional distributions appear to be appropriate are often encountered. Hansen (1992) suggested an approach to allow more flexibility in the conditional distribution within a parametric framework. While conventional ARCH models allow the mean and variance to be time-varying, Hansen argues that skewness and kurtosis should also be time-varying, as well as be a function of the information set. Hansen proposed an autoregressive conditional density model (ARCD), which generalises (3.35) to

$$\eta_t \mid \Psi_{t-1} \sim f(\eta_t, \theta_t), \tag{3.36}$$

where the parameter θ_t, which determines the shape of the conditional density, is itself a function of the elements of the information set Ψ_{t-1}. To illustrate the use of an ARCD model, Hansen estimated a GARCH model with a conditional t distribution and time-varying degrees of freedom for the monthly excess holding yield on short-term US Treasury securities. To allow tail thickness of the conditional distribution to be determined by the information set, the degrees of freedom were parameterised as logistic transformation of a quadratic function of the lagged error and the difference between the one-month yield and the instantaneous yield. A likelihood ratio test rejected a conditional t distribution with constant degrees of freedom in favour of ARCD model.

3.2.5 Multivariate ARCH

As economic variables are interrelated, generalisation of univariate models to the multivariate and simultaneous set-up is quite natural. Apart from possible gains in efficiency in parameter estimation, estimation of a number of financial "coefficients" such as the systematic risk (beta coefficient) and the hedge ratio requires sample values of covariance between relevant variables. The motivation for multivariate ARCH also stems from the fact that many economic variables react to the same information and hence have non-zero covariances conditional on information set.

Let ε_t be the $N \times 1$ vector stochastic process of innovations to the return vector time series $y_t = (y_{1t}, \ldots, y_{Nt})'$, which we write as

$$\varepsilon_t = H_t^{1/2} u_t, \tag{3.37}$$

with u_t being $N \times 1$ i.i.d. vector with $E(u_t) = 0$ and $\text{var}(u_t) = I_N$, and H_t being $N \times N$ covariance matrix given information available at t -1, where N is the number of the assets in the system.

In somewhat general form of the multivariate linear GARCH (p,q) model, Bollerslev, Engle and Wooldridge (1988) assume that H_t is given by a linear function of the lagged cross squared errors and lagged values of H_t:

$$vech(H_t) = A_0 + \sum_{i=1}^{q} A_i vech(\varepsilon_{t-i}\varepsilon_{t-i}') + \sum_{i=1}^{p} B_i vech(H_{t-i}) \tag{3.38}$$

where $vech(.)$ denotes the operator that stacks the lower portion of matrix into a vector with $N(N+1)/2$ elements. This is a direct generalisation of our earlier univarite GARCH (p,q) model in (3.7). Representation (3.38) is called the "vech representation" of the multivariate GARCH (p,q) model. For $N = 2$ and $p = q = 1$, (3.38) takes the form

$$vech(H_t) = \begin{bmatrix} h_{11,t} \\ h_{12,t} \\ h_{22,t} \end{bmatrix} = \begin{bmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{22} \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} \varepsilon^2_{1,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon^2_{2,t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{22,t-1} \end{bmatrix} \quad (3.39)$$

The two main problems concerning the specification of H_t are that it should be positive definite for all realisations and some exclusion restrictions should be imposed so that number of parameters to be estimated is not very large. The total number of the parameters in (3.38) with $p = q = 1$ is $N^2(N+1)^2/2 + N(N+1)/2$, which grows with the fourth power of N. For the special bivariate case (3.39) the number of the parameters amounts to 21.

Engle, Granger and Kraft (1984) published the first paper on multivariate ARCH models. They considered a bivariate ARCH model, which was (3.39) without the lagged h_t components. For that model, they showed necessary conditions for H_t to be positive definite

$$\sigma_{11} > 0, \sigma_{22} > 0, \sigma_{11}\sigma_{22} - \sigma_{12}^2 > 0,$$

$$\alpha_{11} \geq 0, \alpha_{13} \geq 0, \alpha_{31} \geq 0, \alpha_{33} \geq 0,$$

$$\alpha_{11}\alpha_{33} - \alpha_{22}^2 \geq 0,$$

$$\alpha_{11}\alpha_{13} - \tfrac{1}{4}\alpha_{12}^2 \geq 0,\ \alpha_{11}\alpha_{31} - \alpha_{21}^2 \geq 0,$$

$$\alpha_{31}\alpha_{33} - \tfrac{1}{4}\alpha_{32}^2 \geq 0,\ \alpha_{13}\alpha_{33} - \alpha_{23}^2 \geq 0. \quad (3.40)$$

Note that in (3.38) and (3.39), each $h_{ij,t}$ depends on lagged squared residuals and past variances of all variables in the system. One simple assumption that could be made to reduce the number of parameters is to specify that a conditional variance depends only on its own lagged squared residuals and lagged values. The assumption amounts to take A_i and B_i to be diagonal matrices. In that case, conditions in (3.40) (the case without the lagged h_t components) reduce to

$$\sigma_{11} > 0, \sigma_{22} > 0, \sigma_{11}\sigma_{22} - \sigma_{12}^2 > 0,$$

$$\alpha_{11} \geq 0, \alpha_{33} \geq 0,\ \alpha_{11}\alpha_{33} - \alpha_{22}^2 \geq 0. \tag{3.41}$$

For $N = 2$ and $p = q = 1$ the "diagonal representation" can be expressed as

$$vech(H_t) = \begin{bmatrix} h_{11,t} \\ h_{12,t} \\ h_{22,t} \end{bmatrix} = \begin{bmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{22} \end{bmatrix} + \begin{bmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{12} & 0 \\ 0 & 0 & \alpha_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{12} & 0 \\ 0 & 0 & b_{22} \end{bmatrix} \begin{bmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{22,t-1} \end{bmatrix}$$

For the general case with unrestricted N and $p = q = 1$ the "diagonal representation" looks as follows:

$$\{H_t\}_{ij} = h_{ijt},$$

$$h_{ij,t} = \sigma_{ij} + \alpha_{ij}\varepsilon_{i,t-1}\varepsilon_{j,t-1} + b_{ij}h_{ij,t-1} \qquad i,j = 1,2,\ldots,N \tag{3.42}$$

This form was used by Bollerslev, Engle and Wooldridge (1988) for their analysis of returns on bills, bonds, stocks, and by Baillie and Myers (1991) and Bera, Garcia and Roh (1991) for the hedge ratio estimation in commodity markets.[25]

Bollerslev (1988b) suggested a constant correlation specification of multivariate volatility according to which each of the conditional variances follows a univariate GARCH process and the covariance between any two series is given by a constant correlation coefficient, ρ_{ij}, multiplying the conditional standard deviations of returns.

$$\{H_t\}_{ij} = h_{ijt},$$

$$h_{iit} = a_{i0} + a_{i1}\varepsilon_{it-1}^2 + \beta_{i1}h_{iit-1}, \tag{3.43}$$

$$h_{ijt} = \rho_{ij}\sqrt{h_{iit}h_{jjt}}, \tag{3.44}$$

[25] Note that under diagonal representation the positive definiteness of the resulting H_t is still not easy to check and also difficult to impose at the estimation stage.

This model has $N(N+5)/2$ parameters. The necessary and sufficient conditions for the model to be well defined and H_t to be positive definite is that each of the conditional variances is positive and that constant matrix of conditional correlations is positive definite. Many of the recent applications of multivariate GARCH model use this representation. However, it is quite obvious that constant correlation is a strong assumption. Bera et al. (1991) suggested a test for the constant correlation hypothesis and found that the null hypothesis of constant correlation is rejected for many financial data series.[26]

Baba, Engle, Kraft and Kroner (1990) suggested a parameterisation, known as "BEKK", which was modified by Engle and Kroner (1995). The specification is as follows.

$$H_t = A_0^{'}A_0 + \sum_{k=1}^{K} A_{1k}^{'}\varepsilon_{t-1}\varepsilon_{t-1}^{'}A_{1k} + \sum_{k=1}^{K} B_{1k}^{'}H_{t-1}B_{1k}, \tag{3.45}$$

where A_0, A_{1k}, B_{1k} are $N \times N$ parameter matrices with A_0 being restricted to be upper triangular. Summation limit K determines the generality of the process. The BEKK guarantees positive definiteness of the covariance matrix H_t by working with quadratic forms. The H_t is a linear function of its own lagged value, as well as lagged value of the squared unpredictable returns, both of which allow for own-market and cross-market influences in the conditional variances. This specification is relatively flexible, and is also parsimonious in terms of the number of the parameters to be estimated. The total number of the parameters in this system is $(5N^2 + N)/2$. For $N = 2$ and $K = 1$, (3.45) will have only 11 parameters compared to 21 parameters of the vech representation in (3.39).

[26] See also Bera and Higgins (1998).

3.2.6 Estimation of the ARCH type models

GARCH models are usually estimated by the method of maximum likelihood (ML) or quasi-maximum likelihood (QML). In some applications, the generalised method of moments (GMM) has been used.[27]

The log likelihood function for the standard ARCH regression model

$$y_t | \Psi_{t-1} \sim N(x_t^{'}\xi, h_t)$$

is given by

$$l_t = -\ln 2\pi - \frac{1}{2}\ln h_t(\theta) - \frac{1}{2}\varepsilon_t^2(\theta)h_t^{-1}(\theta),$$

and

$$l = \sum_{t=1}^{T} l_t(\theta) \qquad (3.46)$$

where $\theta = (\xi^{'}, \gamma^{'})^{'}$, with $\xi^{'}$ and $\gamma^{'}$ being conditional mean and conditional variance parameters respectively.[28]

Given the initial values for the parameter vector θ and the conditional variance[29], the log-likelihood function can be evaluated by computing h_t, t = 1, 2, ...T recursively and substituting the values in (3.46). The first order condition from maximisation of the log-likelihood for the t^{th} observation

[27] See e.g. Glosten et al. (1993).

[28] In multivariate setting the conditional log likelihood for each time period looks as follows:

$$L_t(\Theta) = -\ln 2\pi - \frac{1}{2}\ln|H_t| - \frac{1}{2}\varepsilon_t^{'}(\Theta)\ln H_t^{-1}(\Theta)\varepsilon_t(\Theta).$$

and

$$L(\Theta) = \sum_{t=1}^{T} L_t(\Theta),$$

where Θ is the vector of all parameters in the system.

[29] In most of the cases the initial value of the conditional variance is taken to be the unconditional value of the process. For example, in case of the GARCH (1,1) process, $h_1 = \sigma^2 = \dfrac{\alpha_0}{1-\alpha_1-\beta_1}$.

with respect to the conditional mean parameters ξ is:

$$\frac{\partial l_t}{\partial \xi} = \frac{\varepsilon_t x_t'}{h_t} + \frac{1}{2h_t}\frac{\partial h_t}{\partial \xi}\left(\frac{\varepsilon_t^2}{h_t} - 1\right). \tag{3.47}$$

Using $\frac{\partial h_t}{\partial \xi} = -2\sum_{i=1}^{q} \alpha_i \varepsilon_{t-i} x_{t-i}'$ for ARCH (q) (3.47) is equivalent to:

$$\frac{\partial l_t}{\partial \xi} = \frac{\varepsilon_t x_t'}{h_t} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\sum_{i=1}^{q} \alpha_i \varepsilon_{t-i} x_{t-i}', \tag{3.48}$$

Using $\frac{\partial h_t}{\partial \xi} = -2\sum_{i=1}^{q} \alpha_i \varepsilon_{t-i} x_{t-i}' + \sum_{j=1}^{p} \beta_j \frac{\partial h_{t-j}}{\partial \xi}$ for GARCH (p,q) (3.47) is equivalent to:

$$\frac{\partial l_t}{\partial \xi} = \frac{\varepsilon_t x_t'}{h_t} + \frac{1}{2h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\left(-2\sum_{i=1}^{q} \alpha_i \varepsilon_{t-i} x_{t-i}' + \sum_{j=1}^{p} \beta_j \frac{\partial h_{t-j}}{\partial \xi}\right) \tag{3.49}$$

The first order condition with respect to the conditional variance parameters γ is:

$$\frac{\partial l_t}{\partial \gamma} = \frac{1}{2h_t}\frac{\partial h_t}{\partial \gamma}\left(\frac{\varepsilon_t^2}{h_t} - 1\right), \tag{3.50}$$

where $\frac{\partial h_t}{\partial \gamma} = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2)$ for the ARCH (q), and $\frac{\partial h_t}{\partial \gamma} = z_t + \sum_{j=1}^{p} \beta_j \frac{\partial h_{t-j}}{\partial \gamma}$ with $z_t \equiv (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, h_{t-1}, \ldots, h_{t-p})$ for the GARCH (p,q) model.

Most of the applied work on ARCH models uses the Berndt, Hall, Hall and Hausman (1974) algorithm (BHHH) to maximise (3.46). In the iterative procedure, the parameters of the $(i+1)^{th}$ iteration of the BHHH algorithm are obtained by:

$$\theta^{(i+1)} = \theta^{(i)} + \left(\sum_{t=1}^{T} \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'}\right)^{-1} \sum_{t=1}^{T} \frac{\partial l_t}{\partial \theta}, \tag{3.51}$$

where the derivatives are evaluated at $\theta^{(i)}$. Under regularity conditions given for example in Crowder (1976), the value of θ which maximises l, $\hat{\theta}_{ML}$, is consistent, asymptotically normally distributed and efficient

$$\sqrt{T}(\hat{\theta}_{ML} - \theta) \overset{a}{\sim} N(0, Var(\hat{\theta}_{ML})). \tag{3.52}$$

The asymptotic covariance matrix of $\hat{\theta}_{ML}$ can be consistently estimated by the inverse of the information matrix, i.e. the negative expectation of the Hessian:

$$Var(\hat{\theta}_{ML}) = -\left[T^{-1} \sum_{t=1}^{T} E \frac{\partial^2 l_t}{\partial \theta \partial \theta'}\right]^{-1}. \tag{3.53}$$

A proof of the consistency and asymptotic normality of the ML-estimator in GARCH (1,1) and IGARCH (1,1) models is given by Lumsdaine (1992) under condition that $E\left[\ln(\alpha_1 \varepsilon_t^2 + \beta_1)\right] < 0$. The existence of finite fourth moments of ε_t is not required.

For most applications it is very difficult to justify the conditional normality assumption in (3.46). Therefore, the log likelihood function may be misspecified. As shown by Weiss (1986) for time series models with ARCH errors, and by Bollerslev and Wooldridge (1992) and Gourieroux (1992) for GARCH process, under a correct specification of the first and the second moment consistent estimates of the parameters of the model can be obtained by maximising a likelihood function constructed under the assumption of conditional normality, even though the true density could be some other. Such estimators are called the quasi-maximum likelihood estimators (QMLE). Under regularity conditions the QML-estimator has the following asymptotic distribution:

$$\sqrt{T}(\hat{\theta}_{QML} - \theta) \overset{a}{\sim} N(0, B^{-1}AB^{-1}), \tag{3.54}$$

where $A = E_0\left[\frac{\partial l_t}{\partial \theta}\frac{\partial l_t}{\partial \theta'}\right]$ is the covariance matrix of the score vector of l and $B = -E_0\left[\frac{\partial^2 l_t}{\partial \theta \partial \theta'}\right]$, where E_0 denotes the expectation conditional on the true probability density function for the data. Of course, if the latter is the normal distribution, the asymptotic distributions in (3.52) and (3.54) will be identical. Lee and Hansen (1994) prove consistency and asymptotic normality of the QML estimator of the Gaussian GARCH (1,1) model. The standardised residuals (the disturbances scaled by their conditional standard deviations) need neither be normally distributed not independent over time. The GARCH process may be integrated $\alpha_1 + \beta_1 = 1$ and even explosive $\alpha_1 + \beta_1 > 1$ provided the conditional fourth moment of the scaled disturbances is bounded.[30] In finite samples, for symmetric departures from conditional normality, the QML has been found close to the exact ML-estimator in a simulation study by Bollerslev and Wooldridge (1992). For non-symmetric conditional true distributions, both in small and large samples the loss of efficiency of QML compared to exact ML can be substantial.[31]

Another attractive way to estimate ARCH models without assuming normality is to apply the generalised method of moments (GMM) approach as advocated by Rich, Raymond and Butler (1991).[32] For simplicity consider an ARCH (1) model and define the following two errors

$$\varepsilon_t = y_t - x_t'\xi$$

[30] For a different set of assumptions for the consistency and asymptotic normality of the GARCH models see Bollerslev and Wooldridge (1992) and Lumsdaine (1992).

[31] Semi-parametric density estimation as proposed by Engle and Gonzaalez-Rivera (1991) using a linear spline with smoothness priors could then be an attractive alternative to QML.

[32] See also Sabau (1987), (1988).

$$v_t = \varepsilon_t^2 - \alpha_0 - \alpha_1 \varepsilon_{t-1}^2$$

$$= (y_t - x_t'\xi)^2 - \alpha_0 - \alpha_1 (y_{t-1} - x_{t-1}'\xi)^2 . \tag{3.55}$$

Then the GMM estimator is obtained from the following two moment conditions

$$E(\varepsilon_t \mid Z_t) = 0 \text{ and } E(v_t \mid Z_t) = 0 \tag{3.56}$$

where Z_t is a set of predetermined variables. The asymptotic distribution of the GMM estimator follows directly from the general formula in Hansen (1982).[33]

All the forms of the conditional variance, h_t, we discussed above are fully parametric. Pagan and Hong (1991) argued that the existing parametric forms are not very convincing due to the lack of optimising theory in their formulation. They suggested non-parametric estimation of both the conditional mean, m_t, and the conditional variance, h_t, since misspecification in the conditional mean might exaggerate the variation in h_t. Pagan and Hong (1991) used the kernel method and Fourier series approximation of Gallant (1982). These procedures estimate the first two conditional moments by relating them to the past values of y_t.

$$\hat{m}_t = \sum_{i=1}^{T} \omega_{it} y_{t-i} \qquad i \neq t \tag{3.57}$$

and

$$\hat{h}_t = \sum_{i=1}^{T} \omega_{it} y_{t-i}^2 - \hat{m}_t^2 \qquad i \neq t \tag{3.58}$$

where ω_{it} are kernel weights. For the Gaussian kernel

$$\omega_{it} = \frac{\kappa_{it}}{\sum_{i=1}^{T} \kappa_{it}} \qquad i \neq t \tag{3.59}$$

[33] Additionally, for the Bayesian approach to estimation of the ARCH models see Geweke (1988a,b, 1989).

where $\kappa_{it} = \exp\{-\frac{1}{2}\sum_{s=1}^{r} h_s^{-2}(y_{i-s} - y_{s-t})^2\}$, where r is the number of lags of y_t chosen, and h_s is the bandwidth.[34] The empirical applications of Pagan and Hong (1991) showed the advantages of the non-parametric approach. The disadvantage is that results from non-parametric analysis are not as easily interpretable in terms of response coefficients as those obtained from a parametric method. However, non-parametric methods can point out deficiencies in the existing parametric models and offer some guidance for modification.[35]

3.3 Empirical Analysis: Analysis of Univariate and Multivariate Volatility in the Transition Equity Markets

3.3.1 Data and Descriptive Statistics

The data used in this chapter are daily stock market index levels from four transition markets described in Chapter 2 of the thesis: Hungarian BUX, Polish WIG20, Czech PX50 and Slovak SAX. These indexes are all value-weighted, and reflect a substantial percentage of total market capitalisation of the considered markets, which could thus minimise the problem of autocorrelation in returns due to non-synchronous trading. While the series from Hungary and Poland start in June 1992, and constitute a total of 1512 observations, Czech and Slovak samples are shorter: they start in April 1994 and constitute 1036 observations. For all markets the sample period ends in March 1998. All data comes from Datastream.

We start by testing the hypothesis that logarithm of each stock market index contains a single unit root. To test this we perform an Augmented Dickey Fuller (ADF) test[36]. The ADF test (with

[34] For details see Pagan and Hong (1991), p. 60.

[35] For example, when Pagan and Hong plotted the non-parametric $\hat{h}_t$ against y_t they found a high degree of non-linearity, which would be difficult to capture by simple parametric models.

[36] The test was first introduced by Dickey and Fuller (1979).

constant and time trend) entails estimating the following regression equation below:

$$\Delta \ln(p_t) = \alpha + \beta \ln(p_{t-1}) + ct + \sum_{i=1}^{k} \delta_i \Delta \ln(p_{t-i}) + \varepsilon_t, \qquad \varepsilon_t \sim NID(0,\sigma^2) \qquad (3.60)$$

where p_t is the relevant stock price index, Δ is the first–difference operator and t is a linear time trend. In our analysis, the ADF test will be performed with and without time trend (by deleting the ct term in (3.60)).

The distribution theory supporting the ADF tests assumes that the errors ε_t are statistically independent. Therefore, we determine the order of augmentation (i.e. the number of lagged differences k) in (3.60) by adding lagged differences of the series until the residuals of the regression are serially independent.[37]

Results given in tables 3.1.A, 3.1.B and 3.1.C show the t-statistics of the ADF test and the MacKinnon critical values for the rejection of a unit root for the whole sample, and for two subsamples of our dataset. Each of the tables presents estimates first for the logarithmic levels of the series and then for their first differences. At 10 % critical value a unit root is found in all series and all subperiods examined, with exception of the Czech series in subperiod II in case of the ADF test with constant and no trend (although the test does not reject a unit root at 5% statistical level). Our results indicate that first-differencing of the series removes the non-stationary components in all cases (for all series and all subperiods). Therefore, we conclude that all four series included in our dataset are characterised by a single unit root.

[37] In our analysis the appropriate number of lagged differences is selected by adding lags until the Ljung-Box test of order 12 fails to reject no serial correlation at 10 per cent level (χ^2_{12} = 18.55).

Table 3.1.A. Tests for a unit root (the whole sample)

	ADFstat. (ln(p_t))				ADFstat. (Δ ln(p_t))			
	constant	*k**	constant + trend	*k*	constant	*k*	constant + trend	*k*
HUN (Obs.: 1-1512)	0.4944	10	-1.8043	10	-10.1250	9	-10.2057	9
POL (Obs.: 1-1512)	-2.2672	2	-1.4408	2	-22.2214	1	-22.3056	1
CZECH (Obs.: 1-1036)	-2.3942	8	-2.2431	8	-11.3730	7	-11.4323	7
SLOV (Obs.: 1-1036)	-0.6085	3	-0.5163	3	-38.8581	1	-38.9171	1
1% Crit.Val.	-3.43		-3.96		-3.43		-3.96	
5% Crit.Val.	-2.86		-3.41		-2.86		-3.41	
10%Crit.Val.	-2.57		-3.12		-2.57		-3.12	

* *k* is the chosen lag length.

Table 3.1.B. Tests for a unit root (subperiod I)

	ADFstat. (ln(p_t))				ADFstat. (Δ ln(p_t))			
	constant	*k**	constant + trend	*k*	constant	*k*	constant + trend	*k*
HUN (Obs.: 1-756)	-0.9272	2	-0.9816	2	-15.4073	1	-15.4029	1
POL (Obs.: 1-756)	-1.5261	1	0.2245	1	-22.3976	0	-22.5217	0
CZECH (Obs.: 1-518)	-1.5135	4	0.0077	4	-10.5206	3	-10.8381	3
SLOV (Obs.: 1-518)	-0.9197	1	-0.6122	1	-28.2471	0	-28.2795	0
1% Crit.Val.	-3.43		-3.96		-3.43		-3.96	
5% Crit.Val.	-2.86		-3.41		-2.86		-3.41	
10%Crit.Val.	-2.57		-3.12		-2.57		-3.12	

* *k* is the chosen lag length.

Table 3.1.C. Tests for a unit root (subperiod II)

	ADFstat. (ln(p_t))				ADFstat. (Δ ln(p_t))			
	constant	*k*	constant + trend	*k*	constant	*k*	constant + trend	*k*
HUN (Obs.: 757-1512)	-0.8007	7	-1.7376	7	-10.5736	6	-10.5797	6
POL (Obs.: 757-1512)	-2.3593	1	-2.5485	1	-22.9535	0	-23.0051	0
CZECH (Obs.: 519-1036)	*-2.6044*	1	-2.4431	2	-16.3314	0	-16.3486	0
SLOV (Obs.: 519-1036)	-1.2688	1	-1.0094	1	-25.5729	0	-25.7164	0
1% Crit.Val.	-3.43		-3.96		-3.43		-3.96	
5% Crit.Val.	-2.86		-3.41		-2.86		-3.41	
10% Crit.Val.	-2.57		-3.12		-2.57		-3.12	

* *k* is the chosen lag length.

Daily returns on national equity indexes are computed as logarithmic price relatives:

$$r_{jt} = \ln(P_{jt} / P_{jt-1}), \qquad j = 1,2,3,4. \tag{3.61}$$

Assuming that investors know the past information set containing the realised values of all relevant variables up to time t - 1, when making their investment decision at time t - 1, we consider the conditional expected return of r_t and the conditional variance of r_t to be the relevant expected return and volatility to the investors. Since our focus is on the conditional variance, rather than the conditional mean, we concentrate on the unpredictable part of stock returns, as obtained through autoregressive regression which removes the predictable part of return series. The daily return r_t is regressed on a constant and the first p lags of r_t to obtain the residual ε_t, which is our *unpredictable* stock return data.[38]

$$r_t = \gamma_0 + \sum_{i=1}^{p} \gamma_i r_{t-i} + \varepsilon_t \tag{3.62}$$

The necessary amount of lags required for the non-autocorrelated residual series ε_t equals 10 for the Hungarian and Slovak series, and 6 for the Polish and Czech series. In the case of the Hungarian market, before the autocorrelation adjustment procedure in (3.62) we eliminate outliers present in the last 100 observations of the series.[39]

Table 3.2 reports summary statistics for the unpredictable stock returns. The evidence for all four series indicates significantly fatter tails than does the stationary normal distribution. The excess kurtosis statistic ranges in value from 3.4965 for the Hungarian to 9.7249 for the Slovak returns. The estimates of the coefficients of skewness reveal that the Czech return data has asymmetric distributions skewed to the left. For the rest of the series the skewness coefficients are insignificant. The examination of the autocorrelation structure for the levels does not indicate any significant

---

[38] In the rest of the text series ε_t will be referred to as *unpredictable* returns.

[39] For the graphs of the nominal and unpredictable return series see Figures 3.1-3.4 at the end of this chapter.

serial correlation left in the series after autocorrelation adjustment procedure in (3.62). However, the application of the Ljung-Box statistic in Table 3.2 to each series of squared residuals strongly suggests the presence of time-varying volatility. This is the so-called stylised fact of *volatility clustering* in return series. Its consequence is leptokurtosis, which we have found above.[40]

Table 3.2 Statistical properties of unpredictable return series

	HUN	POL	CZECH	SLOV
Numb. of observ.	1512	1512	1036	1036
Skewness	-0.02621	-0.0270	-0.5920	-0.03958
Excess Kurtosis	3.49652	4.0545	5.3798	9.72490
Ljung-Box(12) for the levels	3.6778	6.4327	11.3422	9.3826
$\rho(1)$	-0.00147	0.00024	-0.01204	-0.00373
$\rho(2)$	-0.00200	0.00103	0.01447	0.00037
$\rho(3)$	-0.00563	0.00028	-0.00172	-0.00018
$\rho(4)$	-0.00775	-0.00039	-0.04775	-0.00657
$\rho(5)$	-0.00134	-0.00327	0.04911	0.01179
$\rho(6)$	-0.00417	-0.01233	0.00275	0.00404
$\rho(7)$	-0.00055	0.02664	-0.01635	0.01248
$\rho(8)$	0.00274	0.04066	0.00385	0.03998
$\rho(9)$	-0.00470	0.02175	-0.00622	-0.02181
$\rho(10)$	-0.01663	0.02663	0.01229	0.02749
$\rho(11)$	0.04467	-0.02351	-0.02346	0.04779
$\rho(12)$	0.00304	0.00151	-0.01090	-0.05967
Bartlett st. error	0.02572	0.02572	0.03107	0.03107
Ljung-Box(12) for the squares	226.0949	746.7081	219.0603	75.7675

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively. The Bartlett standard error for the autocorrelation equals $1/\sqrt{N}$. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

[40] Bollerslev et al. (1994) show that a series where the variance is not constant but time dependent generates leptokurtosis.

3.3.2 Univariate analysis: model specifications

In the previous section we have established the existence of significant variations from normality and volatility clustering in our data series. We need specifications general enough to capture these characteristics. As it was shown in section 3.2 the Autoregressive Conditional Heteroskedastic (ARCH) class of models has proven to be particularly suited for modelling the behaviour of financial time series. These models are capable of capturing the most common empirical observations in daily return data, such as leptokurtosis and volatility clustering.

We proceed with presenting the family of univariate ARCH models, which will be employed for modelling the volatility in the transition stock markets. The detailed description of most of the models employed here is given in sections 3.2.1 and 3.2.3 of this chapter.

The specification of conditional variance in a GARCH (p,q) model is given in (3.63) below:

Model 1: $\varepsilon_t = h_t^{1/2} u_t$

$$h_t = a_0 + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-1} \tag{3.63}$$

with u_t being iid with $E(u_t) = 0$ and $\operatorname{var}(u_t) = 1$, $a_0 > 0, a_i \geq 0, \beta_i \geq 0$, and $\sum_{i=1}^{p} a_i + \sum_{i=1}^{q} \beta_i < 1$. In practice, numerous studies have demonstrated that GARCH (1,1) specification (which will be employed in our analysis) is most appropriate. The coefficients of the model are easily interpreted, with the estimate of a_1 showing the impact of current news on the conditional variance process and the estimate of β_1 the persistence of volatility to a shock or, alternatively, the impact of 'old' news on volatility.

An alternative specification in (3.64), proposed by Engle and Bollerslev (1986) is the Nonlinear GARCH (NGARCH) model, which implies a reduced response to extreme news if $a_2 < 2$. *Model 1* is nested in *Model 2* for $a_2 = 2$.

Model 2: $$h_t = a_0 + a_1 | \varepsilon_{t-1} |^{a_2} + \beta_1 h_{t-1} \tag{3.64}$$

In the models above positive and negative past values have a symmetric effect on the conditional variance. As it was mentioned in section 3.2.3, many financial series are asymmetric, with negative equity returns being followed by larger increases in volatility than equally large positive returns. Nelson (1991) proposed the exponential GARCH (EGARCH) model[41], which allows for asymmetry.

Model 3: $$\ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1 \left[|\varepsilon_{t-1}| / \sqrt{h_{t-1}} - \sqrt{2/\pi} \right] + \alpha_2 (\varepsilon_{t-1} \sqrt{h_{t-1}}) \tag{3.65}$$

This specification lifts the nonnegativity constraint on the parameters of the model by using logs. The parameter a_2 can generate the leverage effect, as, in contrast to the simple GARCH, the sign of yesterday's shock enters the model.

A frequently used alternative specification for the conditional volatility process is the model proposed by Glosten, Jarannathan and Runkle (1993) (GJR), which extends the simple GARCH model to allow for asymmetric effects by including an indicative dummy as shown in (3.66):

Model 4: $$h_t = a_0 + \beta_1 h_{t-1} + a_1 \varepsilon_{t-1}^2 + a_2 S_{t-1}^- \varepsilon_{t-1}^2 \tag{3.66}$$

The indicative dummy S_{t-1}^- takes the value of 1 if the value of the $\varepsilon_{t-1} < 0$ and 0 otherwise. Note, that while the impact of a piece of positive news is estimated by a_1 alone, the impact of negative news is given by the sum of a_1 and a_2.

Other specifications of GARCH type models, which capture the asymmetric response of volatility to

---

[41] For a detailed description of EGARCH model see section 3.2.3. Here we use slightly different notation from the one presented in 3.2.3 in order to make the parameters of this model easily comparable to the parameters of other ARCH models employed for empirical analysis of the data from transition markets.

news, are presented below:

The Asymmetric GARCH model (AGARCH):

Model 5: $$h_t = a_0 + \beta_1 h_{t-1} + a_1(\varepsilon_{t-1} + a_2)^2 \tag{3.67}$$

The Nonlinear Asymmetric GARCH (NAGARCH):

Model 6: $$h_t = a_0 + \beta_1 h_{t-1} + a_1(\varepsilon_{t-1} + a_2\sqrt{h_{t-1}})^2 \tag{3.68}$$

The VGARCH specification:

Model 7: $$h_t = a_0 + \beta_1 h_{t-1} + a_1(\varepsilon_{t-1} / \sqrt{h_{t-1}} + a_2)^2 \tag{3.69}$$

Engle and Ng (1993) suggest an interesting metric by which to analyse the effect of the news on conditional volatility. Holding constant the information dated *t-2* and earlier, one can examine the implied relation between ε_{t-1} and h_t. Engle and Ng call the curve which characterises this relation, with all lagged conditional variances evaluated at the level of the unconditional variance of the stock return, the *news impact curve*.

The qualitative differences between the models presented above can be compared by contrasting their news impact curves. The standard GARCH model has a news impact curve which is symmetric and centred at $\varepsilon_{t-1} = 0$. This implies that positive and negative return shocks of the same magnitude produce the same amount of volatility. Therefore, if a negative return shock causes more volatility than a positive return shock, the GARCH model underpredicts the amount of volatility following the bad news and overpredicts the amount of volatility following good news. The standard GARCH implies that larger return shocks forecast more volatility at a rate proportional to the square of the size of the return shock. Therefore, if large return shocks cause more volatility than a quadratic function allows, then the standard GARCH model underpredicts volatility after a large return shock and overpredicts volatility after a small return shock. The Nonlinear GARCH (NGARCH) model in

(3.64) is more flexible than the standard GARCH, because the parabola representing the news impact curve of this model is not restricted to be quadratic.

The models 3.65 – 3.69 above add more flexibility to the standard GARCH specification by allowing several types of asymmetry in the impact of news on volatility. The news impact curve of EGARCH and GJR models has a minimum at ε_{t-1}=0, but has asymmetric positive and negative sides. In the conditional variance function AGARCH model, the introduction of the parameter α_2 shifts the parabola horizontally so that the minimum occurs at $\varepsilon_{t-1} = -\alpha_2$. This produces asymmetry because if, for example, $\alpha_2 < 0$, then $h_t = h(-\varepsilon_{t-1})$ exceeds $h_t = h(\varepsilon_{t-1})$ for $\varepsilon_{t-1} > 0$. The shape of the news impact curves of NAGARCH and VGARCH models is symmetric and has a minimum at $\varepsilon_{t-1} = (-a_2)\sqrt{h_{t-1}}$. The difference between them is that the two sides of the VGARCH curve are steeper than that of the NAGARCH. In sum, the news impact curves of the asymmetric models defined above capture the leverage effect by allowing either the slope of the two sides of the curve to differ or the center of the curve to locate at the point where ε_{t-1} is positive.

As it was discussed in the theoretical part 3.2 of this chapter, the GARCH models imply that the conditional distribution of returns is normal, i.e. standardised residuals of these models should be normal. Unfortunately, in practice there is often excess kurtosis in the standardised residuals of GARCH models. To handle this problem, Bollerslev and Wooldridge (1992) studied quasi-maximum likelihood estimation of GARCH models. They showed that under a correct specification of the first and the second moments consistent estimates of the parameters of the model can be obtained by maximising a likelihood function constructed under the assumption of conditional normality, even though the true density could be some other. Consistent standard errors for parameter estimates can then be obtained using a robust covariance matrix estimator. All inferences in the following sections of this chapter will be based on robust standard errors from quasi-maximum likelihood estimation procedure, described in the part 3.2.6 of the thesis.

After estimation of the models presented above we perform a range of diagnostic tests to establish goodness of fit of the employed specifications. We examine whether the standardised residuals of the estimated models display excess skewness and kurtosis. Properly specified GARCH models should be able to significantly reduce the excess skewness and kurtosis present in nominal returns. Second, we examine whether the squared standardised residuals of the models are independent and identically distributed. We also report the Schwartz Information Criterion (SIC) for each specification used, which allows a degrees of freedom free comparison of the models performance.

$$SIC = l - 0.5 * p * \ln(T),$$

where l is the estimated log-likelihood, p is the number of estimated parameters, and T is the number of observations. However, the question of nestedness of the estimated models makes their comparison using SIC criterion difficult.

In a similar vein, given that the parameterisation of the estimated models differs so much, we compare the amount of persistence in the variance predicted by each of these models by regressing h_t on a constant and h_{t-1}, and report slope coefficients of the regressions.[42]

3.3.3. Empirical results

Table 3.3 reports parameter estimates of the alternative volatility models defined in the previous section. The coefficient of lagged conditional variance, β_1, is significantly positive and less than one for all markets and all models estimated. The magnitude of the coefficient is high in all the cases, indicating a long memory (smoothing) in the variance. The coefficient of the lagged squared returns, α_1, is positive for all stocks and statistically significant for almost all countries and specifications,

[42] Another important diagnostic test for GARCH models is out-of-sample forecasting performance of volatility.

with exception of NGARCH for Hungary and Slovakia, and GARCH, NGARCH and GRJ for Czech Republic. Therefore, we conclude that strong GARCH effects are apparent for all four markets examined, probably with exception of Czech Republic where the coefficient of the lagged squared returns is not significant in three out of seven specifications of conditional volatility.

The results of the estimation of the Nonlinear ARCH (NGARCH) model, which implies a reduced response of volatility to extreme news if $a_2 < 2$ as compared to simple GARCH specification, show that the parameter a_2 is close to 2 for all four series. To check whether the simple GARCH specification is nested within the Nonlinear ARCH specification, we perform t-test with null hypothesis $a_2 = 2$ for each market examined. Table 3.3 reports the results of the tests. The null hypothesis is never rejected. We conclude that simple GARCH is preferred to Nonlinear ARCH as it is a more parsimonious specification, which leads to better determined parameters of the model.

Moving to examination of asymmetric impact of shocks on volatility, we find rather weak evidence of asymmetries in the markets. The so-called leverage effects are captured by EGARCH for the Czech and AGARCH for the Slovak returns. In the case of the Czech series the last specification has the highest log-likelihood and Schwartz criterion (SIC) among alternative models as well.

Although the SICs for the standard GARCH specification in Table 3.3 are the highest for the Hungarian, Polish and Slovak markets among the all estimated models, and the SIC for the EGARCH model is just slightly higher in the case of the Czech market, it is not obvious that GARCH specification is the best model chosen on the basis of the SIC criterion. As it was mentioned in the previous section of the chapter, the problem of nonnestedness of the specified models makes their comparison using SIC difficult.[43]

To compare the amount of persistence in variance from one period to the next predicted by

[43] The attempt to set up the estimated volatility models so that they are nested could be an interesting task for further research.

alternative volatility models, the estimates of the parameters for each specification in Table 3.3 are followed by first-order autoregressive coefficient for conditional variance, h_t. To summarise the results, for the Hungarian, Czech and Slovak markets the GARCH (1,1) specification shows the highest persistence in variance. In the case of the Polish series the autoregressive coefficient for conditional variance of the GJR specification is slightly higher than the one for the GARCH model. The persistence in the variance does not follow common trend for the remaining models, and varies from market to market.[44]

The diagnostics in Table 3.4. indicate that though the estimated models reduce coefficients of excess kurtosis present in nominal return series, we still reveal significant nonnormality in the distribution of standardised residuals of the considered specifications for all the series examined.[45] This suggests that, although the GARCH parameterisations we employed can accommodate some of the kurtosis in the data, the use of a fat-tailed conditional distribution might be more appropriate.[46]

The Ljung-Box test statistics and the corresponding autocorrelation coefficients in Tables 3.4.1 and 3.4.2 respectively reveal that all the specifications of conditional volatility used in our study reduce the intertemporal dependence in the squared standardised residuals for the Hungarian, Czech and Slovak series to insignificance at the 5 % level, with exception of the NGARCH for the Hungarian series, where we find a significant $\rho(1)$ coefficient (see Table 3.4.A.2). Diagnostics for the Polish series in Table 3.4.B.2 indicate that volatility clustering present in the returns, while being reduced considerably (see Table 3.2 for comparison), is not removed completely by any of the alternative specifications discussed here.

---

[44] The question why the simplest model, GARCH (1,1), produces the highest persistence in conditional volatility compared to other more elaborate models used in our analysis is another interesting issue to explore further.

[45] Such evidence against normality warrants the use of QML testing procedures.

[46] See section 3.2.4 for a discussion on this topic.

Table 3.3 Results of the estimation of univariate models of volatility

A. Hungary

	a_0	a_1	β_1	a_2	AR(1) on Conditional variance	L	SIC
GARCH	6.11e-05 (2.7020)	0.1292 (3.6971)	0.8286 (19.0786)	_	0.9285 (58.9964)	6021.9209	5963.351
NGARCH	-1.14e-04 (-0.0590)	0.0011 (0.7161)	0.8065 (15.0542)	1.6604* (2.0045)	0.6646 (28.1145)	6014.0001	5951.770
EGARCH	-0.6041 (-2.0822)	0.2725 (3.2870)	0.9314 (28.5008)	-0.02170 (-0.2261)	0.9035 (55.8547)	6017.2387	5955.008
GJR	6.09e-06 (2.4974)	0.1396 (4.2903)	0.8345 (17.6583)	-0.0370 (-1.1600)	0.4419 (15.9336)	6014.3586	5952.128
AGARCH	5.94e-06 (2.4123)	0.1315 (3.7595)	0.8283 (20.1463)	-4.12e-04 (-0.3350)	0.3288 (10.6678)	6014.0011	5951.771
NAGARCH	6.16e-05 (2.4673)	0.1274 (3.5523)	0.8293 (18.2188)	0.0303 (0.1645)	0.1534 (3.8699)	6014.4402	5952.210
VGARCH	-3.17e-07 (-0.2296)	1.44e-05 (5.1193)	0.8870 (43.2559)	0.0153 (0.1272)	0.8871 (59.4915)	6016.2312	5954.001

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood t-values in brackets.
* H0: $a_2 = 2$ t - test = 0.4141

B. Poland

	a_0	a_1	β_1	a_2	AR(1) on Conditional variance	L	SIC
GARCH	4.11e-05 (2.6294)	0.2046 (4.9215)	0.7837 (19.0745)	_	0.9320 (50.5357)	5119.1533	5082.571
NGARCH	4.34e-05 (3.0364)	0.1768 (1.9804)	0.7852 (22.6771)	1.9548* (11.7166)	0.7369 (21.2596)	5119.2028	5078.946
EGARCH	-0.3706 (-3.3981)	0.9490 (5.6832)	0.3761 (64.8704)	0.0048 (0.0690)	0.1312 (3.8141)	5111.7340	5071.493
GJR	8.75e-06 (2.6793)	0.2067 (5.7058)	0.7837 (20.5781)	-0.0000 (-0.0901)	0.9463 (55.2629)	5119.1638	5078.923
AGARCH	4.88e-06 (2.6693)	0.2050 (4.5787)	0.7833 (18.1093)	5.63e-05 (-0.1721)	0.5134 (10.1784)	5119.1845	5078.943
NAGARCH	5.32e-05 (2.3108)	0.2065 (4.3250)	0.7817 (16.7114)	-0.0399 (-0.3410)	0.8141 (21.7034)	5119.2892	5079.048
VGARCH	-4.97e-07 (-5.5061)	3.18e-05 (5.8945)	0.9104 (65.0870)	0.0000 (0.1090)	0.0599 (3.4056)	5096.7606	5056.521

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood t-values in brackets.
* H0: $a_2 = 2$ t - test = 0.2710

C. Czech Republic

	a_0	a_1	β_1	a_2	AR(1) on Conditional variance	L	SIC
GARCH	7.99e-07 (0.8233)	0.1459 (1.2370)	0.7884 (4.2309)	_	0.9320 (50.5357)	4476.1276	4441.446
NGARCH	4.93e-07 (0.5651)	0.0964 (0.8198)	0.7890 (2.9118)	1.8986* (7.8230)	0.5656 (11.3045)	4476.2865	4438.137
EGARCH	-0.1125 (-1.4115)	0.1472 (3.8155)	0.9642 (57.7268)	0.2566 (2.0949)	0.2452 (5.6562)	4479.9891	4441.849
GJR	8.86e-05 (1.1870)	0.1014 (1.3957)	0.8150 (7.7236)	0.0479 (1.1670)	0.3040 (7.1942)	4476.8928	4438.743
AGARCH	4.81e-05 (1.2229)	0.1193 (2.1665)	0.8227 (9.5543)	-0.0000 (-1.3473)	0.2388 (4.6545)	4477.9993	4439.849
NAGARCH	5.76e-06 (1.4263)	0.1074 (2.1994)	0.8322 (12.1952)	-0.3072 (-1.4142)	0.8154 (18.0413)	4478.5453	4440.395
VGARCH	3.59e-07 (0.5580)	4.33e-05 (3.1163)	0.9179 (30.4962)	-0.2400 (-1.2005)	0.0971 (3.3304)	4459.9619	4421.812

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood t-values in brackets.

* H_0: $a_2 = 2$ $\quad$ t - test = 0.4225

D. Slovakia

	a_0	a_1	β_1	a_2	AR(1) on Conditional variance	L	SIC
GARCH	8.57e-06 (1.98936)	0.1655 (2.6036)	0.7014 (3.7997)	_	0.8270 (47.2620)	4018.5251	3967.277
NGARCH	8.75e-05 (1.9666)	0.0336 (0.7025)	0.6592 (4.1283)	1.5891* (4.1202)	0.4129 (14.8306)	4020.0268	3965.118
EGARCH	-1.9140 (-0.9517)	0.2882 (1.86024)	0.8020 (3.8776)	-0.0117 (-0.0516)	0.1623 (2.1830)	4015.9584	3961.050
GJR	9.16e-06 (1.0923)	0.1892 (2.10522)	0.6838 (3.2392)	-0.0286 (-0.1732)	0.4378 (21.6537)	4018.6775	3963.769
AGARCH	4.08e-06 (1.0449)	0.1968 (2.3547)	0.8306 (10.5317)	-0.2418 (-2.0205)	0.4058 (8.4653)	4019.5263	3964.617
NAGARCH	4.20e-06 (0.6125)	0.0860 (2.4162)	0.7271 (9.2253)	-0.2356 (-1.4985)	0.3922 (8.4326)	4019.2122	3964.303
VGARCH	7.71e-06 (1.3441)	9.18e-06 (1.61762)	0.7202 (4.6322)	0.0817 (0.3225)	0.1430 (3.6199)	4012.0474	3957.139

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood t-values in brackets.

* H_0: $a_2 = 2$ $\quad$ t - test = 1.0653

Table 3.4 Properties of standardised residuals of the univariate models of volatility

A. Hungary

A.1 Skewness and excess kurtosis

	GARCH	NGARCH	EGARCH	GJR	AGARCH	NAGARCH	VGARCH
Skewness	0.0568	0.0483	0.0580	0.0319	0.0701	0.0491	0.0345
Excess Kurtosis	2.8935	2.1002	2.6381	2.9519	2.8733	2.9437	2.6317

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively.

A.2 Autocorrelation in the squared standardised residuals

	GARCH	NGARCH	EGARCH	GJR	AGARCH	NAGARCH	VGARCH
$\rho(1)$	0.0399	0.0569*	0.0502	0.0431	0.0363	0.0378	0.0247
$\rho(2)$	-0.0238	-0.0115	-0.0222	-0.0227	-0.0240	-0.0238	-0.0245
$\rho(3)$	0.0104	0.0234	0.0133	0.0138	0.0083	0.0104	0.0118
$\rho(4)$	-0.0244	-0.0190	-0.0237	-0.0224	-0.0269	-0.0255	-0.0245
$\rho(5)$	-0.0369	-0.0413	-0.0436	-0.0380	-0.0366	-0.0370	-0.0312
$\rho(6)$	-0.0074	0.0001	-0.0061	-0.0132	-0.0058	-0.0080	0.0032
$\rho(7)$	-0.0043	0.0012	-0.0007	-0.0051	-0.0042	-0.0038	0.0003
$\rho(8)$	0.0036	0.0128	0.0050	-0.0002	0.0050	0.0031	0.0092
$\rho(9)$	-0.0062	0.0086	-0.0039	-0.0094	-0.0048	-0.0069	0.0138
$\rho(10)$	0.02276	0.0362	0.0240	0.0215	0.0236	0.0224	0.0342
$\rho(11)$	-0.0210	-0.0321	-0.0265	-0.0230	-0.0201	-0.0212	-0.0146
$\rho(12)$	-0.0038	0.0066	-0.0025	-0.0048	-0.0031	-0.0038	0.0073
Ljung-Box(12)	8.0252	12.9736	10.5745	8.7610	7.6847	7.8686	7.0064

Note: The Bartlett standard error for the autocorrelation equals $1/\sqrt{N}$ = 0.0257. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

* denotes statistical significance at the 5 % levels.

B. Poland

B.1 Skewness and excess kurtosis

	GARCH	NGARCH	EGARCH	GJR	AGARCH	NAGARCH	VGARCH
Skewness	0.0064	0.0113	0.0693	0.0025	0.0134	0.0178	0.0093
Excess Kurtosis	2.3697	2.3596	2.4381	2.3780	2.3499	2.3351	1.7954

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively.

B.2 Autocorrelation in the squared standardised residuals

	GARCH	NGARCH	EGARCH	GJR	AGARCH	NAGARCH	VGARCH
$\rho(1)$	-0.0223	-0.0256	-0.0272	-0.0264	-0.0232	-0.0243	0.0164
$\rho(2)$	0.1018**	0.1044**	0.1159**	0.1010**	0.1022**	0.1019**	0.1369**
$\rho(3)$	0.0477	0.0496	0.0437	0.0476	0.0398	0.0378	0.0953**
$\rho(4)$	-0.0525*	-0.0532*	-0.0593*	-0.0524*	-0.0528*	-0.0531*	-0.0396
$\rho(5)$	0.0489	0.0458	0.0431	0.0442	0.0463	0.0496	0.1171
$\rho(6)$	-0.0350	-0.0355	-0.0369	-0.0398	-0.0456	-0.0383	-0.0325
$\rho(7)$	-0.0031	-0.0043	-0.0040	-0.0034	-0.0023	-0.0067	0.0320
$\rho(8)$	-0.0081	-0.0070	-0.0042	-0.0078	-0.0097	-0.0085	0.0546*
$\rho(9)$	-0.0266	-0.0298	-0.0521	-0.0302	-0.0391	-0.0374	-0.0151
$\rho(10)$	0.0128	0.0178	0.0300	0.0155	0.0141	0.0162	0.0664**
$\rho(11)$	-0.0494	-0.0483	-0.0493	-0.0492	-0.0458	-0.0506*	-0.0329
$\rho(12)$	0.0308	0.0321	0.0337	0.0315	0.0322	0.0368	0.0641**
Ljung-Box(12)	36.9106	37.7278	45.6000	36.5360	37.2889	37.7151	89.0350

Note: The Bartlett standard error for the autocorrelation equals $1/\sqrt{N}$ = 0.0257. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

* and ** denote statistical significance at the 5 % and 1 % levels, respectively.

C. Czech Republic

C.1 Skewness and excess kurtosis

	GARCH	NGARCH	EGARCH	GJR	AGARCH	NAGARCH	VGARCH
Skewness	-0.4816	-0.4654	-0.3452	-0.4926	-0.4990	-0.4930	-0.5185
Excess Kurtosis	5.5811	5.6249	5.4807	5.6723	5.7403	5.7098	5.1585

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively.

C.2 Autocorrelation in the squared standardised residuals

	GARCH	NGARCH	EGARCH	GJR	AGARCH	NAGARCH	VGARCH
$\rho(1)$	0.0059	0.0096	0.0589	0.0171	0.0141	0.0195	-0.0185
$\rho(2)$	-0.0190	-0.0165	0.0193	-0.0072	-0.0086	-0.0056	-0.0344
$\rho(3)$	-0.0124	-0.0112	-0.0197	-0.0164	-0.0164	-0.0161	-0.0275
$\rho(4)$	0.0031	-0.0011	0.0055	-0.0014	-0.0011	-0.0005	-0.0100
$\rho(5)$	0.0058	0.0055	-0.0085	-0.0023	0.0012	-0.0044	-0.0002
$\rho(6)$	-0.0227	-0.0175	-0.0291	-0.0205	-0.0211	-0.0233	-0.0254
$\rho(7)$	-0.0228	-0.0211	-0.0268	-0.0257	-0.0249	-0.0263	-0.0291
$\rho(8)$	0.0052	0.0164	0.0299	-0.0046	-0.0037	-0.0047	0.0014
$\rho(9)$	-0.0073	-0.0026	-0.0169	-0.0136	-0.0136	-0.0134	-0.0197
$\rho(10)$	0.0195	0.0224	0.0161	0.0040	0.0072	0.0014	0.0126
$\rho(11)$	-0.0094	-0.0110	-0.0169	-0.0161	-0.0156	-0.0158	-0.0172
$\rho(12)$	-0.0257	-0.0247	-0.0350	-0.0296	-0.0295	-0.0291	-0.0226
Ljung-Box(12)	2.9630	2.9018	9.2347	3.2071	3.1181	3.3715	5.4755

Note: The Bartlett standard error for the autocorrelation equals $1/\sqrt{N}$ = 0.0257. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

D. Slovakia

D.1 Skewness and excess kurtosis

	GARCH	NGARCH	EGARCH	GJR	AGARCH	NAGARCH	VGARCH
Skewness	-0.0771	-0.0745	-0.0549	-0.0691	-0.0750	-0.0661	-0.0643
Excess Kurtosis	6.1860	6.4105	6.3424	6.3205	6.3120	6.4775	6.1486

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively.

D.2 Autocorrelation in the squared standardised residuals

	GARCH	NGARCH	EGARCH	GJR	AGARCH	NAGARCH	VGARCH
$\rho(1)$	0.0297	0.0367	0.0410	0.0327	0.0451	0.0480	0.0365
$\rho(2)$	0.0428	0.0528	0.0491	0.0526	0.0534	0.0599	0.0587
$\rho(3)$	-0.0150	-0.0137	-0.0107	-0.0134	-0.0117	-0.0111	-0.0082
$\rho(4)$	-0.0044	-0.0037	0.0008	-0.0046	-0.0025	-0.0031	0.0044
$\rho(5)$	0.0121	0.0151	0.0172	0.0158	0.0052	0.0049	0.0270
$\rho(6)$	-0.0173	-0.0153	-0.0143	-0.0150	-0.0234	-0.0235	-0.0074
$\rho(7)$	0.0215	0.0301	0.0232	0.0304	0.0189	0.0236	0.0352
$\rho(8)$	0.0227	0.0316	0.0297	0.0272	0.0130	0.0143	0.0312
$\rho(9)$	0.0041	0.0086	0.0072	0.0071	-0.0004	0.0007	0.0138
$\rho(10)$	0.0286	0.0318	0.0339	0.0306	0.0148	0.0136	0.0405
$\rho(11)$	-0.0151	-0.0142	-0.0144	-0.0146	-0.0172	-0.0173	-0.0129
$\rho(12)$	-0.0253	-0.0233	-0.0242	-0.0241	-0.0270	-0.0265	-0.0217
Ljung-Box(12)	6.2782	8.8059	8.3827	8.2157	7.6066	8.8141	10.6651

Note: The Bartlett standard error for the autocorrelation equals $1/\sqrt{N}$ = 0.0257. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

3.3.4 Multivariate Analysis: Model Specifications

In the previous sections we concentrated on the properties of univariate data and statistical models that were useful descriptions of such data. However, financial analysis is also largely concerned with the examination of the multivariate characteristics of data. It is likely that the conditional variance of the return on an asset is related not only to the past history of its own return but also to those on other assets. In an international context, the conditional variance of the return on the national index would be related to the returns on other national indexes. As it was discussed in section 3.2.5 of this chapter a number of authors have proposed using multivariate GARCH models to represent a set of time series whose variances and covariances change over time.

The purpose of this section is to consider the interdependence between conditional second moments of four return series from the transition markets of Central Europe. The returns will be modelled using two alternative specifications of the multivariate GARCH type framework, described in detail in section 3.2.5 of this chapter. Note that in multivariate setting all four series considered in our study start in April 1994 and end in March 1998, constituting a total of 1036 daily observations.

The first multivariate model we employ in this section is the constant correlation specification of volatility suggested in Bollerslev (1988). In this model each of our four unpredictable return series follows a univariate GARCH (1,1) process and the covariance between any two series is given by a constant correlation coefficient, ρ_{ij}, multiplying the conditional standard deviations of returns.

$$\{H_t\}_{ij} = h_{ijt},$$

$$h_{iit} = a_{i0} + a_{i1}\varepsilon_{it-1}^2 + \beta_{i1}h_{iit-1},$$

$$h_{ijt} = \rho_{ij}\sqrt{h_{iit}h_{jjt}}, \tag{3.70}$$

Next, to relax the strong constant correlation assumption, as well as to allow for cross effects in the variance equations, we estimate the BEKK multivariate GARCH specification by Engle and Kroner (1995). As discussed in the section 3.2.5, this model provides cross effects in the variance equations parsimoniously, and also guarantees positive semi-definiteness by working with quadratic forms.

$$H_t = A_0' A_0 + \sum_{k=1}^{K} A_{1k}' \varepsilon_{t-1} \varepsilon_{t-1}' A_{1k} + \sum_{k=1}^{K} B_{1k}' H_{t-1} B_{1k}, \tag{3.71}$$

All inferences in the following section, presenting empirical results from estimation of the multivariate volatility models, are based on robust standard errors from quasi-maximum likelihood estimation procedure described in section 3.2.6 of this chapter.

3.3.5 Empirical Results

Constant correlation specification

Table 3.5 reports the Ljung-Box test statistics for twelfth-order serial correlation in the squared unpredictable return series and their cross-products, $\varepsilon_{it}\varepsilon_{jt}$. We detect significant serial correlation in all the elements of the table.

Table 3.5 Autocorrelation in the cross-products of unpredictable returns[47]

	HUN	POL	CZECH	SLOV
HUN	126.2686			
POL	47.8289	437.2205		
CZECH	32.5557	72.1148	219.0603	
SLOV	28.7669	81.4461	58.6678	75.7675

Note: The values in the table are the Ljung-Box(12) statistics for the autocorrelation in the cross-products of the unpredictable return series. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

[47] As indicated above, in the multivariate setting all four return series constitute 1036 observations. As a result, the Hungarian and Polish series used in this section are shorter than those employed in the univariate section. Consequently, the Ljung-Box statistics for these series vary from those reported in Table 3.2.

The quasi maximum likelihood estimates of the model in (3.70) are given in Table 3.6. All the parameters in the time-varying conditional variances are individually significant at the usual 5 per cent level, with the exception of an insignificant coefficient of the lagged squared return for the Czech data. The estimates for conditional correlations ρ_{ij} in Table 3.6 are significant between three pairs of countries: Hungary and Poland (0.1805), Hungary and Czech Republic (0.1798), and Poland and Czech Republic (0.1273).

Table 3.6 Results of the estimation of the constant correlation specification of multivariate volatility

	HUN	POL	CZECH	SLOV
a_{i0}	1.2940e-005 (3.90486)	2.8950e-005 (1.9978)	3.7658e-005 (1.21783)	9.7844e-005 (2.34095)
a_{i1}	0.1839 (5.30351)	0.2049 (18.59089)	0.0993 (1.39356)	0.1414 (4.47081)
β_{i1}	0.7606 (19.86889)	0.7200 (32.48762)	0.7659 (5.34677)	0.6856 (5.17786)
ρ_{huni}	1.0000 (-)	-	-	-
ρ_{poli}	0.1805 (9.03695)	1.0000 (-)	-	-
ρ_{czechi}	0.1798 (7.30943)	0.1273 (2.49677)	1.0000 (-)	-
ρ_{slovi}	0.0311 (0.95129)	0.0887 (1.02263)	0.0568 (1.00170)	1.0000 (-)

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood t values in brackets.

The diagnostic statistics in Table 3.7 indicate that although the estimated model reduces coefficients of excess kurtosis compared to corresponding statistics for the unpredictable returns in Table 3.2, we still reveal significant nonnormality in the distribution of the standardised residuals of our series.

The Ljung-Box test for twelfth-order serial correlation applied to cross products of the standardised residuals, $(\varepsilon_{it} / \sqrt{h_{iit}})(\varepsilon_{jt} / \sqrt{h_{jjt}})$, fails to detect any serious misspecification of the ARCH components of the model, as both the diagonal $(i = j)$ and off diagonal $(i \neq j)$ elements of the table are considerably smaller than the corresponding statistics for the unpredictable return series in

Table 3.5. However, as it is the case with the univariate GARCH specifications in the section 3.3.3, we still reveal significant autocorrelation in the squares of the Polish standardised residuals. At 10% significance level we find significant correlation in the cross product of the Polish and Slovak standardised residuals as well.

Table 3.7 Properties of the standardised residuals of the constant correlation specification of multivariate volatility

A. Skewness and excess kurtosis

	HUN	POL	CZECH	SLOV
Skewness	-0.05482	-0.03267	-0.70082	-0.07681
Excess Kurtosis	3.16585	1.30712	4.95449	4.54218

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively.

B. Autocorrelation in the cross-products of the standardised residuals

	HUN	POL	CZECH	SLOV
HUN	12.7203	-	-	-
POL	14.7812	31.5630	-	-
CZECH	8.9541	10.0894	3.4358	-
SLOV	12.9996	20.5206	8.7811	11.8659

Note: The values in the table are the Ljung-Box(12) statistics for the autocorrelation in the cross products of the standardised residuals. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

<u>BEKK specification</u>

We started from the estimation of the multivariate BEKK specification defined in (3.71) for the series from four transition markets first setting $K = 3$ (the most general case), and then $K = 1$ (the more restricted specification). In both cases the full multivariate model failed to converge in estimation. Therefore, we decided to restrict estimation to the bivariate BEKK specification for the Hungarian and Polish return series. The choice of these two markets was justified by the highest coefficient of the conditional correlation reported above.

We estimated the bivariate BEKK, first setting $K = 3$. Then we moved to the more parsimonious specification setting $K = 1$. To check whether $K = 1$ sets undesirable restrictions on the general process, we performed a Likelihood Ratio test, which showed no evidence against the null hypothesis of the restricted model. Interpretation of the parameters from the general BEKK model is not obvious, so (given the restriction is accepted) we restrict attention to the parsimonious parametrisation.

The specification of the bivariate BEKK is as follows:

$$H_t = \begin{bmatrix} a_{01} & a_{02} \\ 0 & a_{03} \end{bmatrix}' \begin{bmatrix} a_{01} & a_{02} \\ 0 & a_{03} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}' \begin{bmatrix} \varepsilon^2_{1t-1} & \varepsilon_{1t-1}\varepsilon_{2t-1} \\ \varepsilon_{2t-1}\varepsilon_{1t-1} & \varepsilon^2_{2t-1} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}' \begin{bmatrix} h_{11t-1} & h_{12t-1} \\ h_{21t-1} & h_{22t-1} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \tag{3.72}$$

The off-diagonal parameters in matrix $\boldsymbol{B}$, given by β_{12} and β_{21}, measure the dependence of the conditional return volatility in the Hungarian market on that of the Polish market, and the dependence of the conditional return volatility in the Polish market on that of the Hungarian market in the last period. The absolute size of return shocks originating in one market in previous period, measured by squared value of lagged unpredictable returns, also transmits to the current period's conditional volatility in the other market by means of the off-diagonal elements of matrix $\boldsymbol{A}$, given by parameters a_{12} and a_{21}.

Equations (3.73), (3.74) and (3.75) below solve for the cross effects in the variance equations implied by the BEKK specification.

$$h_{11t} = a_{01}^2 + a_{11}^2\varepsilon_{1t-1}^2 + 2a_{11}a_{21}\varepsilon_{1t-1}\varepsilon_{2t-1} + a_{21}^2\varepsilon_{2t-1}^2 + \beta_{11}^2 h_{11t-1} + 2\beta_{11}\beta_{21}h_{12t-1} + \beta_{21}^2 h_{22t-1} \tag{3.73}$$

$$\begin{aligned} h_{12t} = {} & a_{01}a_{02} + a_{11}a_{12}\varepsilon_{1t-1}^2 + (a_{21}a_{12} + a_{11}a_{22})\varepsilon_{1t-1}\varepsilon_{2t-1} + a_{21}a_{22}\varepsilon_{2t-1}^2 + \\ & + \beta_{11}\beta_{12}h_{11t-1} + (\beta_{21}\beta_{12} + \beta_{11}\beta_{22})h_{12t-1} + \beta_{21}\beta_{22}h_{22t-1} \end{aligned} \tag{3.74}$$

$$h_{22t} = (a_{02}^2 + a_{03}^2) + a_{12}^2\varepsilon_{1t-1}^2 + 2a_{12}a_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + a_{22}^2\varepsilon_{2t-1}^2 + \beta_{12}^2 h_{11t-1} + 2\beta_{12}\beta_{22}h_{12t-1} + \beta_{22}^2 h_{22t-1} \tag{3.75}$$

Table 3.8 shows the results of fitting the BEKK model to the Hungarian (ε_{1t}) and Polish (ε_{2t}) unpredictable return series. Estimates of the model indicate that volatility in each of the markets is affected by events in its own market. However, while volatility in the Polish stock market is affected by return volatility and return shocks originating in the Hungarian market, volatility and shocks originating in the Polish market do not affect volatility in the Hungarian market. The covariance between two return series is significantly affected by the lagged cross products of the return shocks, the lagged covarince term, as well as the return shocks, originating in the Hungarian market. The shocks and volatility from the Polish market do not have any significant impact on the covariance between the series.

Table 3.8 Results of the estimation of the BEKK specification of the conditional volatility for the Hungarian and Polish series

$$h_{11t} = \underset{(2.6298)}{1.57e-05} + \underset{(3.7362)}{0.1756\varepsilon^2_{1t-1}} + \underset{(0.2278)}{0.0151\varepsilon_{1t-1}\varepsilon_{2t-1}} + \underset{(0.0815)}{8.96e-04\varepsilon^2_{2t-1}} + \underset{(10.9930)}{0.7528h_{11t-1}} + \underset{(0.2350)}{0.0161h_{12t-1}} + \underset{(0.4087)}{8.59e-05h_{22t-1}}$$

$$h_{12t} = \underset{(0.6207)}{1.80e-05} + \underset{(1.919403)}{0.0308\varepsilon^2_{1t-1}} + \underset{(2.5901)}{0.1574\varepsilon_{1t-1}\varepsilon_{2t-1}} + \underset{(0.8831)}{0.0111\varepsilon^2_{2t-1}} + \underset{(1.3235)}{0.0110h_{11t-1}} + \underset{(3.0601)}{0.7768h_{12t-1}} + \underset{(0.2748)}{0.0083h_{22t-1}}$$

$$h_{22t} = \underset{(0.9640)}{2.06e-05} + \underset{(2.4909)}{0.0540\varepsilon^2_{1t-1}} + \underset{(1.6194)}{0.0264\varepsilon_{1t-1}\varepsilon_{2t-1}} + \underset{(5.4880)}{0.1372\varepsilon^2_{2t-1}} + \underset{(2.0146)}{0.0005h_{11t-1}} + \underset{(0.6219)}{0.0227h_{12t-1}} + \underset{(22.2240)}{0.8014h_{22t-1}}$$

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood t-values in brackets.

Table 3.9 presents diagnostic statistics for the standardised residuals of the BEKK specification of conditional volatility. Though the estimated model reduces the coefficients of excess kurtosis present in the unpredictable Hungarian and Polish return series (see Table 3.2 for comparison), we still reveal significant nonnormality in the distribution of standardised residuals. This evidence against normality warrants the use of quasi maximum likelihood inferential procedures in our analysis.

The Ljung-Box test statistics in the Table 3.9.B indicate that although the BEKK specification is able to capture volatility clustering present in the Hungarian returns and the cross product of the

Hungarian and Polish series, the intertemporal dependence in the squares of the Polish returns, while being reduced considerably, is not removed completely (see Table 3.5 for comparison).

Table 3.9 Properties of the standardised residuals of the BEKK specification of multivariate volatility

A. Skewness and excess kurtosis

	HUN	POL
Skewness	-0.09727	-0.07687
Excess kurtosis	2.95643	1.57439

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively.

B. Autocorrelation in the cross-products of the standardised residuals

	HUN	POL
HUN	11.6052	-
POL	9.9339	36.5184

Note: The values in the table are the Ljung-Box(12) statistics for the autocorrelation in the cross products of the standardised residuals. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

3.4 Summary

The class of Generalised Autoregressive Conditional Heteroskedastic (GARCH) models has proven to be particularly suited for modelling the behaviour of financial time series. They are capable of capturing the most common empirical observations in daily return data, such as leptokurtosis and volatility clustering. We employ both univariate and multivariate GARCH models of volatility.

In the univariate part of the chapter we specify and contrast several models of time-varying volatility. Five out of the seven estimated models allow different types of asymmetry in the impact of news on volatility. The models are fitted to daily stock returns from four emerging markets of Central Europe. The quasi-maximum estimates indicate that strong GARCH effects are apparent in almost all markets. The examination of asymmetric impact of news on volatility shows rather weak evidence of asymmetries in the markets. The so-called leverage effects, which indicate that negative shocks introduce more volatility than positive shocks, are captured by EGARCH for the Czech and AGARCH for the Slovak returns.

The multivariate part employs two alternative specifications of the parameterisation of the conditional covariance matrix of the series. The results from the constant correlation model indicate significant conditional correlations between three pairs of countries: Hungary and Poland, Hungary and Czech Republic, and Poland and Czech Republic. The estimation of the bivariate BEKK model of multivariate volatility for the Hungarian and Polish returns reveals that while the volatility in the Polish stock market is affected by return volatility and return shocks originating in the Hungarian market, volatility and shocks originating in Polish market do not affect volatility in the Hungarian market.

Examination of the distributional properties of the standardised residuals of the model specifications used in this chapter reveals that, although there is a substantial reduction in the excess kurtosis observed in the return series, nonnormality remains. The direction of subsequent future research includes alternative explanations of the conditional nonnormality. The use of a fat-tailed conditional distribution might be more appropriate. Another possibility is the stochastic volatility specification, which has recently become popular.[48]

Diagnostics for the Polish series indicate that volatility clustering present in the returns, while being reduced considerably, is not removed completely by any of the alternative ARCH specifications employed in this chapter. Our further research will include the search for the alternative specifications of the conditional volatility providing better fit to the series from the Polish stock

---

[48] See Apendix 3.1 for the description of the stochastic volatility model as an alternative to the GARCH specification of time-varying volatility. For the survey of stochastic volatility models see Ghysels et al. (1996).

market.[49]

The ARCH models of time-varying volatility could also be embedded into the tests of explicit financial models.[50]

[49] The results of the diagnostic tests presented in this chapter, indicating that the employed specifications do not provide a perfect fit to our return data, should be taken into account in the next chapter of this thesis, which employs some of the univariate GARCH specifications used here for simulation and forecasting purposes.

[50] See Chapter 5 for precisely such a test of the international CAPM, with the conditional covariance matrix of the assets in the system parameterised as BEKK multivariate GARCH specification.

Figure 3.1.A The Hungarian return series

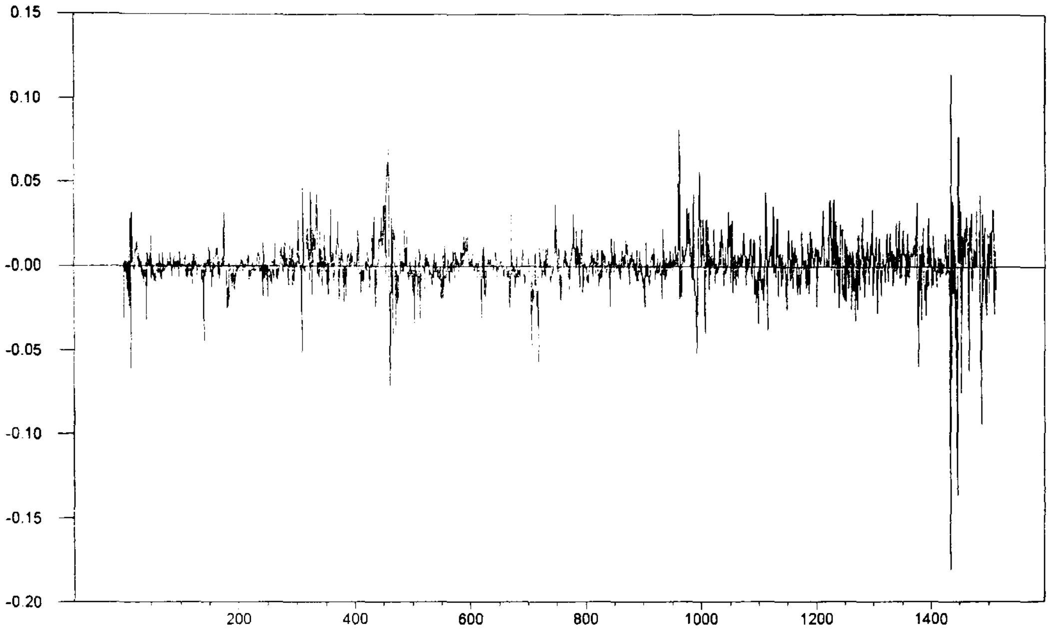


Figure 3.1.B The unpredictable Hungarian return series

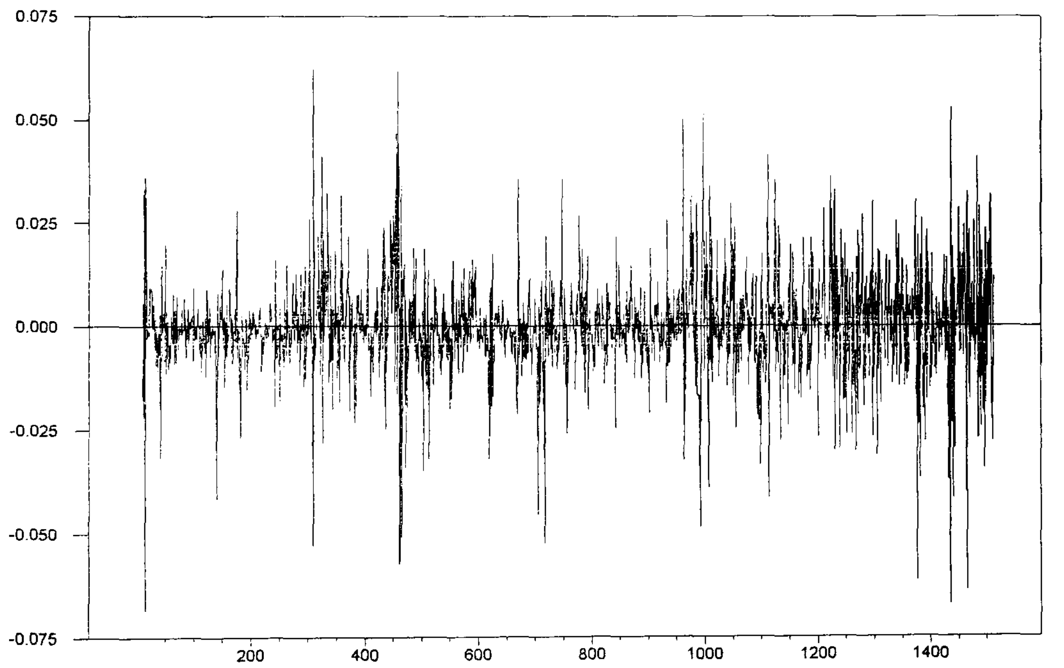


Figure 3.2.A The Polish return series

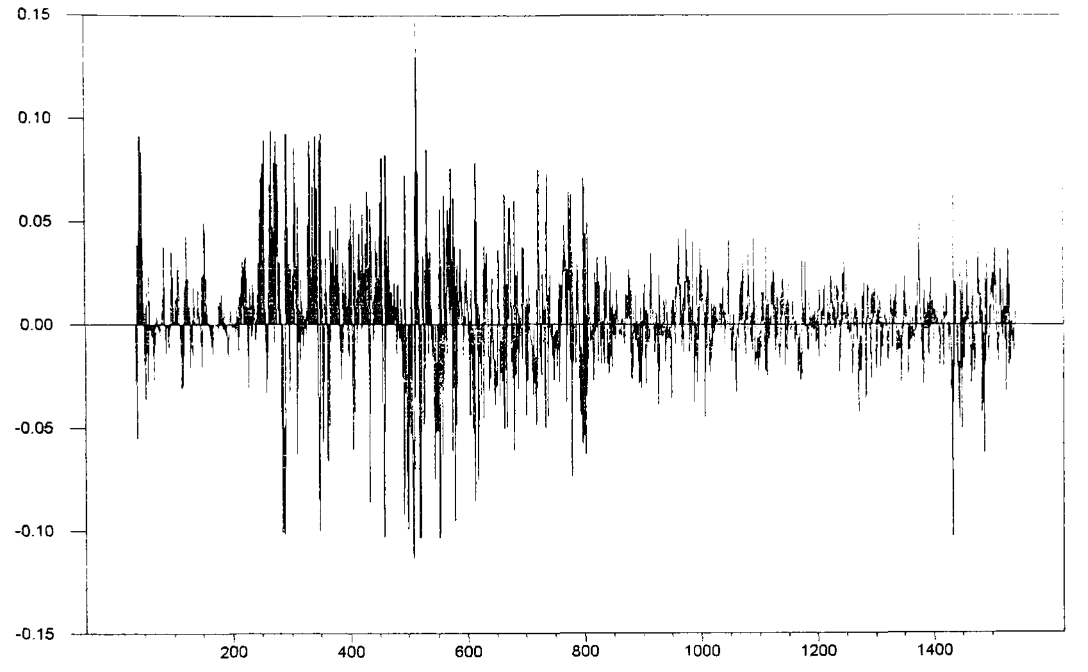


Figure 3.2.B The unpredictable Polish return series

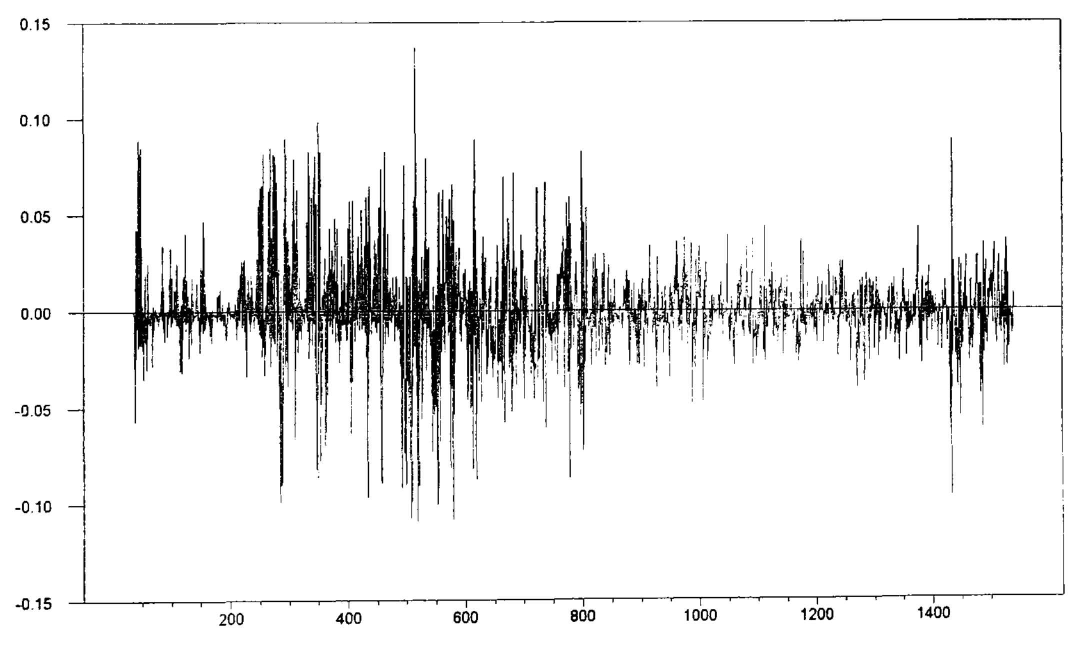


Figure 3.3.A The Czech return series

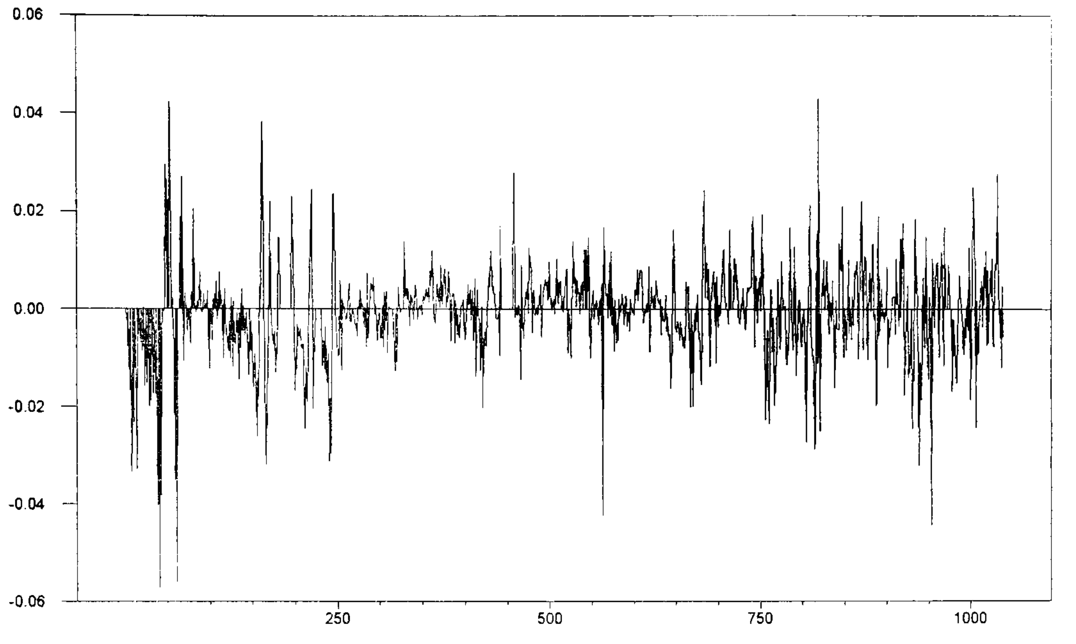


Figure 3.3.B The unpredictable Czech return series

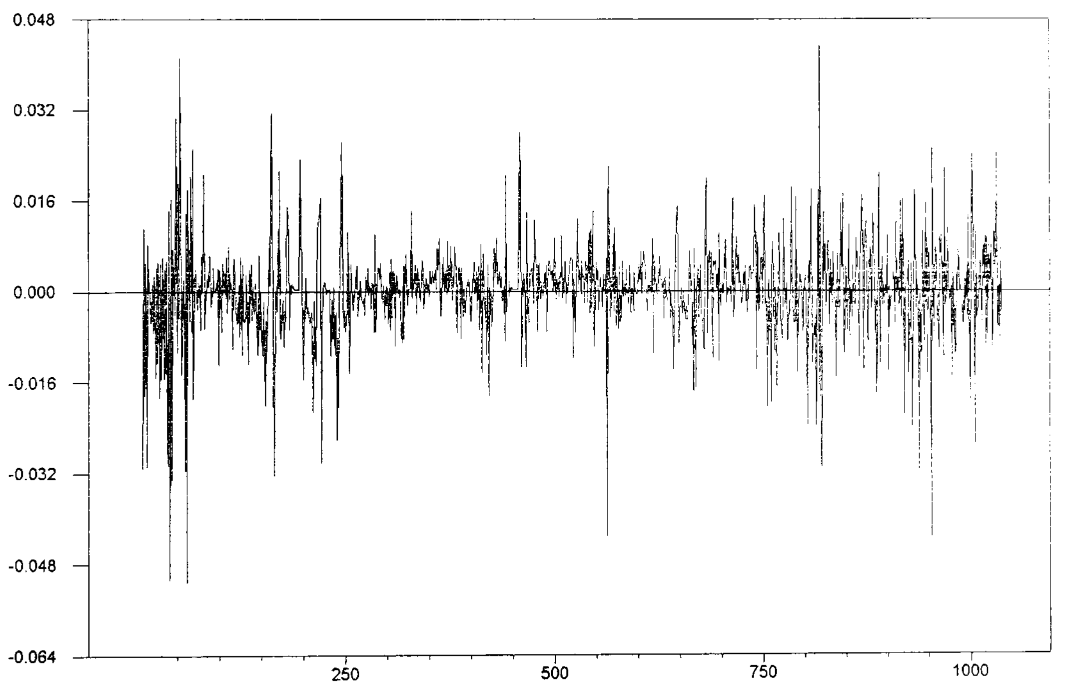


Figure 3.4.A The Slovak return series

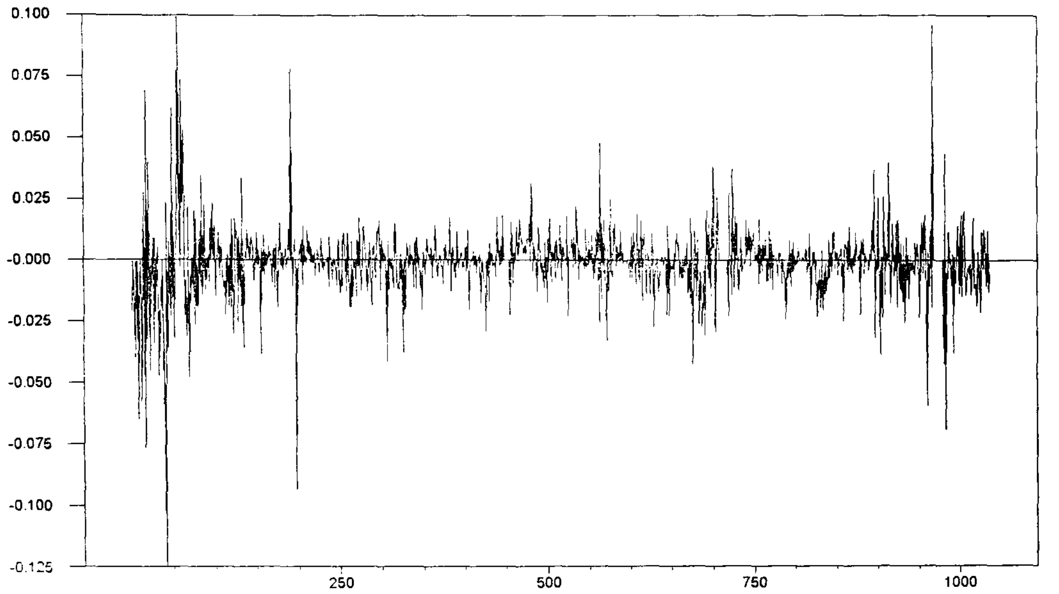
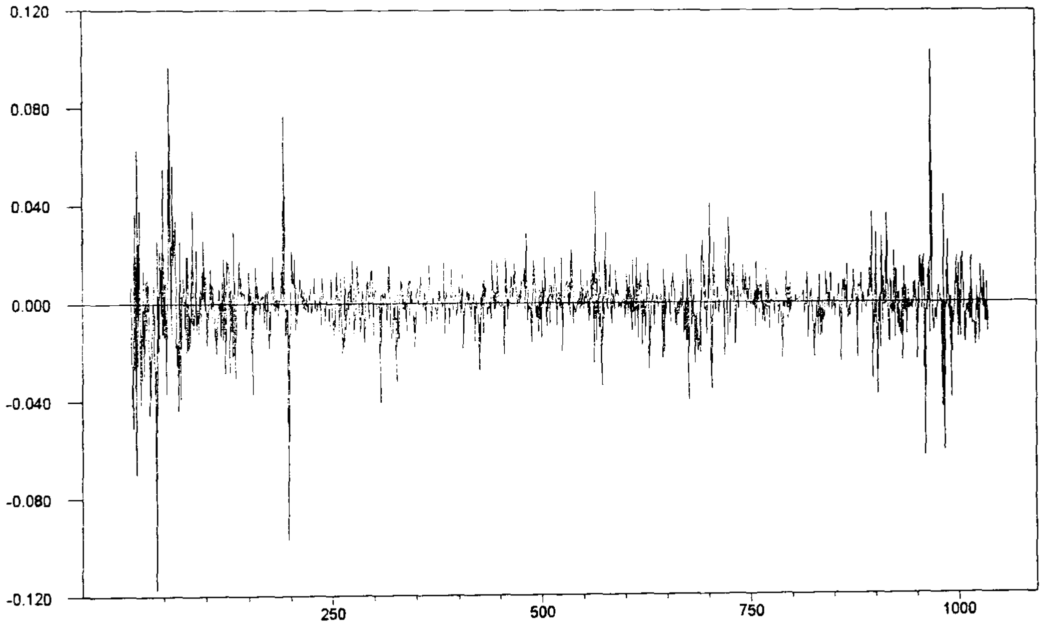


Figure 3.4.B The unpredictable Slovak return series



APPENDIX 3.1

Alternative models of conditional volatility

In this section we consider some of the measures of volatility used in the literature, which are not based on ARCH type specifications.

When deciding on the form of the specification for the conditional variance one has to define the conditioning set of information and to select a functional form for the mapping between the conditioning set and conditional variance. Usually, the conditioning set is restricted to include past values of the series itself. A simple two step estimator of the conditional residual variance can be obtained from a regression of the square residuals against their own lagged values (see Davidian and Carol (1987)). Pagan and Schwert (1990) show that the OLS estimator is consistent although not efficient. This two-step estimator's role is that of a benchmark which can be computed in a straightforward way.

Jump or mixture models possibly combined with a GARCH specification for the conditional variance have been used to describe time-variation in volatility measures, fat-tails and skewness of financial series. In the Poisson jump model it is assumed that upon arrival of abnormal information a jump occurs in the returns. The number of jumps occurring at time t, n_t, is generated by a Poisson distribution. Conditionally on the number of jumps n_t, returns are normally distributed with mean $n_t\theta$ and variance $\sigma_t^2 = \sigma_\xi^2 + n_t\sigma_\psi^2$. The parameter θ denotes the expected jump size. The conditional mean and variance of returns depend on the number of jumps at period t. Time dependency could be introduced by assuming that σ_ξ^2 follows a GARCH-type process. [51]

[51] For the examples of the modelling of stochastic jumps by means of a Poisson process see e.g. Ball and Torous

The alternative way to introduce time dependence is to assume that the probabilities of being in state 1 during period t differ, depending on whether the economy was in state 1 or state 2 in period t-1. This model was first introduced by Hamilton (1989) and applied to exchange rates by Engle and Hamilton (1990), interest rates by Hamilton (1988) and stock returns by Pagan and Schwert (1990). According to the model an unobserved state variable z_t can take the values 0 and 1. The transition probabilities from state j in period t-1 to state i in period t, P_{ij}, are constant and given by the first-order Markov process:

$$P_{11} = prob[z_t = 1 \mid z_{t-1} = 1] = p$$

$$P_{10} = prob[z_t = 0 \mid z_{t-1} = 1] = 1 - p$$

$$P_{00} = prob[z_t = 0 \mid z_{t-1} = 0] = q$$

$$P_{01} = prob[z_t = 1 \mid z_{t-1} = 0] = 1 - q \tag{3.76}$$

In Pagan (1996), the state variable, z_t, evolves as an AR (1) process. Returns, y_t, are assumed to be generated by

$$y_t = \beta_0 + \beta_1 z_t + (\sigma^2 + \phi z_t)^{\frac{1}{2}} \varepsilon_t, \tag{3.77}$$

where $\varepsilon_t \sim NID(0, \sigma^2)$. From (3.77) the expected values of y_t in states 1 and 2 are β_0 and $\beta_0 + \beta_1$ respectively. The variances are σ^2 and $\sigma^2 + \phi$. As a result the model generates states with high and low volatility. The variance of returns conditional on state in the period t-1 is:

$$Var(y_t \mid z_{t-1}) = [\sigma^2 + (1-q)\phi](1 - z_{t-1}) + [p\phi + \sigma^2] z_{t-1}. \tag{3.78}$$

Hamilton and Susmel (1994) generalise the Markov regime switching model by allowing it to be ARCH. Their model is called switching regime ARCH model (SWARCH). As in Hamilton's basic

---

(1985), Hsien (1989), Nieuwland et al. (1991) and Ball and Roma (1993).

model above, y_t depends linearly on the state variable z_t. The disturbance term of y_t follows an autoregressive process of order p with an error $u_t = \sqrt{g_{st}}\tilde{u}_t$, where $\tilde{u}_t$ follows an ARCH (q) process with leverage effects and g_{st} is constant factor which differs across regimes. The innovation $\tilde{u}_t$ is assumed to have a conditional student t-distribution with mean zero. Transitions between regime are governed by an unobserved Markov chain. Using weekly returns data on the value-weighted portfolio of stocks traded on the New York Stock Exchange from 1962 to 1987, the authors compare various ARCH models to SWARCH models allowing for up to four regimes. The SWARCH specification is found to perform best.

Next we consider non-parametric methods for modelling the conditional variance, i.e. methods which do not depend on specific assumptions about the functional form. Pagan and Schwert (1990) and Pagan and Hong (1991) use a non-parametric kernel estimator and non-parametric flexible Fourier form estimator. The kernel estimator of y_t, $g(y_t)$, with a finite number of conditioning variables x_t looks as follows:

$$\hat{E}[g(y_t)\,|\,x_t] = \sum_{s=1}^{T} g(y_s)K(x_t - x_s) / \sum_{s=1}^{T} K(x_t - x_s), \tag{3.79}$$

where K is a kernel function which smoothes the data. One of the popular kernels used in the literature is the normal kernel employed by Pagan and Schwert (1990):

$$K(x_t - x_s) = (2\pi)^{-1/2}\,|H|^{-1/2}\exp[-\frac{1}{2}(x_t - x_s)'H(x_t - x_s)], \tag{3.80}$$

where H is a diagonal matrix with k^{th} diagonal element set equal to the bandwidth $\hat{\sigma}_k T^{-1/(4+q)}$, with $\hat{\sigma}_k$ being the standard deviation of x_{kt}, $k = 1,\ldots q$, and q being the dimension of the conditioning set.

Next non-parametric estimator, the Flexible Fourier Form (FFF) first proposed by Gallant (1981), involves a global approximation of the conditional variance using series expansion. The specification of the conditional variance, σ_t^2, is given by

$$\sigma_t^2 = \sigma^2 + \sum_{j=1}^{L} \{\alpha_j \hat{e}_{t-j} + \beta_j \hat{e}_t^2 + \sum_{k=1}^{2} [\phi_{jk} \cos(k\hat{e}_{t-j}) + \delta_{jk} \sin(k\hat{e}_{t-j})]\}. \qquad (3.81)$$

As a result, σ_t^2 is represented as the sum of low-order polynomial and trigonometric terms constructed from past $\hat{e}_t$ s, which are the residuals from a regression for y_t. The disadvantage of the specification in (3.81) is the possibility that the estimates of σ_t^2 can be negative. This specification for $L=1$ has been applied to stock returns by Pagan and Schwert (1990). They found that the estimates of σ_t^2 are similar for the kernel, GARCH (1,2) and FFF specification across most of the range of $\hat{e}_{t-1}$. Only for large positive and negative values of $\hat{e}_{t-1}$ the estimators exhibit a different behaviour. In FFF case for large negative values of $\hat{e}_{t-1}$, the volatility estimates increase dramatically.

An alternative to using the ARCH framework is to assume the changing variance to follow some latent process. This leads to a stochastic volatility (SV) model.[52] A simple SV model for returns y_t has been proposed by Taylor (1986):

$$y_t = \varepsilon_t \exp(\alpha_t / 2), \quad \varepsilon_t \sim NID(0,1), \qquad (3.82)$$

$$\alpha_{t+1} = \alpha_0 + \phi\alpha_t + \eta_t, \quad \eta_t \sim NID(0,\sigma_\eta^2), \qquad (3.83)$$

where the random variables ε_t and η_t are independent.

[52] See e.g. Ghysels et al. (1996).

This model has been used by Hull and White (1987), for instance, in pricing foreign currency options. Its time series properties are discussed by Taylor (1986, 1994). Taylor (1994) denotes the SV models as autoregressive variance (ARV) models. A major difficulty arises with the estimation of SV models which are nonlinear and not conditionally Gaussian. Many estimation methods such as the generalised method of moments (GMM) or quasi-maximum likelihood method (QML) are inefficient. But methods relying on simulation-based techniques make it possible to perform Bayesian estimation or classical likelihood analysis.[53] Currently, only few studies compare the performance of the GARCH and SV approaches to modelling volatility. Ruiz (1993) compares the GARCH(1,1), EGARCH(1,0) and ARV(1) models when applied to daily exchange rates from 1981 to 1985 for the Pound sterling, Deutsche mark, Yen and Swiss franc vis-à-vis the U.S. dollar. Within sample performance of the three models is very similar. When the models are used to forecast out-of-sample volatility, the ARCH models exhibit severe biases which do not occur for the SV volatilities. For the daily and weekly returns on the S&P500 index over the period from 1962 to 1987 and 1962 to 1992 respectively, Kim and Shephard (1994) conclude that a simple first order SV model fits the data as well as the popular ARCH models. For daily data on the S&P500 index for the years 1980 to 1987, Dannielson (1994) finds that the EGARCH(2,1) model performs better than ARCH(5), GARCH(1,2), IGARCH(1,1,0) models. It also outperforms a simple SV model estimated by simulated maximum likelihood.

[53] See e.g. Kim and Shephard (1994).

CHAPTER 4

PREDICTABILITY IN THE TRANSITION MARKETS OF CENTRAL EUROPE

4.1 Introduction

One of the earliest and most enduring questions of financial econometrics is whether financial assets are forecastable. Perhaps because of the obvious analogy between financial investments and games of chance, mathematical models of asset prices have an unusually rich history that predates virtually every other aspect of economic analysis. Predictability of stock returns has always fascinated practitioners and academics.

Until the early seventies, the assumption of a risk-neutral world (the world is populated by risk neutral agents, all of whom have common and constant time preferences and common beliefs about future states of nature), prevailing in the literature, had very strong implication for financial economics: any predictability in stock returns would necessarily imply that the stock market is informationally inefficient. By the late seventies, the work of LeRoy (1973) and Lucas (1978) had demonstrated the critical role played by risk preferences in the martingale behaviour of stock prices in efficient market.[54] And today most academics realise that predictability is not immediately

[54] See also Hirshleifer (1975).

synonymous with market inefficiencies, because in a risk-averse world rational time-varying risk premia could lead to return-predictability. Nevertheless, one cannot *a priori* rule out the possibility that predictability arises due to irrational mispricing of the securities by agents. There is no evidence so far that unambiguously distinguishes these two competing hypotheses.

Although there is considerable disagreement over the source of predictability, most members of the finance profession seem convinced that predictability of returns documented in the empirical literature is real. Evidence of predictability of stock market returns led researchers to investigate the sources of this predictability.

The literature on financial forecasting documented evidence on the predictability of current returns from past returns. There is also evidence for systematic patterns in stock returns related to various calendar periods such as the weekend effect, the turn-of-the-month effect, the holiday effect, the January effect and predictability originating from bid-ask bounces. The literature also documented significant relationship between expected returns and fundamental variables, such as dividend yields, price-earnings ratios, term structure variables, macroeconomic fundamentals, etc.

Lo and MacKinlay (1988) find that weekly returns on the portfolios of NYSE stocks grouped according to the size show positive autocorrelation. Conrad and Kaul (1988) examine the autocorrelation of Wednesday to Wednesday returns (to mitigate the *nonsynchronous trading* problem[55]) for size-grouped portfolios of stocks that trade on both Wednesdays. Similarly to the findings of Lo and MacKinlay (1988), they find that weekly returns are positively autocorrelated. Cutler et al. (1991) present results from many different asset markets generally supporting the

[55] The *nonsynchronous trading* or *nontrading* effect arises when time series are taken to be recorded at time intervals of one length, when in fact they are recorded at time intervals of other, possibly irregular, lengths. This can create a false impression of predictability in price changes and returns even if true price changes or returns are statistically independent. Perhaps the first to recognise the importance of nonsynchronous prices was Fisher (1966). For a good discussion on this topic, see e.g. Chapter 3 in Campbell, Lo and MacKinlay (1997).

hypothesis that returns are positively correlated at the horizon of several months and negatively correlated at the 3-5 year horizon. Lo and MacKinlay (1990) report positive serial correlation in weekly returns for indices and portfolios and negative serial correlation for individual stocks. Chopra et al. (1992), DeBondt and Thaler (1985), Fama and French (1988) and Poterba and Summers (1988) find negative serial correlation in returns of individual stocks and various portfolios over three to ten year intervals.[56] Jegadeesh (1990) finds negative serial correlation for lags up to two months and positive correlation for longer lags. Lehmann (1990) and French and Roll (1986) report negative serial correlation at the level of individual securities for weekly and daily returns.

In a seminal contribution to the predictability literature, Fama and Schwert (1977) use Treasury bill rates to predict stock and bond returns. Over the last two decades several new fundamental variables have been used to predict stock returns. For example, Campbell and Shiller (1988), Cutler, Poterba , and Summers (1991), and Fama and French (1988, 1989), Flood, Hodrick, and Kaplan (1987), and Keim and Stambaugh (1986), among others, use financial variables such as dividend yield, price-earnings ratios, term structure variables, etc., to predict future stock returns. In a similar vein, Balvers, Cosimano, and McDonald (1990), Chen (1991), Fama (1990), and Schwert (1990) have used macroeconomic fundamentals, such as output and inflation, to predict stock returns. Some recent papers by Ferson and Harvey (1991), Evans (1994), and Ferson and Korajczyk (1995) focus on the relation between predictability based on lagged variables and economic 'factors' similar to those identified by Chen, Roll, and Ross (1986). Ferson and Schadt

[56] One interpretation of negative correlation in longer horizons is that there can be substantial mean-reversion in stock market prices at longer horizons. On the other hand, the obvious concern is that inferences based on long horizon returns are based on extremely small sample size, which makes the mean-reversion hypothesis fragile.

(1995) show that conditioning on predetermined public information removes biases in commonly used unconditional measures of the performance of the mutual fund managers.[57]

In this chapter we address the issues of predictability in the transition markets of Central Europe by employing technical analysis for deciding when to buy and sell in the market. Technical analysis is considered by many to be the original form of investment analysis.[58] Over the years, investors and the technical analysts have devised hundreds of technical market indicators in an effort to forecast the trend of the security market. Technical analysis came into widespread use before the period of extensive and fully disclosed financial information, which in turn enabled the practice of fundamental analysis to develop. Technical traders base their analysis on the premise that the patterns in stock prices are assumed to recur in the future, and thus, these patterns can be used for predictive purposes. The motivation behind the technical analysis is to be able to identify changes in trends at an early stage and to maintain an investment strategy until the weight of the evidence indicates that the trend has reversed.

Historically, technical analysis has never enjoyed the same degree of acceptance by the academic finance community that, for example, fundamental analysis has received. Although some earlier

---

[57] Standard measures of the performance of managed portfolios, designed to detect security selection or market timing ability, are known to suffer from a number of biases. Traditional measures use unconditional expected returns as the baseline. For example, the "*Jensen's alpha*", measuring the average performance of a fund, may be calculated as the average return of a fund in excess of a risk-free rate minus a fixed beta times the average excess return of a benchmark portfolio. However, if expected returns and risks vary over time, such an unconditional approach is likely to be unreliable. Common time variation in risks and returns will be confused with average performance. Traditional measures of average performance are negative more often than positive, which has been interpreted as inferior performance. Ferson and Schadt show that conditioning on public information shifts the distribution of alphas to the right and is centered near zero. Traditional measures of market timing suggest that the typical mutual fund takes more market exposure when stock returns are low. This has been interpreted as perverse market timing ability. Ferson and Schadt find that unconditional versions of the Treynor and Mazuy (1966) and Merton and Henriksson (1981) market timing models are misspecified when applied to naive strategies, and that conditional versions of these models are an improvement. Using the conditional market timing models for U.S. equity funds, the evidence of perverse market timing for the typical fund is removed.

[58] The oldest moving average techniques are attributed to Charles Dow and are traced to the late 1800s.

studies argued that technical analysis is useless[59], the recent literature provides evidence that technical trading rules devised by investors and technical analysts may provide positive profits after accounting for transaction costs.[60] Recent studies by Blume, Easley and O'Hara (1994), Brock, Lakonishok and LeBaron (1992), Brown and Jennings (1989), LeBaron (1996), Neftci (1991), Pau (1991), Taylor and Allen (1992), and Treynor and Ferguson (1985) signal a growing interest in technical analysis among financial academics.

This chapter has three main objectives. First, we assess the ability of the technical trading rules, applied to the data from the four transition markets of Central Europe to predict stock price changes. Next, we employ the bootstrap methodology to check whether patterns uncovered by technical rules are consistent with popular null models of stock returns. And finally, we examine the linear and nonlinear predictability of stock market returns with buy and sell signals generated from technical trading rules.

4.2. Data and Descriptive Statistics

The data used in this chapter is identical to the data we have employed in the previous chapter of the thesis. It includes daily stock price indexes from four transition stock markets: Hungarian BUX, Polish WIG20, Czech PX50 and Slovak SAX. The only difference is that in this chapter for all markets the sample period ends in July 1998, and, therefore, compared to the previous chapter

[59] The earlier work on the technical analysis includes papers by Alexander (1964), Fama and Blume (1966), Levy (1967a, 1967b), Jensen (1967), and Jensen and Bennington (1970).
[60] See e.g. LeBaron (1991). He tests the "economic significance" of the trading rule profits in the foreign exchange market, by accounting for transaction costs and interest rates.

covers a slightly longer period of time. The sample constitutes a total of 1631 observations for the Hungarian and Polish, and 1155 observations for the Czech and Slovak series.[61]

Since the focus of the previous chapter was on conditional volatility, rather than conditional mean, we have presented descriptive statistics of the unpredictable return series. In table 4.1 below we report distributional characteristics of the nominal returns data.

The evidence for all return series indicates significantly fatter tails than does the stationary normal distribution. The excess kurtosis statistic ranges in value from 4.59 for the Polish to 20.39 for the Hungarian returns. The coefficients of skewness, with exception of the Polish returns, indicate that series typically have asymmetric distributions skewed to the left. The autocorrelations coefficients, $\rho(n)$, reported together with the Bartlett standard errors, as well as Ljung-Box test statistic for twelfth-order serial correlation, indicate evidence of significant serial correlation in our sample.

[61] For the graphs of the historical development of the index levels and the corresponding index return series over the sample period employed here, see Figures 4.2.1 and 4.2.2 at the end of the chapter.

Table 4.1 Summary statistics of the returns on four national indexes

	Hungary	**Poland**	**Czech Rep.**	**Slovakia**
Number of observ.	1631	1631	1155	1155
Mean	0.00132	0.00188	-0.00058	-0.00069
Variance	0.00025	0.00064	0.00009	0.00025
Skewness	-1.28739	-0.05119	-0.40777	-0.83615
Excess kurtosis	20.38698	4.58512	4.98394	12.21029
$\rho(1)$	0.15144	0.19694	0.33953	-0.16404
$\rho(2)$	0.06062	0.12631	0.20848	-0.05474
$\rho(3)$	-0.00501	0.04154	0.09507	0.12986
$\rho(4)$	-0.02299	0.03688	-0.01129	-0.00319
$\rho(5)$	-0.01253	-0.01483	-0.01357	0.03402
$\rho(6)$	-0.03897	-0.00680	-0.07704	-0.02515
$\rho(7)$	0.04363	0.02800	-0.05570	0.00847
$\rho(8)$	0.03098	0.04932	0.01443	0.01812
$\rho(9)$	0.05425	0.03826	0.06694	0.03224
$\rho(10)$	0.10135	0.03749	0.08701	0.00775
$\rho(11)$	0.06809	-0.00432	0.09538	0.01172
$\rho(12)$	0.11207	0.01182	0.01193	0.01430
Bartlett st. error	0.02477	0.02477	0.02988	0.02988
Ljung-Box(12) for the levels	101.8537	105.1288	215.4553	49.6788

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively. The Bartlett standard error for the autocorrelation equals $1 / \sqrt{N}$. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively.

4.3 Empirical Analysis: Technical Analysis in the Transition Equity Markets (the moving average strategy)

Technical traders test historical data to establish specific rules for buying and selling securities in the market. In this chapter we explore one of the most popular technical rule when to buy and sell in the market: the moving average rule.[62] Buy and sell signals are generated by two moving

[62] Our selection of trading strategies is influenced by previous work in this area, which has a very long history. The use of moving averages was discussed by Gartley (1930). The examples of early use of moving average techniques have been collected in Coslow et al. (1966).

averages, a long period and a short period. Buy signal is generated when the short moving average rises above the long moving average and the sell signal is generated otherwise. The underlying notion behind this rule is that it provides a means of determining the general direction or trend of the market by examining the recent history.

An n-period moving average (m_t) is computed by adding the n most recent periods of data, then dividing by n.

$$m_t^n = (1/n)\sum_{i=0}^{n-1} p_{t-i} \tag{4.1}$$

This average is recalculated each period by dropping the oldest data and adding the most recent, so the average moves with its data but does not fluctuate as much. The larger the n, the smoother is the moving average rule, and it measures a longer-term trend.

In our analysis we will use four popular moving average rules: 1-150, 1-200, 5-150, 2-200[63], with a band between the short and the long averages equal 0.01, which reduces the number of buy/sell signals by eliminating *whiplash* signals when the short and the long moving averages are close. As a result, the trading rules classify each day in our sample as either buy, sell, or neutral.

If we denote the short moving average ns, and the long one nl, then, taking into account a band of 0.01 mentioned above, an investor gets a buy signal:

$$s_t^{b;ns,nl} = \ln(m_t^{ns}) - \ln(m_t^{nl}), \qquad \text{if } \ln(m_t^{ns}) - \ln(m_t^{nl}) > 0.01 \tag{4.2}$$

and a sell signal:

$$s_t^{s;ns,nl} = \ln(m_t^{ns}) - \ln(m_t^{nl}), \qquad \text{if } \ln(m_t^{ns}) - \ln(m_t^{nl}) < -0.01 \tag{4.3}$$

[63] For the examples of application of these rules to financial data see e.g. LeBaron (1991), Brock et al. (1992), Kim (1994), Karolyi and Kho (1994).

Otherwise the trading day is classified as neutral and $s_t^{n;ns,nl} = 0$.

This rule is designed to replicate returns from trading rule where the trader buys when the short moving average penetrates the long from below and stays in the market until the short moving average penetrates the long moving average from above. In the last case the trader moves out of the market or sells short.[64]

4.3.1 Empirical results

Tables 4.2 A, B, C and D present results from trading strategies based on four moving average rules for the four markets examined. The table reports number of buy and sell signals generated; mean returns during buy and sell periods and corresponding variances of buy and sell returns; the fraction of buy and sell returns greater than zero. The last column lists the differences between the mean daily buy and sell returns. Numbers in parentheses are *t*-statistics testing the difference of the mean buy and sell returns from the unconditional mean, and the mean buys and sells difference from zero. The last rows in the tables report averages over all four rules.

Table 4.2.A reports results for the Hungarian market. For all trading rules the number of buy signals generated is considerably higher than the number of sell signals, which is consistent with upward-trending market (see Figure 4.2.1 at the end of this chapter). All buy returns are positive with an average daily return of 0.28 per cent, compared with the unconditional mean of 0.13 per cent (see Table 4.1). The *t*-statistics testing the difference of the mean buy returns from the

[64] This rule assumes that the market is dominated by positive feedback traders, who buy 'high' and sell 'low'.

unconditional mean reject the null of their equality at the 5 per cent level for the first two trading rules, with the *t*-statistics for the third and fourth rule being marginally significant as well. In contrast to buys all sell returns are negative, with *t*-statistics for the difference from unconditional mean being significant in three out of four cases. The last column of the table shows the spreads between the mean buy and sell returns. All of them are positive with highly significant *t*-statistics, rejecting the null of equality with zero.

Table 4.2.A indicates that not only do the buy signals select periods with higher conditional means, they also pick periods with lower volatilities, which is in contrast to sell periods where while the conditional return is lower, the volatility is higher. Our results are not in accord with the existing equilibrium models. The columns seven and eight of the table present the fraction of the buy and sell returns which are positive. In the case where trading rules do not produce useful signals the fraction of positive buys and positive sells should be the same. However, it is clear from the table that the fraction of positive buys is always higher, constituting on average 56.5 per cent as compared to 45.2 per cent for sells.

Moving to the next three tables presenting results from application of technical trading rules to Polish, Czech and Slovak returns, we discover patterns similar to Hungarian series. For all series buy returns are always positive, with averages over four rules ranging from 0.08 per cent in Czech Republic to 0.37 per cent in Poland. In contrast sell returns are always negative: -0.16, -0.08 and -0.11 per cent in Poland, Czech Republic and Slovakia respectively. And again in all three countries the volatility of buy returns, which have higher conditional means, is lower compared to the volatility of sells, having lower (and always negative) mean returns. The spreads between mean buy and sell returns are positive for all three countries. And finally, as we have found in the case of Hungarian returns, the fraction of buys greater than zero is always higher than the fraction of positive sells for all rules for all markets.

The only difference in the results for four markets we can report from the tables is the relative number of buy and sell signals. While Hungarian and Polish series produce considerably higher number of buy signals, the quantity of buy and sell signals generated for Czech series is almost the same, while Slovakian returns produce more sells than buys. The different trends in the markets are the most obvious explanation to the observed differences. We can see from Figure 4.2.1 at the end of this chapter that while over considered sample period the Hungarian and Polish markets are clearly upward trending, it is difficult to identify some clear trend in the Czech index, with Slovak series being dominated by downward sloping pattern.

Table 4.2 Results from trading rule strategies

A. Hungary

Rule (n1,n2) Band=0.01	Number of buy signals	Number of sell signals	Mean buy returns	Variance of buys	Mean sell returns	Variance of sells	Fraction of positive buys	Fraction of positive sells	Mean buy and sell difference
(1,50)	941	642	0.00365 (3.59935)	0.000207	-0.00212 (-4.66923)	0.000320	60%	40%	0.00577 (7.12898)
(1,200)	1062	371	0.00268 (2.18116)	0.000240	-0.00134 (-2.92462)	0.000267	55%	46%	0.00402 (4.21582)
(5,150)	1032	444	0.00215 (1.31959)	0.000225	0.00017 (1.35140)	0.000280	55%	48%	0.00198 (2.20640)
(2,200)	1062	370	0.00228 (1.53964)	0.000232	-0.00020 (-1.81465)	0.000272	56%	47%	0.00248 (2.59820)
Average	1024	456	0.00281	0.000226	-0.00087	0.000284	56.5%	45,2%	0.00368

B. Poland

Rule (n1,n2) Band=0.01	Number of buy Signals	Number of sell signals	Mean buy returns	Variance of buys	Mean sell returns	Variance of sells	Fraction of positive buys	Fraction of positive sells	Mean buy and sell difference
(1,50)	984	598	0.00515 (3.20219)	0.000592	-0.00358 (-4.51466)	0.000678	46%	37%	0.00873 (6.65532)
(1,200)	1015	417	0.00369 (1.78959)	0.000619	-0.00240 (-3.08307)	0.000854	47%	39%	0.00609 (4.13863)
(5,150)	1015	467	0.00283 (0.93929)	0.000610	-0.00001 (-1.42349)	0.000813	47%	38%	0.00284 (2.00769)
(2,200)	1015	417	0.00299 (1.09748)	0.000638	-0.00070 (-1.85849)	0.000823	48%	39%	0.00369 (2.50764)
Average	1007	474	0.00367	0.000615	-0.00167	0.000792	47%	38%	0.00534

C. Czech Republic

Rule (n1,n2) Band=0.01	Number of buy signals	Number of sell signals	Mean buy returns	Variance of buys	Mean sell returns	Variance Of sells	Fraction of positive buys	Fraction of positive sells	Mean buy and sell difference
(1,50)	471	558	0.00158 (4.12334)	0.000050	-0.00205 (-2.97213)	0.000088	58%	33%	0.00363 (5.91110)
(1,200)	405	476	0.00094 (2.74978)	0.000052	-0.00054 (0.07664)	0.000080	51%	45%	0.00148 (2.30772)
(5,150)	443	488	0.00013 (1.32648)	0.000054	-0.00043 (0.28988)	0.000092	52%	47%	0.00056 (0.89951)
(2,200)	406	475	0.00061 (2.15472)	0.000053	-0.00025 (0.63181)	0.000080	51%	46%	0.00086 (1.34122)
Average	431	499	0.00082	0.000052	-0.00082	0.000085	53%	42%	0.00163

D. Slovakia

Rule (n1,n2) Band=0.01	Number of buy signals	Number of sell signals	Mean buy returns	Variance of buys	Mean sell returns	Variance of sells	Fraction of positive buys	Fraction of positive sells	Mean buy and sell difference
(1,50)	414	615	0.00248 (3.46837)	0.000135	-0.00211 (-1.77780)	0.000127	57%	38%	0.00459 (4.56639)
(1,200)	252	629	0.00125 (1.75385)	0.000102	-0.00087 (-0.22697)	0.000107	55%	41%	0.00212 (1.79847)
(5,150)	291	639	0.00074 (1.36932)	0.000093	-0.00070 (-0.01267)	0.000111	56%	41%	0.00144 (1.28780)
(2,200)	251	630	0.00104 (1.56148)	0.000103	-0.00078 (-0.11354)	0.000107	56%	43%	0.00182 (1.54213)
Average	302	628	0.00138	0.000108	-0.00111	0.000113	56%	40%	0.00249

Note: The t-statistics are in parenthesis. The t-statistics for the mean buy and sell returns equal $\frac{\mu_{b/s} - \mu}{(\sigma^2 / N + \sigma^2 / N_{b/s})^{1/2}}$, where $\mu_{b/s}$ and $N_{b/s}$ are the mean return and number of signals for the buys and sells. The t-statistics for the mean buy and sell returns difference equal $\frac{\mu_b - \mu_s}{(\sigma^2 / N_b + \sigma^2 / N_s)^{1/2}}$, where μ_b and μ_s are mean buy and sell returns, and N_b and N_s are number of buy and sell signals respectively.

4.4 Bootstrap Methodology

The key idea behind the bootstrap procedure, initiated by Efron (1979), is to resample from the original data – either directly or via a fitted model – to create replicate datasets. One of the advantages of the bootstrap is that, in contrast to Monte Carlo simulations, it allows deviations from the Gaussian distribution in the original data series. Bootstrap methods have been widely used in the financial literature during the last decade for a variety of purposes.[65] Empirical analysis in section 4.5 combines the bootstrap methodology with technical analysis for checking the specification of several commonly used models of stock returns, applied to the data from transition equity markets. In this section we review some of the alternative bootstrap methods existing in the literature, and provide a short summary of previous empirical studies using bootstrap and technical trading rules for checking the specification of popular stock price models.

4.4.1 Review of the alternative bootstrap methods

Let $(y_1,\ldots,y_n)$ be a random sample from distribution F, and $T(y_1,\ldots,y_n;F)$ be a random variable of interest which may depend directly upon F. Also let F_n be the empirical distribution function that puts mass $1/n$ on y_i, $i=1,2,\ldots,n$. The bootstrap can be defined as a method that

[65] Some of the purposes of the use of bootstrap methods in financial literature are: (i) to obtain small sample standard errors (e.g. Akgiray and Booth (1988) and Badrinath and Chatterjee (1991)); (ii) to get significance levels for tests (e.g. Hsieh and Miller (1990) and Shea (1989a,b)); (iii) to get significance levels for trading rules profits (e.g. Levich and Thomas (1993) and LeBaron (1994)); (iv) to develop empirical approximations to population distributions (e.g. Bookstaber and McDonald (1987)); (v) to use trading rules on bootstrapped data as a test for model specification (e.g. Brock et al. (1992), LeBaron (1992), Kim (1994) and Karolyi and Kho (1994); (vi) to check the validity of long-horizon predictability (e.g. Nelson and Kim (1993) and Chen (1995); (vii) impulse response analysis in nonlinear models (e.g. Tauchen, Zhang and Liu (1994)).

approximates the distribution of $T(y_1,\ldots,y_n;F)$ under F by the distribution of $T(y_1,\ldots,y_n;F_n)$ under F_n. Bootstrapping is done by taking B bootstrap samples, $Z_b = T(y_{1,b},\ldots,y_{n,b})$, $b = 1,2,\ldots,B$, each of which consists of n random draws from F_n. This procedure is the ***standard bootstrap*** method, which has been extended to classical regression by Freedman (1981a, b). In the case of the classical regression models, it is the residuals that are resampled. When the errors are not IID, one needs to modify this procedure.

To deal with serially correlated errors with a well specified structure (e.g. stationary ARMA (p,q) models with known p and q), one can use the ***recursive bootstrap*** method, first introduced by Freedman and Peters (1984).[66] According to this method one estimates the model by some consistent method, obtains the residuals and resamples them. With the resampled residuals, one next generates the bootstrap samples recursively. For example in the case of regression model with say AR (1) errors, u_t:

$$y_t = \beta x_t + u_t \tag{4.4}$$

$$u_t = \rho u_{t-1} + e_t \tag{4.5}$$

where $e_t \sim IID(0,\sigma^2)$, one estimates (4.4) by OLS (or some other method), then using the estimated residuals $\hat{u}_t$, one estimates $\hat{\rho}$(using e.g. the Cochrane-Orcutt or Prais-Winsen procedures) and obtains $\hat{e}_t$. Then one resamples $\hat{e}_t$ and using a recursive procedure generates $\hat{u}_t$, and the bootstrap sample on y_t.

To deal with general dependent time series data, other approaches, which do not require fitting the data into a parametric form have been developed. Carlstein (1986) first discussed the idea of bootstrapping non-overlapping blocks of observations rather than individual observations. Later,

[66] This method has also been used by Efron and Tibshirani (1986) for bootstrapping the AR(1) and AR (2) models.

Kunsch (1989) and Liu and Singh (1992) independently introduced a more general bootstrap procedure, the ***moving block bootstrap***, which is applicable to stationary time series data. In this method the blocks of observations are overlapping.

Both methods of Carlstein and Kunsch divide the data of n observations of length l and select b of these blocks by resampling with replacement all the possible blocks. For simplicity assume $n = bl$. In the Carlstein procedure there are just b blocks. In the Kunsch procedure there are $n - l + 1$ blocks. The blocks are $L_k = \{x_k, x_{k+1}, \ldots, x_{k+l-1}\}$ for $k = 1, 2, \ldots, (n-l+1)$. For example with n = 6 and $l = 3$ suppose the data are: $x_t = \{7,2,3,6,1,5\}$. The blocks according to Carlstein are $\{(7,2,3),(6,1,5)\}$. The blocks according to Kunsch are $\{(7,2,3),(2,3,6),(3,6,1),(6,1,5)\}$. Now draw a sample of two blocks with replacement in each case. Suppose the first draw gave (7,2,3). The probability of missing the block (6,1,5) is ½ in Carlstein's scheme and ¼ in the moving block scheme. As a result, there is a higher probability of missing entire blocks in the Carlsatein scheme. For this reason, it is not often used.

The literature on blocking methods is mostly on the estimation of the sample mean and variance, although Liu and Singh (1992) consider the applicability of the results to more general statistics, and Kunsch (1989, p.1235) discusses the AR(1) and MA(1) models.

However note that, even if the original series $\{x_t\}$ is stationary, the pseudo times series generated by moving block method is not stationary. For this reason, Politis and Romano (1994) suggest the ***stationary bootstrap*** method. It is very similar to the moving block bootstrap method. However, there is one major difference between the sampling schemes of these two methods. The stationary bootstrap resamples the data blocks of random length, where the length of each block has a geometrical distribution with parameter p, while the moving block bootstrap resamples blocks of data of the same length.

There are, as yet, no applications in the financial literature using the blocking methods.

4.4.2 Bootstrap for model selection using technical trading rules.

Bootstrap methods and trading rules have been used for checking the specification of several commonly used models like the random walk, GARCH processes and the Markov switching regression in application to the data from mature stock markets by LeBaron (1991), Brock et al.(1992), Kim (1994) and Karolyi and Kho (1994). The basic idea is to compare the time series properties of the generated data from the given model with those of the actual data. Trading rule profits are one convenient measure for this purpose. R^2 and other goodness of fit measures do not capture the time series structure of the data. The procedure involves the following steps:

- First get a measure of the profits generated by a trading rule, using the actual data;
- Next estimate the postulated models and bootstrap the residuals and the estimated parameters to generate bootstrap samples;
- Next compute the trading rule profits for each of the bootstrap samples and compare this bootstrap distribution with trading rule profits derived from the actual data.

Brock et al. (1992) tried this procedure with random walk, AR(1), GARCH-M, and EGARCH models on 90 years of daily data on the Dow Jones Industrial average. They found that none of these models replicate the moving average trading rules profits from the actual data. LeBaron (1991) considers random walk, GARCH, regime shifting and interest rate adjusted models and finds that none of them replicates the trading rule profits from the actual data, although GARCH does better than the other models. Besides using trading rule profits as a model specification test, LeBaron also tests the "economic significance" of trading rule profits in the foreign exchange market, by accounting for transaction costs and interest rates and attempting to measure the riskiness of trading strategies in the foreign exchange market relative to the strategies in the other markets. We do not present LeBaron's results in detail but broadly speaking, his conclusion is that the use of technical trading rules in the foreign exchange market generate returns similar to those from a domestic stock portfolio but further tests are necessary to completely answer the question of the "economic significance" of trading rules in the foreign exchange market.

Kim (1994) also uses trading rule profits as a tool for model specification tests. The moving average trading rules are applied to actual and generated data from random walk, GARCH-M, Hamilton's Markov switching model, the SWARCH (ARCH with Markov switching), and the CAPM models. He finds that the random walk, GARCH-M and Hamilton's Markov switching model cannot capture the moving average trading rule profits generated by the actual data. However, the SWARCH performs better than the three other models considered.

Karolyi and Kho (1994) use bootstrap methods in conjunction with trading rules to reexamine the profitability of positive feedback investment strategies, which buy stocks that have performed well in the past and sell stocks that have performed poorly in the past.[67] Karolyi and Kho conclude that their overall findings for NYSE and AMEX stocks from 1965-1989 indicate that the profitability of the relative strategies may simply represent fair compensation for the risks assumed by these strategies. As found by others for moving average trading rules, Karolyi and Kho find that the random walk model cannot explain the significant returns of the positive investment strategy, even

[67] The significant returns to such a strategy were confirmed by Jegadeesh and Titman (1993).

within size- and beta-based subgroups of stocks with similar risk exposures. They, therefore, try to see whether the profitability of the relative strength trading rules is significant after adjusting for time-varying risk. They find that the trading rule profits are consistent with those simulated using a simple conditional CAPM equilibrium model of time-varying expected returns.

4.5 Empirical Analysis: Checking the Validity of the Null Models of Stock Returns

The results in section 4.3 of this chapter are consistent with technical trading rules having predictive power. The purpose of this section is to check whether various null models, addressing the stylised facts normally characterising stock return series, in particular, autocorrelation, leptokurtosis, changing conditional means and variances, are able to account for the predictability in nominal stock price series associated with technical trading rule profits.

Computer simulations of the time series designed to capture the properties of the various null models (in particular, the distribution of the conditional moments under different specifications), will be performed using the estimation-based bootstrap procedure initiated by Freedman (1981a,b). In this procedure each model is fit to the original series to obtain estimated parameters and residuals. The estimated residuals are then redrawn with replacement to form a 'scrambled' residual series, which is then used with the estimated parameters to form a new representative series for a given null model. Next step is to compare the time series properties of the generated data from the given model with those of the actual data. To perform this, we compute trading rule profits for each of the generated bootstrap samples, and compare this bootstrap distribution with trading rule profits derived from the actual data. As described in section 4.4.2 above, the similar procedure was employed for model specification tests by several authors, including LeBaron (1991), Brock et al. (1992), Kim (1994), and Karolyi and Kho (1994).

In our study the price series will be simulated from the three widely used null models for stock prices: AR(p), GARCH-M and EGARCH models. Each of the simulations will be based on 2000 replications of the null model.

In AR(p) model returns follow:

$$r_t = \gamma_0 + \sum_{i=1}^{p} \gamma_i r_{t-i} + \varepsilon_t , \tag{4.6}$$

where ε_t is iid.[68] The simulation experiment with AR(p) process will show whether results from the trading rules could be caused by autocorrelation in the series.

The next two null specifications employed for the simulation of our price series belong to the family of GARCH models. These models allow volatility of returns to change over time. They imply positive serial correlation in the conditional second moments of the return process: periods of high (low) volatility are likely to be followed by the periods of high (low) volatility. The specifications we employ here are GARCH-M and EGARCH models.[69]

Under GARCH-M specification both conditional means and variances are allowed to change over time. According to this model higher returns are expected when conditional volatility is high, which is consistent with efficient market hypothesis.

---

[68] The number of lags, p, in AR model for each of the four markets is choosen to be the same as in the autocorrelation adjustment procedure in the section 3.3.1 of the previous chapter of the thesis. Although the length of the sample in this chapter is slightly longer (about 100 observations more), the number of lags required for iid errors is the same as in the previous chapter for all four series.

[69] See section 3.2 for a review of GARCH models, and, in particular, 3.2.3 for a detailed description of the GARCH-M and EGARCH models.

$$r_t = \gamma_0 + \sum_{i=1}^{p} \gamma_i r_{t-i} + \alpha_1 h_t + \varepsilon_t$$

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} \tag{4.7}$$

where $\varepsilon_t = h_t^{1/2} u_t$ and $u_t \sim NID(0,1)$.

Changing conditional mean and variance could potentially explain results from the trading rules. Many financial series are asymmetric in the sense that negative equity returns are followed by larger increases in volatility than equally large positive returns. The exponential GARCH (EGARCH) model by Nelson (1991) allows for this asymmetry. According to EGARCH specification employed in our analysis, returns, r_t, follow the process as in (4.6), and the conditional variance is specified as follows:

$$\ln(h_t) = \beta_0 + \beta_1\left[\left|\varepsilon_{t-1}\right|\sqrt{h_{t-1}} - \sqrt{2/\pi}\right] + \beta_2 \ln(h_{t-1}) + \beta_3(\varepsilon_{t-1} / \sqrt{h_{t-1}}) \tag{4.8}$$

where $\varepsilon_t = h_t^{1/2} u_t$ and $u_t \sim NID(0,1)$.

This specification lifts the nonnegativity constraint on the parameters of the model by using logs. The parameter β_3 can generate the leverage effect as the sign of yesterday's shock enters the model.

4.5.1 Empirical results

Tables 4.3, 4.4, 4.5 below present means and variances of buy and sell returns[70], as well as difference of buy and sell returns generated by 2000 simulations for three models and four markets

[70] These are returns associated with buy and sell signals from the market described in the section 4.4 of this chapter.

used in our study. To conserve space we present only the average simulated results over the four technical trading rules.

Table 4.1 in the section 4.2 of this chapter indicates presence of significant autocorrelation in our sample. This fact could potentially cause the technical strategies producing *abnormal* returns. The purpose of the AR simulations in Table 4.3 is to detect whether the results from trading rules could be caused by daily autocorrelation in the series. The simulation estimates indicate that the AR processes do generate a positive buy-sell spread between buys and sells for all four countries. This spread is relatively large in the case of the Hungarian and Slovak series. However, even for these two markets the magnitude of the spread is small when compared with the original series (see rows 'Average' in Table 4.2 for comparison).

Tables 4.2 indicates that for all actual return series considered, the market is less volatile during buy periods relative to sell periods. The averaged variances over four moving average rules of buys and sells for Hungary are 0.0226 and 0.0284, for Poland 0.0615 and 0.0792, for Czech Republic 0.0052 and 0.0085, and for Slovakia 0.0108 and 0.0113 per cent respectively. However, Table 4.3 reveals that for the simulated series variances of buys and sells basically do not differ, being even marginally larger for buy periods.

Table 4.4 displays the results of the simulation of GARCH-M model. Potentially, a changing conditional mean could explain some of the differences between buy and sell returns. We see that compared to AR specifications GARCH-M does slightly increase the spread for all four countries. The most significant spread between buys and sells is generated for the Slovak returns (0.00111 compared with 0.00249 for the original series). Moving to the analysis of volatility estimates we find that for the Hungarian and Polish series this specification does not move us any closer to the results for the original returns in Table 4.2. The simulation results indicate that in the case of the Slovak market the variance of sells exceeds the variance of buys. However, the values of the variances obtained from GARCH-M simulation of the Slovak series are still much closer to the variance of the unconditional returns than those in Table 4.2.

We conclude, that the autoregressive and GARCH-M specifications are not very successful in explaining the moving average rules profits generated by the actual data, and account only for a fraction of the differences between actual buy and sell returns.

The ability of the EGARCH specification to capture the asymmetric effects of the positive and negative shocks on volatility could to some extent explain the spread between variances of the buy and sell returns generated by application of technical trading rules to our dataset. Results in Table 4.5 reveal that, in spite of the fact that EGARCH specification does not add much to the explanation of the difference between buy and sell returns, it does create some difference between variances of buys and sells for three out of four markets, with the variances of sells higher than the variances of buys. As a result, the EGARCH model fails to match conditional means in the data. However, it performs better than the AR and GARCH-M models in predicting volatility.

In sum, our findings show that stock return-generating process is probably more complex than suggested by various linear conditional mean models of past returns.

Table 4.3 Simulated averages over four rules for AR model

	Uncon. returns	Uncon. var-ce	Mean buy returns	Variance of buys	Mean sell returns	Variance of sells	Mean buy and sell difference
Hungary	0.00132	0.00025	0.00175	0.00019	0.00071	0.00017	0.00104
Poland	0.00188	0.00064	0.00242	0.00064	0.00155	0.00064	0.00087
Czech R.	-0.00058	0.00009	3.42e-05	0.00009	-0.00025	0.00008	0.00025
Slovakia	-0.00069	0.00025	0.00038	0.00025	-0.00060	0.00023	0.00098

Table 4.4 Simulated averages over four rules for GARCH-M model

	Uncon. returns	Uncon. Var-ce	Mean buy returns	Variance of buys	Mean sell returns	Variance of sells	Mean buy and sell difference
Hungary	0.00132	0.00025	0.00196	0.00021	0.00081	0.00018	0.00115
Poland	0.00188	0.00064	0.00198	0.00063	0.00155	0.00064	0.00090
Czech R.	-0.00058	0.00009	0.00006	0.00007	-0.00031	0.00008	0.00037
Slovakia	-0.00069	0.00025	0.00043	0.00020	-0.00068	0.00022	0.00111

Table 4.5 Simulated averages over four rules for EGARCH model

	Uncon. returns	Uncon. Var-ce	Mean buy returns	Variance of buys	Mean sell returns	Variance of sells	Mean buy and sell difference
Hungary	0.00132	0.00025	0.00172	0.00022	0.00064	0.00020	0.00108
Poland	0.00188	0.00064	0.00235	0.00064	0.00168	0.00070	0.00067
Czech R.	-0.00058	0.00009	2.11e-05	0.00004	-0.00023	0.00007	0.00023
Slovakia	-0.00069	0.00025	0.00039	0.00011	-0.00062	0.00014	0.00101

4.6 Nonlinear Predictability: the Artificial Neural Networks (ANNs)

The asymmetric behaviour of returns and volatility over the buy and sell periods, as well as the failure of linear conditional mean estimators to characterise the temporal dynamics of stock returns in the previous sections of this chapter, suggest the existence of possible non-linearities. Gencay (1998) indicates that the linear conditional mean specifications with past buy-sell signals as predictors of current returns provide forecast improvements over linear models of past returns. Based on this results, in this section we examine linear and nonlinear predictability of stock returns from four transition markets with buy and sell signals generated from the moving average rules. We use the AR and EGARCH models with past buy and sell signals in the conditional mean specification as linear mean estimators, and the single-layer feedforward networks as non-linear.

Broadly speaking Artificial Neural Networks (ANNs) are nonlinear nonparametric models. ANNs are data-driven modelling approach, which allows the data determine the structure and parameters of a model without any restrictive parametric modelling assumptions. As shown in Hornik et al. (1989) and Cybenko (1989), neural networks are more parsimonious models than linear subspace methods such as polynomial, spline, and trigonometric series expansions in approximating unknown functions. Thus, if the behaviour of economic variables exhibits nonlinearity, a suitably constructed neural network can serve as a useful tool to capture such regularity. ANNs are appealing in financial area because of the abundance of high quality financial data and the insufficient amount of testable financial models. As the speed of computers increases and the cost of computing declines exponentially, this computer intensive method becomes attractive. The past decade has seen an explosive growth in studies of neural networks, which has been brought about largely by realisation that ANNs have powerful pattern recognition properties that may outperform other existing modelling techniques in many applications. Many different networks, such as single- and multilayer feedforwrd networks, recurrent and statistical networks, associate memory networks and self-organisation networks, etc., have been developed for different purposes. Wide-ranging introductions to neural networks theory can be found in Hecht-Nielsen (1990), Hertz, Krogh and Palmer (1991), Wasswerman (1993) and Bose and Liang (1996). White, Gallant, Hornik, Stinchcombe and Wooldridge (1992) present a collection of papers that carry out mathematical analysis of the approximation and learning abilities of ANNs. Gately (1996) provides a very non-technical approach to neural network for beginners.

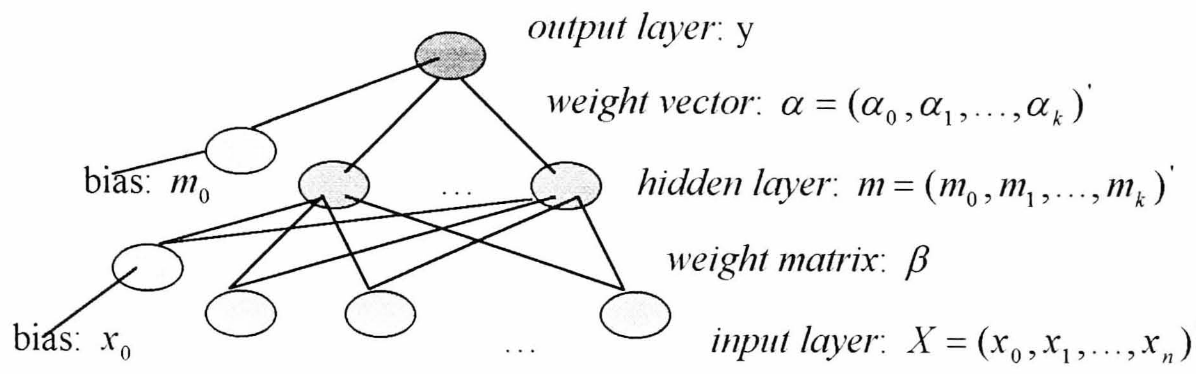
In the next section we give a brief review of the issues on the structure, estimation and financial applications of the Artificial Neural Networks (ANN).

4.6.1 ANN structure

The feedforward backpropagation networks are the most popular ones in financial applications, and will be used in the empirical part of this chapter. Therefore, we focus on the structure of the feedforward network. Individual units of the network are organised in layers: the input, hidden and output layers. Feedforward networks map inputs into outputs with signals flowing in one direction only, from the input layer to the hidden layer and then the output layer. Each unit in the hidden and output layers has a transfer function, which transfers the signal it receives. The input layer units do not have a transfer function, but they are used to distribute input signals to the network. Each connection has a numerical weight, which modifies the signals that pass through it.

Let us consider an example of a feedforward network with a single output unit, a single hidden layer with k hidden units, and n input units.

Figure 4.1 Feedforward neural network



Any hidden layer unit receives the weighted sum of all inputs and a bias term, x_0, and produces an output signal

$$m_j = F(\sum \beta_{ij} x_i) = F(X' \beta_j), \qquad j = 1,2,\ldots,k, \qquad i = 0,1,2,\ldots,n \tag{4.9}$$

where F is the transfer function, x_i is the i^{th} input signal, and β_{ij} is the weight of the connection from the i^{th} input unit to the j^{th} hidden layer unit. In the same way, the output unit receives the weighted sum of the output signals of the hidden layer units, and produces a signal

$$y = G(\sum \alpha_j m_j), \qquad j = 0,1,2,\ldots,k, \tag{4.10}$$

where G is the transfer function, α_j is the weight of the connection from the j^{th} hidden layer unit to the output unit, and $j = 0$ indexes a bias unit m_0. Combining (4.9) and (4.10) we get

$$y = G(\alpha_0 + \sum_{j=1}^{k} \alpha_j F(\sum \beta_{ij} x_i)) = f(X, \theta), \tag{4.11}$$

where X is the vector of inputs, and

$\theta = (\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k, \beta_{01}, \beta_{02}, \ldots, \beta_{0k}, \beta_{11}, \beta_{12}, \ldots \beta_{1k}, \ldots, \beta_{n1}, \beta_{n2}, \ldots, \beta_{nk})'$ is the vector of network weights.[71]

A major advantage of ANNs is their ability to provide a flexible mapping between inputs and outputs. Based on a series of studies by Kolmogorov (1957), Sprecher (1965), Lorentz (1976), and Hecht-Nielsen (1987, 1990), any continuous function can be computed using linear summations and a single properly chosen nonlinear function. Therefore, the arrangement of the simple units into multilayer framework as displayed, for example, in Figure 4.1 produces a mapping between inputs and outputs that is consistent with any underlying functional relationship regardless of its "true" functional form. Having a general mapping between the input and output vectors eliminates the need for unjustified a priori restrictions which are needed in common statistical and econometric modelling.

To implement a perfectly general mapping between inputs and outputs, correct transfer functions are needed. Studies by Cybenko (1989), Funahashi (1989), Hecht-Nielson (1990), Hornik et al. (1989) have shown that sigmoid transfer functions could serve this purpose. Stinchcombe and White (1989) show that some non-sigmoid functions can also be used. F and G in (4.11) could be, for example, the sigmoid (or logistic) functions which produces output between 0 and 1: $F(\alpha) = G(\alpha) = 1/(1+\exp(-\alpha))$, or $F(\alpha) = \alpha$ and $G(\alpha) = 1/(1+\exp(-\alpha))$; or the threshold functions, which produce binary (± 1) or $(0/1)$ output.

[71] The basic ANN structure represented by (4.11) can be generalised in many different ways. For example, Poli and Jones (1994) introduce a feedforward network ANN with observation noise and random connection between units. Based on some distributional assumptions of the noise and randomness of the connections, such an ANN can be estimated by a Kalman filtering procedure.

As a result, an ANN can be viewed as a "universal approximator", i.e. a flexible functional form that can approximate an arbitrary function arbitrary well, given sufficiently many hidden layer units and properly adjusted weights.

4.6.2 ANN estimation

A good discussion of various estimation methods of the ANN is given in Kuan and White (1994). The most widely used estimation method (or so-called learning rule) is error backpropagation.[72] Backpropagation is a recursive gradient descent method that minimises the sum of the squared errors of the system by moving down the gradient of the error curve. More specifically, network weight vector θ is chosen to minimise the loss function

$$\min_{\theta} L = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2 , \tag{4.12}$$

where N is a sample size, y_t is desired (or actual) output value and $\hat{y}_t$ is the calculated output value,

$$\hat{y}_t = f(X_t, \theta) = G(\alpha_0 + \sum_{j=1}^{k}\alpha_j F(\sum \beta_{ij} x_{it})) . \tag{4.13}$$

Then the iterative step of the gradient algorithm takes θ to $\theta + \Delta\theta$, and

$$\Delta\theta = -\eta \nabla f(X_t, \theta)(y_t - f(X_t, \theta)), \tag{4.14}$$

where $\eta > 0$ is the step size, or learning rate, and $\nabla f(X_t, \theta)$ is the gradient of $f(X_t, \theta)$ with respect to θ (a column vector).

[72] For detailed presentation see Rumelhart, Hilton and Williams (1986a, b).

The error surface of ANN is multidimensional and may contain many local minima. Therefore, training the network often requires experimentation with different starting weights, adjusting the learning rate, or adding a momentum term to avoid getting stuck in local optima or slow convergence. For the studies that aim to compare the ANN with some alternative models, as long as the ANN performs slightly better than its counterpart, it is not necessary to search for global minima. For studies that try to search for global minima, a grid search method is often used.[73] White, Gallant, Hornik, Stinchcombe and Wooldridge (1992) have a good discussion about different methods of the global optimisation.

4.6.3 Model selection

Though ANNs can be universal approximators, the, optimal network structure is not determined automatically. Failures in applications are sometimes due to a suboptimal ANN structure. To develop the optimal network in any financial application, one needs to:

1) identify the relevant inputs and outputs;
2) choose the appropriate network structure including the necessary number of hidden layers and hidden layer units;
3) use proper model evaluation criteria.

1) ANN inputs and outputs

It is a common practice to use independent variables as network inputs and use dependent variables as network outputs in a model. In order to minimise the effect of magnitude among the inputs and

[73] See. for example Gorr, Nagin and Szczypula (1994).

outputs and increase the effectiveness of the learning algorithm, the data set is often normalised (or scaled) to be within a specific range depending on the transfer function. For example, if ANN has sigmoid transfer function in the output unit, the output needs to be scaled to fall on the range of [0, 1]. Otherwise, a target output which falls outside that range will constantly create large backpropagation errors, and the network will be unable to learn the input-output relationship that is implied by the particular training pattern.

2) ANN architecture

The next step, after specifying the network input and output layers, is to determine the necessary number of hidden layers and hidden layer units. Cybenko (1988) proves that ANN with at most two hidden layers can approximate a particular set of functions with arbitrary accuracy given enough units per layer. It has also been proved that only one hidden layer is enough to approximate any continuous function (see Cybenko (1989), Hornik, Stinchcombe and White (1989)).[74]

The choice of the number of hidden units represents a compromise. If the number of units is too small, an ANN may not approximate at the desired accuracy. However, if the number of units is too large, an ANN may overfit and can not forecast out of sample. A useful method is cross-validation, by which the number of hidden layer units is selected to optimise the out-of-sample performance (see White (1990)). Another related model selection criterion, the predictive stochastic complexity (PSC) can also be used (see (4.24) below).[75]

[74] Many empirical studies, such as Collins, Ghosh and Scofield (1988), Dutta and Shekhar (1988), Salchenberger, Cinar and Lash (1992), to name few, have confirmed this.
[75] Other common methods for optimal network design have been reviewed by Refenes (1995b). These methods fall into three groups. The first is analytic techniques in which algebraic or statistical analysis is used to determine a priori hidden unit size. The main problem with these techniques is that they perform static analysis and can only provide a very rough estimate for hidden unit size. The second type is constructive techniques, such as cascade correlation (Fahlman and Lebiere (1990)), tiling algorithm (Mezard and Nadal (1989)), neural decision tree (Gallant (1986)) and upstart algorithm (Frean (1989)). These methods construct hidden units in layers one by one as they are

3) *ANN evaluation criteria*

Let $(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_N)$ denote the predicted values and $(y_1, y_2, \ldots, y_N)$ be the actual values, where N is the sample size. Some of the commonly used criteria to compare the performance of alternative models are listed below.[76]

(1) Mean square error (MSE) (sometimes called mean square prediction error (MSPE)) and root mean square error (RMSE):

$$MSPE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2, \tag{4.15}$$

$$RMSE = \sqrt{MSPE}. \tag{4.16}$$

(2) Mean absolute error (MAE) and mean absolute percentage error (MAPE)[77]:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|, \tag{4.17}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right|. \tag{4.18}$$

(3) Coefficient of determination (R^2), which measures the proportion of the total variation in y that is accounted for by variation in the regressors:

needed. Though these techniques guarantee the network convergence, generalisation (or forecasting) and stability are not guaranteed. The last type, network pruning, operates in the opposite direction by pruning the network and removing "redundant" or least sensitive connections (see Sietsma and Dow (1991) and Hergert, Finnoff and Zimmermann (1992)). However, optimal pruning is not always possible.

[76] For a discussion on criteria for forecast evaluation see e.g. Chapter 3 in Clements and Hendry (1998).

[77] MAPE is not available for samples, in which y_i has zero actual values.

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^2}, \tag{4.19}$$

where $\bar{y} = \frac{1}{N}\sum y_i$.

(4) Pearson correlation (ρ), which measures the linear correlation between predicted values and actual values:

$$\rho = \frac{\Sigma(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\Sigma(y_i - \bar{y})^2}\sqrt{\Sigma(\hat{y}_i - \bar{\hat{y}})^2}}\text{[78]}. \tag{4.20}$$

(5) Theil's coefficient of inequality (U), which gives prediction performance relative to the random walk prediction,

$$U = \frac{RMSE}{\sqrt{\frac{1}{N-1}\sum(y_i - y_{i-1})^2}}. \tag{4.21}$$

(6) Akaike information criterion (AIC): AIC adjusts MSE to account for the model complexity,

$$AIC = MSE\left(\frac{N+k}{N-k}\right), \tag{4.22}$$

where k is the number of free parameters in the model, or the number of free weights in an ANN.

(7) Schwarz information criterion (SIC), or Bayesian information criterion (BIC), is another way of adjusting MSE to account for model complexity,

[78] Note that R^2 can actually be expressed as the square of Pearson's ρ (see Greene (1993), p.192).

$$SIC = BIC = \ln(MSE) + \frac{\ln(N)}{N} k . \tag{4.23}$$

(8) Predictive stochastic complexity (PSC) due to Rissanen (1986a,b)[79] is a criterion to regularise network complexity computed as the average of squared, '*honest*' prediction errors:

$$PSC = \frac{1}{N-k} \sum_{i=k+1}^{N} (y_i - \hat{y}_{ii})^2 , \tag{4.24}$$

where $\hat{y}_{ii}$ is the predicted value based on parameters obtained from the data up to the i-1 observation. The prediction error $(y_i - \hat{y}_{ii})$ is '*honest*' in the sense that no information at time t or beyond is used to calculate $\hat{y}_{ii}$.[80]

(9) Direction accuracy (DA) and confusion rate (CR):

$$DA = \frac{1}{N} \sum a_i , \tag{4.25}$$

where $a_i = \begin{cases} 1 & if \quad (y_{i+1} - y_i)(\hat{y}_{i+1} - y_i) > 0, \\ 0 & otherwise \end{cases}$

$$CR = 1 - DA . \tag{4.26}$$

It is worth noting that the in-sample performance of any properly designed and well trained ANNs, evaluated by the above measures, is usually much better than from their traditional statistical counterparts. This is not surprising given the universal approximation properties of ANNs. To

---

[79] See also Rissanen (1987).

[80] A model is selected if it has the smallest PSC within a class of models. Clearly, the PSC criterion is based on *forward* validation, which is important in forecasting. For a thorough discussion of the notion of stochastic complexity we refer to Rissanen (1989).

avoid spurious fit or overfit, it is important to test the trained ANN using hold-out sample, i.e. to evaluate the trained ANN using data not been used in training the ANN. Whether the selected ANN model is useful or not depends primarily on the *out-of-sample performance* (see Swanson and White (1995a, b)).

4.6.4 Financial applications of ANNs

ANNs have been successfully applied in several financial areas, such as option pricing, bancruptcy prediction, exchange rate and stock market forecasting. We review some of the empirical studies in this area.

There are only a few published studies regarding neural networks and *option pricing*. The first well-known experiment of option pricing using neural networks is by Hutchinson, Lo and Poggio (1994). The potential value of neural network pricing formula has been shown by the fact that neural networks can discover the Black-Scholes formula from a two-year training set of simulated daily option prices.[81] The resulting network formula has been shown successful in pricing and delta-hedging options out-of-sample. When the network is applied to the pricing and delta-hedging of S&P 500 futures options from 1987 to 1991, the results show that neural networks outperform the Black-Scholes formula.

However, Hutchinson et al. (1994) assume constant risk-free interest rate and constant volatility of the underlying asset. They also assume that the return of the underlying asset is independent of the level of the stock price, so that the option pricing formula is homogeneous of degree one in both S,

[81] The option prices are simulated based on all the Black-Scholes assumptions, such as geometric Brownian motion with constant mean and volatility, constant interest rates, etc.

the asset price, and X, the exercise price. Thus their network has only two inputs, S/X and T (time to maturity), with the output, C/X, the ratio of the call price to the exercise price. It is reasonable to doubt whether such a network can capture all the option price variations.

Qi and Maddala (1995a), unlike Hutchinson et al. (1994), use variables that are believed to be important in determining option prices as network inputs, and use option prices as network output. The input variables are the underlying asset price (S), exercise price (X), risk-free rate (r), time-to-maturity (T), and open interest (V). Such a network provides superior performance to the Black-Scholes formula both in and out of sample for S&P 500 index call options, and the results are better than those reported by Hutchinson et al. (1994).

In contrast to option pricing, a lot of studies have been done in *bankruptcy prediction*. The standard tools are discriminant analysis and the logit model.

Tam and Kiang (1992) compare the neural network approach with linear discriminant analysis, the logit model and other approaches in predicting the failure of Texas banks from 1985 to 1987 using 19 financial ratios. The empirical results show that neural networks offer better predictive accuracy than the alternative approaches. Salchenberger, Cinar and Lash (1992) present a neural network developed to predict the probability of failure for savings and loan associations, using financial variables that signal an institution's deteriorating financial condition. Unlike Tam and Kiang (1992) who use 19 financial ratios as network inputs, Salchenberger et al. (1992) reduce the data dimensions from 29 to 5 by stepwise regression, and use these five variables as network inputs.

Nevertheless, the results are similar: the ANN has performed as well as or better than the best logit model with their data.[82]

Exchange rates are well known for their unpredictability. Most of the unpredictable conclusions are drawn from linear time series techniques, thus the linear unpredictability of exchange rates may be due to limitations of linear models. As a class of flexible functional form nonlinear models, ANNs may provide improvements in forecasting accuracy.

Kuan and Liu (1995) investigate the out-of-sample forecasting ability of neural networks on five exchange rates against the US dollar, including the British pound, the Canadian dollar, the Deutsche mark, the Japanese yen and the Swiss franc. For the Japanese yen and British pound, ANNs are found to have significant market timing ability and significantly lower out-of-sample MSE relative to the random walk model in different testing periods; for the Canadian dollar and Deutsche mark, however, the selected networks do not perform as good. The results show that nonlinearity in exchange rates may be exploited to improve both point and sign forecasts. These results are different from those in Tsibouris (1993), in which ANN is found to be useful in forecasting the direction of the exchange rate change, but not the magnitude.

Other applications of ANN to exchange rates, which show a statistically significant improvement in performance of the models using ANNs are those by Abu-Mostafa (1995) and Hsu, Hsu and Tenorio (1995).

[82] Other studies of bankruptcy prediction using ANNs are Altman, Marco and Varetto (1994), Tam and Kiang (1990). Odom and Sharda (1990), Coats and Fant (1992), Huang (1993) and Poddig (1995).

Traditional models, such as the CAPM and APT, have been very useful in expanding the understanding of stock price behaviour. However, their practical use is often limited given their limit success in forecasting *stock returns*. Because of the inductive, adaptive, and robust nature of ANNs, a great deal of effort has been devoted to developing ANNs for predicting stock returns.

White (1988) investigates the forecastability of IBM daily stock returns using historical data. Though the surprisingly good fit ($R^2 = 0.175$) has been found in-sample, the out-of-sample correlation between actual and forecasted returns is –0.0699 (the in-sample correlation is 0.0751). Such results do not provide evidence for the forecastability of ANN. However, the question of forecastability remains open because of the simple network used in White's study.

Chuah (1993) uses ANN to forecast stock index returns of NYSE using data from January 1963 to December 1988, and compares the predictability and profitability of the network forecasts with those from a benchmark linear model using the same data. The predictability tests show that the forecast errors of the network are not significantly different from those of the benchmark linear model. The profitability test examines profits generated from a trading simulation over a five year forecast period, in comparison with a benchmark buy-and-hold strategy. The nonlinear network generated a total return of 116% versus 94% from the buy-and-hold strategy, while the linear network (with linear transfer function) generated only a 38% total return. Similar results have been obtained by Qi and Maddala (1995b) on S&P 500 index returns using data from January 1959 to June 1995.

Gencay (1996 and 1998) uses the daily Dow Jones Industrial Index Average to examine the linear and nonlinear predictability of stock market returns with simple technical trading rules. Using buy and sell signals from moving average trading rules as inputs to the network, and index returns as the output, the results show significant improvement in forecast performance of the feedforward network over autoregressive and EGARCH specifications with past buy and sell signals in linear mean equation.

Refenes, Zapranis and Fransis (1995) show that neural networks are superior substitute for linear regression in a dynamic multi-factor model of stock returns, a dynamic version of APT. [83]

4.7 Empirical Analysis: Comparison of the Out-of-Sample Forecast Performance of Linear and Nonlinear (ANN) Estimators with Technical Trading Signals in the Transition Equity Markets

In this section we present linear and nonlinear mean estimators with past buy and sell signals from the market, which will be employed for examination of predictability in the transition stock markets of Central Europe. The evidence of predictability should result from a careful analysis of out-of-sample forecasts. In our study the 20-day out-of-sample mean square prediction error is used as a measure of the performance of our linear and nonlinear estimators.

The first conditional mean specifications we employ is the autoregressive (AR) model[84]:

$$r_t = \alpha_0 + \sum_{i=1}^{4} \alpha_i s_{t-1}^{i,ns,nl} + \varepsilon_t , \qquad \text{with } \varepsilon_t \sim ID(0, \sigma_t^2) \tag{4.27}$$

[83] Other studies in this area are Kamijo and Tanigawa (1990), Refenes, Zapranis and Francis (1994), Haefke and Helmenstein (1995).

[84] The model is linear in parameters with the pre-determined variable s. s itself is a nonlinear function of the moving averages; this is spelt out in equations (4.2) and (4.3). Note, that throughout this chapter under *linear* mean specification we understand mean specification linear in parameters.

where $s_{t-1}^{i,ns,nl}$ are the lagged trading rules signals defined in (4.2) and (4.3), and i = 1, 2, 3, 4 corresponds to four trading rules 1-150, 1-200, 5-150, 2-200 respectively.

The next linear conditional mean specification belongs to the family of GARCH models, which allow volatility of returns to change over time. The model we employ here is EGARCH, which, on the basis of results from bootstrapping, appears to be the most successful among other considered models in replicating trading rules volatility. The EGARCH specification of conditional volatility with past trading signals in mean equation is specified as follows:

$$r_t = \alpha_0 + \sum_{i=1}^{4} \alpha_i s_{t-1}^{i,ns,nl} + \varepsilon_t$$

$$\ln(h_t) = \beta_0 + \beta_1\left[|\varepsilon_{t-1}|\sqrt{h_{t-1}} - \sqrt{2/\pi}\right] + \beta_2 \ln(h_{t-1}) + \beta_3(\varepsilon_{t-1}/\sqrt{h_{t-1}}) \tag{4.28}$$

where $\varepsilon_t = h_t^{1/2} u_t$ and $u_t \sim NID(0,1)$.

The nonlinear specification of the conditional mean we employ here is the feedforward regression model:

$$r_t = \alpha_0 + \sum_{j=1}^{d} \eta_j F(\alpha_j + \sum_{i=1}^{4} \beta_{ij} s_{t-1}^{i,ns,nl}) + \varepsilon_t \tag{4.29}$$

where $\varepsilon_t \sim ID(0, \sigma_t^2)$, d is the number of hidden units or nodes in the network, *F(.)* is nonlinear transfer (or activation) function, $\{\alpha_j\}$ and $\{\beta_{ij}\}$ are the vectors of network weights, and i = 1, 2, 3, 4 corresponds to four moving average rules. The network architecture we employ here is a single layer feedforward network. The activation function F is chosen to be the following sigmoidal function:

$$F(u) = \tanh(u / 2)$$

where u is a linear function of the input values and the network weights. This filter scales the outputs of the network between -1 and 1.

4.7.1 Empirical Results

In this section we examine the out-of-sample predictive performance of models defined in (4.27) – (4.29) for each of the four transition stock markets. The forecast horizon is chosen to be 20 days. The out-of-sample prediction is done by excluding the last 20 observations from in-sample estimation. The out-of-sample forecasts with the buy and sell signals as the conditioning set are computed recursively by estimating $E(r_t | s_{t-1}^{i,ns,nl})$, then $E(r_{t+1} | s_t^{i,ns,nl})$, etc., until the sample is exhausted. This will result in the sequence of 20 ex ante one-step-ahead forecasts. As a measure of the forecast performance we use the out-of-sample mean square prediction error (MSPE), defined in (4.15) of section 4.5.3.

The mean square prediction errors of the autoregressive (AR), the EGARCH, and the feedforward network (NN) models with past buy and sell signals in conditional mean equations are presented in table 4.7. The AR specification is used as a benchmark model. Columns 4 and 6 of the table show the ratio of the MSPEs of the EGARCH and NN models to the MSPE of the AR specification respectively.

Starting from comparison of our results for linear specifications of the conditional mean, we document that out-of-sample forecast performance of the EGARCH model does outperform the benchmark AR specification in three out of the four markets examined. EGARCH provides about 19.4%, 15.4% and 17.3% improvement over the AR model for Hungarian, Polish and Czech markets respectively. In case of Slovakia the ratio of the MSPEs of EGARCH and AR specifications is close to the unitary value, which indicates the lack of improvement in forecast performance. On average across four markets, the MSPE of the EGARCH model is 11.7% smaller than the MSPE of the benchmark AR model.

The results from the feedforward network regression[85] with past buy-sell signals indicate significant improvement in forecast performance over the benchmark AR model. The difference between the MSPEs of NN and AR specifications varies from 11.5% for Slovak to 41.6% for Hungarian data. On average, the improvement in the MSPE of the feedforward network model is about 28.4%.

The forecast improvements of the feedforward network regression are substantial and dominate the results from EGARCH linear conditional mean specification. This evidence is confirmed by the Figures 4.3 – 4.6 at the end of this chapter.[86]

[85] The optimal number of hidden layer units for each of the series is selected by method of cross-validation, which optimises the out-of-sample performance of the network. The chosen number of hidden units varies from 5 to 8 across the markets.

[86] The out-of-sample forecasting period considered in this chapter covers 20 days only, and is too short to draw robust conclusions about forecasting performance of the models employed in our analysis. However, we would like to point to the fact that the problem with convergence of the neural network we faced at the beginning of the work on this chapter induced us to undertake the out-of sample forecasting experiments for several subsamples of our dataset. This at the same time enabled us to compare the performance of the models employed in our analysis under different market conditions. In most of the cases the results confirmed the substantial forecast improvements of the feedforward network regression over the models with the linear conditional mean specification. This fact supports robustness of the results presented in Table 4.7.

Table 4.6 Comparison of the MSPEs of linear and nonlinear models of conditional mean

	MSPE of AR	**MSPE of EGARCH**		**MSPE of FNN**	
	in levels ($\times 10^{-4}$)	in levels ($\times 10^{-4}$)	as a ratio to MSPE of AR	in levels ($\times 10^{-4}$)	As a ratio to MSPE of AR
Hungary	0.5432173	0.437841	**0.806014 (19.4%↓)**	0.317121	**0.583783 (41.6%↓)**
Poland	1.3898767	1.175244	**0.845574 (15.4%↓)**	0.994630	**0.715625 (28.4%↓)**
Czech Rep.	0.6306617	0.521301	**0.826594 (17.3%↓)**	0.428318	**0.679157 (32.1%↓)**
Slovakia	1.691337	1.782788	**1.054018 (5.4%↑)**	1.495385	**0.884843 (11.5%↓)**
Average forecast improvement over AR			**11.7%↑**		**28.4%↑**

4.8 Summary

Recent research provides evidence that technical trading rules devised by investors and technical analysts may provide positive profits. Most contributions to this literature deal with mature markets. In this chapter we work with the data from four new stock markets of Central Europe. Our results provide strong support for the technical strategies applied to our dataset. They are consistent with technical rules having predictive power.

The results from the application of the moving average rules to our series document several important stylised facts. Firstly, buy signals consistently generate higher returns than sell signals. In addition we find that returns associated with sell signals are negative.[87] This fact cannot be explained by various seasonalities since it is based on a large fraction of trading days. Many previous studies found as we did that returns are predictable. This predictability can reflect either changes in expected returns that result from an equilibrium model, or market inefficiency. Although rational changes in expected returns are possible, it is hard to imagine an equilibrium model that predicts negative returns over such a large fraction of trading days. Our further results reveal that the differences between buy and sell returns are not easily explained by changing risk levels. We document that the returns associated with buy signals are less volatile than returns associated with sell signals, which is inconsistent with existing models of changing risk. And finally, the last stylised fact we obtain from the application of technical strategies to our data is that for all four markets the fraction of positive returns associated with buy periods is at about 10 per cent or more larger than the fraction of positive returns generated from sell signals.

In the second part of the chapter we use the bootstrap methodology to compare the returns conditional on trading rules signals from the actual series to the returns from the simulated series generated from three popular models for stock prices. These null models are the AR, the GARCH-M of Engle, Lilien and Robins (1987), and the EGARCH by Nelson (1991). We find that the employed models are not very successful in explaining the moving average rules profits generated by the actual data, and account only for a fraction of the differences between actual buy and sell returns. Our findings are quite in accord with the results of previous research, which uses bootstrap

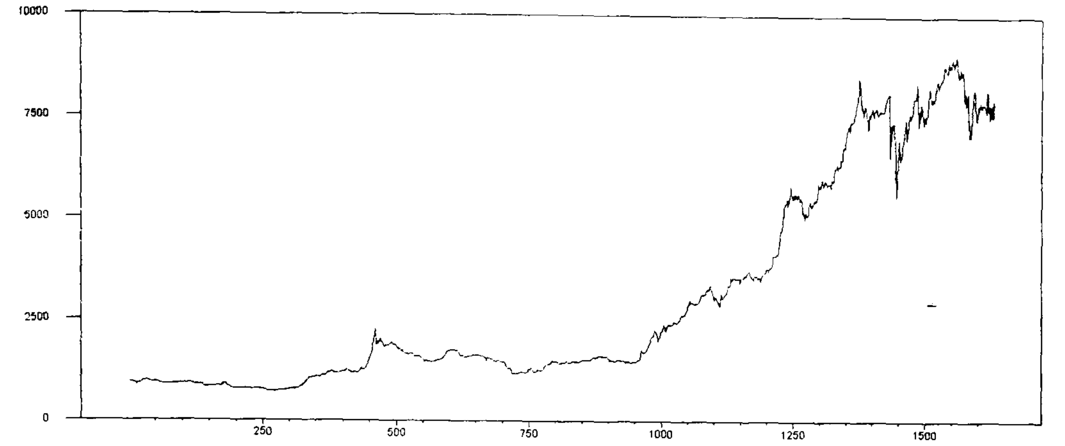
[87] We want to emphasise that our analysis focuses on the simpliest trading rules. The subsequent future research could involve the investigation of more elaborate technical strategies, which could potentially increase the magnitude of our results.

methods and trading rules for the purpose of checking the adequacy of different commonly used models of stock prices in application to more mature markets.

Our conclusion about the asymmetric nature of returns and volatility over the periods of buy and sell signals, as well as the fact that linear conditional mean estimators with past returns fail to characterise the temporal dynamics of security returns, suggest the existence of possible non-linearities. Previous studies indicate that the linear conditional mean specifications with past buy-sell signals as predictors of current returns provide forecast improvements over linear models of past returns. We therefore use the AR and EGARCH models with past buy and sell signals in the conditional mean specification as linear mean estimators, and single-layer feedforward networks as nonlinear estimators. The 20 day out-of-sample mean square prediction error is used as a measure of the performance of the two approaches. The results indicate strong evidence of nonlinear predictability in the stock market returns by using the past buy and sell signals of the moving average rules. The improvement in forecast performance of the feedforward network model over the AR specification with past buy-sell signals varies from 11.5% for the Slovak to 41.6% for the Hungarian data. As a result, one of the main findings of our analysis indicates that nonlinearities in the stock return series play an important role in modelling the conditional mean of the security returns.

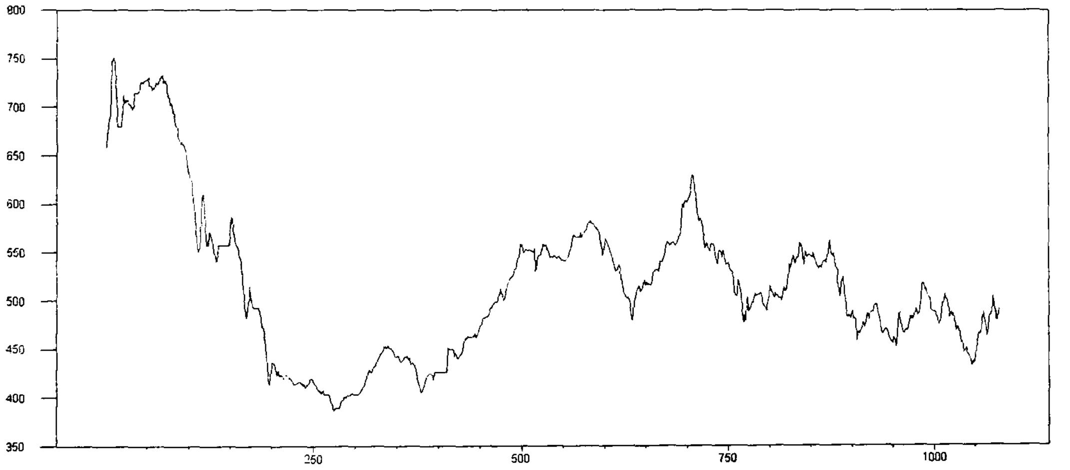
Figure 4.2.1

A. Hungarian index



B. Polish index

C. Czech index



D. Slovak index

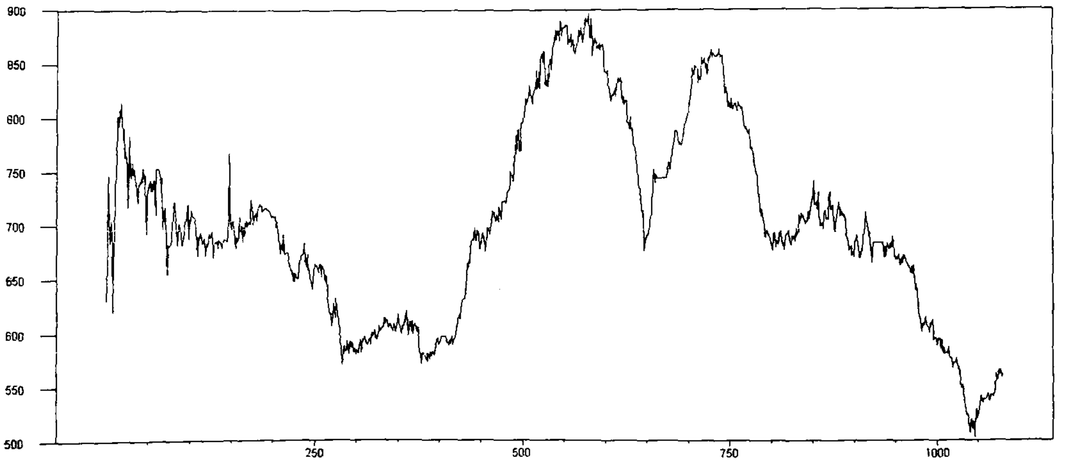
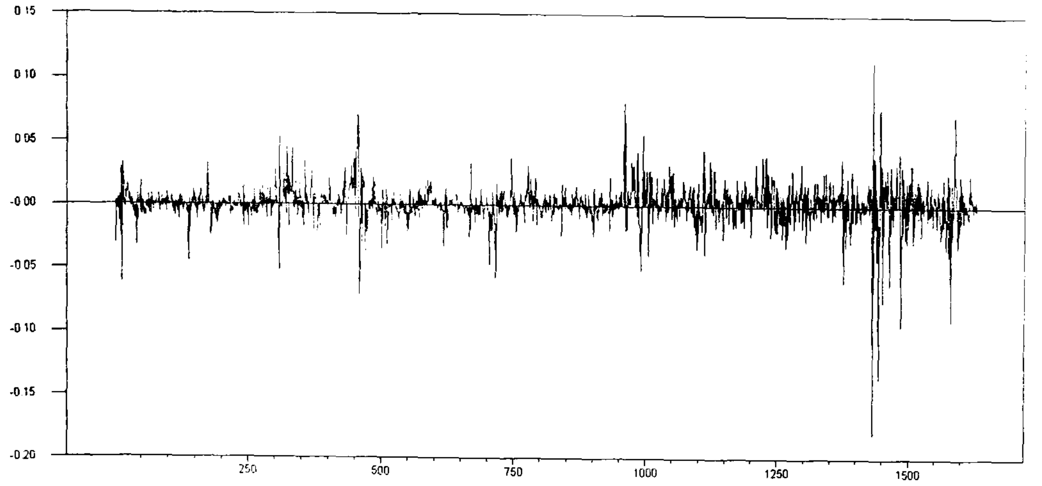
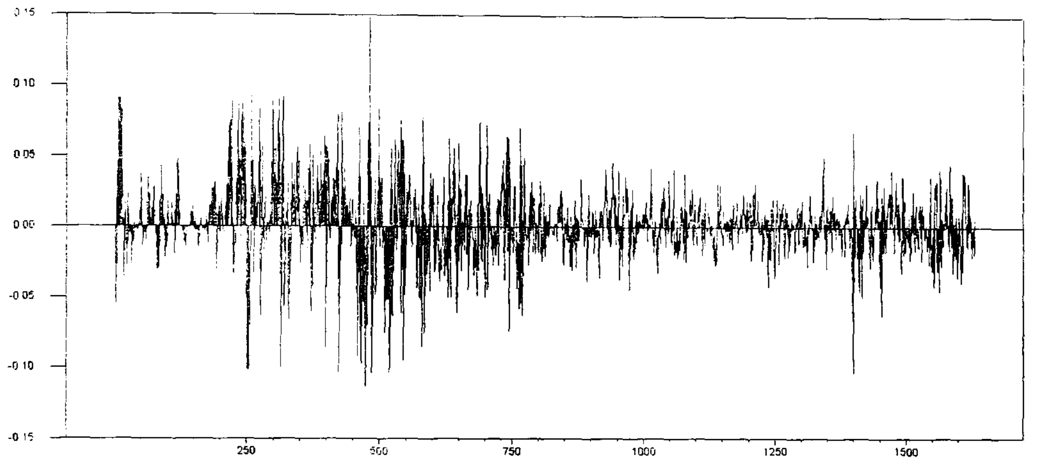


Figure 4.2.2

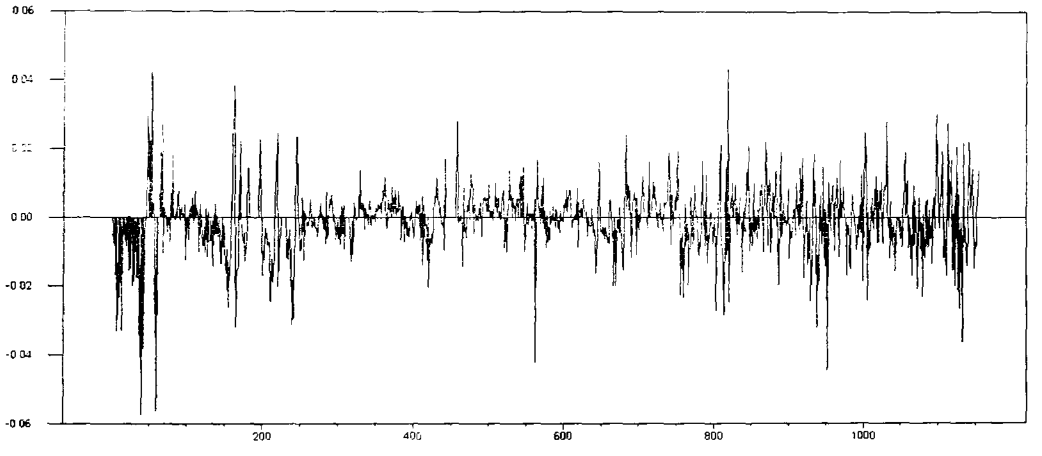
A. The Hungarian return series



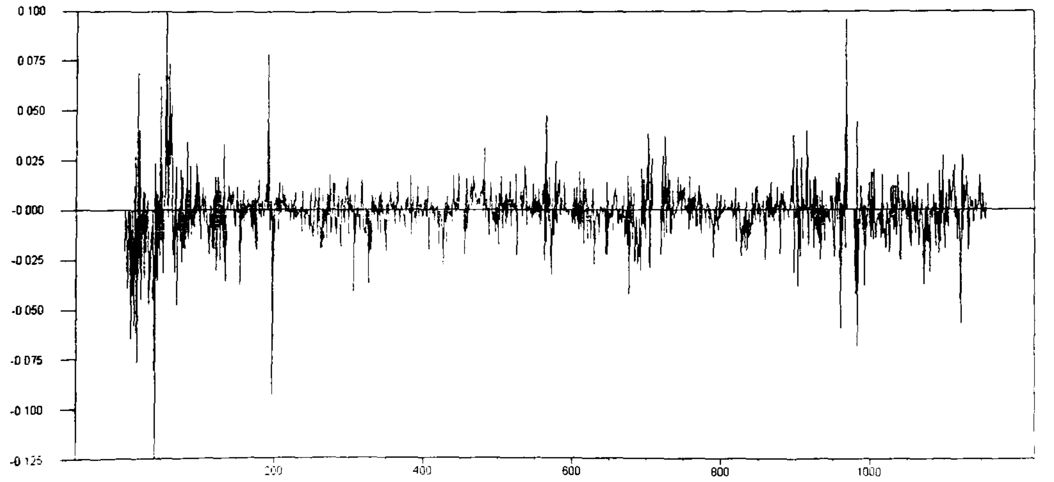
B. The Polish return series



C. The Czech return series



D. The Slovak return series



'igure 4.3 **Linear vs Nonlinear Predictability of Hungarian Returns**

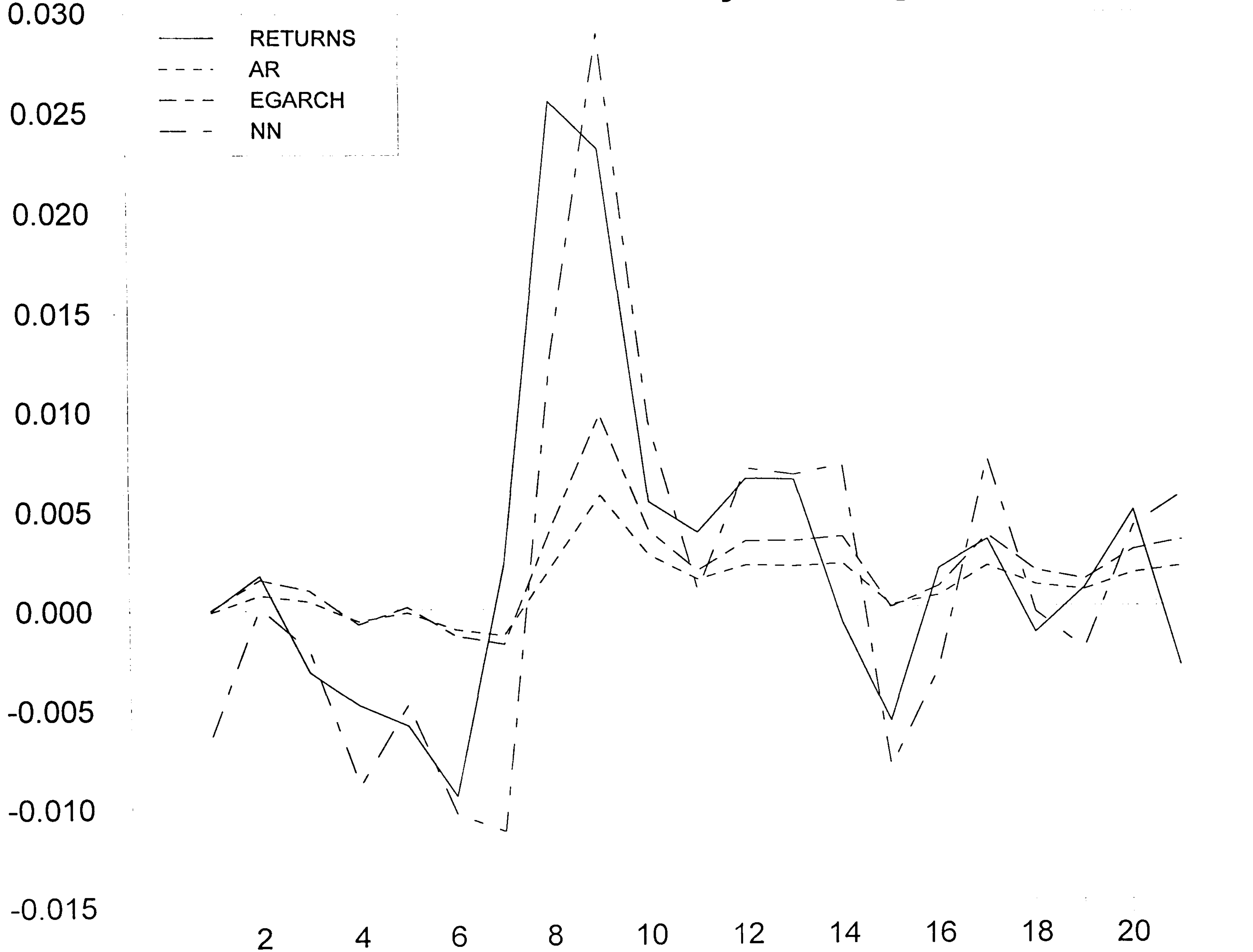


Figure 4.4 **Linear vs Nonlinear Predictability of Polish Returns**

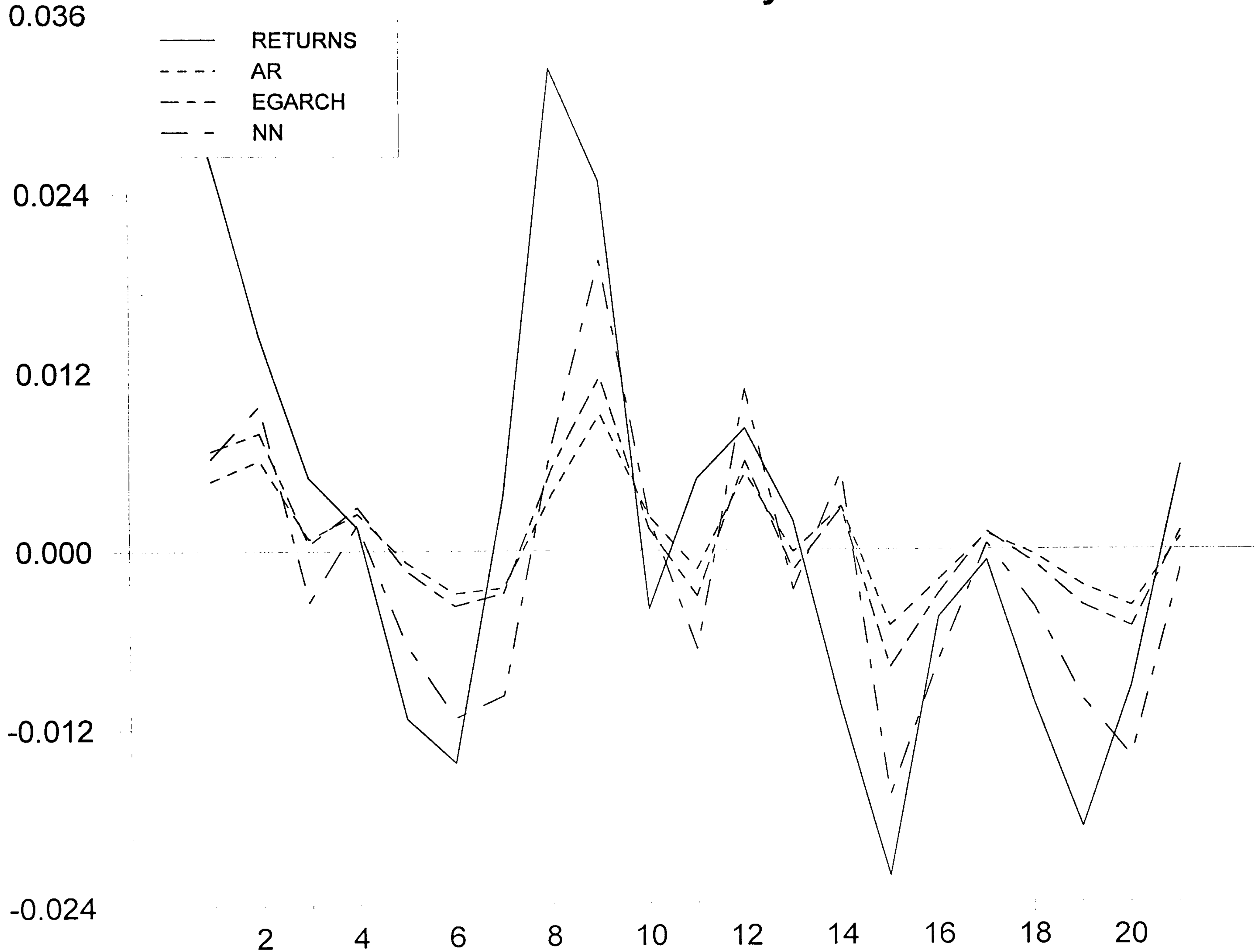


Figure 4.5 **Linear vs Nonlinear Predictability of Czech Returns**

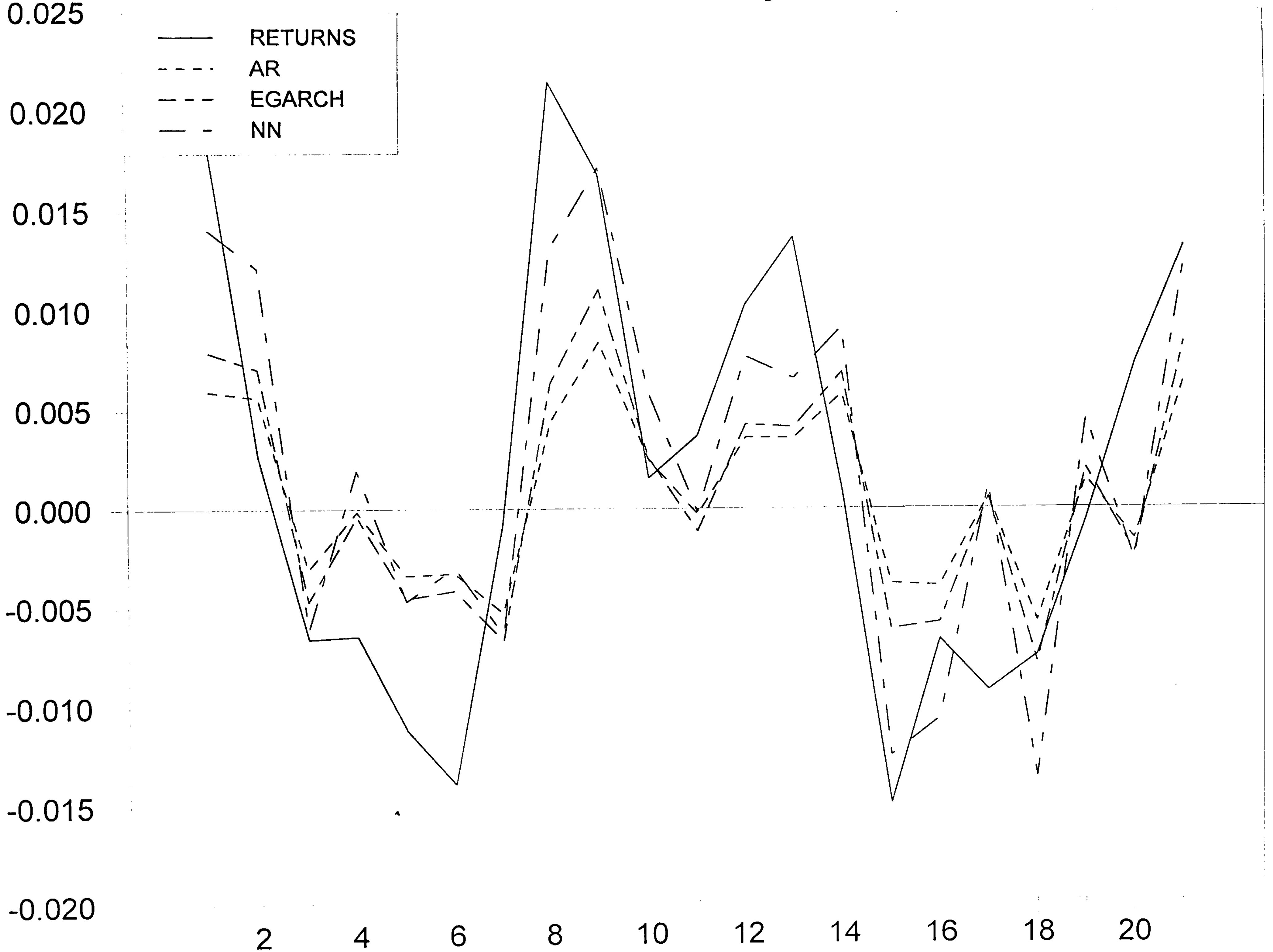
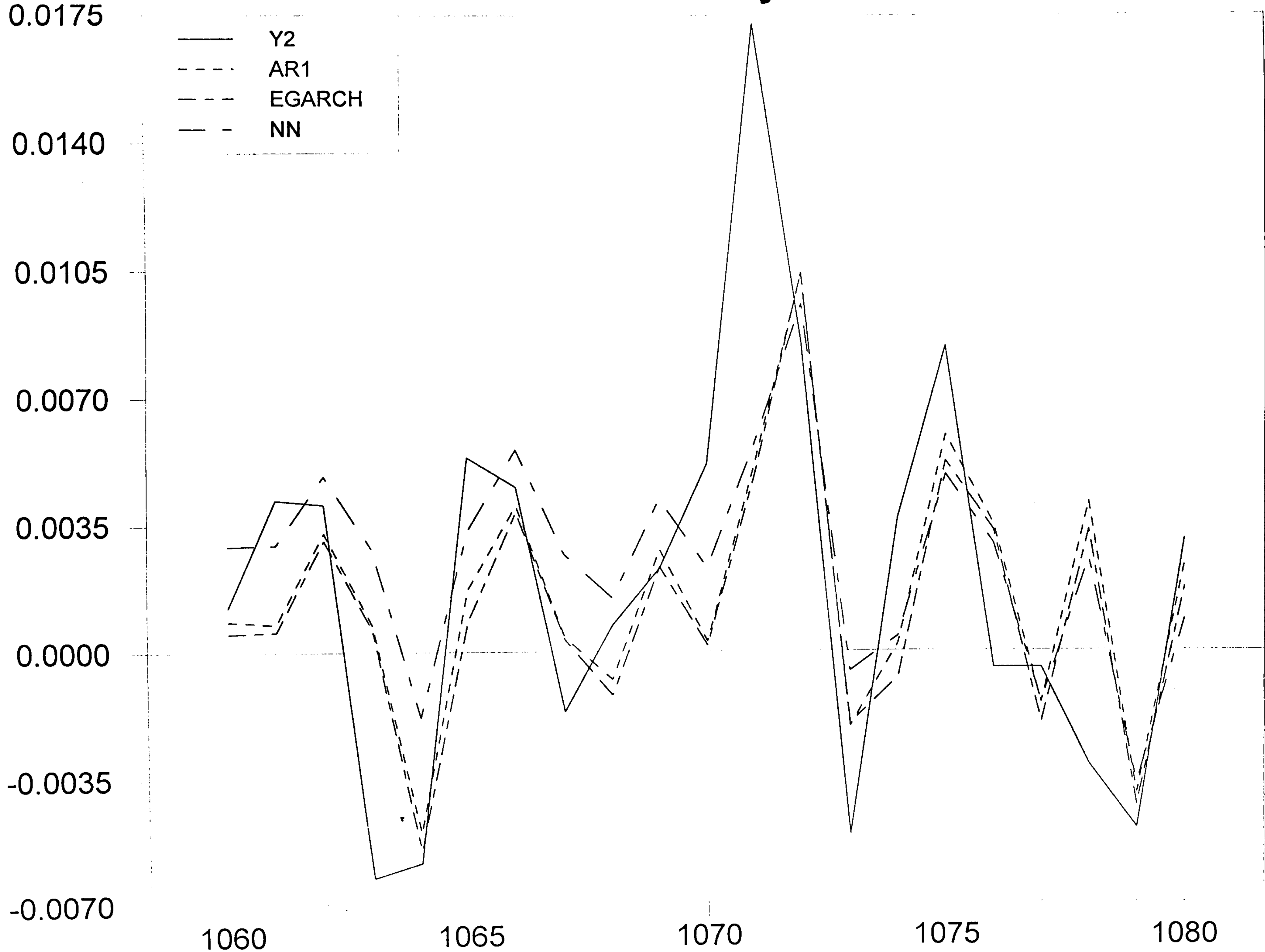


Figure 4.6 Linear vs Nonlinear Predictability of Slovak Returns



CHAPTER 5

INTERNATIONAL CAPITAL ASSET PRICING MODEL AND THE INTEGRATION OF THE TRANSITION MARKETS OF CENTRAL EUROPE INTO INTERNATIONAL CAPITAL MARKETS

5.1 Introduction

A major part of the research effort in finance is directed towards understanding why we observe a variety of financial assets with different expected rates of return. To appreciate the magnitude of this difference, note that in 1926 a nice dinner for two in New York would have cost about $10. If the same $10 had been invested in Treasury bills, by the end of 1991 it would have grown to $110, still enough for a nice dinner for two. Yet $10 invested in stocks would have grown to $6,756. The point is that the average return differentials among financial assets are both substantial and economically important.

A variety of asset pricing models have been proposed to explain this phenomenon. Differences among these models arise from differences in their assumptions that restrict investors' preferences, endowments, production and information sets; the stochastic process governing the arrival of news in the financial markets; and the type of frictions allowed in the markets for real and financial assets.

While there are differences among asset pricing models, there are also important commonalties. All asset pricing models are based on one or more of three central concepts. The first is the law of one price, according to which the prices of any two claims, which promise the same payoff must be the same. The law of one price arises as an implication of the second concept, the no-arbitrage principle. Arbitrage opportunities tend to be eliminated by trading in financial markets, because prices adjust as investors attempt to exploit them. The law of one price follows from the no-arbitrage principle, when it is possible to buy or sell two claims to the same future payoff. If the two claims do not have the same price, and if transaction costs are smaller than the difference between their prices, then an arbitrage opportunity is created. The third central concept behind asset pricing models is financial market equilibrium. Investor's desired holdings of financial assets are derived from an optimisation problem. A necessary condition for financial market equilibrium in a market with no frictions is that the first-order conditions of the investor's optimisation problem are satisfied. This requires that investors be indifferent at the margin to small changes in their asset holdings. Equilibrium asset pricing models follow from the first-order conditions for the investor's portfolio choice problem and from a market-clearing condition. The market clearing condition states that the aggregate of the investor's desired asset holdings must equal the aggregate "market portfolio" of securities in supply.

The earliest of the equilibrium asset pricing models is the capital asset pricing model (CAPM) developed in the early 1960s. The credit for the development of the CAPM is assigned to a number of individuals, including Sharpe, Lintner, Treynor, Mossin, Fama, and Black, with essential groundwork laid by Markowitz and Tobin. In his seminal research Markovitz (1959) considered the investor's portfolio selection problem in terms of expected return and variance of return. He argued that investors would optimally hold a mean-variance efficient portfolio, that is, a portfolio with the highest expected return for a given level of variance.[87] Sharpe (1964) and Lintner (1965)

[87] If the risk-averse agent is presented with a portfolio 'A' (of n securities) and portfolio 'B' (of a different set of n securities), then according to *mean variance criterion* (MVC), portfolio A is preferred to portfolio B if (i) $E_A(R) \geq E_B(R)$ and (ii) $\mathrm{var}_A(R) \leq \mathrm{var}_B(R)$. Portfolios that satisfy the MVC are known as the set of *efficient portfolios*.

built on Markowitz's work to develop economy-wide implications. Their model, called the Capital Asset Pricing Model, is based on some simplifying assumptions:

1. Agents can borrow and lend as much as they like at the risk free rate.
2. All individuals have homogeneous expectations about expected returns and the variances and covariances between the various returns, and optimally hold mean-variance efficient portfolios.
3. Capital markets are perfectly competitive and frictionless (i.e. no taxes, transaction costs or restrictions on short sales, and all assets are perfectly divisible and marketable).

According to Sharpe and Lintner, given the assumptions above, the portfolio of all invested wealth, or the market portfolio, will be a mean-variance efficient portfolio. The Capital Asset Pricing Model is a direct implication of the mean-variance efficiency of the market portfolio. In its simplest form the CAPM predicts that the expected return on an asset above the risk-free rate is proportional to the nondiversifiable risk, measured by the asset's *beta*, which is the covariance of the asset return with the market portfolio.[88] Merton (1973) extended the CAPM, which is a single period model, to an economic environment where investors make consumption, savings, and investment decisions repetitively over time.

In our analysis we test an international version of the CAPM (ICAPM). In the international CAPM investors are presumed to hold a diversified portfolio of equities from all national markets, that is, a world market portfolio. The ICAPM implies that if the particular market is the part of a global market (or integrated in the global market) then this market's expected return in excess of the safe rate should be proportional to the expected excess return on a world market portfolio.

[88] The CAPM is a single-beta model, where the expected return on an asset is a linear function of its market beta. The well-known alternative to the CAPM is multi-beta model by Ross (1976), the arbitrage pricing theory (APT), based on no-arbitrage principle discussed above. For a brief review of this alternative asset pricing model see Appendix 5.1 at the end of this chapter.

In this chapter the ICAPM is applied to a dataset including daily stock price data from both developed (U.S. and Germany) and relatively new transition equity markets (Hungary and Poland). This will allow us to address the issue of integration of the transition markets of Central Europe into the global capital market. The dynamics of the conditional covariance matrix of the assets in our system will be parameterized using multivariate GARCH specification, which allows volatility spillover effects between the markets. Methodologically our specification extends the multivariate GARCH model of Engle and Kroner (1995) ("BEKK") to accommodate GARCH-M effects. We compare the performance of the standard ICAPM (the benchmark model) to two alternative specifications (the partial segmentation and partial integration models) which can provide useful insights, in case the standard model is misspecified.

5.2 The Capital Asset Pricing Model: A Review

The aim of this section is to provide a review of the Capital Asset Pricing Model, including its derivation, extensions, tests and empirical applications.

5.2.1 Derivation of the CAPM

The expected return, ER_p, and standard deviation, σ_p, for any portfolio 'p' consisting of n risky asset and a risk free asset are:

$$ER_p = \sum_{1}^{n} x_i ER_i + [1 - \sum_{1}^{n} x_i] r \tag{5.1}$$

$$\sigma_p = \sqrt{\sum_{i=1}^{n} x_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij}} \qquad i \neq j \tag{5.2}$$

where x_i is proportion of wealth held in asset i, i =1, 2, ...n, ER_i is the expected return on asset i, and r is a risk free rate. The CAPM is the solution to the problem of minimising σ_p subject to a given level of expected return ER_p. The Lagrangian is:

$$L = \sigma_p + F\left[ER_p - \Sigma x_i ER_i - (1 - \Sigma x_i) r\right] \tag{5.3}$$

Choosing x_i to minimise L gives a set of first-order conditions of the form:

$$\partial L / \partial x_1 = \frac{1}{2}(\sigma_p^2)^{-1/2}[2x_1\sigma_1^2 + 2\sum_{j=2}^{n} x_j \sigma_{1j}] - F(ER_1 - r) = 0$$

$$\partial L / \partial x_2 = \frac{1}{2}(\sigma_p^2)^{-1/2}[2x_2\sigma_2^2 + 2\sum_{\substack{j=1 \\ j\neq 2}}^{n} x_j \sigma_{2j}] - F(ER_2 - r) = 0$$

$$\partial L / \partial x_n = \frac{1}{2}(\sigma_p^2)^{-1/2}[2x_n\sigma_n^2 + 2\sum_{\substack{j=1 \\ j\neq n}}^{n} x_j \sigma_{nj}] - F(ER_n - r) = 0$$

(5.4)

Differentiation with respect to F gives:

$$\partial L / \partial F = ER_p - \sum_{i=1}^{n} x_i ER_i - (1 - \sum_{i=1}^{n} x_i) r = 0 \tag{5.5}$$

Multiplying the first equation in (5.4) by x_1, the second by x_2, etc., summing over all equations and using (5.2) gives:

$$\sum_{i=1}^{n} x_i^2 \sigma_i^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij} = \sigma_p^2 = F\sigma_p[\sum_{i=1}^{n} x_i ER_i - \sum_{i=1}^{n} x_i r] \qquad i \neq j \tag{5.6}$$

From (5.6) we get

$$\sigma_p = F[\sum_{i=1}^{n} x_i ER_i - \sum_{i=1}^{n} x_i r] = F[\sum_{i=1}^{n} x_i ER_i + (1 - \sum_{i=1}^{n} x_i) r - r] \tag{5.7}$$

At the point where $\sum_{i=1}^{n} x_i = 1$ (5.7) gives:

$$\sigma_m = F(ER_m - r) \tag{5.8}$$

$$1/F = (ER_m - r)/\sigma_m \tag{5.9}$$

where $ER_m = (\sum x_i ER_i)$ and m denotes the market portfolio. From (5.9) we see that the term 1/*F* is the slope of the capital market line. It measures the price of a unit risk, which is the same for all investors. Define the *price of market risk*, λ, as follows:

$$\lambda = 1/F\sigma_m = \frac{ER_m - r}{\sigma_m^2} = \frac{ER_m - r}{\text{var}(R_m)} \tag{5.10}$$

Since the first order conditions in (5.4) hold for all investors independent of the "tastes" we can use this to derive the CAPM expression for equilibrium returns on each individual asset in the portfolio.

The *i*th equation in (5.4) at the point $\sum_{i=1}^{n} x_i = 1$ is:

$$ER_i = r + \frac{1}{F\sigma_m}[x_i\sigma_i^2 + \sum_{j=1}^{n} x_j\sigma_{ij}] \qquad i \neq j \tag{5.11}$$

Substitute from (5.10) into (5.11)

$$ER_i = r + \frac{ER_m - r}{\sigma_m^2}[x_i\sigma_i^2 + \sum_{j=1}^{n} x_j\sigma_{ij}] \qquad i \neq j \tag{5.12}$$

Note that:

$$\sigma_{im} = \text{cov}(R_i, R_m) = \text{cov}[R_i, (x_1R_1 + x_2R_2 + \ldots + x_nR_n)] = x_i\sigma_i^2 + \sum_{j=1}^{n} x_j\sigma_{ij} \qquad i \neq j \tag{5.13}$$

and substituting (5.13) in (5.12) we obtain the CAPM expression for the equilibrium expected return on asset *i*:

$$ER_i = r + (ER_m - r)\beta_i \tag{5.14}$$

where

$$\beta_i = \sigma_{im} / \sigma_m^2 . \tag{5.15}$$

The one-period CAPM in (5.14) predicts that in equilibrium all investors will hold the market portfolio (i.e. all risky assets will be held in the same proportions in each individual's portfolio). Merton (1973) developed this idea in an intertemporal framework, when agents face investment decisions at discrete time steps t, $t = 1, 2, \ldots$, and showed that the expected return on the market portfolio in excess of risk free rate is proportional to the expected variance of returns on the market portfolio

$$E_t R_{mt+1} = r_t + \lambda E_t \sigma_{mt+1}^2 . \tag{5.16}$$

The Merton's intertemporal CAPM implies that to be willingly held as part of a diversified portfolio the expected return on asset (or portfolio of assets which is a subset of the market portfolio) i is given by:

$$E_t R_{it+1} = r_t + \lambda E_t \sigma_{imt+1_t} = r_t + \beta_{it}(E_t R_{mt+1} - r_t) \tag{5.17}$$

where

$$\beta_{it} = E_t(\sigma_{imt+1} / \sigma_{mt+1}^2) . \tag{5.18}$$

As a result, according to the CAPM, the market only rewards investors for nondiversifiable systematic risk. The required rate of return to willingly hold an individual stock as part of a wider portfolio is equal to the risk-free rate plus a reward for risk or risk premium.

5.2.2 Critique of the Tests of the CAPM.

One of the most controversial papers written on the CAPM is that by Roll (1977). A full presentation of Roll's argument need not be given here, but his major points are discussed.

Following Roll's argument, the CAPM is a general equilibrium model based on the existence of the portfolio that is defined as the value weighted portfolio of all investment assets. Furthermore, the market portfolio is defined as *ex ante* mean-variance efficient. Roll shows that the only true test of the CAPM is whether the market portfolio is in fact ex ante mean-variance efficient. However, the true market portfolio has not yet been observed, since it includes all investment assets (e.g., stocks, bonds, real estate, art objects, and human capital).

The main points of the argument are summarised below:

1. Tests of the CAPM are extremely sensitive to which market proxy is used, even though returns on most market proxies (e.g., the S&P 500 and the NYSE index) are highly correlated.
2. A researcher cannot unambiguously discern whether the CAPM failed the test because the true market portfolio was ex ante mean-variance inefficient or because the market proxy was inefficient. The researcher also cannot unambiguously discern whether the test supported the CAPM because the true market portfolio was ex ante mean-variance efficient or because the market proxy was efficient.
3. The effectiveness of variables such as dividend yield in explaining risk-adjusted asset returns is evidence that the market proxies used to test the CAPM are not ex ante mean-variance efficient.

Hence, Roll submits that the CAPM is not testable until the exact composition of the true market portfolio is known, and the only valid test of the CAPM is to observe whether the ex ante true market portfolio is mean-variance efficient. As a result of his findings, Roll does not believe that there ever will be an unambiguous way to test the CAPM owing to the nonobservability of the true market portfolio and its characteristics.

Despite this critique researchers have continued to explore the empirical validity of the CAPM even though their proxy of the market portfolio could be incorrect. This is because it is still of interest to see how far a particular empirical model, even an imperfect one, can explain equilibrium returns.

5.2.3 Direct tests of the CAPM

Direct tests of the CAPM use a *two-pass regression technique*. Assuming that β_i is constant over time, in the *first-pass* time-series regression $(R_{it} - r_t)$ is regressed on $(R_t^m - r_t)$ and a constant term:

$$(ER_{it} - r_t) = \alpha_i + \beta_i (ER_t^m - r_t) + \varepsilon_{it} \tag{5.19}$$

On the basis of the finance theory underlying the CAPM we expect the estimates of the constant term, α_i, to be close to zero for each security or for a portfolio of securities (i = 1, 2,... k). In the *second-pass* cross-section regression the sample mean returns $\overline{R_i}$ are regressed on the estimates of *beta*s, $\hat{\beta}$s, for each security from the first-pass regression:

$$\overline{R_i} = \psi_0 + \psi_1 \hat{\beta}_i + v_i \tag{5.20}$$

Given the CAPM relation we expect:

$$\psi_0 = \overline{r_t}, \qquad \psi_1 = \overline{R^m} - \overline{r_t}$$

An even stronger test of the CAPM in the second-pass regression is to note that if there is an unbiased estimate of β_i, then only β_i should influence R_i. Then the second-pass can be based on the following cross-section regression:

$$\overline{R_i} = \psi_0 + \psi_1\beta_i + \psi_2\beta_i^2 + \psi_3\sigma_{\varepsilon i}^2 + \eta_i \qquad (5.21)$$

where $\sigma_{\varepsilon i}^2$ is an unbiased estimate of the variance of security i from the first-pass regression. The null hypothesis is:

$$H_0 = \psi_2 = \psi_3 = 0$$

1. $\psi_2 \neq 0$ would indicate the presence of non-linearities in the security market line.

2. $\psi_3 \neq 0$ would mean that the diversifiable risk affects the expected return on a security.

Both 1. and 2. above are violations of the CAPM.

There are a number of econometric problems with the use of the two step approach.

1. The CAPM does not rule out the possibility that the variances of the error terms are not constant (i.e. the error terms may be heteroskedastic). In this case the OLS standard errors are incorrect and other estimation techniques need to be used.

2. The CAPM does not rule out the error terms being serially correlated over time. In this case one could use, for instance, the GMM estimator.

3. The estimate $\hat{\beta}_i$ from the first-pass regression may be unbiased, but measured with error. This would mean that in the second-pass regression one would face so-called 'error in variables' problem, which would mean that the OLS coefficient of ψ_1 is downward biased and ψ_0 is upward biased. Besides if the true β_i is positively correlated with the variance of the error $\sigma_{\varepsilon i}^2$, then the latter serves as a proxy for the true β_i and as a result, if $\hat{\beta}_i$ is measured with error, then $\sigma_{\varepsilon i}^2$ may be significant in the second-pass regression.

4. In case the distribution of the error term ε_{it} is non-normal then estimation technique based on normality will produce biased estimates. For example, positive skewness in the residuals of the cross-section regressions will show up as an association between residual risk and return even though in the true model there is no association.

As an illustration of the early studies addressing the problems listed above can be considered that of Black et al. (1972) using monthly rates of return in the first-pass time series regressions. They have minimized the heteroskedasticity problem and the error in estimating the betas by grouping all stocks into a set of ten portfolios based on the size of the betas for individual securities. In the first-pass time series regression the return R_{it}^{p} for each of the ten portfolios is regressed on R_t^m. The results of the second pass regression (the R^2 of the regression is 0.98), in general support the CAPM.

Fama and MacBeth (1974) using the same approach of grouping all stocks into a set of 20 portfolios based on the size of the betas for individual securities extended the second-pass cross-section regression used in Black et al (1972) by including β_i^2 and $\sigma_{\varepsilon i}^2$ from the first-pass regression to get (5.21) above.[89] The second-pass regression is repeated for each month over 1935-1968 to see how the ψ s vary over time. They find that the average of the ψ_2 and ψ_3 estimates are not significantly different from zero and hence support the standard CAPM. They also find that η_i is not serially correlated, which is again in support of the CAPM.

[89] More recent empirical studies incorporate the other cross-sectional variables such as the relative size of firms, measured by the market value of their equity, the ratio of book-to-market-equity, and related variables. Fama and French (1992) is a prominent recent example of this approach. Berk (1995) provides a justification for using relative market value and book-to-price ratios as measures of expected returns.

5.2.4 Consumption CAPM

An alternative view on the determination of the equilibrium returns in a well-diversified portfolio is provided by the intertemporal consumption CAPM (C-CAPM). This view was developed by Lucas in the paper 'Asset Prices in an Exchange Economy' in 1978 and Mankiw and Shapiro in the paper 'Risk and Return: Consumption Beta Versus Market Beta' in 1986. If in the standard one-period CAPM the individual investor's objective function is assumed to be fully determined by the standard deviation and return on the portfolio, then in the case of the C-CAPM the investor maximises expected utility which depends only on current and future consumption.

Following Mankiw and Shapiro (1986), suppose that the individual investor maximises:

$$E_t\left[\sum_{j=0}^{\infty}\theta^j U(C_{t+j})\right], \qquad 0<\theta<1 \tag{5.22}$$

where $U(C)$ is the utility function and $\theta=(1+t_p)^{-1}$ where t_p is the individual's subjective rate of time preference for the consumption today versus consumption tomorrow assumed to be constant. The first order condition for maximising utility has the agent equating the utility loss from the reduction in current consumption, with additional gain in consumption next period. A \$1 reduction in consumption today reduces utility by $U'(C_t)$ but results in an expected payout of $\$(1+E_tR_{it+1})$ next period. When spent on the next period's consumption, the extra utility per \$ next period discounted at the rate θ_t is $\theta_t U'(C_{t+1})$. As a result in equilibrium we have:

$$U'(C_t)=E_t\left[(1+E_tR_{it+1})\theta U'(C_{t+1})\right] \tag{5.23}$$

Suppose, we define S_{t+1} as the marginal rate of substitution of current for future consumption:

$$S_{t+1}=\theta U'(C_{t+1})/U'(C_t), \tag{5.24}$$

and use the specific form of the utility function:

$$U(C_t) = \frac{C_t^{1-a}}{1-a}, \qquad 0 < a < 1 \tag{5.25}$$

where a is a constant Arrow-Pratt measure of relative risk aversion. Then one can derive the expression for the C-CAPM in the form similar to the standard CAPM[90]:

$$E_t(R_{it+1}) = \gamma_{0t} + \gamma_{1t}\left[\text{cov}(R_{it+1}, g_{t+1}^{\ c})\right] \tag{5.26}$$

where

$$g_{t+1}^{\ c} = \frac{C_{t+1}}{C_t} \text{ (the growth in consumption)} \tag{5.27}$$

$$\gamma_{0t} = 1 - ES_t / ES_t \tag{5.28}$$

$$\gamma_{1t} = a\theta / ES_t \tag{5.29}$$

As a result, the interpretation of the C-CAPM is as follows. If the return on the asset i has a low covariance with consumption growth then it will be willingly held, even, if its expected return is relatively low. This is because the asset on average has a high return when consumption is low and hence can be sold to finance current consumption, which has high marginal utility.

The C-CAPM model presented above can be tested either using cross-section or time series data. *Cross-section test* is based on the following rearrangement of (5.26), which represents security market line for this model.

$$ER_i = \alpha_0 + \alpha_1 \beta_{ci} \tag{5.30}$$

where

$$\alpha_0 = (1 - ES) / ES \tag{5.31}$$

$$\alpha_1 = a\theta\,\text{cov}(R^m, g^c) \tag{5.32}$$

[90] See Mankiw and Shapiro (1986) for a detailed derivation.

$$\beta_{ci} = \mathrm{cov}(R_i, g^c) / \mathrm{cov}(R^m, g^c) \tag{5.33}$$

The expected value of the marginal rate of substitution ES and $\mathrm{cov}(R^m, g^c)$ are assumed to be constant over time. Note that α_1 is the same for all stocks, as it depends only on market return R^m and consumption growth g^c. Sample mean value $\overline{R_i}$ can be used as a proxy for ER_i.

Mankiw and Shapiro (1986) test the basic CAPM and the C-CAPM using cross section data on 464 companies listed on the NYSE. They find that the basic CAPM clearly outperforms the C-CAPM, since when the mean value $\overline{R_i}$ is regressed on both β_{mi} (from the basic CAPM) and β_{ci} (from C-CAPM), β_{mi} is statistically significant while the latter is not.

Equation (5.24) suggests a way in which we can perform *time-series test* of the

C-CAPM. As it shown in Scott (1991), if we assume a constant relative risk aversion utility function (like e.g. in (5.25) above), and rational expectations ($R_{it+1} = ER_{it+1} + \varepsilon_{it+1}$), then (5.24) for two assets *i* and *j* becomes:

$$\left[(1 + R_{it+1})\theta(g_{t+1})^{-a} - 1\right] = \varepsilon_{it+1}$$

$$\left[(1 + R_{jt+1})\theta(g_{t+1})^{-a} - 1\right] = \varepsilon_{jt+1} \tag{5.34}$$

Equations in (5.34) imply a set of cross-section restrictions, since the parameters (θ, α) appear in equation for any asset or portfolio *i*. The restrictions suggest that expected returns on all assets are proportional. If we estimate these equations jointly using time-series data, these cross-section restrictions provide a strong test of the C-CAPM. However, the studies of Hodrick (1987), Cumby (1990) and Smith (1993) show that this model does not appear to perform well for a wide range of alternative assets included in the portfolio by the authors. The Smith's studies show that the considered parameters (θ, α) are not constant over time. Hence the model or some of its assumptions appear to be invalid.

Flood et al. (1987) apply (5.34) to the aggregate stock market return, and, noting that (5.34) applies to any return horizon $t+j$, have for $j=$ 1,2,... :

$$\left[(1+R_{t+j}^{m})\theta(g_{t+1})^{-a}-1\right]=\varepsilon_{t+j} \tag{5.35}$$

Equation (5.35) can be estimated on time series data and because it holds for horizons $j=$ 1,2,... , we again have a system of equations with common parameters (θ,α). Flood et al find that the C-CAPM represented by (5.35) performs worse, in statistical terms, as the time horizon is extended. In general, the C-CAPM does not appear to perform well in empirical tests.

5.2.5 Changing Perception of Risk: CAPM with Time-Varying Risk Premia

The evidence suggests that equilibrium returns on an aggregate stock market index are influenced by agents' changing perception of risk. Hence the CAPM with time varying risk-premia would appear to be an improvement on the assumptions of the standard CAPM. Modelling of time-varying risk-premia has been one of the recent growth areas in empirical research on asset prices. It may seem strange that financial economists have only recently focused on most obvious attribute of holding assets, namely that they are risky and that perceived riskiness is likely to vary substantially over different historical periods. In this section we present a number of alternative models, which examine the validity of the CAPM under the assumption that equilibrium returns on assets depend on time-varying risk premium determined by conditional variances and covariances.

Poterba and Summers (1988) investigate whether changes in investor's perception of risk are large enough to account for the very sharp movements in stock prices that are actually observed. According to the rational valuation formula (RVF) for stock prices:

$$P_t = E_t\left[\sum_{j=1}^{\infty} \gamma_{t+j} D_{t+j}\right] \tag{5.36}$$

and

$$\gamma_{t+j} = 1 / \prod_{i=1}^{J} (1 + r_{t+i} + rp_{t+i}) \tag{5.37}$$

where γ_t is discount rate, D_t are dividends paid, r_t is risk-free rate and rp_t is risk premium. As a result, the stock prices may vary if forecasts of expected dividends are revised or if the discount factor changes: Poterba and Summers concentrate on the behaviour of the latter. They argue that growth in future dividends is fairly predictable and concentrate instead on volatility and persistence of the risk premium rp_t.

The authors of the article treat all variables in the RVF in real terms and consider P_t as the real stock price of the market portfolio.[91] The assumptions of the model are:

1. Dividends, D_t, grow at a constant rate g, such that $E_t(D_{t+j}) = (1+g)^j D_t$.

2. The risk-free rate is constant, $r_t = r$.

The CAPM suggests that the risk premium on the market portfolio is proportional to the conditional variance of forecast errors

$$E_t R_{t+1} = r_t + \lambda E_t \sigma_{t+1}^2 = r_t + rp_t \tag{5.38}$$

where $\lambda E_t \sigma_{t+1}^2 = rp_t$ is the risk premium. In Merton's intertemporal CAPM the market price of risk, λ, depends on different consumers' relative risk aversion parameters which are assumed to be constant. Poterba and Summers assume and later verify empirically that volatility can be represented by ARMA models and consider an AR(1) process:

$$\sigma_{t+1}{}^2 = \alpha_0 + \alpha_1 \sigma_t{}^2 + \nu_{t+1} \quad \text{with } 0 > \alpha_i > 1 \tag{5.39}$$

[91] As the approximation of the market portfolio index they use S&P index over the period 1928-1984.

where v_{t+1} is a white noise error. If α_1 is small, the degree of persistence in volatility is small and vice versa. From (5.38) if σ_t^2 follows an AR(1) process then so will rp_t:

$$rp_t = \lambda\alpha_0 + \lambda\alpha_1 rp_{t-1} \tag{5.40}$$

Poterba and Summers linearise (5.36) around the mean value of the risk premium $\overline{rp}$ and calculate percentage response of P_t to the percentage change in σ_t^2:

$$\frac{\partial[\ln P_t]}{\partial[\ln(\sigma_t^2)]} = \frac{-\overline{rp}}{[1+\bar{r}+\overline{rp}-\alpha_1(1+g)]} \tag{5.41}$$

As a result, we see that the response of P_t to a change in volatility σ_t^2 increases with the degree of persistence α_1. Poterba and Summers compute an unconditional volatility measure for the variance of monthly stock returns based on average daily change in the S&P composite index over a particular month. For the month t:

$$\hat{\sigma}_t^2 = \sum_{i=1}^{m} s_{ti}^2 / m \tag{5.42}$$

where s_{ti} is the daily change in the stock index in month t and m is the number of trading days in the month. Then they investigate the persistence in $\hat{\sigma}_t^2$ using the number of alternative ARMA models. For example, for AR(1) model they obtain a value of α_1 in the range 0.6 - 0.7. Substituting the values of the average return on Treasury bills ($\bar{r}$ = 0.4 percent per annum = 0.035 percent per month), average growth rate of real dividends (g = 0.01 percent per annum = 0.00087 percent per month), mean risk premium: average value of the excess return on the market portfolio $(ER^m - r)$ ($\overline{rp}$ = 0.006 per month), they find that a 50 percent increase in volatility ($\hat{\sigma}_t^2$) depresses the share prices by only 0.7 percent. As a result, according to Poterba and Summers the measured persistence in volatility is too low to explain the observed sharp movements in stock prices.

Chou (1988) argues that the Poterba-Summers estimation of the time varying variance σ_t^2 is not a correct measure of conditional variance, as it remains constant within a given horizon (i.e. a month) and is then assumed to vary over a longer horizons. Besides Pagan and Ullah (1988) show that the Poterba-Summers two step estimation technique yields inconsistent estimates for parameters. Chou repeats Poterba-Summers analysis using GARCH(1,1) model for conditional variance. Chou, as in Poterba and Summers, assumes expected returns on the market portfolio are given by the CAPM (plus RE):

$$(R_{t+1} - r_t) = \lambda E_t \sigma_{t+1}^2 + \varepsilon_{t+1} \tag{5.43}$$

Taking expectations it is easy to see that the best forecast of the excess return depends on the best forecast of the conditional variance:

$$(E_t R_{t+1} - r_t) = \lambda E_t \sigma_{t+1}^2 \tag{5.44}$$

Therefore, the conditional forecast error is:

$$R_{t+1} - E_t R_{t+1} = \varepsilon_{t+1} \tag{5.45}$$

where we assume ε_{t+1} is $N(0, \sigma_{t+1}^2)$

The conditional variance of the forecast error:

$$E_t (R_{t+1} - E_t R_{t+1})^2 = E_t(\varepsilon_{t+1}^2) = E_t(\sigma_{t+1}^2) \tag{5.46}$$

According to the CAPM the expected excess return varies directly with the time varying variance of the forecast errors: more risk requires compensation in the form of higher expected returns. To describe the time-path of the conditional variance Chou assumes a GARCH(1,1) model, where the conditional variance is the weighted average of last periods conditional variance σ_t^2 and the squared forecast error ε_t^2:

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \alpha_2 \sigma_t^2 \tag{5.47}$$

Chou estimates (5.43) and (5.47) simultaneously (GARCH-M model) using the maximum likelihood method. He uses the data for weekly returns on the NYSE index over the period 1962-1985. Chou finds that the estimate of the market price of risk λ over various periods is not well determined statistically, however it has plausible point estimates in the range 3 – 6.[92] The value of $(\alpha_1 + \alpha_2)$ is very stable over subperiods and is around 0.98. This fact indicates substantial persistence in the conditional variance. Chou, which uses Poterba-Summers framework, shows that when $\alpha_1 + \alpha_2 = 1$ (which he founds to be largely acceptable on statistical grounds), then stock prices move tremendously and the elasticity $\partial(\ln P_t) / \partial(\ln \sigma_t^2)$ can be as high as -60.

When Chou estimates his GARCH(1,1) model using returns over N = 5, 20, 50 and 250 trading days, he finds that estimates of $\alpha_1 + \alpha_2$ are very stable at around 0.95 in all cases. For the sake of comparison he calculates the Poterba-Summers measure of variance (5.42) using different values of N. He then estimates $\sigma^2_{Nt+1} = \alpha_0 + \alpha_1 \sigma^2_{Nt} + v_{t+1}$ for these various values of N. He finds that the degree of persistence, α_1, varies a lot for different N ranging from near zero to 0.9. This suggests that Poterba-Summers method may not have correctly captured the true degree of persistence.

Attanasio and Wadhwani (1990) start with the empirical observation that, from previous empirical work, it is known that the expected return on the aggregate stock market index (the proxy of the market portfolio) is often significantly related on the previous period's dividend price ratio. This fact violates the EMH under constant expected excess returns. The literature in this area often assumed a constant risk premium but then sometimes interpreted the presence of the dividend yield as indicative of a time-varying risk premium. Attanasio and Wadhwani suggest that if we explicitly model the time-varying risk premium then we may find that the dividend yield (Z_t) does not influence expected returns. This would support the CAPM version of the EMH. The authors

[92] Poterba and Summers obtain the value of 3.5.

assume that the time-varying conditional variance of the market return is given by the GARCH(1,2) model with the added dividend yield. They use monthly data on an aggregate US stock price series over period 1953-1988, and get the following results:

$$\left[R_{t+1}^{m} - r_t\right] = -0.035 + 0.55Z_t - 4.05r_t + 22.3\sigma_{t+1}^2 \qquad (5.48)$$
$$(0.025) \quad (0.39) \quad (1.05) \quad (11.3)$$

$$\sigma_{t+1}^2 = \alpha_0 + 0.87\sigma_t^2 + 1.5(10^{-2})\varepsilon_t^2 + 2.2(10^{-2})\varepsilon_{t-1}^2 + 5.3(10^{-2})Z_t \qquad (5.49)$$
$$(0.06) \quad (2.9.10^{-2}) \quad (4.0.10^{-2}) \quad (2.4.10^{-2})$$

where figures in parentheses are standard errors.

As we see in the CAPM relationship (5.48) the conditional variance is found to be significant and the dividend yield is not, thus supporting the CAPM. However, the results in (5.48) are reported when the authors also include the short rate, which is statistically significant. The latter rejects the CAPM model of equilibrium returns. In the GARCH equation (5.49), the dividend yield has a statistically significant effect on the conditional variance, and this would explain why previous studies, which assumed a constant risk premium, found dividend yield significant in the CAPM equation.

Sentana and Wadhwani (1992) introduce noise traders into the market, who do not base their asset decisions on the fundamental value. Positive feedback traders buy after a price rise and sell after a price fall. This causes positive serial correlation in returns, since price rises are followed by an increase in demand and further price rises next period. Negative feedback traders pursue the opposite strategy, they buy ‘low’ and sell ’high’. Hence, if market is dominated by the negative feedback traders, then a price fall would be followed by the price rise. The demand for stocks by noise traders, N_t, as a proportion of the total market value of stocks, may be presented

$$N_t = \gamma R_{t-1} \ , \tag{5.50}$$

where $\gamma > 0$ in case of positive feedback traders and $\gamma < 0$ in case of negative feedback traders, and R_{t-1} is the holding period return in the previous period. Sentana and Wadhwani assume that the demand for shares by the smart money is given by a simple mean-variance model

$$S_t = \left[E_t R_t - \alpha\right] / \mu_t , \tag{5.51}$$

where S_t = proportion of stock value held by smart money, α = expected rate of return for which demand by the smart money is zero, μ_t = measure of perceived riskiness of shares, which is a positive function of the conditional variance of stock returns ($\mu = \mu(\sigma^2)$). If the smart money holds all the stocks, then $S_t = 1$ and rearranging (5.51) we get the CAPM for the market portfolio: the excess return on the portfolio, $(E_t R_t - \alpha)$, depends on the risk premium, which is proportional to the conditional variance of stock prices ($\mu_t = c\sigma_t^2$). In equilibrium:

$$S_t + N_t = 1 \tag{5.52}$$

Now if we substitute (5.50) and (5.51) into (5.52) and use rational expectations we get:

$$R_t = \alpha + \mu(\sigma_t^2) - \gamma\mu(\sigma_t^2)R_{t-1} + \varepsilon_t \tag{5.53}$$

Thus in a market with smart money and noise traders the serial correlation in R_t will depend on the type of noise traders. Somewhat paradoxically a positive feedback trader ($\gamma > 0$) results in negative serial correlation in R_t, while negative ($\gamma < 0$) in positive serial correlation in R_t. A linear form for $\mu\sigma_t^2$ in equation (5.53) gives:

$$R_t = \alpha + \theta\sigma_t^2 + (\gamma_0 + \gamma_1\sigma_t^2)R_{t-1} + \varepsilon_t \tag{5.54}$$

The direct impact of the feedback traders is given by the sign of γ_0. Now suppose γ_0 is positive (i.e. positive serial correlation in R_t) but γ_1 is negative. In this case, as risk σ_t^2 increases the coefficient on R_{t-1} could change sign and the serial correlation in stock returns would move from positive to negative. This suggests that as volatility increases the market becomes more dominated by positive feedback traders, who interact in the market with the smart money, resulting in overall negative serial correlation in returns. According to the model in (5.54) the switch point for the change from positive correlation in returns to negative serial correlation is $\sigma_t^2 > (-\gamma_0 / \gamma_1)$.

The model in (5.54) is estimated together with a complex GARCH model, which allows the number of non-trading days to influence the conditional variance. Using US daily stock market data, Sentana and Wandhwani find $\gamma_0 = 0.09$, $\gamma_1 = -0.01$ and the switch point is $\sigma_t^2 > 5.8$. Hence when volatility is low stock returns exhibit positive serial correlation, but when volatility is high returns exhibit negative autocorrelation. This model therefore provides some statistical support for the view that the relative influence of positive and negative feedback traders may vary with the degree of risk but it doesn't explain why this might happen.

The authors find that θ in (5.54) is not statistically different from zero, so that the influence of volatility on the mean return on stocks only works through non-linear interaction variable $\gamma_1 \sigma_t^2 R_{t-1}$. Thus the empirical results are not in complete conformity with the theoretical CAPM model.

5.3 International Asset Pricing and CAPM

Asset pricing studies can be classified in three broad categories: *segmented markets*, *integrated markets*, and *partially segmented markets*. An example of an asset pricing study that assumes markets are segmented is one that 'tests' a model like the CAPM of Sharpe (1964), Lintner (1965) and Black (1972), using one country's data (see examples in sections 5.2.3 - 5.2.5 of this chapter). Many of the seminal U.S. asset pricing studies assume that the United States is a completely segmented market – or that the market proxy represents a broader world market return. While this might have been a reasonable working assumption through the 1970s, in the 1980s the U.S. equity capitalisation dropped below 50 percent of the world market capitalisation. Japan's market capitalisation exceeded the United States (albeit briefly) in 1989.

The second class of asset pricing studies assumes that world capital markets are perfectly integrated. These include studies of a world CAPM (see e.g. Buckberg (1995), De Santis and Gerard (1997), Harvey (1991) and references therein), a world CAPM with exchange risk (see e.g. Dumas and Solnik (1995), Dumas (1994), and De Santis and Gerard (1998)), a world consumption-based CAPM (see e.g. Wheatley (1988)), world arbitrage pricing theory (see e.g. Solnik (1983) and Cho, Eun, and Senbet (1986)), world multibeta models (see e.g. Ferson and Harvey (1993, 1994)) and world latent factor models (see e.g. Campbell and Hamao (1992), Bekaert and Hodrick (1992), and Harvey, Solnik, and Zhou (1994)). Rejection of these models can be viewed as a rejection of the fundamental asset pricing model, inefficiency in the market, or rejection of market integration.

Yet another strand of the literature falls in between segmentation and integration – the so-called mild segmentation models (see Errunza, Losq, and Padmanabhan (1992) and references therein). The advantage of these models is that the polar segmented/integrated cases are not assumed. The disadvantage of these models is that in most of these studies the degree of market segmentation is fixed through time. This does not allow for possibility that some markets have become more integrated/segmented through time.

The international CAPM (or world CAPM) belongs to the class of asset pricing models, which assume that international capital markets are fully integrated. In the international CAPM (ICAPM) investors are presumed to hold a diversified portfolio of equities from all national markets, that is, a world market portfolio. In the presence of exact purchasing power parity (PPP), the classic

CAPM would hold internationally, i.e. the risk of the individual market would be measured by its contribution to (or covariance with) the world market portfolio.

There are many possible sources of statistical rejection of the CAPM, and the ICAPM, in particular. Among these, we may single out four. First, the fundamental assumptions that provide the building blocks for this model, such as utility specification, information environment, or distributional assumptions, could be violated. Second, the benchmark portfolio that is used to measure risk could be improperly specified. Third, there could be problems with the returns data caused by infrequent trading of the component stocks. Fourth, capital markets may not be integrated.

In analysis involving relatively new transition equity markets this last factor may be crucial. The notion that risk can be defined as the sensitivity to the changes in world market returns is contingent on the assumption of complete market integration. As the amount of segmentation increases, risk takes on a new definition as a security's sensitivity to local-market factors. In integrated world capital markets the sensitivity to many local events can be hedged by a diversified portfolio. That is, a negative event in one country may be offset by positive news in another country. However, if capital markets are segmented, the sensitivity to local events can have significant effects on the required returns for the securities that trade in the local markets.

An example of application of the ICAPM to the emerging equity markets is that by Buckberg (1995). The author uses the so-called 'conditional' or 'expectational' asset pricing. 'Conditional' refers to the use of conditioning information - some information set Z_{t-1} - to calculate expected moments and to test the ICAPM as a relation between expected returns and ex ante risk. The conditional formulation restricts the conditionally expected return on an asset (based on Z_{t-1}) to be proportional to the asset's covariance with the market portfolio, yet allows expected returns to vary over time. Following notation in Buckberg (1995), define the excess returns in the market of the country *j*: $r_{jt} = ER_{jt} - r$, and the excess return on the world portfolio: $r_{wt} = ER_t^w - r$. Then according to the conditional ICAPM:

$$E[r_{jt} / Z_{t-1}] = \beta_j E[r_{wt} / Z_{t-1}] \tag{5.55}$$

where

$$\beta_j = \frac{\text{cov}(r_{jt}, r_{wt} / Z_{t-1})}{\text{var}(r_{wt} / Z_{t-1})}$$

Buckberg uses the set of instrumental variables, which are supposed to replicate the information the investors use to predict the prices. For each period, actual rates of return set during the previous period serve as conditioning information. In addition to this, the common-instrument set, containing information about developed markets only, includes a constant, the lagged return on the world index less the return on a thirty-day Treasury bill, a dummy for the month of January, the lagged differential between the return to holding a ninety-day Treasury bill for a month and the return on a Treasury bill thirty days to maturity, the lagged differential between the yield on a Moody's Baa bond and the yield on a Moody's Aaa bond, and the lagged dividend yield on the world portfolio less the return on a thirty-day Treasury bill.

Buckberg tests the model for two time periods, 1977-1984 and 1985-1991, using single market and multimarket tests[93] (with and without imposing the cross-section restrictions, respectively). Both tests reveal that many emerging markets were integrated into the global market during 1985-1991, but that many of the same markets reject the model when data for 1977-1884 is used. This fact suggests that rising capital inflows from developed economies served as the means of integration.

[93]The model is estimated by Hansen's (1982) generalised method of moments (GMM). For details of GMM estimation see technical appendix in Buckberg (1995).

In Harvey (1991) the specification of the conditional ICAPM, very similar to one in Buckberg (1995), successfully explains both time-series variation and cross-sectional differences in OECD market returns. Harvey (1991) finds returns consistent with the ICAPM in fourteen of seventeen mature equity markets - an 82 percent success rate - during 1970-1989; only Austria, Denmark, and Japan reject the model. Moreover, the betas he obtains for each market correspond in ranking to the ranking of mean returns across these markets, with the exception of underestimating the Japanese return. The fact that the model fails for Austria and Denmark, the two smallest OECD markets, suggests that illiquidity or other small market effects may impede the ICAPM relationship. Harvey also runs a multimarket test over all Group of Seven nations and is unable to reject the model.

As we have noted above, in the presence of exact purchasing power parity (PPP), the classic CAPM would hold internationally. However, if purchasing power parity is violated, investing in foreign markets entails exposure to exchange rate risk.[94] Any investment in a foreign asset is a combination of an investment in the performance of the foreign asset and an investment in the performance of the domestic currency relative to the foreign currency. The presence of this additional source of risk raises two related issues. The first one is whether currency risk is a priced factor in international financial markets. Second, if currency risk is priced, it becomes important to measure the compensation that investors can expect from bearing such risk.

So far the empirical studies on international portfolio diversification have paid relatively little attention to the role of foreign exchange risk. Most existing tests of the ICAPM incorporating this

[94] The idea that national investors are identified by deviations from PPP was originally proposed by Solnik (1974a).

additional source of risk are limited to the unconditional version of the model[95] and deliver inconclusive results, as they fail to account for the new information that periodically becomes available to investors, who then use it to adjust their investment strategies.

De Santis and Gerard (1998) analyse the conditional version of an international asset pricing model that incorporates currency risk as well as market risk.[96] The model is derived under the assumption that PPP is violated and, therefore, investors from different countries have a different appreciation for the real returns from any given asset. Following De Santis and Gerard, deviations from PPP have two important implications for portfolio choice and asset pricing in equilibrium. First, optimal portfolios differ across countries. Second, the expected return on any asset in addition to a market premium must include a currency premium as well. In this sense, if the tests reveal that the premium for currency risk is statistically and economically significant, one can infer that international asset prices are consistent with PPP violations.

De Santis and Gerard consider a world economy with $L+1$ countries and a set of M securities. Formally, the pricing restrictions on asset i imposed by their model are[97]

$$E_{t-1}(r_{it}) = \delta_{m,t-1}\operatorname{cov}_{t-1}(r_{it}, r_{mt}) + \sum_{c=1}^{L}\delta_{c,t-1}\operatorname{cov}_{t-1}(r_{it}, \pi_{ct}), \qquad i = 1, \ldots, M \tag{5.56}$$

where r_{it} denotes the excess returns on asset i; r_{mt} denotes the excess return on the world portfolio of all traded stocks; and π_{ct} indicates the inflation of country c. All returns in (5.56) are measured in the reference currency and in excess of the risk-free rate.

---

95 See, for example, Solnik (1974b, c), Stehle (1977), Stulz (1981), Adler and Dumas (1983), Anderson and Danthine (1983), Korajczyk and Viallet (1989). For a survey see Stulz (1995).

96 The model by De Santis and Gerard is a modification of the model originally proposed by Adler and Dumas (1983).

97 For the restrictions imposed on the parameters of the model see De Santis and Gerard (1998), p. 378.

The covariance between return i and the return on the worldwide portfolio represents the market risk component which appears in the standard CAPM of Sharpe (1964) and Lintner (1965). The covariances between the return on asset i and the rate of inflation of each of the countries in the model represent the additional sources of risk induced by PPP deviations. Specifically, the term $\mathrm{cov}_{t-1}(r_{it}, \pi_{ct})$ measures the exposure of asset i to both the inflation risk and the currency risk associated with country c. De Santis and Gerard simplify the model by assuming that domestic inflation is nonstochastic. In this case, the only random component in π_{ct} is the relative change in the exchange rate between the reference currency and the currency of country c. Therefore, $\delta_{c,t-1}$ can be interpreted as the price of risk of currency c. To complete the parameterisation of the model, De Santis and Gerard assume that the conditional second moments follow a diagonal GARCH process.[98] Empirical evidence presented by the authors using data from Germany, Japan, the United Kingdom, and the United States during the period 1973-1994 supports the specification of the CAPM that includes both market risk and foreign exchange risk.

Another test of the conditional version of the international capital asset pricing model is contained in Dumas and Solnik (1995). Similar to De Santis and Gerard (1998) they find that the price of currency risk is significantly different from zero. Their test uses the generalised method of moments (GMM) of Hansen (1982) and is derived as a variation of the methodology originally proposed by Harvey (1991). One of the appealing features of their test is that it can be computed without pre-specifying the dynamics of the conditional moments. Although very general, their approach has several limitations. First, because they do not specify the dynamics of the conditional second moments, they cannot evaluate the economic magnitude of the exchange risk premiums relative to the market premium. Second, without second moments, several quantities of interest to

[98] For description of the diagonal GARCH see section 3.2.5. in Chapter 3.

the investor, such as correlations, betas, and hedge ratios, cannot be measured. Lastly, their test should be interpreted as a test of some of the unconditional implications of the conditional model rather than as a direct test of the conditional model.

5.4 Empirical Analysis: ICAPM with Time-Varying Conditional Covariance Matrix, Testing for Integration and Segmentation in the Transition Markets of Central Europe

In this section we present the theoretical framework for testing the multi-market ICAPM for the dataset which along with equity portfolios from mature markets includes data from transition markets of Central Europe. We focus on two new markets of Central Europe: Hungary and Poland, as well as two mature markets: U.S. and Germany. All returns are measured in the reference currency, the U.S. dollar. Our approach assumes that investors do not cover their exposure to exchange rate risk. To present justification for ignoring exchange rate premia in our application of the ICAPM, we point to the fact that during the period covered by our data (1994-1998) the value of the Hungarian and Polish currencies were determined within the official currency baskets.[99] The Hungarian forint was defined in terms of a two-currencies basket (USD 30% , DEM 70%), while the Polish zloty in terms of a five-currencies basket (USD 45%, DEM 35%, GBP 10%, FRF 5%, CHF 5%). Since the exchange rates of these currencies were allowed to move within a rather narrow fluctuation band, the importance of modelling exchange rate premia in our application of the ICAPM diminishes.

[99] Both countries followed the so-called crawling-peg regimes.

For simplicity of notation assume that y_{it} denotes the excess return on the national stock index of country i, and y_{mt} denotes the excess return on the worldwide portfolio. In completely integrated world and in the absence of exchange risk, a conditional international CAPM imposes a restriction:

$$E(y_{it} | \Psi_{t-1}) = \lambda_{t-1} \operatorname{cov}(y_{it}, y_{mt} | \Psi_{t-1}) + \varepsilon_{it} \qquad \forall i \tag{5.57}$$

where λ_{t-1} can be interpreted as the *price of world market risk*. The model requires this equation to hold for every asset, including the world market portfolio.

Now consider the world economy with (N - 1) risky national equity portfolios plus the world market portfolio. Let $y_t = \{y_{1t}, \ldots y_{N-1t}, y_{mt}\}$ denote the $(N \times 1)$ vector of the returns on the national and world-wide portfolios, H_t denote the $(N \times N)$ conditional covariance matrix of the returns on the portfolios, and h_{Nt} the Nth column of H_t, which contains the covariance of the return on each national index (or portfolio) with the world market portfolio. Then the following system must hold:

$$y_t = \lambda_{t-1} h_{Nt} + \varepsilon_t, \qquad \varepsilon_t | \Psi_{t-1} \sim N(0,\ H_t) \tag{5.58}$$

based on a system of pricing restrictions:

$$\begin{aligned}
y_{1t} &= \lambda_{t-1} \operatorname{cov}(y_{1t}, y_{mt} | \Psi_{t-1}) + \varepsilon_{1t} \\
&\quad \ldots\ldots. \\
y_{N-1t} &= \lambda_{t-1} \operatorname{cov}(y_{N-1t}, y_{mt} | \Psi_{t-1}) + \varepsilon_{N-1t} \\
y_{mt} &= \lambda_{t-1} \operatorname{var}(y_{mt} | \Psi_{t-1}) + \varepsilon_{mt}
\end{aligned}$$

We also specify the dynamics of the conditional covariance matrix of the assets in the system. H_t follows the BEKK multivariate GARCH specification[100], which allows return volatility spillover effects between the markets in the system.

$$H_t + A_0^{'}A_0 + A_1^{'}\varepsilon_{t-1}\varepsilon_{t-1}^{'}A_1 + B_1^{'}H_{t-1}B_1 \tag{5.59}$$

The system (5.58) and (5.59) will be referred to as a *benchmark* model. According to international CAPM, if international markets are fully integrated, the price of the covariance risk λ_{t-1} in (5.58) should be positive and equal across all markets. To test pricing restrictions of the benchmark model, we define two alternative specifications, which could provide useful insights into pricing of the assets in the considered markets. The first alternative is *the partial segmentation model*, which allows for some level of market segmentation:

$$y_{it} = \alpha_i + \lambda_{t-1}\operatorname{cov}(y_{it}, y_{mt}|\Psi_{t-1}) + \gamma_i \operatorname{var}(y_{it}|\Psi_{t-1}) + \varepsilon_{it} \qquad \forall i \tag{5.60}$$

In addition to world-wide risk, it adds to the standard ICAPM:

- the *local conditional variance* to measure country-specific risk

- a *constant* to accommodate other forms of segmentation

Under hypothesis of full market integration both coefficients α_i and γ_i in (5.60) must equal to zero for all markets.

The second alternative is the *partial integration model*:

$$y_{it} = \alpha_i + \theta z_{t-1} + \lambda_{t-1}\operatorname{cov}(y_{it}, y_{mt}|\Psi_{t-1}) + \varepsilon_{it} \qquad \forall i \tag{5.61}$$

[100] See section 3.2.5 for details. In our analysis summation limit K is set to 1.

This model introduces a $(k \times 1)$ vector of market-wide information variables, z_{t-1}, included in Ψ_{t-1}. This specification gives us an opportunity to check whether the variation in conditional expected returns is fully explained by world market risk. In case this is not true, some of the variables in z_{t-1} could have statistical power in explaining the behaviour of the expected return series.

The dynamics of the conditional second moments of both equation (5.60) and (5.61) is restricted to the multivariate GARCH specification in (5.59).

The parameters of the models defined above will be estimated using quasi-maximum likelihood (QML) estimation procedure by Bollerslev and Wooldridge (1992) described in section 3.2.6 of Chapter 3.

5.4.1 Data and Descriptive Statistics

Our sample of national stock markets includes daily data from both developed and relatively new developing equity markets. We focus on two new markets of Central Europe: Hungary and Poland, as well as two mature markets: U.S. and Germany[101]. The indices used in our analysis are American Dow Jones, German DAX, Hungarian BUX and Polish WIG 20. The Morgan Stanley Capital International (MSCI) World Index will approximate the world market portfolio. The return on the US 3-month Treasury bill is used as a measure of the risk-free rate. All the data is measured in a common currency, the US dollar.

[101] Our original idea was to estimate the system including return series from all four transition markets considered in the previous chapters of this thesis. However, the seven-dimensional system (i.e. system including return series from four transition markets, U.S., Germany, as well as the world market) failed to converge in estimation. To keep original combination of the developed and transition markets in the system, we decided to restrict our attention to two transition markets only. Our choice of the Hungarian and Polish markets is explained by the advancement of their stock markets and a higher market capitalisation compared to the Czech and Slovak markets.

'he set of the conditioning information variables used in this article includes the change in the US erm premium (YB), measured by the yield on the 10-year US Treasury note in excess of the risk-ree rate; the change in the US interest rates (INTR), measured as the change in 3-month US Treasury bill; and the US default premium (BAA-AAA), measured by the yield difference between Moody's Baa and Aaa rated bonds. All the data comes from Datastream and covers the period rom April 1994 to December 1998.

Table 5.1 presents summary statistics on the US dollar returns on the four national indices and world market index in excess of the risk free rate. The evidence for all return series indicates significantly fatter tails than does the stationary normal distribution. The excess kurtosis statistic ranges in value from 5.39 for Polish to 17.63 for world excess returns. The coefficients of skewness indicate that the series have asymmetric distribution skewed to the right in case of the world, US and German excess returns, and skewed to the left in case of Hungarian and Polish series. The Ljung-Box test statistic for twelfth-order serial correlation in levels reveals the lack of statistically significantly autocorrelations in the excess return series.[102] However, we detect autocorrelations in the squared returns, which suggests that a GARCH specification of the second moments might be appropriate.

Table 5.2 displays cross-correlations of the excess return series. And finally, Table 5.3 contains cross-correlations of the conditioning variables specified above. The low correlations of information variables imply that they carry nonredundant information.

[102] The results in the previous chapter (see section 4.2) have shown that nominal Hungarian and Polish returns in local currency terms are autocorrelated.

Table 5.1 Summary statistics of the excess returns

	Poland	Hungary	Germany	US	World
Mean	0.000117	0.000172	0.000532	0.000586	0.000269
Variance	0.000697	0.000500	0.000232	0.000172	0.000140
Skewness	-0.15152	-1.11529	0.23492	0.59338	1.16346
Excess kurtosis	5.39398	17.03238	9.41929	10.57795	17.63387
Ljung-Box (12) for the levels	25.1305	21.9645	16.3703	18.5148	22.1850
Ljung-Box (12) for the squares	264.4949	242.9491	69.4253.	56.1101	51.3930

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively

Table 5.2 Unconditional correlation of the excess returns

	Poland	Hungary	Germany	US
Hungary	0.44			
Germany	0.36	0.53		
USA	0.43	0.39	0.77	
World	0.45	0.55	0.59	0.90

Table 5.3 Correlation of conditioning variables

	INTR	BAA-AAA
YB	0.28	-0.36
INTR		-0.11

5.4.2 Empirical results[103]

The results of this chapter are divided into two parts. First, we introduce the additional assumption that the price of market risk, λ_{t-1}, is constant, i.e. the slope of the capital market line is fixed.[104] However, if excess returns are considerably more variable than their conditional variance with the market, a specification with a constant price of risk may not have enough power to fully explain the dynamics of the risk premia. Therefore, in the second part, we allow for time variation in the price of risk by parameterising the dynamics of λ_{t-1}.

5.4.2.A. Constant price of market risk

Tables 5.4 and 5.5 present results of the estimation of the benchmark model in (5.58) and (5.59). The estimate of the constant price of market risk equals 2.59 and is highly significant. The χ^2_3 Likelihood Ratio test statistic to check the hypothesis that price of risk is equal across the markets equals 3.79, which implies that the null of common price cannot be rejected at any statistically reasonable level. This is an important result, as it indicates that the assets in the transition markets command the same level of return for a given level of risk as the assets in the developed countries. It is consistent with ICAPM, and points to integration of the Hungarian and Polish markets into the global capital market.

[103] The data employed in our analysis covers a period of only about three and a half years, some three years after Hungarian and Polish markets begun operating. It is, therefore, reasonable to assume that results of this chapter are robust over sub-periods of the available dataset.
[104] For the examples of the estimation of the conditional CAPM with this restriction see, e.g., Giovannini and Jorion (1989) and Chan, Karolyi and Stulz (1992).

Table 5.4 Constant price of market risk: the benchmark model

Price of market risk (λ)	2.593340 (0.720759)
χ^2_3 (H_0: $\lambda_i = \lambda \quad \forall i$) *	3.792596

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood standard errors in brackets.

* χ^2 is the Likelihood Ratio statistic to test the hypothesis that price of risk is equal across markets. The upper 1 and 5 percentile points in the χ^2 distribution for three degrees of freedom are 11.34 and 7.81 respectively.

The BEKK multivariate GARCH specification of the conditional covariance matrix in (5.59) employed in our analysis allows return volatility spillover effects between the markets in the system. To conserve space in Table 5.5 we present estimates of the diagonal elements of the covariance matrix only: the conditional variances of the considered markets.

Table 5.5 Volatility spillovers across the markets

A. Polish volatility

$$\begin{aligned}
h_{55t} =\; & \underset{(0.0004)}{0.0002} + \underset{(4.19e\text{-}07)}{9.36e-04}\varepsilon_{1t-1}^2 + \underset{(8.43e\text{-}07)}{0.0006}\varepsilon_{1t-1}\varepsilon_{2t-1} + \underset{(2.23e\text{-}06)}{0.0033}\varepsilon_{1t-1}\varepsilon_{3t-1} - \underset{(4.57e\text{-}07)}{0.0002}\varepsilon_{1t-1}\varepsilon_{4t-1} + \underset{(4.29e\text{-}07)}{0.0126}\varepsilon_{1t-1}\varepsilon_{5t-1} + \\
& + \underset{(8.91e\text{-}08)}{1.03e-04}\varepsilon_{2t-1}^2 + \underset{(2.03e\text{-}06)}{0.0018}\varepsilon_{2t-1}\varepsilon_{3t-1} - \underset{(1.15e\text{-}07)}{6.31e-05}\varepsilon_{2t-1}\varepsilon_{4t-1} + \underset{(4.98e\text{-}05)}{0.0035}\varepsilon_{2t-1}\varepsilon_{5t-1} + \\
& + \underset{(1.71e\text{-}06)}{0.0029}\varepsilon_{3t-1}^2 + \underset{(1.86e\text{-}06)}{0.0003}\varepsilon_{3t-1}\varepsilon_{4t-1} + \underset{(5.72e\text{-}05)}{0.0184}\varepsilon_{3t-1}\varepsilon_{5t-1} + \underset{(2.42e\text{-}08)}{9.67e-06}\varepsilon_{4t-1}^2 - \underset{(4.99e\text{-}05)}{0.0011}\varepsilon_{4t-1}\varepsilon_{5t-1} + \underset{(4.87e\text{-}05)}{0.0297}\varepsilon_{5t-1}^2 + \\
& + \underset{(5.93e\text{-}012)}{6.06e-08}h_{11t-1} + \underset{(1.76e\text{-}08)}{9.17e-04}h_{12t-1} + \underset{(4.90e\text{-}07)}{1.49e-05}h_{13t-1} - \underset{(1.11e\text{-}05)}{2.52e-05}h_{14t-1} - \underset{(0.0007)}{0.0005}h_{15t-1} + \\
& + \underset{(1.58e\text{-}07)}{3.47e-04}h_{22t-1} + \underset{(8.37e\text{-}07)}{0.0011}h_{23t-1} + \underset{(1.33e\text{-}05)}{0.0019}h_{24t-1} + \underset{(0.0007)}{0.0358}h_{25t-1} + \underset{(4.89e\text{-}07)}{9.20e-04}h_{33t-1} + \underset{(0.0015)}{0.0031}h_{34t-1} + \underset{(0.0007)}{0.0582}h_{35t-1} + \\
& + \underset{(0.0005)}{0.0206}h_{44t-1} + \underset{(0.0009)}{0.0984}h_{45t-1} + \underset{(0.0007)}{0.9199}h_{55t-1}
\end{aligned} \tag{5.62}$$

B. Hungarian volatility

$$h_{44t} = -0.0001 + 0.0002\varepsilon_{1t-1}^2 + 0.0004\varepsilon_{1t-1}\varepsilon_{2t-1} + 0.0002\varepsilon_{1t-1}\varepsilon_{3t-1} + 0.0136\varepsilon_{1t-1}\varepsilon_{4t-1} + 0.0024\varepsilon_{1t-1}\varepsilon_{5t-1} +$$

(0.0012) (1.72e-07) (1.20e-06) (3.92e-07) (0.0015) (2.28e-05)

$$+0.0003\varepsilon_{2t-1}^2 + 0.0002\varepsilon_{2t-1}\varepsilon_{3t-1} + 0.0171\varepsilon_{2t-1}\varepsilon_{4t-1} + 0.0030\varepsilon_{2t-1}\varepsilon_{5t-1} +$$

(5.08e-07) (7.34e-07) (0.0015) (2.32e-05)

$$+5.72e-05\varepsilon_{3t-1}^2 + 0.0079\varepsilon_{3t-1}\varepsilon_{4t-1} + 0.0014\varepsilon_{3t-1}\varepsilon_{5t-1} + 0.3018\varepsilon_{4t-1}^2 + 0.0943\varepsilon_{4t-1}\varepsilon_{5t-1} + 0.0082\varepsilon_{5t-1}^2 +$$

(8.00e-08) (0.0015) (0.0001) (0.0015) (0.0015) (2.09e-05)

$$+0.0095h_{11t-1} + 0.0586h_{12t-1} - 0.0283h_{13t-1} - 0.0850h_{14t-1} - 0.1083h_{15t-1} +$$

(0.0001) (0.0014) (0.0005) (0.0056) (0.0043)

$$+0.0907h_{22t-1} - 0.0876h_{23t-1} - 0.2629h_{24t-1} - 0.0353h_{25t-1} + 0.0212h_{33t-1} + 0.1270h_{34t-1} + 0.0620h_{35t-1} +$$

(0.0009) (0.0017) (0.0088) (0.0067) (0.0002) (0.0059) (0.0049)

$$+0.2705h_{44t-1} + 0.0859h_{45t-1} + 0.0998h_{55t-1}$$

(0.0047) (0.0012) (0.0035)

(5.63)

C. German volatility

$$h_{33t} = -0.0001 + 0.0170\varepsilon_{1t-1}^2 + 0.0393\varepsilon_{1t-1}\varepsilon_{2t-1} + 0.0779\varepsilon_{1t-1}\varepsilon_{3t-1} - 0.0031\varepsilon_{1t-1}\varepsilon_{4t-1} + 0.0001\varepsilon_{1t-1}\varepsilon_{5t-1} +$$

(0.0001) (4.31e-05) (0.0003) (0.0006) (5.05e-05) (0.0001)

$$+0.0227\varepsilon_{2t-1}^2 + 0.0899\varepsilon_{2t-1}\varepsilon_{3t-1} - 0.0036\varepsilon_{2t-1}\varepsilon_{4t-1} + 0.0234\varepsilon_{2t-1}\varepsilon_{5t-1} +$$

(0.0001) (0.0006) (0.0001) (0.0002)

$$+0.0892\varepsilon_{3t-1}^2 - 0.0072\varepsilon_{3t-1}\varepsilon_{4t-1} + 0.0463\varepsilon_{3t-1}\varepsilon_{5t-1} + 0.0001\varepsilon_{4t-1}^2 - 0.0019\varepsilon_{4t-1}\varepsilon_{5t-1} + 0.0060\varepsilon_{5t-1}^2 +$$

(0.0004) (0.0005) (0.0007) (4.41e-06) (0.0001) (7.02e-05)

$$+0.0135h_{11t-1} - 0.0204h_{12t-1} + 0.1450h_{13t-1} - 0.0582h_{14t-1} + 0.0147h_{15t-1} +$$

(7.18e-05) (0.0002) (0.0096) (0.0014) (0.0002)

$$+0.0077h_{22t-1} + 0.1096h_{23t-1} + 0.0440h_{24t-1} + 0.0111h_{25t-1} + 0.3904h_{33t-1} + 0.3132h_{34t-1} - 0.0794h_{35t-1} +$$

(4.96e-05) (0.0083) (0.0014) (0.0002) (0.0083) (0.0098) (0.0084)

$$+0.0628h_{44t-1} + 0.0318h_{45t-1} + 0.0040h_{55t-1}$$

(0.0011) (0.0014) (7.50e-05)

(5.64)

D. ***US volatility***

$$h_{22t} = 0.0001 + 0.0011\varepsilon_{1t-1}^2 + 0.0057\varepsilon_{1t-1}\varepsilon_{2t-1} + 0.0093\varepsilon_{1t-1}\varepsilon_{3t-1} - 0.0715\varepsilon_{1t-1}\varepsilon_{4t-1} - 0.0325\varepsilon_{1t-1}\varepsilon_{5t-1} +$$

(0.0003) (1.11e-05) (0.0001) (0.0003) (0.0642) (0.0060)

$$+0.0974\varepsilon_{2t-1}^2 + 0.02414\varepsilon_{2t-1}\varepsilon_{3t-1} - 0.18519\varepsilon_{2t-1}\varepsilon_{4t-1} - 0.08423\varepsilon_{2t-1}\varepsilon_{5t-1} +$$

(0.0005) (0.0004) (0.0644) (0.0059)

$$0.0198\varepsilon_{3t-1}^2 - 0.3038\varepsilon_{3t-1}\varepsilon_{4t-1} - 0.1282\varepsilon_{3t-1}\varepsilon_{5t-1} + 1.1661\varepsilon_{4t-1}^2 + 1.0607\varepsilon_{4t-1}\varepsilon_{5t-1} + 0.2412\varepsilon_{5t-1}^2 +$$

(0.0002) (0.0675) (0.0072) (0.8640) (0.5967) (0.2276)

$$+0.0567h_{11t-1} + 0.0770h_{12t-1} + 0.8197h_{13t-1} + 0.4076h_{14t-1} + 0.2543h_{15t-1} +$$

(0.0011) (0.0030) (0.0665) (0.0414) (0.0192)

$$+0.5162h_{22t-1} + 0.5572h_{23t-1} + 0.2771h_{24t-1} + 0.1728h_{25t-1} +$$

(0.0700) (0.0609) (0.0391) (0.0178)

$$+0.4248h_{33t-1} + 0.9488h_{34t-1} + 0.0382h_{35t-1} + 0.2132h_{44t-1} + 0.0157h_{45t-1} + 0.0053h_{55t-1} +$$

(0.0563) (0.1258) (0.0766) (0.0384) (0.0821) (0.0160)

(5.65)

E. World volatility

$$h_{11t} = 0.0020 + 0.3778\varepsilon_{1t-1}^2 + 0.1025\varepsilon_{1t-1}\varepsilon_{2t-1} + 0.1166\varepsilon_{1t-1}\varepsilon_{3t-1} - 0.5001\varepsilon_{1t-1}\varepsilon_{4t-1} - 0.1386\varepsilon_{1t-1}\varepsilon_{5t-1} +$$

(0.0006) (0.3404) (0.0050) (0.2844) (0.1005)

$$+0.0695\varepsilon_{2t-1}^2 + 0.1582\varepsilon_{2t-1}\varepsilon_{3t-1} - 0.67840\varepsilon_{2t-1}\varepsilon_{4t-1} - 0.1881\varepsilon_{2t-1}\varepsilon_{5t-1} +$$

(0.0022) (0.0060) (0.2984) (0.1119)

$$+0.0900\varepsilon_{3t-1}^2 - 0.7719\varepsilon_{3t-1}\varepsilon_{4t-1} - 0.2140\varepsilon_{3t-1}\varepsilon_{5t-1} + 1.6551\varepsilon_{4t-1}^2 + 0.9177\varepsilon_{4t-1}\varepsilon_{5t-1} + 0.1272\varepsilon_{5t-1}^2 +$$

(0.0030) (0.3111) (0.1069) (1.2861) (0.5699) (0.8077)

$$+0.9424h_{11t-1} + 1.9356h_{12t-1} + 6.7433h_{13t-1} + 1.0761h_{14t-1} + 0.0111h_{15t-1} +$$

(0.0244) (0.0635) (0.7824) (0.1230) (0.0246)

$$+0.3072h_{22t-1} + 3.8498h_{23t-1} + 1.1050h_{24t-1} + 0.0063h_{25t-1} +$$

(0.0178) (0.6169) (0.1045) (0.0179)

$$+0.6023h_{33t-1} + 1.9247h_{34t-1} + 0.0396h_{35t-1} + 0.2338h_{44t-1} + 0.0114h_{45t-1} + 3.01e-05h_{55t-1}$$

(0.0828) (0.7512) (0.0833) (0.0720) (0.0624) (3.20e-06)

(5.66)

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood standard errors in brackets.

The results of fitting the multivariate GARCH model to our international dataset show that most of the volatility and return shock spillover effects are highly significant. Starting from the analysis of the results for the conditional volatility in the Polish market in this multivariate setting, we note that while most of the estimated coefficients are significant, the magnitude of the parameters reflecting cross-market transmissions of lagged shocks and volatility is very small. Considering the effects of the local GARCH terms we see that lagged conditional variance, h_{55t-1}, has a considerably stronger impact on the current Polish return volatility compared to the lagged local absolute shock term, ε_{5t-1}^2 (compare 0.9199 to 0.0297). Moving to the Hungarian conditional variance we note that the size of many of the estimated parameters is larger than in the case of the Polish volatility. The strongest significant effects on the Hungarian volatility are those from the local lagged volatility and return shock terms, as well as from the lagged conditional covariance of the Hungarian returns with the US and German returns.

The significant lagged GARCH terms of a relatively large magnitude affecting the German conditional volatility, are the local lagged volatility, as well as the lagged covariances of the German returns with the returns on the world wide portfolio, the US returns and the Hungarian returns. Looking at the results for the US and world market we find that the local GARCH effects are stronger on the world portfolio volatility compared to the US one. Most of the international volatility and shock terms are significantly priced in these two markets, with exception of some of the terms related to the Polish lagged conditional variance and return shocks, as well as the Hungarian lagged absolute shocks.

Table 5.6 includes diagnostic statistics to evaluate the considered model specification. Though the estimated model reduces coefficients of skewness and excess kurtosis present in the excess return series (see Table 5.1 for comparison), we still reveal significant nonnormality in the distribution of standardised residuals. This evidence against normality warrants the use of quasi maximum likelihood inferential procedures in our analysis. The Ljung-Box test statistics in the Table 5.6 indicates that the employed specification reduces the intertemporal dependence in squares of the standardised residuals, $\varepsilon_{it}^2 / h_{it}$, to insignificance at the 5% level for all return series, though not for Polish excess returns, where volatility clustering, while being reduced considerably, is not removed completely.

Table 5.6 Diagnostics for the standardised residuals of the benchmark model with constant price of market risk

	World	US	Germany	Hungary	Poland
Excess kurtosis	2.96384	2.70202	2.51901	6.62064	3.52420
Skewness	0.10087	0.01563	-0.11054	-0.55842	-0.15413
Ljung-Box (12) for the sq. stand. residuals	7.3588	7.5167	12.1816	18.6917	35.0598

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively

To compare the benchmark ICAPM to the alternative models specified in (5.60) and (5.61) we conduct some diagnostic tests. The results of these tests are presented in Tables 5.7 and 5.8 below. In the partial segmentation model in (5.60), the excess return for each market depends not only on the covariance with the world portfolio, but also on the local conditional variance to measure country-specific risk, and a constant to account for other forms of segmentation. The Likelihood Ratio statistic to test whether the country-specific intercepts are jointly equal to zero supports the conditional ICAPM. However, the statistic to test the null that the prices of country-specific risk are jointly equal to zero is significant at 5 percent level only (but not at 1 per cent level). To show that this result is influenced by the individual significance of the Hungarian conditional variance, priced in the market after accounting for world-market risk[105], we test the hypothesis that the prices of country-specific risks excluding the Hungarian risk are jointly equal to zero. The resulting test statistic confirms that the general estimate is influenced by the significance of the Hungarian variance, and points to the possibility of some degree of segmentation of the Hungarian market.

Table 5.7 Constant price of market risk: the partial segmentation model

$\chi_4^2 \ (H_0: \alpha_i = 0 \quad \forall i)$ *	7.059864
$\chi_4^2 \ (H_0: \gamma_i = 0 \quad \forall i)$ **	11.32698
$\chi_3^2 \ (H_0: \gamma_i = 0 \quad \forall i \neq Hungary)$ ***	7.188425

* χ^2 is the Likelihood Ratio statistic to test the hypothesis that country-specific intercepts are jointly equal to zero. ** χ^2 is the Likelihood Ratio statistic to test the hypothesis that the prices of country-specific risk are jointly equal to zero. The upper 1 and 5 percentile points in the χ^2 distribution for four degrees of freedom are 13.28 and 9.49 respectively. *** χ^2 is the Likelihood Ratio statistic to test the hypothesis that the prices of country-specific risk excluding the Hungarian risk are jointly equal to zero. The upper 1 and 5 percentile points in the χ^2 distribution for three degrees of freedom are 11.34 and 7.81 respectively.

In the partial integration model in (5.61) we test the hypothesis that the variation in conditional expected returns is fully explained by world market risk by introducing a set of conditioning world information variables, described in section 5.4.1. The Likelihood Ratio statistic in Table 5.8 shows that this hypothesis is strongly rejected by the data.

[105] The parameters of the estimated models are not reported here, however are available on request.

Table 5.8 Constant price of market risk: the partial integration model

χ_3^2 (H_0: $\theta_k = 0 \quad \forall k$) *	39.2441

* χ^2 is the Likelihood Ratio statistic to test the hypothesis that the information variables in z_{t-1} are orthogonal to the risk-adjusted excess returns. The upper 1 and 5 percentile points in the χ^2 distribution for three degrees of freedom are 11.34 and 7.81 respectively.

Summarising results for the specifications with constant price of risk, we find that a number of diagnostic tests support the pricing restrictions of the CAPM. The estimates reveal a statistically significant positive relation between excess returns and covariance risk. The price of covariance risk is equal across the markets. However, while country-specific constant factors add no explanatory power to the model, the Hungarian idiosyncratic risk is individually significant. The results of the partial integration model show that the variation in the excess returns is still predictable after accounting for market-wide risk. To check whether this result is in part due to the auxiliary assumption about constancy of the price of risk, in the next section of the chapter we allow λ to vary through time.

5.4.2.B Time-varying Price of Market Risk

Ferson (1989) and Ferson, Foerster and Keim (1993) parameterise the dynamics of the price of risk by using linearity assumption to test latent variable models. In the absence of additional restrictions, the disadvantage of this approach is that a linear price of market risk can become negative, which is inconsistent with the theory. In our analysis we approximate the price of market risk by an exponential function of the instruments.[106]

$$\lambda_{t-1} = \exp(\kappa' z_{t-1}), \tag{5.67}$$

[106] De Santi and Gerard (1997) and Bekaert and Harvey (1995) use a similar parametrization.

where the instruments in z_{t-1} are information variables described in section 5.4.1. As a result our benchmark model will look as follows:

$$y_{it} = \exp(\kappa' z_{t-1}) \operatorname{cov}(y_{it}, y_{mt} | \Psi_{t-1}) + \varepsilon_{it} \tag{5.68}$$

with the dynamics of the conditional covariance matrix specified by the BEKK multivariate GARCH specification in (5.59).

Table 5.9 presents estimates of the parameters of the model in (5.68). The results show that the dynamics of time varying price of market risk is driven by the change in the US term premium (k_2), as well as the US default premium (k_3). To compare the estimate of the time-varying price of market risk with the corresponding estimate of the constant price of risk in Table 5.4, we evaluate the average value and the standard error of the time-varying price of risk, $\lambda_{t-1} = \exp(\kappa' z_{t-1})$, conditional on the sample means of the instruments in z_{t-1}. The value of the average time-varying risk, reported in the 5th column of Table 5.9 is close to the constant price of world covariance risk, although both the magnitude and significance of the time-varying coefficient exceed the corresponding estimates for the model with constant risk. In spite of this fact, Table 5.10 indicates that there is no significant improvement in the standardised residuals of the time-varying risk specification over the model with constant price of risk (see Table 5.6 for comparison). As before, though the coefficients of the skewness and excess kurtosis are reduced compared to the corresponding statistics for the nominal excess returns, there is still evidence of significant nonnormality in the distribution of the standardised residuals of the model. And again, as is the case with the model with constant price of risk, the Ljung-Box statistics indicate that the time-varying risk specification does capture volatility clustering in all series with exception of the Polish excess returns (although again reduces it considerably). Thus, the diagnostic tests do expose some limitations of the estimated models.

Table 5.9 Time-varying price of market risk: the benchmark model

k_0 (constant)	k_1 (interest rate)	k_2 (term premium)	k_3 (default premium)	$\lambda_{t-1} = \exp(\kappa' z_{t-1})$ average risk	χ_3^2 (H_0: $k_i = 0$ $\forall i \geq 1$) *
1.3560 (0.6855)	-1.4784 (5.0085)	-9.0186 (2.3086)	-14.8448 (3.1948)	3.9185 (0.9553)	32.4535

Note: Bollerslev and Wooldridge (1992) quasi-maximum likelihood standard errors in brackets.

* χ^2 is the Likelihood Ratio statistic to test the hypothesis that the price of market risk is constant. The upper 1 and 5 percentile points in the χ^2 distribution for three degrees of freedom are 11.34 and 7.81 respectively.

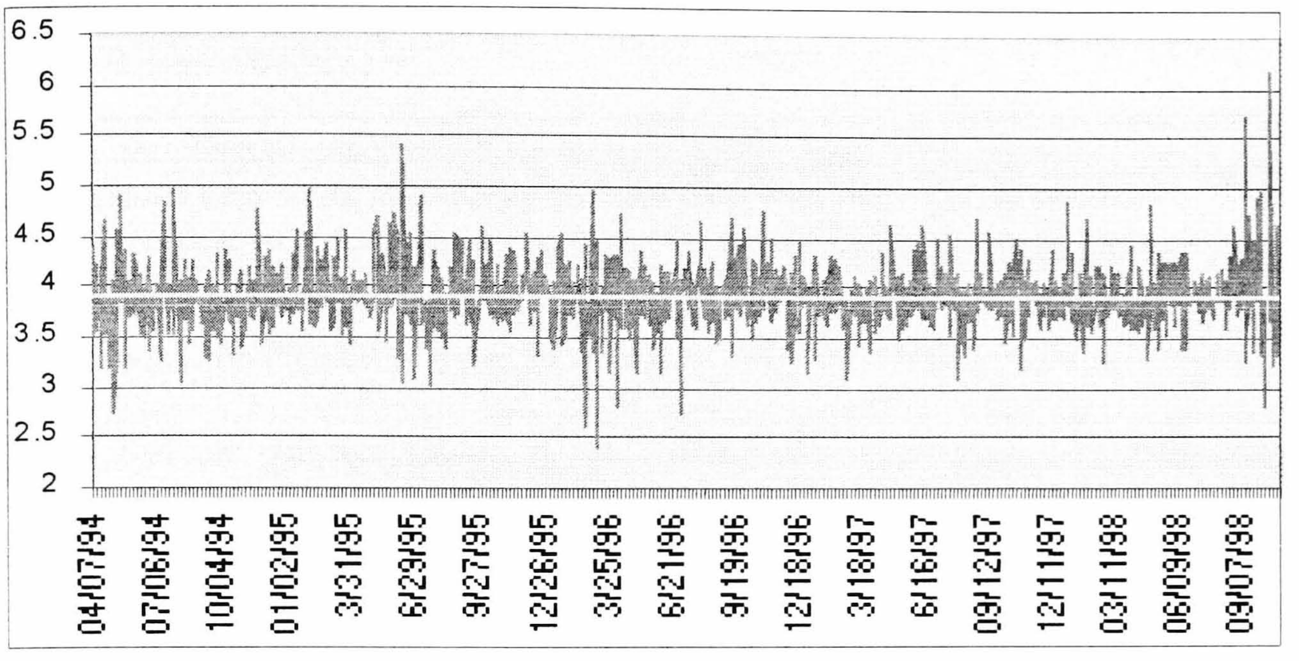
Table 5.10 Diagnostics for the standardised residuals of the benchmark model with time-varying price of market risk

	World	US	Germany	Hungary	Poland
Excess kurtosis	3.4572	2.7121	2.01437	5.2352	3.6477
Skewness	0.0901	0.01666	-0.09891	-0.5599	-0.1255
Ljung-Box (12) for the sq. stand. residuals	6.2256	5.9834	14.5455	17.9189	32.8111

Note: The upper and lower 1 percentile points in the distribution of the skewness statistics are .058 and -.058, respectively. The upper and lower 1 percentile points in the distribution of the excess kurtosis are .13 and -.11, respectively. The upper 1 and 5 percentile points in the χ^2 distribution for the Ljung-Box(12) statistic are 26.22 and 21.03, respectively

Consistent with the evidence presented in Harvey (1991) and Bekaert and Harvey (1995), the Likelihood Ratio statistic in the last column of Table 5.9 for the hypothesis that the world price of risk is constant is strongly rejected at any conventional level. This fact is also confirmed in Figure 5.1, which plots the fitted time-varying price of the covariance risk, as well as the average risk (the white line).

Figure 5.1 Time-varying price of world covariance risk



The estimates of the elements of the conditional covariance matrix are essentially not significantly affected when the price of market risk is allowed to vary. Therefore, to conserve space, we do not present the structure of the conditional second moments here, and refer the reader to the corresponding estimates for the specification with the constant price of market risk (see (5.62) - (5.66) above).

Next we evaluate the two alternative specifications of the conditional CAPM in (5.60) and (5.61) allowing the price of market risk to vary through time. The results for the partial segmentation model in Table 5.11 indicate that the relaxation of the constancy condition does not have any significant effect on the results from the previous testing for the specification with the constant price of risk (see Table 5.7 for comparison).

Table 5.11 Time-varying price of market risk: the partial segmentation model

χ_4^2 (H_0: $\alpha_i = 0 \quad \forall i$) *	7.4699
χ_4^2 (H_0: $\gamma_i = 0 \quad \forall i$) **	10.9881

* χ^2 is the Likelihood Ratio statistic to test the hypothesis that country-specific intercepts are jointly equal to zero. ** χ^2 is the Likelihood Ratio statistic to test the hypothesis that the prices of country-specific risk are jointly equal to zero. The upper 1 and 5 percentile points in the χ^2 distribution for four degrees of freedom are 13.28 and 9.49 respectively.

The results for the partial integration model with time-varying price of market risk in Table 5.12 show that the null hypothesis that the information variables z_{t-1} are orthogonal to the excess returns is strongly rejected as in the case of the model with the constant price of risk.

Table 5.12 Time-varying price of market risk: the partial integration model

χ_3^2 (H_0: $\theta_k = 0 \quad \forall k$) *	35.2188

* χ^2 is the Likelihood Ratio statistic to test the hypothesis that the information variables in z_{t-1} are orthogonal to the risk-adjusted excess returns. The upper 1 and 5 percentile points in the χ^2 distribution for three degrees of freedom are 11.34 and 7.81 respectively.

Therefore, the expected excess returns remain predictable after accounting for world market risk. One possible reason for that is the inability of our specification of the ICAPM to accommodate negative risk premia (see (5.68)). De Santi and Gerard (1997), who apply ICAPM to data from OECD countries, show that predictability disappears when they relax the nonnegativity restriction and consider linear parameterisations of the time-varying price of the covariance risk.[107] Another reason for the predictability after accounting for market risk could be the fact that the conditioning information variables we have considered for parameterisation of the time-varying risk (see section 5.4.1) are not sufficient to explain the variation in the price of the covariance risk. And finally, it is possible and, perhaps, highly probable that the national stock indices considered contain

[107] Richardson and Smith (1993) find that for the U.S. market negative risk premia are associated with periods of high expected inflation and downward sloping term structures.

characteristics which are not accounted by the world market risk, and a model with multiple sources of risk could perform better than a single factor model.

5.5 Summary

In this chapter we test pricing restrictions of the international CAPM simultaneously for four national equity markets: two developed markets (U.S. and Germany) and two new transition markets (Hungary and Poland). Methodologically we extend the multivariate GARCH specification of Engle and Kroner (1995) ("BEKK") to accommodate GARCH-M effects, and propose an alternative specification of the conditional CAPM, which allows return volatility transmissions between the markets in the system.

Using two alternative specifications of the conditional ICAPM (the partial segmentation and partial integration models) we carry out some diagnostic tests of the benchmark ICAPM, and at the same time consider the issues of integration of the markets in our system into world capital markets.

The specified models are first estimated with the additional assumption that the price of the market risk is constant. The evidence supports some of the pricing restrictions of the ICAPM and rejects the others. Market risk, measured by the conditional covariance between the return on each national index and a world-wide index, is positive and equally priced across the countries. This is consistent with ICAPM and supports the hypothesis of international market integration. However, while country-specific intercepts add no explanatory power to the model, the Hungarian idiosyncratic risk is individually significant. This fact is against the hypothesis of the full market integration of all countries in our system, and points to the possibility of some degree of segmentation of the Hungarian market. The results of the partial integration model, which introduces a vector of world-wide information variables in the return equations, shows that the variation in the excess returns is still predictable after accounting for market-wide risk.

To check whether our results are due to the auxiliary assumption on the constancy of the price of risk, we allow the price of market risk to vary through time. We find that although the specification with time-varying price of risk explains a larger fraction of the variation in excess national returns compared to the model with constant risk, excess returns are still predictable after accounting for the market-wide risk. This fact could have different possible explanations. One possible reason is our parameterisation of time-varying risk, which in accord with the theory imposes nonnegativity restriction on the risk premia. Another reason is the possibility of the fact that the conditioning information variables we have considered for parameterisation of the time-varying risk are not sufficient to explain the variation in the price of the covariance risk. It is also possible that the considered national stock indices contain characteristics, which are not accounted by the world market risk.

Our general conclusion is that even though the conditional version of the traditional ICAPM applied to the dataset including new markets of Central Europe provides useful information on the dynamics of market premia, a more adequate model of international asset pricing should probably include additional factors.

Appendix 5.1

An Alternative to the CAPM: the Arbitrage Pricing Theory

The alternative to the CAPM in determining the expected rate of return on individual stocks and on portfolio of stocks is the arbitrage pricing theory (APT). The theory is called an arbitrage theory because Ross (1976) was able to show that if, in equilibrium, expected asset returns were not a linear function of the *k* factors, an investor could construct a portfolio using short-selling that requires zero net investment and has a greater-than-zero expected return. This result would violate the 'no free lunch' law of the financial theory.

Broadly speaking, the APT implies that the return on the asset can be broken down into an expected return, R_{it}^e, and an unexpected or surprise component, u_{it}.

$$R_{it} = R_{it}^e + u_{it} \tag{5.69}$$

For any individual asset this surprise (or news) component itself can be broken down into 'general news' that affect all assets (*systematic or market risk*), m_t, and 'specific news' which affects only this particular assets (*idiosyncratic or specific risk*), ε_{it}.

$$u_{it} = m_t + \varepsilon_{it} \tag{5.70}$$

As in the case of the CAPM, the systematic risk cannot be diversified away because this element of news affects all assets, however, unsystematic or specific risk may be diversified away.

The market risk at time *t* can be presented in the form:

$$m_t = \sum_j b_{ij}(F_j - EF_j)_t \tag{5.71}$$

where market-wide factors, F_j s, may have different effects on different securities, reflected in the different values for the coefficients b_{ij} s for each individual security *i*.

The APT does not require any assumptions about utility function or that the mean and variance of a portfolio are the only two elements in the investor's objective function. The model is a mechanism that allows one to derive an expression for the expected return on a security (or a portfolio of securities) based on the idea that riskless arbitrage opportunities will be instantaneously eliminated.

Assuming that agents have homogeneous expectations, and based on several orthogonality conditions, the APT can be summed up in two equations[108]:

$$R_{it} = \alpha_i + \sum_{j=1}^{k} b_{ij} F_{jt} + \varepsilon_{it} \tag{5.72}$$

$$ER_i = \lambda_0 + \sum_{j=1}^{k} \lambda_j b_{ij} \tag{5.73}$$

which implies that return on any security is linearly related to a set of *k* factors F_{jt}, and the expected return on any security *i* may be written as a linear combination of the factor weightings b_{ij}. As it was noted above that b_{ij} s are specific to security *i*. The expected return on security *i* weights these security-specific betas by a weight λ_j, which is the same for all securities. λ_j can be interpreted as the extra expected return required because of securities sensitivity to the *j*th factor.

Testing the APT

One of the procedures of the estimation of the APT is known as *factor analysis*. A subset of all factors F_j is chosen so that the covariance between each expected return equation's residuals is zero ($E(\varepsilon_i \varepsilon_j) = 0$, $\forall \; i \neq j$, i.e. the unsystematic risk is uncorrelated across securities), which is consistent with theoretical assumption that the portfolio is fully diversified. We stop adding factors

[108] See Ross (1976).

F_j when the next factor adds little additional explanation. Thus the appropriate number of F_j s and their corresponding $b_{ij}s$ are estimated simultaneously. The λ_j s are then estimated from the cross-section regression (5.73).[109]

Roll and Ross (1980) applied factor analysis to 42 groups of stocks using daily data between 1962 and 1972. The results of the first pass regression show that for most groups of the stocks around five factors provide a sufficiently good statistical explanation of returns. In the second pass regression they find that three factors are sufficient. Dhrymes et al (1984) show that one of the problems of the interpreting the results from factor analysis is that as more securities are included in the analysis the number of significant factors appears to increase.

Sharpe(1982) , Chen (1983), Roll and Ross (1984), and Chen et al (1986) specify a wide variety of factors that might influence expected returns in the first-pass time series regression. Among these factors are dividend yield, a long-short yield spread on government bonds and changes in industrial production. The $b_{ij}s$ estimated from the first pass regression are then used as cross-section variables in the second pass regression. The authors find that that several λ_j are statistically significant thus supporting the multifactor APT. In Sharpe's results the securities' CAPM beta was significant whereas in Chen et al the CAPM beta when added to the factors already included in the APT contributed no additional explanatory power to the second pass regression. However, in all the studies mentioned above the second pass regressions have a maximum R^2 of 0.1, which indicates the presence of a great deal of noise in asset returns.

[109] There are, however, problems in interpreting the results from factor analysis. First, the signs on the b_{ij} and λ_j are arbitrary and could be reversed (e.g. a positive b_{ij} and negative λ_j is statistically indistinguishable from a negative b_{ij} and positive λ_j). Second, there is a scaling problem in that the results still hold if the b_{ij} are doubled and the λ_j halved. Finally, if the regressions are repeated on different samples of data there is no guarantee that the same factors will appear in the same order of importance.

Clare and Thomas (1994), using the APT, present empirical evidence of the pricing of the macroeconomic factors in the UK stockmarket between 1983 and 1990 grouping stocks into portfolios (both to eliminate diversifiable risk and to reduce the so-called error-in-variables problem.) before estimation. For the US data similar analysis has been formalised by Chan, Chen and Hsieh (CCH) (1985) and Chen, Roll and Ross (CRR) (1986), who find that a number of economic factors are significantly priced in the stockmarket: the changing state of the economy, inflation, yield curve shifts, and a measure of the market risk premium. Clare and Thomas attach particular importance to extending the range of economic variables beyond those considered by CCH and CRR to reflect the 'small, open economy' nature of the UK, such as exchange rate and the current account balance. The conclusions of the authors are as follows: over the period 1983-1990, using CAPM beta-sorted portfolios and time-varying risk premia associated with macroeconomic variables, they find that a number of factors have been priced in the UK stockmarkets; these are oil prices, two measures of corporate default or 'market risk', the retail price index, UK private sector bank lending, the current account balance and the redemption yield on an index of UK corporate debentures and loans. Further, the return on the market has no additional role to play once these macroeconomic factors are included.

The alternative approach to testing the APT, applied in the literature and, in particular, by McElroy et al (1985) on US portfolios, uses two basic equations:

$$R_{it} = E(R_{it}) + \sum_{j=1}^{k} b_{ij} F_{ij} + \varepsilon_{ij} \tag{5.74}$$

$$E(R_{it}) = \lambda_0 + \sum_{j=1}^{k} b_{ij} \lambda_j \tag{5.75}$$

to get:

$$R_{it} = \lambda_0 + \sum_{j=1}^{k} b_{ij}\lambda_j + \sum_{j=1}^{k} b_{ij}F_{ij} + \varepsilon_{ij} \tag{5.76}$$

If we set $\lambda_0 = r$, the risk-free rate, and compare (5.76) with regressions with unrestricted constant term α_i, then:

$$(R_{it} - r_t) = \alpha_i + \sum_{j=1}^{k} b_{ij}F_{ij} + \varepsilon_{ij} \tag{5.77}$$

And the comparison of (5.76) and (5.77) gives $\alpha_i = \sum_{j=1}^{k} b_{ij}\lambda_j$: the restrictions that should hold if the APT is the correct model. The restrictions can be tested by applying non-linear least squares. The λ_j s and b_{ij} s are jointly estimated in (5.76). The results of McElroy et al suggest that the restrictions do not hold.

The CAPM and the APT

The standard CAPM may be shown to be a special case of the APT, namely a single-factor version of the APT, where the single factor is the expected return on the market portfolio. If the return generating equation for security *i* is hypothesised to depend on only one factor and this factor is taken to be the return on the market portfolio, then the APT gives:

$$R_i = \alpha_i + b_i R_m + \varepsilon_i . \tag{5.78}$$

This single index APT equation can be shown to imply that the expected return is given by:

$$ER_i = r + b_i(ER_m - r), \tag{5.79}$$

which conforms with the equilibrium return equation for the CAPM.

The standard CAPM may also be shown to be consistent with a multi-index APT. Consider, for example, the two-factor APT:

$$R_i = \alpha_i + b_{i1}F_1 + b_{i2}F_2 + \varepsilon_i \tag{5.80}$$

and

$$ER_i = r + b_{i1}\lambda_1 + b_{i2}\lambda_2 \tag{5.81}$$

As it was mentioned above, the λ_j can be interpreted as the extra expected return required because of a securities sensitivity to the *j*th factor. Since λ_j is an excess return, then if the CAPM is true:

$$\lambda_1 = \beta_1(ER_m - r) \tag{5.82}$$

$$\lambda_2 = \beta_2(ER_m - r) \tag{5.83}$$

where the β_i are the CAPM betas.[110] Substituting for λ_j from (5.82) and (5.83) in (5.81) and rearranging we get:

$$ER_i = r + \beta_i^*(ER_m - r) \tag{5.84}$$

where $\beta_i^* = b_{i1}\beta_1 + b_{i2}\beta_2$. Thus, the two-factor APT model with the λ_j determined by the betas of the CAPM yields an equation for the expected return on asset *i*, which is of the form given by the standard CAPM itself. Hence, in testing the APT, if one finds that more than one factor λ_j is significant in (5.81), then this may or may not reject the CAPM. If the estimate of λ_j is such that this just equals $\beta_j(ER_m - r)$, where β_j is the CAPM beta, then the APT is consistent with the CAPM. If the latter restriction on λ_j does not hold, then APT and CAPM are not consistent with each other.[111]

The CAPM is more restrictive than the APT, but it has a more immediate intuitive appeal. The APT contains arbitrary elements when its empirical implementation is considered (for example, what are the appropriate factors F_j?) and may be difficult to interpret (see footnote 109).

[110] Note that the β_i is not the same as the b_{ij}.

[111] The argument is easily generalised to the *k* factor case.

CHAPTER 6

CONCLUSION

6.1 Concluding remarks

The last few decades have seen an extraordinary growth in the use of quantitative methods in financial markets. A range of models was developed to account for empirical regularities in financial data. The present work adds evidence from the transition equity markets of Central Europe to the body of work on the econometric modelling of financial time series by addressing the issues of volatility, predictability and international asset pricing in these markets. We focus on a sample of four Central European stock markets, including Hungarian, Polish, Czech and Slovak markets.

Volatility is a key variable which permeates most financial instruments and plays a central role in many areas of finance. For example, volatility is crucially important in asset pricing models and dynamic hedging strategies, as well as in determination of option prices. From an empirical standpoint, it is therefore of utmost importance to model carefully any temporal variation in the volatility process. Chapter 3 of the thesis focuses on modelling of univariate and multivariate volatility of the transition equity markets of Central Europe. We start from examination of the statistical properties of our dataset. The results reveal that it has previously documented characteristics of the unconditional distribution of financial time series, which are used to justify the GARCH specification of the time-varying volatility. The estimation of univariate GARCH specifications indicates that GARCH effects are apparent in all four series examined. Five out of the seven estimated univariate GARCH models allow different types of asymmetry in the impact of news on volatility of the transition markets. The results reveal rather weak evidence of asymmetries in the markets. The so-called leverage effects, which indicate that negative shocks introduce more volatility than positive shocks, are captured by EGARCH for the Czech and AGARCH for the Slovak returns.

The growing internationalisation of financial markets has prompted several recent empirical studies to examine the mechanism through which stock market movements are transmitted around the world. In this chapter our analysis focuses on the dynamic relationship between stock return volatility of four transition markets of Central Europe. Two multivariate GARCH models of time-varying volatility employed in our study allow us to examine the conditional correlation and the pattern of volatility spillover effects between the markets in the system. The constant correlation specification indicates significant conditional correlation between three pairs of countries: Hungary and Poland, Hungary and Czech Republic, and Poland and Czech Republic. The BEKK model of multivariate volatility fitted to the Hungarian and Polish series shows evidence of return volatility spillovers from the Hungarian to Polish market, but no volatility spillover effects are found in the opposite direction.

The <u>predictability of financial asset prices</u> is one of the earliest and most enduring problems of financial econometrics. Chapter 4 addresses the issue of return predictability in the transition markets of Central Europe by employing technical trading rules for buying and selling in the market. Technical analysis is an approach to investment management based on belief that historical price series, trading volume and other market statistics exhibit regularities that can be profitably exploited to extrapolate future price movements. In our analysis we explore one of the most popular technical rules when to buy and sell in the market: the moving average rule.

The results from application of the moving average trading strategies to our series are intriguing. We document several important stylised facts. The estimates indicate that buy signals consistently generate higher returns that sell signals. In addition we find that returns associated with sell signals are negative. This fact cannot be explained by various seasonalities since it is based on a large fraction of trading days. There are two major opinions on the ability to predict returns. Firstly, predictability arises from market inefficiencies; secondly, there may be rational variation in expected returns, resulting from an equilibrium model. Although rational changes in expected returns are possible, it is hard to imagine an equilibrium model that predicts negative returns over such a large fraction of trading days. Our further results indicate that the second moments of the distributions of buy and sell signals behave quite differently. Returns associated with buy signals are less volatile than returns associated with sell signals. This result is inconsistent with existing models of changing risk. Finally, we find that for all four markets the fraction of positive returns associated with buy periods is significantly larger, than the fraction of positive returns generated from sell signals.

The next part of the chapter employs bootstrap methodology to check whether various null models of stock returns are able to account for the predictability in nominal stock price series associated with technical trading rule profits documented in the first part of the chapter. The results reveal that none of these models (including AR, GARCH-M and EGARCH) is very successful in explaining the moving average rules profits generated by the actual data.

The inability of the linear conditional mean estimators to characterise the temporal dynamics of stock returns, as well as the asymmetric nature of returns and volatility over the periods of buy and sell signals, suggest the existence of possible non-linearities. In the next part of the chapter we compare the forecast performance of linear (AR and EGARCH) and nonlinear (feedforward networks) conditional mean estimators with past trading signals in the conditional mean equation.

The forecast improvements of the feedforward network regression over the linear conditional mean estimators are substantial. We conclude that nonlinearities play an important role in modelling the conditional mean of stock returns.

Having analysed the issues related to volatility and predictability of the stock market indices of four transition economies, in Chapter 5 we address the question of integration of these relatively new markets into the global capital market. The Sharpe-Lintner-Mossin-Black mean-variance model of exchange, the Capital Asset Pricing Model (CAPM), is the first equilibrium asset pricing model, and it remains one of the foundations of financial economics. In our analysis we test a conditional version of the CAPM in an international setting, using parsimonious GARCH parameterisation of the conditional covariance matrix of assets in the system. Methodologically, we extend the 'BEKK' multivariate GARCH specification of Engle and Kroner (1995) to accommodate GARCH-M effects. The model is estimated for a dataset including returns on the national indices from both developed (U.S. and German) and relatively new developing (Hungarian and Polish) equity markets.

International CAPM (ICAPM) implies that if international markets are fully integrated and purchasing power parity holds, then the world market risk is the only relevant pricing factor, and the assets with the same risk have identical expected return irrespective of the market. To test pricing restrictions of the benchmark standard ICAPM, we compare its performance to two alternative specifications, the partial segmentation and partial integration models, which could provide useful insights should the standard ICAPM fail to hold.

The specified models are first estimated with additional assumption that the price of market risk is constant. The results reveal that market risk, measured by the conditional covariance between the return on each national index and a world-wide index, is positive and equally priced across the countries. This is an important result, as it indicates that the assets in the transition markets command the same level of return for a given level of risk as the assets in the developed countries. It is consistent with ICAPM, and points to integration of the Hungarian and Polish markets into the global capital market. However, the estimation of the partial segmentation model shows individual significance of the Hungarian idiosyncratic risk, indicating some level of segmentation of the Hungarian market. Finally, the results of the partial integration model, which introduces a vector of world-wide information variables in the expected return equations, show that the variation in the excess returns is still predictable after accounting for market-wide risk. These results may be due, at least partially, to the auxiliary assumption of the constancy of market risk. In the next part of the chapter the price of risk is allowed to vary through time. We find that the relaxation of the constancy condition does not have any significant effect on the results from the previous testing for the specification with the constant price of risk, and the information variables introduced into the system by partial integration model are still priced in the markets.

The results of this thesis suggest directions for the further research of a range of issues related to the econometric analysis of the data from transition stock markets, and financial data in general. Examination of the distributional properties of the standardised residuals of the various univariate and multivariate ARCH specifications of time-varying volatility employed in our work reveals that, although, there is a substantial reduction in the excess kurtosis observed in nominal returns, nonnormality remains. The subsequent research includes alternative explanations of the conditional nonnormality. One possibility is the new approach emerged from the work of Taylor (1986) who formulated a discrete time stochastic volatility (SV) model as an alternative to ARCH models. Until recently estimating the SV model remained almost unfeasible. Recent advances in econometric theory have made estimation of the SV models much easier.

The results in chapter 4 are consistent with technical trading rules having predictive power in transition equity markets of Central Europe. Our analysis was focused on the simplest trading rules. It would be interesting to investigate other more elaborate technical strategies, which could potentially generate even larger differences between conditional returns.

The results in chapter 5 indicate that even though the conditional version of the traditional ICAPM applied to the dataset including new markets of Central Europe provides useful information on the dynamics of market premia, a more adequate model of international asset pricing should probably include additional factors. Alternatively, a nonlinear pricing model may be needed to explain the cross-section of returns we analyse.

REFERENCES

Abu-Mostafa, Y.S., 1995, Financial markets applications of learning from hints, in *Refenes A.P. (eds.) Neural Networks in the Capital Markets*, John Wiley & Sons, Chichester, pp. 221-232.

Adler, M. and B. Dumas., 1983, International portfolio selection and corporation finance: a synthesis, *Journal of Finance* 38, pp. 925-984.

Akgiray, V. and G.G. Booth, 1988, Mixed diffusion – jump process modelling of exchange rate movements, *Review of Economic Statistics* 70, pp. 631-637.

Alexander, S.S., 1964, Price movements in speculative markets: Trends or random walks, in *Coother. P. (eds.) The Random Character of Stock Market Prices*, MIT Press, Cambridge, Mass.

Altman, E., G. Marco and F. Varetto, 1994, Corporate distress diagnosis: Comparison using linear discrimination analysis and neural networks, *Journal of Banking and Finance* 18, pp. 505-529.

Anderson, R. and J.P. Danthine, 1983, Cross hedging, *Journal of Political Economy* 89, pp. 1182-1196.

Attanasio, O. and S. Wadhwani, 1990, Does the CAPM Explain Why the Dividend Yield Helps Predict Returns?, London School of Economics Financial Group Discussion Paper No. 4.

Baba, Y., R.F. Engle, D.F. Kraft and K.F. Kroner, 1990, Multivariate simultaneous generalised ARCH, unpublished manuscript, Department of Economics, University of California, San Diego.

Bachelier, L., 1900, Theorie de la speculation, *Annales de l'Ecole Normale Superieure* 17, pp. 21-86.

Badrinath, S. G. and S. Chaterjee, 1991, A data-analytical look at skewness and elongation in common-stock return distributions, *Journal of Business Economic Statistics* 9, pp. 223-233.

Bailie, R. and T. Bollerslev, 1987, A multivariate simultaneous generalised ARCH, unpublished manuscript, Department of Economics, University of California, San Diego.

Bailie, R. and T. Bollerslev, 1989, The message in daily exchange rates: a conditional variance tale, *Journal of Business and Economic Statistics* 7, pp. 297-305.

Bailie, R. and T. Bollerslev, 1992, Prediction in dynamic models with time-dependent conditional variances, *Journal of Econometrics* 52, pp. 91-113.

Baillie, R. and R.J. Myers, 1991, Bivariate GARCH estimation of optimal commodity futures hedge, *Journal of Applied Econometrics* 16, pp. 109-124.

Ball, C.A. and A. Roma, 1993, A jump diffusion model for the European Monetary System, *Journal of International Money and Finance* 12, pp. 475-492.

Ball, C.A. and W.N. Torous, 1985, On jumps in common stock prices and their impact on call option pricing, *Journal of Finance* 40, pp. 155-173.

Balvers, R.J., T.F. Cosimano and B. McDonald, 1990, Predicting stock returns in an efficient market, *Journal of Finance* 45, pp. 1109-1128.

Bekaert, G. and C. Harvey, 1995, Tome-varying World market integration, *Journal of Finance* 50, pp. 403-444.

Bekaert, G. and R. Hodrick, 1992, Characterising predictable components in excess returns on equity and foreign exchange markets, *Journal of Finance* 47, pp. 467-509.

Bera, A.K. and M.L. Higgins, 1998, Testing constancy of correlation with an application to international equity returns, paper presented at the Economic Society winter meeting, Chicago.

Bera, A.K., P. Garcia and J.S. Roh, 1991, Estimation of time-varying hedge ratios for agricultural commodities: BGARCH and random coefficient approaches, unpublished manuscript, Department of Economics, University of Illinois at Urbana-Champaign.

Berk, J.B., 1995, A critique of size-related anomalies, *Review of Financial Studies* 8, pp. 275-286.

Berndt, E.K., B.H. Hall, R.E. Hall and J. Hausman, 1974, Estimation and inference in nonlinear structural models, *Annals of Economic and Social Measurement* 4, pp. 653-665.

Black, F., 1976, Studies in stock volatility changes, Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association, pp. 177-181.

Black, F., Jensen, M.C. and M. Scholes, 1972, The Capital Asset Pricing Model: Some Empirical Tests, *in M.C. Jensen (ed.) Studies in the Theory of Capital Markets*, New York, Praeger.

Blume, L., D. Easley, and M. O'Hara, 1994, Market statistics and technical analysis: the role of volume, *Journal of Finance* 49, pp. 153-181.

Bollerslev, T., 1986, Generalised autoregressive conditional heteroscedasticity, *Journal of Econometrics* 31, pp. 307-327.

Bollerslev, T., 1987, A conditionally heteroskedastic time series model for speculative prices and rates of return, *Review of Economics and Statistics* 69, pp. 542-547.

Bollerslev, T., 1988a, On the correlation structure of the GARCH process, *Journal of Time Series Analysis* 9, pp.121-131.

Bollerslev, T., 1988b, A multivariate generalised ARCH model with constant conditional correlation for a set of exchange rates, unpublished manuscript, Northwestern University.

Bollerslev, T., R.Y. Chou and K.F. Kroner, 1992, ARCH modelling in finance: A review of the theory and empirical evidence, *Journal of Econometrics* 52, pp. 5-59.

Bollerslev,T., R. Engle and J. Wooldridge, 1988, A capital asset pricing model with time varying covariances, *Journal of Political Economy* 96, pp. 116-131.

Bollerslev, T. and J. Wooldridge, 1992, Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometric Reviews* 11, pp. 143-172.

Bollerslev, T., R. Engle and D. Nelson, 1994, ARCH models, *in R. Engle and D. McFadden (eds.) Handbook of Econometrics*, Vol. 4, North Holland, pp. 2961-3031.

Bookstaber, R. M. and J. B. McDonald, 1987, A general distribution for describing security price returns, *Journal of Business* 60, pp. 401-424.

Bose, N.K. and P. Liang, 1996, *Neural Networks Fundamentals with Graphs, Algorithms, and Applications*, McGraw-Hill, New-York.

Brock, W., Lakonishok and B. LeBaron, 1992, Simple technical trading rules and the stochastic properties of stock returns, *Journal of Finance* 47, pp. 1731-1764.

Brown, D., and R. Jennings, 1989, On technical analysis, *Review of financial studies* 2, pp. 527-552.

Buckberg, E., 1995, Emerging stock markets and international asset pricing, *The World Bank Economic Review* 9, pp. 51-74.

Campbell, J.Y. and Y. Hamao, 1992, Predictable bond and stock returns in U.S. and Japan: A study of long-term capital market integration, *Journal of Finance* 47, pp. 43-70.

Campbell, J.Y., A.W. Lo and A.C. MacKinlay, 1997, The Econometrics of Financial Markets, Princeton University Press, Princeton, NJ.

Campbell, J.Y. and R.J. Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, pp. 661-676.

Carlstein, E., 1986, The use of subseries values for estimating the variance of a general statistic from a stationary sequence, *Annals of Statistics* 14, pp. 1171-1179.

Chan, K., K.C. Chan, and A. Karolyi, 1991, Intraday volatility in the stock index futures markets, *The Review of Financial Studies* 4, pp. 657-684.

Chan, K.C., N. Chen and D.A. Hsieh, 1985, An exploratory investigation of the firm size effect, *Journal of Financial Economics*, Vol. 14, pp. 451-71.

Chan, K.C., G.A. Karolyi and R.M. Stulz, 1992, Global financial markets and the risk premium on U.S. equity, *Journal of Financial Economics* 32, pp. 137-168.

Charemza, W. and L. Syczewska, 1999, The application of the Dickey-Fuller and KPSS tests to the east European time series, Working Paper No. 99/3, University of Leicester.

Chen, N., 1983, Some empirical tests of the theory of arbitrage pricing, *Journal of Finance* 38, pp. 1393-1414.

Chen, N., 1991, Financial investment opportunities and macroeconomy, *Journal of Finance* 46, pp. 529-554.

Chen, J., 1995, Long-horizon Predictability of Foreign Currency Prices and Excess Returns: Alternative Procedures for Estimation and Inference, Ph.D. dissertation, Ohio State University.

Chen, N., R. Roll and S.A. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, pp. 383-403.

Cho, C., C.S. Eun and L.W. Senbet, 1986, International APT: An empirical investigation, *Journal of Finance* 41, pp. 313-330

Chobanov, G., P. Mateev, S. Mittnik and S. Rachev, 1996, Modelling the distribution of highly volatile exchange-rate time series, Arbeiten aus dem Institut fur Statistik und Okonometrie der Universitat Kiel 90.

Chopra, N., J. Lakonishok and J. Ritter, 1992, Performance measurement methodology and the question of whether stocks overreact, *Journal of Financial Economics* 31, pp. 235-268.

Chou, R.Y., 1988, Volatility persistence and stock valuations: some empirical evidence using GARCH, *Journal of Applied Econometrics* 3, pp. 279-294.

Christi, A., 1982, The stochastic behaviour of common stock variances: value, leverage and interest rate effects, *Journal of Financial Economics* 10, pp. 407-432.

Chuah, K.L., 1993, A Nonlinear Approach to Return Predictability in the Securities Markets Using Feedforward Neural Networks, Ph.D. Dissertation, Washington State University.

Clare, A.D. and S.H. Thomas, 1994, Macroeconomic factors, the APT and the UK stock market, *Journal of Business Finance and Accounting* 21, pp. 309-330.

Clements, M.P. and D.F. Hendry, 1998, *Forecasting Economic Time Series*, Cambridge University Press, Cambridge.

Coats, P. and L. Fant, 1992, A neural network approach to forecasting financial distress, *Journal of Business Forecasting* 10, pp. 9-12.

Collins, E., S. Ghosh and C. Scofield, 1988, An application of a multiple neural network system to emulation of mortgage underwriting judgements, Proceedings of IEEE International Conference on Neural Networks 2, pp. 459-466.

Conrad, J. and G. Kaul, 1988, Time-variation in expected returns, *Journal of Business* 61, pp. 409-425.

Coslow, S. and H. D. Schultz, 1966, *A Treasury of Wall Street Wisdom*, Investors Press Inc., Pallisades Park, N.Y.

Crowder, M.J., 1976, Maximum likelihood estimation with dependent observations, *Journal of the Royal Statistical Society* 38, pp. 45-53.

Cumby, L., 1990, Consumption risk and international equity returns: some empirical evidence, *Journal of International Money and Finance* 9, pp. 182-192.

Cuthbertson, K., 1996, *Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange*, John Wiley & Sons, Chichester.

Cutler, D.M., J.M. Poterba and L.M. Summers, 1991, Speculative dynamics, *Review of Economic Studies* 58, pp. 529-546.

Cybenko, G., 1988, Continuous valued neural networks with two hidden layers are sufficient, Technical Report, Department of Computer Science, Tufs University, Medford, MA.

Cybenko, G., 1989, Approximation by superposition of a sigmoidal function, *Mathematics of Control, Signals and Systems* 2, pp. 303-314.

Dannielson, J., 1994, Stochastic volatility in asset prices: Estimation with simulated maximum likelihood, *Journal of Econometrics* 64, pp. 375-400.

Davidian, M. and R.J. Carol, 1987, Variance function estimation, *Journal of American Statistical Association* 82, pp. 1079-1091.

Davidson, J., 2000, *Econometric Theory*, Blackwell Publishers.

DeBondt, W.F. and R.H. Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, pp. 793-805.

De Santis, G. and B. Gerard, 1997, International asset pricing and portfolio diversification with time-varying risk, *Journal of Finance* 52, pp. 1881-1912.

De Santis, G. and B. Gerard, 1998, How big is the premium for currency risk? *Journal of Financial Economics* 49, pp. 375-412.

Dhrymes, P.T., I. Friend and N.B. Gultekin, 1984. A Critical re-examination of the empirical evidence on the APT, *Journal of Finance* 39, pp. 323-346.

Dickey, D.A. and W.A. Fuller, 1979, Autoregressive time series with a unit root, *Journal of American Statistical Association* 74, pp. 427-431.

Dumas, B., 1994, A test of International CAPM using business cycles indicators as instrumental variables, *in Jeffrey F. (ed.) The Internatiolisation of Equity Markets*, University of Chicago Press, Chicago, pp. 23-50.

Dumas, B. and B. Solnik, 1995, The world price of the exchange rate risk, *Journal of Finance* 50, pp. 445-479.

Dutta, S. and S. Shekhar, 1988, Bond rating: A non-conservative application of neural networks, Proceedings of IEEE International Conference on Neural Networks 2, pp. 443-450.

Efron B., 1979, Bootstrap methods: another look at the jackknife, *The Annals of Statistics* 7, pp.1-26.

Efron, B. and R. Tibshirani, 1986, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Statistical Science* 1, pp. 54-77.

Emmerson, R., S. Hall and A. Zelweska-Mitura, 1996, Evolving market efficiency with an application to some Bulgarian shares, unpublished manuscript, University of Leicester.

Engle, R., 1982, Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation, *Econometrica* 50, pp. 987-1008.

Engle, R., 1983, Estimates of the variance of US inflation based upon the ARCH model, *Journal of Money, Credit and Banking* 15, pp. 286-301.

Engle, R., 1990, Discussion on Schwert (1990), *Review of Financial Studies* 3, pp. 103-106.

Engle, R. and T. Bollerslev, 1986, Modelling the persistence of conditional variances, *Econometric Reviews* 5, pp. 1-50.

Engle, R. and G. Gonzaalez-Rivera, 1991, Semiparametric ARCH models, *Journal of Business Economics and Statistics* 9, pp. 345-359.

Engle, R., C.W. Granger and D. Kraft, 1986, Combining competing forecasts of inflation using a bivariate ARCH model, *Journal of Economic Dynamics and Control* 8, pp. 151-165.

Engle, C. and J.D. Hamilton, 1990, Long swings in exchange rates: Are they in the data and do markets know it?, *American Economic Review* 80, pp. 689-713.

Erb, C., C. Harvey and T.Viskanta, 1996, Expected returns and volatility in 135 countries, unpublished manuscript.

Engle, R.F. and D. Kraft, 1983, Multiperiod forecast error variances of inflation estimated from GARCH models, *in A. Zellner (ed.) Applied Time Series Analysis of Economic Data*, Bureau of the Census, Washington DC.

Engle, R. and K. Kroner, 1995, Multivariate simultaneous generalised GARCH, *Econometric Theory* 11, pp. 122-150.

Engle, R., D.M., Lilien, and R.P. Robbins, 1987, Estimating time-varying risk premia in the term structure, *Econometrica* 55, pp. 391-407.

Engle, R. and V. Ng, 1993, Measuring and testing the impact of news on volatility, *Journal of Finance* 48, pp. 1749-1778.

Errunza, V., E. Losq and P. Padmanabhan, 1992, Tests of integration, mild segmentation and segmentation hypothesis, *Journal of Banking and Finance* 16, pp. 949-972.

Evans, M.D., 1994, Expected returns, time-varying risk, and risk premia, *Journal of Finance* 49, pp.655-679.

Fahlman, S.E. and C. Lebiere, 1990, The cascade-correlation learning algorithm, *in Touretzky, D.S. (ed.) Advances in Neural Information Processing Systems* 2, Morgan Kaufmann, San Mateo, CA, pp. 525-532.

Fama, E., 1965, The behaviour of stock market prices, *Journal of Business* 38, pp. 34-105.

Fama, E., 1990, Stock returns, expected returns, and real activity, *Journal of Finance* 45, pp. 1089-1108.

Fama, E., 1991, Efficient capital markets: II, *Journal of Finance* 46, pp. 1575-613.

Fama, E. and K.R. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, pp. 246-273.

Fama, E. and K.R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, pp. 23-49.

Fama, E. and K.R. French, 1992, The cross section of expected returns, *Journal of Finance* 47, pp. 427-465.

Fama, E.F. and MacBeth, J.D., 1973, Risk return and equilibrium: empirical tests, *Journal of Political Economy* 38, pp. 607-36.

Fama, E.F. and MacBeth, J.D., 1974, Tests of the multiperiod two-parameter model, *Journal of Financial Economics* 1, pp.43-66.

Fama, E. and G.W. Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, pp. 115-146.

Ferson, W.E., 1989, Changes in expected security returns, risk and the level of interest rates, *Journal of Finance* 44, pp. 1191-1218.

Ferson, W.E., S.R. Foerster and D.B. Keim, 1993, General tests of latent variable models and mean-variance spanning, *Journal of Finance* 48, pp. 131-156.

Ferson, W.E. and C.R. Harvey, 1991, The variation of economic premiums, *Journal of Political Economy* 99, pp. 385-415.

Ferson, W.E. and C.R. Harvey, 1993, The risk and predictability of international equity markets, *Review of Financial Studies* 6, pp. 527-566.

Ferson, W.E. and C.R. Harvey, 1994, Sources of risk and expected returns in the global equity markets, *Journal of Banking and Finance* 18, pp. 775-803.

Ferson, W.E. and R.A. Korajczyk, 1995, Do arbitrage pricing model explain predictability of stock returns?, *Journal of Business* 68, pp.309-349.

Ferson, W.E. and R.W. Schadt, 1995, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, pp. 425-461.

Fisher, F.M., 1966, *A Priori Information and Time Series Analysis: Essays in Economic Theory and Measurement*, North-Holland, Amsterdam.

Flood, K., R.J. Hodrick and P. Kaplan, 1987, An evaluation of the recent evidence on stock market bubbles, Working Paper No. 1971, National Bureau of Economic Research, Cambridge MA.

Frean, M.R., 1989, The upstart algorithm: A method for constructing and training feedforward neural networks, *Neural Computation* 2, pp. 198-209.

Freedman, D.A., 1981a, Bootstrapping regression models, *The Annals of Statistics* 9, pp. 1218-1228.

Freedman, D.A., 1981b, Bootstrapping regression models, *The Annals of Statistics* 9, pp. 1229-1238.

Freedman, D.A. and S. Peters, 1984, Bootstrapping a regression equation: Some empirical results, *Journal of American Empirical Society* 79, pp. 97-106.

French K.R. and R. Roll, 1986, Stock return variances: The arrival of information and the reaction of traders, *Journal of Financial Economics* 17, pp. 5-26.

French K. R., G. W. Schwert and R. F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, pp. 3-29.

Funahashi, K.-L., 1989, On the approximate realisation of continuos mappings by neural networks, *Neural Networks* 2, pp. 183-192.

Gallant, A.R., 1981, On the bias in flexible functional forms and an essentially unbiased form: the Fourier flexible form, *Journal of Econometrics* 15, pp. 211-244.

Gallant, A.R., 1986, Three constructive algorithms for neural learning, Proceedings of the 8th Annual Conference of Cognitive Science Society.

Gallant, A. R., and H. White, 1988, *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*, Basil Blackwell, New York.

Gallant, A. R., and H. White, 1992, On learning the derivatives of unknown mapping with multilayer feedforward networks, *Neural Networks* 5, pp. 129-138.

Gartley, H. M., 1930, *Profits in the Stock Market*, H.M. Hartley Inc., New York, N.Y.

Gately, E., 1996, *Neural Network for Financial Forecasting*, John Wiley & Sons, New York.

Gencay, R., 1996, Non-linear prediction of security returns with moving average rules, *Journal of Forecasting* 15, pp. 165-174.

Gencay, R., 1998, The predictability of security returns with simple technical trading rules, *Journal of Empirical Finance* 5, pp. 347-359.

Geweke, J., 1986, Modelling the persistence of conditional variances: comment, *Econometric Reviews* 5, pp. 57-61.

Geweke, J., 1988a, Comments on Poirier: operational Bayesian methods in econometrics, *Journal of Economic Perspectives* 2, pp. 159-166.

Geweke, J., 1988b, Exact inference in models with ARCH, *in Berndt, E., H.White, and W. Barnett (eds) Dynamic Econometric Modelling*, Cambridge University Press, Cambridge.

Geweke, J., 1989, Exact predictive densities in linear models with ARCH disturbances, *Journal of Econometrics* 44, pp. 307-325.

Ghysels, E., A. Harvey and E. Renault, 1996, Stochastic volatility, *Handbook of Statistics* 14.

Giovannini, A. and Jorion, P., 1989, The time variation of risk and return in the foreign exchange and stock markets', *Journal of Finance* 44, pp. 307-325.

Glosten, L., R. Jarannathan and D.Runkle, 1993, Relationship between the expected value and volatility of the nominal excess returns on stocks, *Journal of Finance* 48, pp. 1779-1802.

Gorr, W.L., D. Nagin and J. Szczypula, 1994, Comparative study of ANN and statistical models for predicting student grade point average, *International Journal of Forecasting* 10, pp. 1-34.

Gourieroux, C., 1992, *Modeles ARCH et Application Financieres*, Paris, Economica.

Gouroeroux, C., and A. Monfort, 1992, Qualitative threshold ARCH models, *Journal of Econometrics* 52, pp. 159-199.

Greene, W., 1993, *Econometric Analysis*, Prentice Hall.

Haefke, C. and C. Helmenstein, 1995, Predicting stock market averages to enhance profitable trading strategies, Proceedings of the Conference on Neural Networks in the Capital Markets.

Hamao, Y., Masulis, R. and Ng, V., 1990, Correlation in price changes and volatility across international stock markets, *The Review of Financial Studies* 3, pp. 281-307.

Hamilton, J.D., 1988, Rational expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates, *Journal of Economic Dynamic Control* 12, pp. 385-423.

Hamilton, J.D., 1989. Analysis of the time series subject to changes in regime, *Journal of Econometrics* 64, pp. 307-333.

Hamilton, J.D. and R. Susmel, 1994, ARCH and changes in regime, *Journal of Econometrics* 64, pp. 307-333.

Hansen, L.P., 1982, Large sample properties of the method of moment estimation, *Econometrica* 50, pp. 1029-1054.

Hansen, B.E., 1992, Autoregressive conditional density estimation, Rochester Centre for Economic Research, Working Paper No. 332.

Harvey, C., 1991, The World Price of Covariance Risk, *Journal of Finance* 46, pp.111-57.

Harvey, C., B. Solnik and G. Zhou, 1994, What determines international asset returns?, unpublished manuscript, Duke University, Durham, N.C.

Hecht-Nielsen, R., 1987, Kolmogorov's mapping neural network existence theorem, Proceedings of IEEE 1st International Conference on Neural Networks 3, pp.11-14.

Hecht-Nielsen, R., 1990, Theory of backpropagation neural networks, Proceedings of the International Joint Conference on Neural Networks, Washington, DC, I.593-I.605, New York IEEE Press.

Hergert, F., W. Finnoff and H.G. Zimmermann, 1992, A comparison of weight elimination methods for reducing complexity in neural networks, International Joint Conference on Neural Networks, Baltimore, III, pp. 980-987.

Hertz, J., A. Krogh and R.G. Palmer, 1991, *Introduction to the Theory of Neural Computation*, Addison Welsey, Redmont City.

Higgins, M.L, and A.K. Bera, 1992, A class of nonlinear ARCH models, *International Economic Reviews* 33, pp. 137-158.

Hirshleifer J., 1975, Speculation and equilibrium: Information, risk, and markets, *Quarterly Journal of Economics* 89, pp. 519-542.

Hong, P.Y., 1991, The autocorrelation structure for the GARCH-M process, *Economic Letters* 37, pp. 129-132.

Hodrick, R.J., 1987, *The Empirical Evidence of the Efficiency of Forward and Futures Foreign Exchange Markets*, Harwood, Chur, Switzerland.

Hornik, K., M. Stinchcombe and H. White, 1989, Multilayer feedforward networks are universal approximators, *Neural Networks* 2, pp. 359-366.

Hsieh, D.A., 1989, Testing for nonlinear dependence in daily foreign exchange rate changes, *Journal of Business* 62, pp. 339-368.

Hsieh, D. A. and M. H. Miller, 1990, Margin regulation and stock market volatility, *Journal of Finance* 45, pp. 3-29.

Hsu, W., L.S. Hsu and M.F. Tenorio, 1995, A neural network procedure for selecting predictive indicators in currency trading, *in Refenes A.P. (eds) Neural Networks in the Capital Markets*, John Wiley & Sons, Chichester, pp. 245-257.

Huang, C.S., 1993, Neural Networks in Financial Distress Prediction, Ph.D. dissertation, University of Mississippi.

Hull, J. and A. White, 1987, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42, pp. 281-300.

Hutchinson, J., A. Lo and T. Poggio, 1994, A nonparametric approach to pricing and hedging derivative securities via learning networks, *Journal of Finance* 99, pp. 851-889.

Jegadeesh, N., 1990, Evidence of predictive behavior of security returns, *Journal of Finance* 45, pp. 881-898.

Jensen, M.C., and G. Bennington, 1970, Random walks and technical theories: some additional evidences, *Journal of Finance* 25, pp. 469-482.

Kamijo, K.I. and T. Tanigawa, 1990, Stock price recognition – a reccurent neural network approach, Proceedings of the International Joint Conference on Neural Networks, San Diego, CA.

Karolyi, A., 1995, A Multivariate GARCH model of international transmission of stock returns and volatility: the case of the United States and Canada, *Journal of Business and Economic Statistics* 13, pp. 11-25.

Karolyi, G. A. and B-C Kho, 1994, Time-varying risk premia and returns to buying winners and selling losers: Caveat emptor et venditor, unpublished manuscript, Ohio State University.

Keim, D. and R.F. Stambaugh, 1986, Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, pp. 357-390.

Kim, B., 1994, A Study of Risk Premiums in Foreign Exchange Market, PhD dissertation, Ohio State University.

Kim, S. and N. Shephard, 1994, Stochastic volatility: Likelihood inference and comparison with ARCH models, unpublished manuscript, Nuffield College, Oxford.

Klien, B., 1977, The demand for quality adjusted cash balances: price uncertainty in the US demand for money function, *Journal of Political Economy* 85, pp. 692-715.

Kolmogorov, A.N., 1957, On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition, *Doklad Akademik Nauk USSR* 114, pp. 953-956.

Kon, S., 1994, Alternative models for the conditional heteroscedasticity of stock returns, *Journal of Business* 67, pp. 563-598.

Korajczyk, R. and C. Viallet, 1989, An empirical investigation of international asset pricing, *Review of Financial Studies* 2, pp. 553-585.

Kuan, C.M. and T. Liu, 1995, Forecasting exchange rates using feedforward and recurrent neural networks, *Journal of Applied Econometrics* 10, pp. 347-364.

Kuan, C.M. and H. White, 1994, Artificial neural networks: An econometric perspective, *Econometric Reviews* 13, pp. 1-91.

Kunsch, H.R., 1989, The jackknife and the bootstrap for general stationary observations, *Annals of Statistics* 17, pp. 1217-1241.

Lamoureux, G.C., and W.D. Lastrapes, 1990b, Persistence in variance, structural change, and the GARCH model, *Journal of Business and Economic Statistics* 8, pp. 225-234.

LeBaron, B., 1991, Technical trading rules and regime shifts in foreign exchange, unpublished manuscript, University of Wisconsin.

LeBaron, B., 1992, Do moving average trading rule results imply non-linearities in foreign exchange markets, SSRI, University of Wisconsin, Working Paper No. 9222.

LeBaron, B., 1996, Technical trading rules profitability and foreign exchange intervention, SSRI, University of Wisconsin, Working Paper No. 9445.

Lee, S.W. and B.E. Hansen, 1994, Asymptotic theory for the GARCH (1,1) quasi-maximum likelihood estimator, *Econometric Theory* 10, pp.29-52.

Lee, T.K.Y. and Y.K., Tse, 1991, Term structure of interest rates in the Singapore Asian dollar market, *Journal of Applied Econometrics* 6, pp. 143-152.

Lehmann,B.N., 1990, Fads, martingales and market efficiency, *Quarterly Journal of Economics* 105, pp. 1-28.

LeRoy, S. F., 1973, Risk aversion and the martingale property of stock returns, *International Economic Review* 14, pp. 436-446.

Levich, R. M., and L.R. Thomas, 1993, The significance of trading-rule profits in the foreign exchange market: A bootstrap approach, *Journal of International Money Finance* 12, pp. 451-474.

Levy, R. A., 1967a, Random walks: Reality or myth, *Financial Analysts Journal* 23, pp. 69-77.

Levy, R. A., 1967b, Relative strength as criterion for investment selection, *Journal of Finance* 22, pp. 595-610.

Lin, W., R. Engle and T. Ito, 1994, Do bulls and bears move across borders?, *The Review of Financial Studies* 7, pp. 507-538.

Linne, T., 1999, The integration of the central and east European equity markets into the international capital market: A kalman filter approach, Working Paper No. 99/2, University of Leicester.

Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, pp. 13-37.

Liu, R.Y. and K. Singh, 1992, Moving blocks jackknife and bootstrap capture week dependence, *in LePage, R. and L. Billard (eds) Exploring the Limits of Bootstrap*, John Wiley & Sons, New York, pp.225-248.

Lo, A.W. and A.C. MacKinlay, 1988, Stock market prices do not follow random walk: Evidence from a simple specification test, *Review of Financial Studies* 1, pp. 41-66.

Lo, A.W. and A.C. MacKinlay, 1990, When are constrain profits due to stock market overreaction?, *Review of Financial Studies* 3, pp. 175-205.

Lorentz, G.G., 1976, The 13th problem of Hilbert, *American Mathematical Society* 28.

Lucas, R. E., 1978, Asset prices in exchange economy, *Econometrica* 46, pp. 1429-1446.

Lumsdaine, R.L., 1992, Asymptotic properties of the quasi-maximum likelihood estimator in GARCH (1,1) and IGARCH (1,1) models, unpublished manuscript, Department of Economics, Princeton University.

Mandelbrot, B., 1963a, The variation of certain speculative prices, *Journal of Business* 36, pp. 394-419.

Mandelbrot, B., 1963b, New methods in statistical economics, *Journal of Political Economy* 71, pp. 421-440.

Mandelbrot, B., 1967, The variation of some other speculative prices, *Journal of Business* 40, pp. 393-413.

Mankiw, N.G. and M.D. Shapiro, 1986, Risk and return: consumption beta versus market beta, *Review of Economics and Statistics* 68, pp. 453-569.

Markovitz, H., 1959, *Portfolio Selection: Efficient Diversification of Investment*, John Wiley, New York.

McElroy, M.B., E. Bulmeister and K.D. Wall, 1985, Two estimators for the APT model when factors are measured, *Economic Letters* 19, pp. 271-275.

Merton, R., 1973, An intertemporal Capital Asset Pricing Model, *Econometrica* 41, pp. 867-887.

Merton, R. and R. Henriksson, 1981, On market timing and investment performance II: Statistical procedures for evaluating forecasting skills, Journal of Business 54, pp. 512-534.

Meszaros, D., 1996, On efficiency and anomalies of the Slovak capital market, unpublished manuscript.

Mezard, M. and J. Nadal, 1989, Learning in feedforward layered network: the tiling algorithm, *Journal of Physics A* 22, pp. 2191-2203.

Milhoj, A., 1987a, A multiplicative parametrisation of ARCH models, Research Report No. 101, Institute of Statistics, University of Copenhagen.

Neftci, D., 1991, Naive trading rules in financial markets and Wiener-Kolmogorov prediction theory: a study of 'technical analysis', *Journal of Business* 64, pp. 549-572.

Nelson, D., 1990a, Stationarity and persistence in GARCH (1,1) model, *Econometric Theory* 6, pp. 318-334.

Nelson, D., 1990b, ARCH models as diffusion approximations, *Journal of Econometrics* 45, pp.7-38.

Nelson, D., 1991, Conditional heteroskedasticity in asset returns: a new approach, *Econometrica* 59, pp. 347-70.

Nelson, D.B., and C.Q. Cao, 1992, Inequality constraints in univariate GARCH models, *Journal of Business and Economic Statistics* 10, pp. 229-235.

Nelson, C.R. and M.J. Kim, 1993, Predictable stock returns: the role of small sample bias, *Journal of Finance* 48, pp. 201-218.

Ng, L., 1989, Tests of the CAPM with time-varying covariances: a multivariate GARCH approach, MS, University of Texas at Austin, Graduate School of Business.

Nieuwland, F.G., W.F. Verschoor and C.C. Wolff, 1991, EMS exchange rates, *Journal of International Financial Markets, Institutions and Money* 2, pp.21-42.

Odom, M. and R. Sharda, 1990, A neural network for bankruptcy prediction, Proceedings of the International Joint Conference on Neural Networks 2, San Diego, CA, pp. 163-168.

Pagan, A., 1996, The econometrics of financial markets, *Journal of Empirical Finance* 3, pp.15-102.

Pagan, A.R., and Y.S. Hong, 1991, Non-parametric estimation and the risk premium, *in Barnett, W.A., J. Powell and G. Tauchen (eds) Non-parametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press, Cambridge.

Pagan, A. and G. Schewert, 1990, Alternative models for conditional stock volatility, *Journal of Econometrics* 45, pp. 267-290.

Pagan A.R. and A. Ullah, 1988, The econometric analysis of models with risk terms, *Journal of Applied Econometrics* 3, pp. 87-105.

Pau, L., 1991, Technical analysis for portfolio trading by syntactic pattern recognition, *Journal of Economic Dynamics and Control* 15, pp. 715-730.

Poddig, T., 1995, Bankruptcy prediction: A comparison with discriminant analysis, *in Refenes A.P. (eds) Neural Networks in Capital Markets*, John Wiley & Sons, Chichester, pp. 311-323.

Politis, D.N. and J.P. Romano, 1994, The stationary bootstrap, *Journal of American Statistical Association* 89, pp. 1303-1313.

Poon, S. and S.J. Taylor, 1991, Macroeconomic factors and the UK stock market, *Journal of Business and Accounting* 18, pp. 619-636.

Poterba, J.M. and L.H. Summers, 1988, Mean reversion in stock prices: evidence and implications, *Journal of Financial Economics* 22, pp. 27-59.

Qi, M. and G.S. Maddala, 1995a, Option pricing using ANN: the case of S&P 500 index call options, Neural Networks in Financial Engineering; Proceedings of the 3rd International Conference on Neural Networks in Capital Markets, London, pp. 78-91.

Qi, M. and G.S. Maddala, 1995b, Economic factors and the stock market: A new perspective, Working Paper, Department of Economics, The Ohio State University.

Roll, R. and S.A. Ross, 1984, A critical re-examination of the empirical evidence on the APT: A reply, *Journal of Finance* 39, pp. 347-350.

Refenes, A.P., 1995, Methods for optimal network design, *Neural Networks in the Capital Markets*, John Wiley & Sons, Chichester, pp. 33-54.

Refenes, A.P., A.D. Zapranis and G. Fransis, 1994, Stock performance modeling using neural networks: A comparative study with regression models, *Neural Networks* 7, pp. 375-388.

Refenes, A.P., A.D. Zapranis and G. Fransis, 1995, Modeling stock returns in the framework of APT: A comparative study with regression models, *in Refenes A.P. (eds) Neural Networks in the Capital Markets*, John Wiley & Sons, Chichester, pp. 101-125.

Rich, R.W., J. Raymond and J.S. Butler, 1991, Generalised instrumental variables estimation of ARCH models, *Economic Letters* 35, pp. 179-185.

Richardson, M.P. and T. Smith, 1993, Tests of financial models in presence of overlapping observations, *Review of Financial Studies* 4, pp. 227-257.

Rissanen, J., 1986a, A predictive least-squares principle, *IMA Journal of Mathematical Control and Information* 3, pp. 211-222.

Rissanen, J., 1986b, Stochastic complexity and modelling, *Annals of Statistics* 14, pp. 1080-1100.

Rissanen, J., 1987, Stochastic complexity (with discussions), *Journal of the Royal Statistical Society* B 49, pp. 223-239 and 252-265.

Rissanen, J., 1989, *Stochastic Complexity in Statistical Inquiry*, World Science Publishing Co., Singapore.

Roll, R., 1977, A Critique of Asset Pricing Theory's Tests, *Journal of Financial Economics* 4, pp. 1073-1103.

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

Roll, R. and S.A. Ross, 1980, An empirical investigation of the APT, *Journal of Finance* 35, pp. 1073-1103.

Ross, S.A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory*13, pp.341-360.

Ruiz, E., 1993, Stochastic Volatility versus ARCH, Universidad Carlos III de Madrid, Working Paper.

Sabau, H.C.L., 1987, The structure of GMM and ML estimators in conditionally heteroskedastic models, Working Paper No. 153, Faculty of Economics, Australian National University.

Sabau, H.C.L., 1988, Some Theoretical Aspects of Econometric Inference with Heteroskedastic Models, Ph.D. dissertation, Australian National University.

Salchenberger, L., E. Cinar and N. Lash, 1992, Neural Networks: A new tool for predicting bank failures, *Decision Sciences* 23, pp. 899-916.

Schwert, G.W., 1990, Stock returns and real activity: A century of evidence, *Journal of Finance* 45, pp. 1237-1257.

Scott, L.O., 1991, Financial Markets Volatility, *IMF Staff Papers* 38, pp. 347-365.

Sentana, E. And S. Wadhwani, 1992, Feedback traders and stock return autocorrelations: evidence from the century of daily data', *Economic Journal* 102, pp. 415-425.

Shanken, J., 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, pp.1-33.

Shanken, J. and M. Weinstein, 1990, Macroeconomic variables and asset pricing: estimation and tests, Working Paper, University of Rochester.

Sharpe, W., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, pp. 425-442.

Sharpe, W., 1982, Factors on NYSE Security Returns 1931-79, *Journal of Portfolio Management* 8, pp.5-19.

Shea, G. S., 1989a, Ex-post rational price approximations and the empirical reliability of the present-value relation, *Journal of Applied Econometrics* 4, pp. 139-159.

Shea, G. S., 1989b, A re-examination of excess rational price approximations and excess volatility in the stock market, *in R.C. Guimaraes et al. (eds.) A Re-appraisal of the Efficiency of Financial Markets*, pp. 469-494.

Sietsma, J. and R.F. Dow, 1991, Creating ANNs that generalise, *Neural Networks* 4, pp. 67-79.

Smith, P.N., 1993, Modeling risk premia in international equity markets, *European Economic Review* 37, pp. 159-176.

Solnik, B., 1974a, An equilibrium model of the international capital market, *Journal of Economic Theory* 8, pp. 500-524.

Solnik, B., 1974b, The international pricing of risk: an empirical investigation of the world capital market structure, *Journal of Finance* 29, pp. 365-378.

Solnik, B., 1974c, Why not diversify internationally rather than domestically? *Financial Analysts Journal* 30, pp. 48-54.

Solnik, B., 1983, International APT, *Journal of Finance* 38, pp. 449-457.

Sprecher, D.A., 1965, On the structure of functions of several variables, *Transatlantic American Mathematical Society* 115, pp. 340-355.

Stehle, R., 1977, An empirical test of the alternative hypotheses of national and international pricing of risky assets, *Journal of Finance* 32, pp. 493-502.

Stulz, R., 1981, A model of international asset pricing, *Journal of Financial Economics* 9, pp. 383-406.

Stulz, R., 1995, International portfolio choice and asset pricing: an integrative survey, *in Maksimovic, V., Ziemba, W. (eds.), The Handbook of Modern Finance*, Noth-Holland, Amsterdam.

Stinchcombe, M. and H. White, 1989, Universal approximation using feedforward networks with non-sigmoidal hidden layer activation function, *Proceedings of International Joint Conference on Neural Networks* 1, San Diego, IEEE Press, pp. 612-617.

Swanson, N.R. and H. White, 1995a, A model selection approach to assessing the information in the term structure using linear models and artificial neural networks, *Journal of Business and Economic Statistics* 13, pp. 265-275.

Swanson, N.R. and H. White, 1995b, A model selection approach to real time macroeconomic forecasting using linear models and artificial neural networks, Working Paper , Department of Economics, Pennsylvania State University.

Tam, K.Y. and Y.M. Kiang, 1990, Predicting bank failures: A neural network approach, *Applied Artificial Intelligence* 4, pp. 265-282.

Tam, K.Y. and Y.M. Kiang, 1992, Managerial application of neural networks: The case of bank failure predictions, *Management Science* 38, pp. 926-947.

Tauchen, G., H. Zhang and M. Liu, 1994, Volume volatility and leverage analysis, unpublished manuscript, Duke University.

Taylor, S. J., 1986, *Modelling Financial Time Series*, John Wiley, Chichester.

Taylor, S.J., 1994, Modelling stochastic volatility: A review and comparative study, *Mathematical Finance* 4, pp. 183-204.

Taylor, M., and H. Allen, 1992, The use of technical analysis in the foreign exchange market, *Journal of International Money and Finance* 11, pp. 304-314.

Thomas, S.H. and M.R. Wickens, 1989, International CAPM: why has it failed, CEPR Discussion Paper No. 354.

Treynor, J., and R. Ferguson, 1985, In defense of technical analysis, *Journal of Finance* 40, pp. 757-773.

Treynor, J. and K. Mazuy, 1966, Can mutual funds outguess the markets? *Harward Business Review* 44, pp. 131-136.

Tsibouris, G.C., 1993, *Essays on Nonlinear Models of Foreign Exchange*, Ph.D. Dissertation, University of Wisconsin-Madison.

Wajerowska, E., 1999, Validity of the optimal portfolio allocation model with price constraints on the example of the Warsaw Stock Exchange, Working Paper No. 99/5, University of Leicester.

Wasswerman, P., 1993, *Advanced Methods in Neural Computing*, Van Nostrand Reinhold, New York.

Weiss, A.A., 1986, Asymptotic theory for ARCH models: Estimation and testing, *Econometric Theory* 2, pp.107-131.

Wheatley, S., 1988, Some tests of International Equity Integration, *Journal of Financial Economics* 21, pp. 177-212.

White, H., 1988, Economic prediction using neural networks: The case of the IBM daily stock returns, *Proceedings of the IEEE International Conference on Neural Networks*.

White, H., 1990, Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings, *Neural Networks* 3, pp. 535-549.

White, H., A.R. Gallant, K. Hornik, M. Stinchcombe and J. Wooldridge, 1992, *Artificial Neural Networks: Approximation and Learning Theory*, Blackwell Publishers, Cambridge.

Wickens, M.R. and S.H. Thomas, 1989, Non-parametric estimates of the equity and foreign exchange risk premia, *CEPR Discussion Paper* No. 356.

Zakoian, J. M., 1990, Threshold heteroskedastic model, unpublished manuscript, INSEE, Paris.

Zalewska-Mitura, A. and S.G. Hall, 1999, Do market participants learn? The case of the Budapest Stock Exchange, Working Paper No. 99/4, University of Leicester.