MASTER THESIS

The Predictability of Asset Returns: An Empirical Analysis of Central-European Stock Markets

Filip Žikeš

Charles University in Prague
Faculty of Social Sciences
Institute of Economic Studies

July 18, 2003

Abstract

The aim of the thesis is to investigate the predictability of Central-European common stock returns. In the first chapter, using weekly data on the Czech, Hungarian and Polish major value-weighted indices in the period 1996:1 to 2002:12, it is shown that the index returns contain predictable components and that both the mean and volatility can be forecasted from the time-series of historical returns. In Chapter 2, the Johansen cointegration analysis is applied to the weekly data on the Czech, Hungarian, Polish and German equity market index prices in the period 1998:1 to 2002:12. The results indicate that the four considered markets are cointegrated when prices are expressed in local currencies, whereas no cointegration was found for prices in terms of Euro. In both case, there is significant cross-country predictability, i.e. lagged returns from one market can be used to predict returns from at least one other market. The third chapter is dedicated to studying the impact of nonsynchronous trading on the predictability of stock returns. The Lo and MacKinlay (1990) econometric model is generalized to allow for an autocorrelated common factor. Finally, in Chapter 4 the economic significance of stock return predictability is evaluated. A dynamic trading strategy based on a maximally predictable portfolio is developed for monthly returns on Polish stocks in the period 2000:1 to 2002:12 and its performance evaluated using various market timing measures. The results imply economically significant predictability in the Polish stock returns.

Keywords: emerging markets, predictability of stock returns, cointegration, conditional heteroskedasticity, nonsynchronous trading.

JEL Classification: C22, C32, E44, G14, G15

Abstract

[in Czech] Cílem této diplomové práce je analýza prediktibility výnosů akcií na středoevropských kapitálových trzích. V první kapitole je ukázáno, že v období 1996:1 až 2002:12 bylo možno predikovat výnosy i volatilitu hlavních akciových indexů českého, polského a maďarského trhu pomocí časových řad historických výnosů. V druhé kapitole je aplikována Johansenova kointegrační analýza na týdenní ceny akciových indexů českého, maďarského, polského a německého kapitálového trhu v období 1998:1 až 2002:12. Z analýzy vyplývá, že čtyři výše uvedené trhy jsou kointegrované, pokud používáme ceny vyjádřené v národních měnách, přičemž kointegraci nebylo možno prokázat pro ceny vyjádřené v Euru. V obou případech lze ale predikovat výnosy národních indexů pomocí zpožděných hodnot výnosů alespoň jednoho z ostatních indexů. Třetí kapitola se zabývá vlivem nesynchronního obchodování na prediktibilitu výnosů akcií. Na základě zobecněného ekonometrického modelu nesynchronního obchodování (Lo a MacKinlay, 1990) je ukázáno, že nesynchronní obchodování vnáší do časových řad výnosů akcií zdánlivou autokorelaci. Poslední čtvrtá kapitola je zaměřena na posouzení ekonomického významu prediktibility výnosů polských akcií. Porovnává se zde výkonnost dynamické obchodní strategie založené na maximálně predikovatelném portfoliu s pasivní investiční strategií. Z analýzy polských vyplývá, že prediktibilita výnosů byla v období 2000:1 až 2002:12 ekonomicky významná, tj. bylo možno dosáhnout nadprůměrného výnosu po započtení transakčních nákladů.

Klíčová slova: výnosy akcií a jejich predikce, lokální heteroskedasticita, kointegrace, nesynchronní obchodování.

JEL klasifikace: C22, C32, E44, G14, G15

Contents

Acknowledgments

First and foremost, I would like to thank Miloslav S. Vošvrda of the Faculty of Social Sciences, Charles University for supervising my work on this thesis. I have greatly benefited from numerous discussions on various topics in finance and econometrics we have had over the past two years. I owe a great deal to Dr. Vošvrda for many useful suggestions and comments regarding the topic of this thesis. I am also grateful to: Martin Netuka of the Faculty of Social Sciences, Charles University for his reading the earlier drafts and providing valuable comments and insights; Marketa Elisova of the Faculty of Mathematics and Physics, Charles University, for her help with *Scientific WorkPlace* and LaTeX by which this thesis was typeset; Marketa Fiserova of Bloomberg for providing the data used in my research. The financial support of the Grant Agency of Charles University under grant no. 287/2003/A-EK/FSV is also greatly appreciated.

Introduction

Until the late 1980's, the Random Walk Model has dominated the finance literature as a theory best describing the time-series behavior of stock prices in an efficient capital market. In his famous paper, Samuelson (1965) argues that the active participation of many profit-seeking investors must result into random fluctuation of stock prices, provided that the prices properly incorporate the expectations and information of all investors. In other words, if all investors form rational expectations about future stock price movements based on all currently available information, then any sign of stock return predictability will be immediately exploited and thus ceases to exist. The Samuelson version of the Efficient Market Hypothesis, however, fails to take into account the risk associated with holding stocks. We will show below that random walk is not a necessary condition for an informationally efficient market and that market efficiency may hold even if stock returns are forecastable. The confusion of the Efficient Market Hypothesis with the Random Walk Hypothesis is, however, not uncommon in the empirical literature and hence, readers beware[1].

Grossman and Stiglitz (1980) provide an even stronger argument against the Random Walk Model for stock prices. They assert that a stock market cannot be informationally efficient in equilibrium, because acquiring information is costly. If all information is already incorporated in the stock prices then there would be no private return for collecting information and hence no incentive for investors to trade. As a result, the market would collapse. Only if there is some degree of informational inefficiency would investors exert effort to gather information and trade, but this situation is clearly inconsistent with completely unpredictable stock price fluctuations, i.e. the Random Walk Model.

Despite these sound theoretical arguments against the Random Walk Model, it was not until the seminal paper by Lo and MacKinlay (1988), where the null hypothesis of random walk was decisively rejected for portfolios of American stocks. Subsequently, voluminous research emerged indicating that stock prices in various developed equity markets around the world contain predictable components. New methods and techniques have been developed for detecting predictabilities in stock returns and evaluating their statistical and economic significance. A state-of-the-art reference in this field is the collection of seminal papers by Andrew W. Lo and A. Craig MacKinlay published together under the title "A Non-Random Walk Down Wall Street"[2].

The purpose of this thesis is to investigate the predictability of stock returns on Central-European equity markets. In particular, we focus on the Czech, Hungarian and Polish

[1] For example, Hanousek and Filer (1996, pp. 2) write: "A market is weekly efficient if [stock] prices fully reflect all information contained in historical [stock] price series. Such efficiency implies that stocks follow a random walk..."

[2] Lo and MacKinlay (1999).

capital markets and consider the German stock market as a benchmark. The thesis is divided into four chapters. With a slight modification, each chapter can be considered a separate article or working paper, but combining them together in a single work provides a more accurate and comprehensive picture of the sources and significance of stock return predictability. Although the thesis is an empirical one, we always provide a brief overview of the underlying theory and describe the methodology being applied to the time-series of stock returns under study. Basic knowledge of modern finance and time-series econometrics is, however, assumed.

We start by time-series analysis of the Central-European stock index returns in Chapter 1. The Random Walk Hypothesis is put to test and an alternative parametric model (ARIMA) is proposed for modeling and forecasting stock returns. We also consider forecasting the volatility of returns using a GARCH model with t-distributed innovations. This approach allows for explicit modeling of the 'fat tails' usually observed in the stock return distributions.

In Chapter 2, we study stock return predictability in a multivariate context. We apply the Johansen (1988, 1991) cointegration analysis to the Czech, Hungarian, Polish and German stock index prices to investigate whether these equity markets 'move together' or if they evolve independently. We analyze the short-term and long-term dynamics of the index returns and its implications for cross-country predictability, i.e. the forecast power of lagged returns from one market in predicting future returns from other markets.

Chapter 3 is dedicated to studying the effect of nontrading and nonsynchronous trading on the time-series properties of stock returns. Within the framework of the Lo-MacKinlay (1990) econometric model of nonsynchronous trading we show that the empirically observed autocorrelation of stock returns may be spurious, induced by infrequent trading. We generalize the model by letting the common factor driving stock returns follow a stationary first-order autoregressive process. This specification allows for a direct decomposition of the estimated first-order autocorrelation coefficient of stock returns into the (spurious) part induced by nonsynchronous trading and the (real) part inherent to the returns.

Finally, Chapter 4 is aimed at evaluating the economic significance of stock return predictability in Central-European equity markets. We first present the methodology of maximizing predictability of stock returns in a context of a multifactor forecasting model, where various macroeconomic and term-structure variables are considered as factors. On the basis of the maximally predictable portfolio, we follow a simple dynamic asset allocation strategy and asses its out-of-sample performance using various measures of market-timing skills. The analysis is performed on Czech and Polish stock returns.

The thesis concludes with a brief summary of the empirical results and suggestions for further research.

Chapter 1

Time-Series Predictability

Time-series predictability of asset returns is primarily concerned with forecasting future returns from the time series of past returns[1]. It has always attracted a great deal of attention from academics but also from the professional community due to its apparent simplicity and intuitive appeal. To forecast future returns from past returns, all an investor needs is the time series of historical prices and a powerful econometric software package that will enable her to perform the necessary tests and estimation to uncover possible regular patterns in the stock prices. And since the marginal costs of obtaining the data and econometric software package are negligible in these days, potential gains from correctly predicting future market movements can be enormous.

This chapter presents modern econometric methods and tools aimed at detecting the presence of regular patterns in asset returns. These are not confined to stock returns but can be readily applied to other financial instruments, real estate, etc. The chapter is organized as follows. In Section 1.1 we briefly introduce the Efficient Market Hypothesis (EMH) which spurred the research on return predictability. We define the Fair Game Model and the Random Walk Model and discuss the impossibility of testing the EMH per se. In Section 1.2 we present the methodology of a simple specification test of the random walk hypothesis. Next, in Section 1.3 comes an alternative to the random walk, the autoregressive integrated moving average process, which can be used to forecast the conditional mean of a time series when the random walk hypothesis does not hold. Since it is useful to forecast not only the mean but also the variance of asset returns[2], Section 1.4 describes the widely used model of time-varying volatility, the generalized autoregressive conditional heteroskedastic process. Finally, the empirical results for the Czech, Polish and Hungarian stock returns will be reported in Section 1.5 followed by a short discussion and suggestions for future research in Section 1.6.

---

[1] In Fama(1991), the field of time-series predictability also includes forecasting future returns from variables like dividend yields (D/P), price/earnings ratios (P/E), and term structure variables. But since we are concerned with the Central-European stock markets, where the data on D/P and P/E are generally not available, we will not consider such possibility. The forecast power of the term structure variables will be investigated in Chapter 4 of the thesis.

[2] The variance of returns is, for example, a key parameter in derivative pricing models. Hence an accurate estimate of a derivatives price requires an efficient estimate of future volatility. See Hull (2000), Chapter 15, for details.

1.1 The Efficient Market Hypothesis

One of the most important concepts in modern finance has been the Efficient Market Hypothesis. In his classic review, Fama (1970) summarizes that "a market in which prices always 'fully reflect' available information is called 'efficient'." The quotation marks suggest that it has to be further defined what 'fully reflects' means. Bellow, we will focus on two models of market efficiency, the Fair Game Model and the Random Walk Model as outlined in Fama (1970). We also have to specify what subset of available information is to be 'fully reflected' in stock prices. The classical taxonomy of information sets, due to Roberts (1967), distinguishes among

- **Weak-Form Efficiency:** The information set includes only historical prices of returns.
- **Semistrong-Form Efficiency:** The information set includes all publicly available information (i.e. information available to *all* agents).
- **Strong-Form Efficiency:** The information set includes all privately available information (i.e. information available to *any* agent).

In our work, we will be concerned with the Weak-Form Efficiency only since our goal is to investigate the predictability of stock returns from the time series of historical returns. We now turn to formally defining what "fully reflects" means.

1.1.1 The Fair-Game Model

The Fair Game Model is based on the assumption that the conditions of market equilibrium can be stated in terms of expected returns and that expected returns are formed on the basis of the information set, which we denote Φ_t. Formally,

$$E\left[P_{jt+1} \mid \Phi_t\right] = (1 + E\left[r_{jt+1} \mid \Phi_t\right])P_{jt}, \tag{1.1}$$

where P_{jt} is the price of security j at time t and r_{jt} is its return. The expected equilibrium return $E\left[r_{jt+1} \mid \Phi_t\right]$ usually comes from a particular equilibrium theory (such as CAPM, APT, etc.) but regardless of what equilibrium theory used the information set Φ_t is fully utilized when determining the expected return. This rules out the possibility of a trading strategy based on Φ_t that would yield systematically returns above expected equilibrium returns. To see this, let z_{jt+1}denote the excess return at time t on security j, i.e.

$$z_{jt+1} \equiv r_{jt+1} - E\left[r_{jt+1} \mid \Phi_t\right]. \tag{1.2}$$

But if the information set Φ_t is fully utilized when forming expectations then

$$E\left[z_{jt+1} \mid \Phi_t\right] = 0,$$

and hence the sequence of excess returns is a fair game with respect to the information set Φ_t. Note that the fair game model does not imply that returns must be serially uncorrelated.

In the weak-form efficiency tests, Φ_t contains only the time series of historical returns. The covariance between successive returns can be then written as

$$Cov\left[r_{jt}, r_{jt+1}\right] = \int_{r_{jt}} (r_{jt} - E\left[r_{jt}\right])(E\left[r_{jt+1} \mid r_{jt}\right] - E\left[r_{jt+1}\right]) f(r_{jt}) dr_{jt} \tag{1.3}$$

where $f(r_{jt})$ denotes the marginal distribution of r_{jt}. The fair-game property (1.2) does not imply that $E\left[r_{jt+1} \mid r_{jt}\right] = E\left[r_{jt+1}\right]$ and thus the covariance between successive returns can be non-zero. What is important for market efficiency, however, is that the excess return z_{jt} be serially uncorrelated. Hence the presence of serial correlation in returns does not necessarily imply weak-form inefficiency. The expected return can well be serially correlated in such a way that the excess returns is unpredictable. Elton and Gruber (1995) provide an example of such situation. Consider a stock of a company successively increasing its debt/equity ratio. This results into successive increases in risk associated with holding the stock and hence in increases in expected return. The realized returns will exhibit serial correlation but since the expected return has also successively increased, this information cannot be used to earn positive excess return.

1.1.2 The Random Walk Model

The Random Walk Model can be viewed as a special case of the Fair Game Model. Suppose that market environment is such that the process generating new information along with the evolution of investors tastes produces equilibria in which return distributions repeat themselves through time (Fama, 1970). Furthermore, if we assume that the fact that a security's price fully reflects available information implies independence of successive returns, we arrive at the Random Walk Model. Formally,

$$f(r_{jt+1} \mid \Phi_t) = f(r_{jt+1}), \tag{1.4}$$

where the density function f is the same for all t. The serial covariance of returns is zero at all leads and lags and the expected return is equal to the unconditional mean of the distribution f at all times. If r_{jt} denotes the one-period continuously compounded rate of return[3], then the log-price process, p_{jt}, can be written as

$$p_{jt} = \mu + p_{jt-1} + \varepsilon_{jt}, \tag{1.5}$$

where ε_{jt} is an *iid* zero-mean random variable[4] and μ denotes the mean of the distribution f. Equation (1.5) is the common form of the random walk model for stock prices used in the literature. The reason for formulating the random walk in natural logarithms is due to the limited liability property of stock prices (stock prices cannot be negative).

Clearly, the excess return in the Random Walk Model is a fair game with respect to the information set Φ_t, since

$$E\left[z_{jt+1} \mid \Phi_t\right] = \mu - \mu = 0.$$

[3]The one-period continuously compounded rate of return is the solution to the equation $P_t = e^r P_{t-1}$, i.e. $r = p_t - p_{t-1}$, where $p_t \equiv \ln P_t$.

[4]Throughout the thesis, we use *iid* to denote independently identically distributed random variable.

The Random Walk Model is, however, more restrictive than the Fair Game Model since it requires successive returns to be independently identically distributed. It follows that random walk is a sufficient but not necessary condition for market efficiency and hence rejecting the null hypothesis of random walk does not imply that the market is inefficient.

The hypothesis of identity of distributions of returns can be relaxed to allow for time-varying volatility. There has been mounting evidence in the literature that stock returns are conditionally heteroskedastic[5] and since we are primarily concerned with the predictability of the mean of returns, rejection of the random walk hypothesis due to heteroskedasticity is not of much interest.

From the discussion above follows that the only true test of market efficiency is a test based on the excess return. To perform such a test one needs to assume a particular equilibrium theory that generates expected returns, $E\left[r_{jt+1} \mid \Phi_t\right]$. But this inevitably leads to the problem of joint hypothesis of market efficiency and equilibrium theory: when we find that the behavior of returns is not in line with the implications of the EMH, it remains ambiguous in what way it should be split between market inefficiency and an inappropriate model of market equilibrium (Fama, 1991). Hence the Efficient Market Hypothesis per se is not testable.

At this point the reader may wonder why we have spent so much time discussing market efficiency when, after all, it is not testable in practice. The reasons are twofold. First, the EMH provides some theoretical framework for studying asset return predictability. The Random Walk Model is clearly a natural starting point of all tests of return predictability: when the null hypothesis of random walk cannot be rejected then it is very unlikely that the subsequent tests would reveal any regular patterns in stock returns. Moreover, the failure to reject the null provides evidence in favor of the EMH. Second, as Fama (1991) writes, '[the empirical literature on efficiency] has changed out views about the behavior of returns, across securities and through time'. Although researchers do not agree about the implications of the tests for market efficiency (due to the joint-hypothesis problem), they agree on the fact that these tests have substantially helped to improve out knowledge of the behavior of stock returns. Thus the Efficient Market Hypothesis provides the background for now a more general area of asset return predictability.

1.2 A Random Walk Test

This section presents a simple specification test due to Lo and MacKinlay (1988) aimed at testing the random walk hypothesis outlined above. Recall that the random walk for a log-stock price p_t can be written as[6]

$$p_t = \mu + p_{t-1} + \varepsilon_t, \tag{1.6}$$

where μ is the expected one-period continuously compounded rate of return on the stock and $\{\varepsilon_t\}$ is a sequence of independently but not necessarily identically distributed random variables with finite variances $\sigma^2_{\varepsilon_t}$. Note that we do not assume that the ε_t's are Gaussian. To test the null hypothesis of random walk, Lo and MacKinlay (1988) exploit the fact that

[5] See Section 1.4 for more details on conditional heteroskedasticity.

[6] We drop the subscribt j for simplicity.

the variance of a q-th difference of the process (1.6) is equal to the sum of the corresponding q first-difference variances. To see this, note that the q-th difference of (1.6) can be written as

$$\begin{aligned} p_t - p_{t-q} &= (p_t - p_{t-1}) + (p_{t-1} - p_{t-2}) + \cdots + (p_{t-q+1} - p_{t-q}), \\ &= \mu q + \sum_{i=t-q+1}^{t} \varepsilon_i, \end{aligned}$$

and thus the variance of $p_t - p_{t-q}$ is equal to

$$\begin{aligned} Var\left[p_t - p_{t-q}\right] &= \sum_{i=t-q+1}^{t} \sigma^2_{\varepsilon_i}, \\ &= \sum_{i=t-q+1}^{t} Var\left[p_i - p_{i-1}\right]. \end{aligned}$$

It follows that the ratio $Var\left[p_t - p_{t-q}\right] / \sum_{i=t-q+1}^{t} Var\left[p_i - p_{i-1}\right]$ must be equal to one under the null hypothesis and the test based on this idea is called the variance ratio test.

To build a test statistics suppose we collect $nq+1$ observations on the log-stock price p_t. Define the estimators:

$$\begin{aligned} \hat{\mu} &\equiv \frac{1}{nq} \sum_{k=1}^{nq} (p_k - p_{k-1}) = \frac{1}{nq}(p_{nq} - p_0), \\ \hat{\sigma}^2_a &\equiv \frac{1}{nq-1} \sum_{k=1}^{nq} (p_k - p_{k-1} - \hat{\mu})^2, \\ \hat{\sigma}^2_b(q) &\equiv \frac{1}{m} \sum_{k=q}^{nq} (p_k - p_{k-q} - q\hat{\mu})^2, \end{aligned}$$

where $m = q(nq - q + 1)(1 - 1/n)$, and define the statistic:

$$\hat{M}_r(q) = \frac{\hat{\sigma}^2_b(q)}{\hat{\sigma}^2_a} - 1. \tag{1.7}$$

Then Lo and MacKinlay (1988) show that under the null hypothesis of random walk with fairly general forms of heteroskedasticity[7] in the sequence $\{\varepsilon_t\}$, the test statistic

$$z^*(q) \equiv \sqrt{nq}\hat{M}_r(q)/\sqrt{\hat{\theta}} \tag{1.8}$$

converges asymptotically to the standard normal distribution, i.e. $z^*(q) \overset{a}{\sim} N(0,1)$. The variable $\hat{\theta}$ in (1.8) is a heteroskedasticity-consistent estimator of the asymptotic variance of $\hat{M}_r(q)$. See Lo and MacKinlay (1988) for details on the formula for $\hat{\theta}$.

[7] For a detailed description of the heteroskedastic null hypothesis, see Lo and MacKinlay (1988). For our purposes, it is sufficient to note that this null hypothesis allows for GARCH processes in the error term ε_t, and hence the test is robust to conditional heteroskedasticity in stock returns.

To develop some intuition for the variance ratio (1.7), it can be shown that the following equality obtains asymptotically[8]:

$$\hat{M}_r(q) \stackrel{a}{=} \sum_{j=1}^{q-1} \frac{2(q-j)}{q} \hat{\rho}(j), \tag{1.9}$$

where $\hat{\rho}(j)$ is the j-th order autocorrelation coefficient of the first-differences of p_t. From (1.9) we see, that the variance ratio can be expresses as a weighted sum of the first $q-1$ autocorrelation coefficients of the first differences of p_t with positive arithmetically declining weights. Under the null hypothesis of random walk, all autocorrelation coefficients are equal to zero and hence $\hat{M}_r(q)$ is also equal to zero. The sequence of the continuously compounded rate of return is simply a white noise process with possibly time-varying variances.

1.3 An Alternative to Random Walk: ARIMA

There has been mounting evidence in the literature that stock prices do not follow random walk. To give just a few examples, Lo and MacKinlay (1988) apply their variance ratio test to the weakly returns on size-sorted portfolios of NYSE-AMEX stocks and decisively reject the null hypothesis of random walk. Gilmore and McManus (2001) study the random walk hypothesis using weakly data on Central-European equity markets (Czech Republic, Poland and Hungary) and find significant autocorrelation in stock returns, thereby rejecting the random walk model. Similarly, Filacek et al. (1998) investigate the weak-form efficiency of the Prague Stock Exchange using daily data on the main value-weighted market index PX-50 and conclude that the returns are significantly positively autocorrelated.

1.3.1 The ARIMA(p,1,q) Model

A natural generalization of the Random Walk Model better approximating the data generating process of stock returns is to allow for some pattern of serial correlation in the error process $\{\varepsilon_t\}$. Consider an infinite-order moving average process for $\{\varepsilon_t\}$:

$$p_t = \kappa + p_{t-1} + \psi(L)\varepsilon_t, \tag{1.10}$$

where $\psi(L)$ denotes an infinite-order polynomial in the lag operator. Although this representation is fairly general, it is impossible in practice to fit an infinite number of parameters to the data. Moreover, we usually prefer a parsimonious model over a too complicated one. It is thus useful to make an additional assumption regarding $\psi(L)$, namely that it can be expressed as a ratio of two finite-order polynomials in the lag operator:

$$\psi(L) = \frac{\theta(L)}{\phi(L)} = \frac{1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q}{1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p}. \tag{1.11}$$

This expression is meaningful provided that all roots of the polynomial $\phi(L)$ lie outside the unit circle. Substituting (1.11) into (1.10) and multiplying by $\phi(L)$ we obtain

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)(1-L)p_t = \mu + (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)\varepsilon_t, \tag{1.12}$$

[8] See Lo and MacKinlay (1988) for the derivation of (1.9).

where $\mu \equiv (1 - \phi_1 - \phi_2 - \cdots - \phi_p)\kappa$. Expression (1.12) is the *autoregressive integrated moving average process,* denoted $p_t \sim$ARIMA$(p, 1, q)$. The order of integration is equal to one in this particular case, since from (1.10) follows that first-differencing the time series of log-stock price is sufficient to achieve stationarity. The usual way to estimate a model of the type of (1.12) is to apply the methodology developed by Box and Jenkins (1972). This involves four steps. First, the time series under study has to be differenced to achieve stationarity. In our case, first-differencing should suffice, but the null hypothesis of stationarity of the differenced time series should be tested rigorously before proceeding further in the analysis. We will return to this problem later. Second, the autocorrelation and partial autocorrelation functions are examined to make an initial guess about the values of p and q. Then the model (1.12) is estimated[9] and finally, diagnostic checks are performed to determine the appropriateness of the model. When it fails, steps three and four are repeated with different values of p and q. Since the Box-Jenkins methodology is well-known and widely used we will not discuss is in greater detail here. The interested reader is urged to consult Box and Jenkins (1972) or, for example, Enders (1996).

A common objection among economists to the class of models introduced above and the methodology used to estimate them is that they are usually not based on any rigorous economic theory. Rather, the idea "let the data speak" is invoked when selecting the parameters p and q. This approach thus necessarily leads to problems with interpreting the results. Suppose, for example, that a researcher finds that a time series of monthly inflation appears to follow an ARIMA$(3, 1, 2)$. Although this may be quite useful for short-term predictions of future inflation, no economic theory implies or predicts that this should be indeed the case. Hence we do not learn much about the way inflation is determined in the real economy with these models. A similar argument clearly holds for the application of the Box-Jenkins methodology to stock prices. But this must not imply that ARIMA modelling of stock prices should be completely discarded. As we show in a moment, there are several more or less well theoretically founded specifications of an ARIMA model for stock prices that guide our selection of the parameters p and q and help improve our understanding of the behavior stock market prices over time.

1.3.2 Market 'Fads'

The empirical findings of significant autocorrelation in stock returns have led some authors to proposing an alternative theory to the Random Walk Hypothesis. Summers (1986), Poterba and Summers (1988), and Fama and French (1988), for example, propose the following representation of the log-stock price process $\{p_t\}$:

$$p_t = p_t^* + u_t, \tag{1.13}$$

where p_t^* is the fundamental value of the security at time t, and u_t is a stationary component that reflects the departure of the price of a security from its fundamental value. Summers (1986) argues that although this model violates the Efficient Market Hypothesis

[9] The model is usually estimated using non-linear least squares or maximum likelihood. We will return to the estimation method in the next chapter, where we show that an ARIMA model can be simultaneously estimated with a model of time-varying volatility for the error process $\{\varepsilon_t\}$.

by allowing the price to systematically differ from its fundamental value there are theoretical explanations for this behavior. One such explanation is what he calls market 'fads' or 'animal spirits'. Investors do not always price assets perfectly rationally but tend to be subject to temporary enthusiasm or pessimism about future economic conditions and market development. This behavior clearly drives the price apart from its fundamental value but since the irrational enthusiasm of pessimism is assumed to be only temporary, the stock price will eventually return to its fundamental. To be specific, Summers (1986) assumes that the fundamental value and the irrational component evolve according to the following stochastic processes:

$$p_t^* = \mu + p_{t-1}^* + \varepsilon_t, \quad (1.14)$$

$$u_t = \alpha u_{t-1} + v_t, \quad (1.15)$$

where $\{\varepsilon_t\}$ and $\{v_t\}$ are independent white-noise processes with variances σ_ε^2 and σ_v^2, respectively, and α is assumed to be close to but less than one. Hence, the fundamental value p_t^* represents a permanent component of the stock prices (is follows a random walk), whereas u_t is a transitory though fairly persistent component. The closer the parameter α is to one, the larger and longer and the swings of the stock price from its fundamental value. Summers (1986) then argues that as α approaches one, it is virtually impossible to reject the null hypothesis of random walk for the stock price p_t since in such circumstances the process for u_t is indistinguishable from random walk in finite samples. Consequently, the failure to reject the null hypothesis of random walk (and hence market efficiency) is not evidence in favor of its acceptance: the parameter α may be very close to but less than one and hence the market is extremely inefficient due to very persistent (though temporary) deviations from the fundamental value. The available statistical tests cannot simply detect this pattern. Nevertheless, the log-stock price process defined in (1.13),(1.14) and (1.15) has interesting time-series implications. One can show that equations (1.13),(1.14) and (1.15) imply that p_t follows an ARIMA$(1, 1, 1)$ process of the form[10]

$$(1 - \rho L)(1 - L)p_t = c + (1 - \theta L)e_t, \quad (1.16)$$

[10] To see this, write the first-difference of (1.13)

$$p_t - p_{t-1} = p_t^* - p_{t-1}^* + u_t - u_{t-1}.$$

Now lag this equation one period, multiply it by α and subtract it from the above equation. This yields after rearranging

$$(1 - \alpha)\Delta p_t - (1 - \alpha)\mu = \varepsilon_t - \alpha\varepsilon_{t-1} + v_t - v_{t-1}.$$

The expression on the right-hand side of this equation is a sum of two independent MA(1) processes and hence it is an MA(1) process (see Hamilton,1994, for a proof of this result). The coefficient θ and the variance σ_e^2 can then be computed by equating the variance and first-order covariance of $e_t - \theta e_{t-1}$ and $\varepsilon_t - \alpha\varepsilon_{t-1} + v_t - v_{t-1}$. For computational simplicity, we made the assumption (without loss of generality) that $\sigma_\varepsilon^2 = 1$.

where $\{e_t\}$ is a white-noise process and

$$\begin{aligned} \rho &= \alpha, \\ c &= (1-\alpha)\mu. \\ \theta &= \left\{1+\alpha^2+2\sigma_v^2+(1-\alpha)\left[(1+\alpha^2+4\sigma_v^2\right]^{1/2}\right\}/(2\alpha+2\sigma_v^2), \\ \sigma_e^2 &= (\alpha+\sigma_v^2)/\theta. \end{aligned}$$

This ARIMA specification for stock price is, however, fairly restrictive because it implies that the stock returns for all time intervals (e.g. days, months, years, etc.) are negatively autocorrelated. To see this, note that the continuously compounded rate of return from time t to $t+T$, which we denote $r_{t,t+T}$, implied by (1.13) is given by

$$\begin{aligned} r_{t,t+T} &= p_{t+T}-p_t, \\ &= (p^*_{t+T}-p^*_t)+(u_{t+T}-u_t). \end{aligned}$$

Fama and French (1988) show that the first-order autocorrelation coefficient of the T-period continuously compounded rate of return, which is the slope coefficient in the regression of $r_{t,t+T}$ on $r_{t-T,t}$ can be written as

$$\rho_1(T) = \frac{Cov\left[r_{t,t+T}, r_{t-T,t}\right]}{Var\left[r_{t-T,t}\right]}, \tag{1.17}$$

$$= \frac{-(1-\alpha^T)^2 Var\left[u_t\right]}{Var\left[u_{t+T}-u_t\right]+Var\left[p^*_{t+T}-p^*_t\right]}, \tag{1.18}$$

provided that the random walk and stationary components of the stock price are uncorrelated. Note that $\rho_1(T)$ is negative for all T regardless of the sign of α. It is interesting to examine the behavior of (3.18) as T approaches infinity. Fama and French (1988) demonstrate that as $T \to \infty$, the stationary component tends to push $\rho_1(T)$ towards -0.5 (in the absence of the random-walk component $\rho_1(T) \to -0.5$ as $T \to \infty$). The variance of the random-walk component explodes, however, as $T \to \infty$ and this tends to push $\rho_1(T)$ towards zero. To sum up, if the stock price has both components, $\rho_1(T)$ might have a U-shaped pattern, starting close to zero for small T, then becoming more negative as T grows larger, but eventually the variance of the random walk begins to dominate and $\rho_1(T)$ moves back to zero. This property of the random-walk-plus-stationary-AR(1) model, however, has not observed empirically: although both Fama and French (1988) and Poterba and Summers (1988) find that the returns of longer holding periods (larger than one year) appear to be negatively serially correlated[11], Poterba and Summers (1988) also examine shorter holding periods and find significant positive autocorrelation in monthly returns. Their findings are similar to those of Lo and MacKinlay (1988) who detect significant positive autocorrelation in weekly returns. Thus the empirical results for American stocks are inconsistent with the model given in equations (1.13),(1.14) and (1.15).

The market fads hypothesis provides a candidate theoretical framework for estimating an ARIMA model for stock prices in that it specifies a particular source of return predictability

[11] Both studies examine American stocks only.

and identifies the possible values of p and q. Of course, one may argue that it is not a rigorous theory. But in light of the recent burst of the technology bubble, which is a prominent example of irrational enthusiasm of investors, along with the fact that the 2002 Nobel Prize in economics was awarded to Daniel Kahneman for his work on psychological models of irrational behavior, the hypothesis of market 'fads' or 'animal spirits' has definitely gained in importance. It remains the task of future research to formulate a sound theory of irrational behavior of stock market agent that will allow for more general pattern of return autocorrelation that the simple model of Summers (1986).

1.3.3 Time-Varying Expected Return

A second example of an ARIMA model that has been applied in the empirical literature which will be briefly discussed here is based on Conrad and Kaul (1988). Recall from Section 1.1 that market efficiency does not impose any conditions on the time-series properties of the expected return. All that is required from an efficient market is that the excess return be unpredictable from the time series of the past returns. Also, as Conrad and Kaul (1988, pp. 410) note, "existing asset-pricing theories do not specify any particular a priori restrictions on the variations through time in expected returns". Usually, an asset pricing model would specify the cross-sectional structure of expected returns (CAPM, APT, etc.) without concretely addressing the question of how the expected return should evolve over time. These considerations led Conrad and Kaul (1988) to try to characterize the time-series properties of expected returns empirically, given market efficiency.

They propose a first-order autoregressive process for the expected return. Their model is thus given by

$$r_t = E_{t-1}[r_t] + \varepsilon_t, \tag{1.19}$$

$$E_{t-1}[r_t] = \phi E_{t-2}[r_{t-1}] - u_{t-1}, \tag{1.20}$$

where r_t is the realized return on a stock in period from $t-1$ to t, $E_{t-1}[r_t]$ is the expected return for a security over period from $t-1$ to t based on the information available at $t-1$, $\{\varepsilon_t\}$ and $\{u_t\}$ are independent white-noise processes with zero means and variances σ_ε^2, σ_u^2, respectively and $\phi \leq 1$. Note that although the expected return is serially correlated, the excess return

$$z_t = r_t - E_{t-1}[r_t] = \varepsilon_t,$$

is independent through time and hence the model in equations (1.19) and (1.20) is consistent with the weak-form Efficient Market Hypothesis. The choice of the AR(1) process for the expected returns is based on empirical evidence on cross-sectional stock return predictability. The economic variables used in such studies (see Fama, 1991, for an overview)[12] to explain the cross sectional behavior of the realized stock returns are usually highly autocorrelated themselves implying that the expected returns should reflect this variation through time. The choice of the first-order autoregression and not a higher-order autoregression is then made for simplicity and parsimony.

[12]The cross-sectional behavior of stock returns will be discussed in greater detail in Chapter 4 of the thesis, where the predictive power of various economic variables will be investigated.

Without discussing the estimation procedure[13] used by Conrad and Kaul (1988) to extract the unobserved expected return and estimate the autoregressive parameter ϕ, we merely note that the model in (1.19) and (1.20) implies that the realized returns follow an ARIMA(1, 1, 1) model:

$$(1 - \phi L)(1 - L)p_t = (1 - \theta L)a_t,$$

where $\{a_t\}$ is a white-noise process with variance σ_a^2. The parameters θ and σ_a^2 can be computed in a similar way as those in (1.16)[14].

The two examples of an ARIMA model for stock returns described above provide some theoretical guidance for identifying the model. Although they both imply that stock returns contain predictable components, they differ fundamentally in addressing the question of whether this fact can be used to attain systematically positive access return. If the Summers's (1986) hypothesis holds, i.e. if the market is extremely inefficient, then clearly there exists a trading strategy that yields superior returns. If, on the other hand, market efficiency hypothesis holds and the serial correlation in returns is induced by autocorrelated risk factors driving the expected stock returns as suggested by Conrad and Kaul (1988), then the fact that returns are predictable cannot be used to achieve abnormal returns. It remains therefore to be determined, which of the two hypothesis is supported by empirical results. We will touch upon this question in Section 1.5 and in Chapter 4. Now, we will give two more examples of an ARIMA model for stock prices that have become fairly common in empirical studies of developed stock markets. These are not based on the concept of market efficiency or expected return but rather on market microstructure arguments.

1.3.4 Nonsynchronous Trading

It is the common practice in the vast majority of stock market empirical studies to implicitly assume that the data used for the study were sampled at equidistant points in time. Usually, a researcher chooses a particular data frequency (daily, weekly, monthly, etc.) that best suits the purpose of his/her analysis and takes for granted that the data were really sampled at times implied by the data frequency. To give an example, it is common to use daily data on stock prices for testing various hypothesis about emerging stock markets. This is because the history of these stock markets is short which rules out the possibility of obtaining a sample of statistically plausible size with lower frequency. The daily stock price is typically defined as the closing security price on that day. That is a price prevailing at the stock exchange close. Hence it may appear that recording all closing stock prices simultaneously implies that daily prices are sampled simultaneously and at equidistant points in time (the time of the stock exchange closing is the same for all trading days). But this argument is clearly false. The closing stock price, though recorded at the time of the stock exchange close, was actually sampled at the time when the last trade in that security on that particular day was executed. For stocks of large corporations traded on developed stock markets these two times usually almost coincide since trading in these securities takes place almost continuously ('almost' continuously here means within intervals of seconds or at maximum tens of seconds). For small stocks or for stocks traded at emerging markets

[13] They use a Kalman filter to estimate their model. See Conrad and Kaul (1988, p.411) for details.

[14] See footnote 9.

this need not be, however, the case. It is not uncommon for such stocks not to trade at all during some trading days.

To formalize these considerations, Lo and MacKinlay (1990) develop an econometric model of nonsynchronous trading. Since a generalized version of their model and its empirical application will be the topic of Chapter 3 of the thesis, we will introduce this model here only briefly and demonstrate what implications for the time-series properties of stock returns the effect of nonsynchronous trading has.

Assume that the virtual unobserved return on a security j at time t, denoted r_{jt}, is generated by the following stochastic process:

$$r_{jt} = \mu_j + \beta_j \Lambda_t + \varepsilon_{jt}, \tag{1.21}$$

where μ_j is the security's expected return, Λ_t is an *iid* zero-mean common factor and ε_{jt} is an *iid* zero-mean idiosyncratic noise. Assume further that Λ_t and ε_{jt} are independent across securities and through time. This specification of the return-generating process implies that the unobserved stock price follows random walk[15]. In each period there is some positive probability p_j that security j does not trade. Then the observed return, denoted r_{jt}^0, is equal to zero, although the unobserved return r_{jt} is given by (1.21). In the next period security j again does not trade with probability p_j and the process continues. If the security trades in period t and did not trade for the past n consecutive periods then the observed return for period t is defined as the sum of the n past virtual returns and the virtual return for period t. Under these assumptions, along with the assumption that the binary process governing the trading-nontrading behavior is a sequence of *iid* Bernoulli variables, Lo and MacKinlay (1990) show that the moments of the observed return r_{jt}^0 are given by:

$$E\left[r_{jt}^0\right] = \mu_j, \tag{1.22}$$

$$Var\left[r_{jt}^0\right] = \sigma_j^2 + \frac{2p_j}{1-p_j}\mu_j^2, \tag{1.23}$$

$$Cov\left[r_{jt}^0, r_{jt+n}^0\right] = -\mu_j^2 p_j^n, \tag{1.24}$$

where $\sigma_j^2 \equiv Var\left[r_{jt}\right]$. From (1.22), (1.23) and (1.24) we see that nontrading increases the variance of returns and induces negative spurious autocorrelation into the observed return time-series that decays geometrically. Similar results can be derived for portfolios of securities with identical nontrading probabilities. The important message of this model is that ignoring the effect of non-synchronicity in the data may result into completely false inferences: when the researcher neglects this effect he/she may find that stock returns are autocorrelated and hence predictable. But from the model we learn that the autocorrelation is not a symptom of genuine predictability (the unobserved returns are unpredictable) but rather it is caused by infrequent trading and hence it is spurious. This in turn implies that this type of autocorrelation of returns cannot be used as a basis for a trading strategy aimed at achieving abnormal returns. Suppose, for example, that an investor finds significant

[15]The unobserved return, r_{jt} is a sum of a constant and two independent white-noise processes, hence it is independent through time. But since r_{jt} is the first-difference of the log-price process, this implies that the log-price process follows random walk.

serial correlation in stock returns. He may (falsely) attribute this evidence to genuine predictability and base his decision as to whether to buy the security or sell it short on the forecasted return from his (ARMA) model. But since there is some positive probability that the security does not trade each day, the investor may find that he/she is indeed unable to buy or sell the security on some days and hence his/her trading strategy brakes down. Thus just like in the model of Conrad and Kaul (1988) this trading strategy does not produce systematically positive excess return.

The fact that the autocorrelation of observed returns, r^0_{jt}, declines geometrically is consistent with an ARIMA$(1,1,0)$ process for the observed returns. This specification is also supported by empirical studies (see, for example, Lo and MacKinlay (1988)). Hence the effect of nonsynchronous trading provides another candidate theoretical framework for ARIMA modelling of stock returns. And since this effect is likely to be particularly important in the Central-European stock market, where liquidity is extremely low for the majority of traded stocks, we will dedicate an entire chapter to examining its impacts in greater detail.

1.3.5 The Bid-Ask Spread

Another example of spurious autocorrelation induced by market microstructure is a model of the bid-ask spread due to Roll (1984). The bid price of a security is the highest price at which investors are willing to buy the security and the ask price is the lowest price at which investors are willing to sell the security. The difference between these two prices is the bid-ask spread. It is reasonable to assume that the true value of the security is somewhere in between the bid and ask prices. Roll (1984), for example, assumes that the true price is in the middle of the bid-ask interval. He then shows that even if the true value of the security follows a random walk, the bid-ask spread induces spurious autocorrelation into the stock returns. To see this, let $\{p_t\}$ denote the log-price process and $\{p^*_t\}$ denote the unobserved process for the log-true value of the security. Define p_t and p^*_t as follows:

$$p_t = p^*_t + \frac{s}{2} I_t + \varepsilon_t, \quad (1.25)$$

$$p^*_t = \mu + p^*_{t-1} + v_t, \quad (1.26)$$

$$I_t = \begin{cases} 1 \text{ with probability } 1/2, \\ -1 \text{ with probability } 1/2, \end{cases} \quad (1.27)$$

where s is the bid-ask spread, $\{\varepsilon_t\}$ and $\{v_t\}$ are independent zero-mean white-noise processes with variances σ^2_ε and σ^2_v, respectively, and $\{I_t\}$ is a sequence of *iid* Bernoulli random variable with probability $1/2$. The continuously compounded rate of return in the model (1.25), (1.26) and (1.27) then obeys

$$r_t = \Delta p_t = \Delta p^*_t + \frac{s}{2}\Delta I_t + \Delta \varepsilon_t, \quad (1.28)$$

$$= \mu + v_t + \varepsilon_t - \varepsilon_{t-1} + \frac{s}{2}(I_t - I_{t-1}). \quad (1.29)$$

The term on the right-hand side of (1.28) is a sum of a) a constant, b) a white-noise process, c) an MA(1) process, and d) an MA(1) process. Since ε_t, v_t, and I_t are assumed independent, it follows that r_t is an MA(1) process and hence p_t follows an ARIMA$(0, 1, 1)$ process that can be written as

$$(1 - L)p_t = \mu + (1 - \theta L)a_t, \tag{1.30}$$

where $\{a_t\}$ is a zero-mean white-noise process with variance σ_a^2. The parameters θ and σ_a^2 can be derived in a similar way as in the Summer's model (see footnote 9).

The model of the bid-ask spread completes our exposition of the possible theoretical arguments for ARIMA modelling of stock prices. We have seen that various theoretical considerations can explain the time-series properties of stock returns. It remains the task of empirical research to determine which of the models discussed above are likely to account for the empirically found predictable components of stock prices. But before we turn to empirical applications we present a model of time-varying volatility, which can substantially enhance our knowledge of the time-series behavior of stock returns and increase the efficiency of estimation of the ARIMA-type models.

1.4 Time-Varying Volatility: GARCH

One of the main principles of modern finance is that expected return is somehow related to risk. There are numerous equilibrium models specifying this relationship (see Elton and Gruber (1995) for details). For instance, in the standard Sharpe-Lintner-Mossin CAPM, expected return is proportional to systematic risk, i.e. to risk that cannot be diversified away. Although there is not a general agreement in the literature as to what model best explains the reality, it is widely accepted that when agents are risk-averse, expected return is an increasing function of risk.

A widely used measure of the risk of an asset is the standard deviation of returns from the unconditional mean. This measure can be loosely interpreted as long-run volatility, since it appears to be determined by various economic fundamentals characteristic for a particular asset and is usually assumed constant for the period under study. It has been observed, however, in numerous studies that returns are not homoskedastic and that "large price changes tend to be followed by large changes of either sign and small changes tend to be followed by small changes of either sign" (Mandelbrot, 1963, p.418). Also it has been observed empirically that the unconditional distributions of stock returns exhibits fatter tails (see, for example, Fama, 1965) than what would be consistent with the normal distribution. But the distributional properties of stock returns have important implications for asset pricing: in the above mentioned CAPM model, for example, the variances and covariances are used as a measure of risk, but as Bollerslev (1987, p.542) notes, "depending on the distribution of the returns, the variance may not be a valid or sufficient statistic to use". Another example of the importance of the distributional assumptions is the well-known Black-Scholes option pricing model, which rests on the hypothesis of normality of the underlying stock returns and assumes constant volatility of returns over the life of the option. Clearly, correctly specifying the distribution of asset returns is a necessary condition for rational asset pricing.

In this section, we present a model of time-varying volatility originally introduced by Engle (1982) and later generalized by Bollerslev (1986). This model will allow us to pin down the distributional properties of stock returns and it will also provide a means of predicting future volatility from the time-series of historical returns[16].

1.4.1 The GARCH(r, p) Model

Consider a standard linear regression model

$$y_t = x_t^T \beta + \varepsilon_t, \tag{1.31}$$

where x_t is a vector of explanatory variables, possibly including lagged dependent variable and moving average terms, β is a vector of unknown parameters to be estimated and $\{\varepsilon_t\}$ is a white-noise process. Although ε_t is uncorrelated with ε_{t-j} for all j, the observation that large changes tend to be followed by large changes of either sign cited above suggests that the same need not be true for ε_t^2 and ε_{t-j}^2. In his pioneering work, Engle (1982) proposes the following model of the error process:

$$\varepsilon_t = u_t \sqrt{h_t}, \tag{1.32}$$
$$h_t = \zeta + \alpha_1^2 \varepsilon_{t-1}^2 + \alpha_2^2 \varepsilon_{t-2}^2 + \cdots + \alpha_p^2 \varepsilon_{t-p}^2, \tag{1.33}$$

where $\{u_t\}$ is a white-noise process with zero mean and unit variance and h_t is the conditional variance of ε_t, conditioned on the information available at time $t-1$. In particular, the conditional variance h_t is assumed to linearly depend on the past p squared disturbances from the regression model in (1.31). A white-noise process satisfying (1.32) and (1.33) is the *autoregressive conditional heteroskedastic process*, denoted $\varepsilon_t \sim$ARCH(p).

In a generalization of the process (1.33) introduced by Bollerslev (1986), the conditional variance of ε_t depends on an infinite number of lagged squared disturbances:

$$h_t = \zeta + \psi(L)\varepsilon_t^2, \tag{1.34}$$

where $\psi(L)$ is an infinite-order polynomial in the lag operator. By parametrizing $\psi(L)$ as a ratio of two finite-order polynomials, $\alpha(L)/\ \delta(L)$, and assuming that the roots of the polynomial $\delta(L)$ lie outside the unit circle, the following expression for the conditional variance can be derived:

$$h_t = \omega + \delta_1 h_{t-1} + \delta_2 h_{t-2} + \cdots \delta_r h_{t-r} + \alpha_1^2 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2, \tag{1.35}$$

where $\omega \equiv (1-\delta_1-\delta_2-\cdots-\delta_r)$. Equation (1.35) is the *generalized autoregressive conditional heteroskedastic process*, denoted $\varepsilon_t \sim$GARCH(r, p). The reader may note the resemblance of the ARCH and GARCH processes to the ARMA models. Indeed, it can be shown that if ε_t follows an ARCH(p) process then ε_t^2 follows an ARMA$(p, 0)$. Similarly, if $\varepsilon_t \sim$GARCH(r, p) then $\varepsilon_t^2 \sim$ARMA(m, r), where $m = \max(p, r)$[17]. These facts may prove to be useful for identification purposes as we will see in the subsequent empirical application.

[16] For a more rigorous treatment of the ARCH family of models, see Hamilton (1994) or Engle (1995). An extensive review of theory and empirical evidence from developed stock markets is provided in Bollerslev et al. (1992).

[17] See Hamilton (1994, Ch.21) on the proof of these results.

To develop some intuition for the GARCH model, consider a simple GARCH(1,1) specification[18]:

$$\begin{aligned} h_t &= \omega + \delta_1 h_{t-1} + \alpha_1 \varepsilon_{t-1}^2, \\ &= \omega + (\delta_1 + \alpha_1) h_{t-1} + \alpha_1 (\varepsilon_{t-1}^2 - h_{t-1}). \end{aligned} \quad (1.36)$$

The expected variance h_t is a weighted sum of the variance expected for the past period, h_{t-1}, and the unexpected shock to the variance in the past period, $\varepsilon_{t-1}^2 - h_{t-1}$. The parameter α_1 measures the impact of the unexpected shock on the expected next period's volatility and $(\delta_1 + \alpha_1)$ measures the rate at which the unexpected shock dies out over time. The process for the conditional variance is covariance-stationary provided that $\delta_1 + \alpha_1 < 1$. The closer $\delta_1 + \alpha_1$ is to 1, the more persistent the volatility is. A general GARCH(r, p) process is covariance-stationary if and only if $\delta(1) + \alpha(1) < 1$. To ensure non-negativity of volatility, ω, δ_i,and α_j are assumed positive.

The models of autoregressive conditional heteroskedasticity have important implications for the distributional properties of the unconditional disturbances, ε_t. One of the measures used to describe the shape of the density function of a random variable is the standardized fourth moment, i.e. the kurtosis (denoted K). For a random variable that is normally distributed, kurtosis is equal to 3. When $K > 3$ there is more probability mass in the tails of the density function then in the tails of a normal density function. Such distribution are called *leptokurtic*. When, on the other hand, $K < 3$, the opposite holds and such distributions are called *platykurtic*. To see how conditional heteroskedasticity effects the unconditional distribution of the error terms in equation (1.31), assume that $u_t \overset{iid}{\sim} N(0,1)$. Then we have[19]

$$K(u_t) = \frac{E\left[\varepsilon_t^4\right]}{\left(E\left[h_t\right]\right)^2} = \frac{E\left[u_t^4\right] E\left[h_t^2\right]}{\left(E\left[h_t\right]\right)^2} \geq \frac{3(E\left[h_t\right])^2}{\left(E\left[h_t\right]\right)^2} = 3. \quad (1.37)$$

where the second equality follows from the independence of u_t and h_t, and the inequality follows from the Jensen's inequality. Thus the unconditional distribution of the error term ε_t will have fatter tails even though the conditional distribution is normal. Recall that excess kurtosis for unconditional stock returns is consistent with some empirical findings (e.g. Fama,1965). Although the conditional normality assumption in the GARCH model generates some degree of unconditional excess kurtosis (equation (1.37)), this is usually insufficient to account for the fat-tailed properties of the data: the standardized residuals from a GARCH model for stock returns, $\hat{u}_t = \hat{\varepsilon}_t / \sqrt{\hat{h}_t}$, often appear to be leptokurtic[20]. Bollerslev (1987) therefore suggests an alternative to the conditional normality assumption: the standardized Student's t-distribution with the number of the degrees of freedom to be estimated. There are also other alternatives, such as the generalized exponential distribution or the stable Paretian distribution[21]. In our work we will focus on the Student's t-distribution since it is fairly simple to implement in empirical applications.

[18] The GARCH(1,1) model is the most widely used model in the empirical finance literature.

[19] We assume throughout this section that the fourth moment of ε_t exists. This need not be generally the case. See Hamilton (1994, Ch.21) for details on the conditions for existence of the fourth moment of ε_t.

[20] See, for example, Bollerslev (1987) or Connolly (1989).

[21] See Bollerslev et al. (1992) for a full list of alternatives that have been applied in the literature.

1.4.2 Estimation by Maximum Likelihood

The distributional considerations bring us to the problem of estimation of a GARCH model. The vast majority of empirical studies use the method of maximum likelihood to estimate the parameters of the model given in equations (1.31), (1.32) and (1.35). The usual assumption is made that the error term is conditionally normally distributed, i.e. $\varepsilon_t \sim N(0, h_t)$. Then the maximum likelihood estimator is consistent and asymptotically normally distributed even if the true distribution is non-normal. The standard errors of the estimates, however, are inconsistent if the true distribution is not normal, but Bollerslev and Wooldridge (1988) derive a robust estimator of the covariance matrix yielding consistent estimates of the standard errors under such circumstances. This method is called the quasi-maximum likelihood estimation (QMLE). Alternatively, one can assume a different conditional distribution for the error term. Following Bollerslev (1987), we assume that ε_t is conditionally t-distributed with v degrees of freedom, i.e.,

$$\begin{aligned}(\varepsilon_t \mid h_t) &\sim f_v(\varepsilon_t \mid h_t) = \\ &= \Gamma\left(\frac{v+1}{2}\right)\Gamma\left(\frac{v}{2}\right]^{-1}((v-2)h_t)^{-1/2}\left(1+\frac{\varepsilon_t^2}{h_t(v-2)}\right)^{-(v+1)/2}, v > 2 \qquad (1.38)\end{aligned}$$

where $f_v(\varepsilon_t \mid h_t)$ denotes the conditional density function and $\Gamma(\cdot)$ is the gamma function. It is well-known property of the Student's t-distribution that as the number of the degrees of freedom increases without bound, the t-distribution approaches a normal distribution with zero mean and variance h_t, but for finite v, the t-distribution has fatter tails than the corresponding normal distribution. Hence assuming the conditional t-distribution in connection with a GARCH model may better account for the excess kurtosis found in the stock market returns. Moreover, the estimated number of degrees of freedom, $\hat{v}$ may indicate the source of the excess kurtosis in the unconditional stock returns: according to Connolly (1992), if $\hat{v} < 10$, both non-normality and conditional heteroskedasticity explain the excess kurtosis in returns, whereas if $\hat{v} > 30$ conditional heteroskedasticity is the only source of fat tails in the unconditional distribution of returns.

To estimate the model given in (1.31), (1.32), (1.35) and (1.38) by the method of maximum likelihood we collect the unknown parameters (including v) in a vector θ. The log-likelihood function is given by

$$\ell(\theta) = \sum_{t=1}^{T} \log f_v(\varepsilon_t \mid h_t). \qquad (1.39)$$

To maximize (1.39) we employ the algorithm developed in Berndt, Hall, Hall and Hausman (1974). The obtained parameter estimates are asymptotically normally distributed, consistent and asymptotically efficient under the null hypothesis that ε_t is conditionally t-distributed. To test against the null hypothesis of conditionally normal errors (i.e. $1/v = 0$) we use the usual likelihood ratio test statistic[22]. The test statistic is asymptotically χ^2 distributed with one degree of freedom. Similarly, we can test the null hypothesis that the

[22] According to Bollerslev (1987), the usual test statistic will likely be more concentrated towards the origin that a χ_1^2 distribution because $1/v$ is on the boundary of the admissible parameter space. He notes, however, that for sample sizes of one hundred and more the bias is very small. In the subsequent empirical application, we will neglect this bias because our sample contains over 300 observations.

disturbances ε_t are unconditionally Student's t-distributed (i.e. $\delta_1 = \delta_2 = \cdots = \delta_r = \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$) by referring the LR test statistic to the critical values of the χ^2 distribution with $r + p$ degrees of freedom.

As an alternative to MLE, the GARCH(r, p) model can be estimated by the Generalized Method of Moments (GMM)[23]. An advantage of this approach is that one does not need to assume any particular distribution of the error term ε_t. This advantage, however, comes at a cost: if the true distribution of the error term is the Student's t, then MLE outlined above produces asymptotically efficient estimates, whereas GMM does not.

1.4.3 Testing for ARCH effects

Engle (1982) derives a Lagrange multiplier test for the null hypothesis of no ARCH effects in the residuals from the regression (1.31). This test is run by regressing the estimated squared residuals $\hat{\varepsilon}_t^2$ on the past q values of $\hat{\varepsilon}_t^2$, i.e.

$$\hat{\varepsilon}_t^2 = \gamma_1 \hat{\varepsilon}_{t-1}^2 + \gamma_2 \hat{\varepsilon}_{t-2}^2 + \cdots \gamma_q \hat{\varepsilon}_{t-q}^2 + \upsilon_t. \tag{1.40}$$

Provided that the residuals in (1.31) were estimated consistently, the TR^2 from the regression (1.40) is asymptotically χ^2 distributed with q degrees of freedom. An asymptotically equivalent alternative testing procedure to the LM test is to run the Ljung-Box Q-test[24] for serial correlation on the squared residuals $\hat{\varepsilon}_t^2$. The Q-statistic is given by

$$Q(p) = T(T+2) \sum_{j=1}^{p} \frac{r_j^2}{T-j},$$

where r_j^2 denotes the squared j-th order autocorrelation coefficient of $\hat{\varepsilon}_t^2$. Under the null hypothesis that there is no serial correlation in $\hat{\varepsilon}_t^2$ up to lag p the $Q(p)$ statistic is asymptotically χ^2 distributed with p degrees of freedom.

1.4.4 Diagnostics: The BDS test

Once a GARCH-type model has been estimated, it needs to be established whether the model is correctly specified, i.e. if it captures all serial correlation in the squared residuals. A widely applied approach is to run the ARCH-LM test on the standardized residuals, $\hat{u}_t = \hat{\varepsilon}_t / \sqrt{h_t}$. If the model is correctly specified the squared standardized residuals should be serially uncorrelated and hence the ARCH-LM test should fail to reject the null hypothesis of homoskedasticity of $\hat{u}_t$.

Some authors have suggested applying the correlation integral-based test statistic developed by Brock et al. (1987)[25]. The correlation integral is the probability that a randomly selected pair of points is close in the sense of the L^∞-norm. The points are n-dimensional vectors formed from the data as follows:

$$x_t^n = (x_{t-n+1}, \ldots, x_t).$$

[23]See Campbel et al. (1999, Ch.12) for details.

[24]Ljung and Box (1978).

[25]See Bollerslev et al. (1992, pp.23) for an overview.

The parameter n is called the embedding dimension. The proximity parameter, which we denote k, defines 'how close' to each other the points should be. The correlation integral, $C_n(k)$, clearly depends on the two parameters. Under the null hypothesis that the data are *iid*, Brock et al. (1987) show that for any n, $C_n(k) = [C_1(k)]^n$ and they exploit this property to build a test statistic for testing the null hypothesis:

$$J_{n,T}(k) = \sqrt{T}\frac{C_{n,T}(k) - [C_{1,T}(k)]^n}{\hat{\sigma}_{n,T}(k)}, \tag{1.41}$$

where $C_{n,T}(k)$ and $C_{1,T}(k)$ are the sample counterparts to $C_n(k)$ and $C_1(k)$, respectively, and $\hat{\sigma}_{n,T}(k)$ is an estimator of the asymptotic standard deviation of $\{C_{n,T}(k) - [C_{1,T}(k)]^n\}$ [26]. Under H_0, the BDS test statistic $J_{n,T}(k)$ is asymptotically standard normal. Although the BDS test is usually applied to test against the alternative of deterministic chaos, it has also power against nonlinear dependencies. It can be thus applied to the standardized residuals from a GARCH model to test whether these residuals are *iid* and hence the GARCH model captures all nonlinearities in the variance of residuals[27].

1.4.5 Other GARCH-type Models

Since the introduction of the basic ARCH and GARCH models by Engle (1982) and Bollerslev (1986) many alternative specifications based on these models have emerged[28]. Two particular models are interesting from the point of view of the finance theory, the threshold ARCH, denoted TARCH developed independently by Zakoian (1994) and Glosten et al. (1993), and the GARCH-in-mean model due to Engle et al. (1987).

In the standard GARCH model given in equation (1.35), the effects of random shocks, ε_t on the conditional variance h_{t+1} are symmetric, i.e. a positive shock ($\varepsilon_t > 0$) increases expected volatility in the same magnitude as a negative shock ($\varepsilon_t > 0$). Now consider the value of common stock of a highly levered company, i.e. a company with a high debt/equity ratio. An unexpected decrease in the price of the stock results into a decrease in the value of equity and hence, other things unchanged, into an increase in leverage and financial risk associated with holding the stock. As a result, future volatility may increase more after a negative price shock than after a positive one. This asymmetry can be modeled using the threshold ARCH (TARCH) specification for the conditional variance:

$$h_t = \omega + \delta_1 h_{t-1} + \delta_2 h_{t-2} + \cdots \delta_r h_{t-r} + \lambda d_{t-1}\varepsilon_{t-1}^2 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \cdots + \alpha_p\varepsilon_{t-p}^2, \tag{1.42a}$$

where

$$d_{t-1} = \begin{cases} 1 \text{ if } u_{t-1} < 0, \\ 0 \text{ otherwise.} \end{cases}$$

A negative price shock ($\varepsilon_t > 0$) has an impact of $(\alpha_1 + \lambda)$ on h_t, whereas a positive price shock influences h_t though α only. If the estimate of λ is positive and statistically

[26] See Campbell, Lo and MacKinlay (1997, Ch.12) for details on the formulas for $C_{n,T}(k), C_{1,T}(k)$ and $\sigma^2_{n,T}(k)$.

[27] A correction factor is required when applying the BDS test statistic given in (1.41) to standardized residuals from estimated GARCH models, see Brock et al. (1990).

[28] See Engle (1995) for an overview.

significant[29] we say that the leverage effect exists. Whenever $\lambda \neq 0$, positive and negative price shocks influence the conditional volatility asymmetrically. The TARCH model can be estimated by the method of maximum likelihood in the same way as the standard GARCH model.

The concept of risk is closely related to the second application of the GARCH-type models in finance, the GARCH-in-mean model. We have noted above that the models of market equilibrium imply that expected return is a function of risk. Some equilibrium model are explicitly derived under the hypothesis of risk-aversion (such as the Sharpe-Lintner-Mossin CAPM) but other models do not require this strong assumption (APT). To measure the degree of risk aversion of investors empirically one can explicitly model the relationship between expected return and risk by adding the conditional volatility h_t into the regression equation for stock return:

$$r_t = x_t^T \beta + \pi h_t + \varepsilon_t,$$

where h_t is given in (1.35) or (1.42a), x_t is a vector of explanatory variables and $\theta = (\beta, \pi)$ is a vector of unknown parameters. When the estimate of π is positive, expected return is an increasing function of expected volatility and hence investors are risk-averse. On the contrary, $\pi < 0$ implies that investors are risk-seeking. Finally, risk-neutrality is consistent with $\pi = 0$. Although the GARCH-in-mean can be also estimated by ML, to obtain consistent parameter estimates the conditional distribution of the error term ε_t must be correctly specified (Bollerslev et al., 1992). Hence if the true conditional distribution is not normal, QMLE does not yield consistent estimates.

1.5 Data Description

To perform the analysis of predictability of stock returns on the Central-European capital markets we focus on the 365-week period starting in January, 1996 and ending in December, 2002. We use weekly closing prices of the value-weighted indices WIG, BUX and PX-50 of the Warsaw, Budapest and Prague stock exchanges, respectively. These value-weighted indices are widely used in empirical studies on emerging markets and are believed to well describe the evolution of the corresponding equity markets. For an overview of the institutional framework, trading system, number of listed companies and other details on the Central-European stock markets, see Hanousek and Filer (1996) and Gilmore and McManus (2001). For comparison, we also perform the analysis on the DAX index of Deutsche Boerse. The data were downloaded from Bloomberg.

The choice of weekly data is important for two reasons. First, the effect of nonsynchronous trading induces spurious autocorrelation into the index returns, but the lower the data frequency the less important this effect is[30]. Thus to mitigate this effect we choose to use weekly rather than daily data. Second, almost all statistical inferences drawn in the subsequent section are based on asymptotic theory and hence require a sufficient number of observations to be reliable. Using monthly data would therefore not be appropriate due to the short history of the stock markets under study.

[29] The likelihood ration test can be used to test the null hypothesis that $\lambda = 0$.

[30] See Chapter 3 for details.

We define the continuously compounded rate of return in week t on an market index MI as

$$r_t^{MI} \equiv \ln MI_t - \ln MI_{t-1},$$

where $MI_t = \{W_t, B_t, P_t, D_t\}$ and W, B, P and D stand for WIG, BUX, PX-50 and DAX, respectively. The descriptive statistics for the weekly returns on W, B, P and D are summarized in Table 2.1.

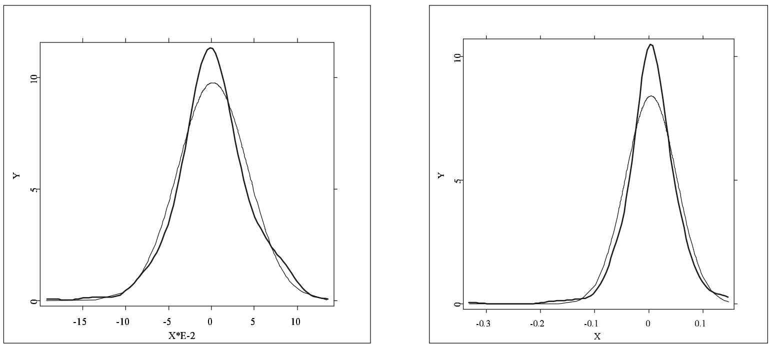
Table 2.1. *Descriptive Statistics*

	WIG	BUX	PX-50	DAX
Mean	0.00176	0.00449	0.00017	0.00063
Variance	0.00167	0.00225	0.00093	0.00124
Maximum	0.13470	0.14736	0.11580	0.12887
Minimum	-0.19244	-0.33016	-0.14045	-0.14079
Skewness	-0.23585	-1.06877	-0.12109	-0.36527
Kurtosis	4.80766	11.67206	4.52814	4.45511
Jarque-Bera	53.0793**	964.6577**	36.4067**	40.3178**

**significant at the 1% level

The Jarque-Bera test[31] indicates significant departures from normality for all indexes. To learn more about the shape of the unconditional density functions we employ a non-parametric kernel density estimator[32] rather than reporting the usual histogram. Using Epanechnikov kernel and setting the bandwidth to the rule-of thumb value proposed by Silverman (1986) we obtain estimates of the unconditional densities given in Figures 2.1a,b.[33].

Figure 2.1a. *Kernel Density Estimates*



WIG BUX

The estimated density functions of weekly returns on the Central-European stock market indexes are similar to those usually obtained for indexes of developed stock markets (e.g. Fama, 1965). The empirical densities are leptokurtic which has important implications for

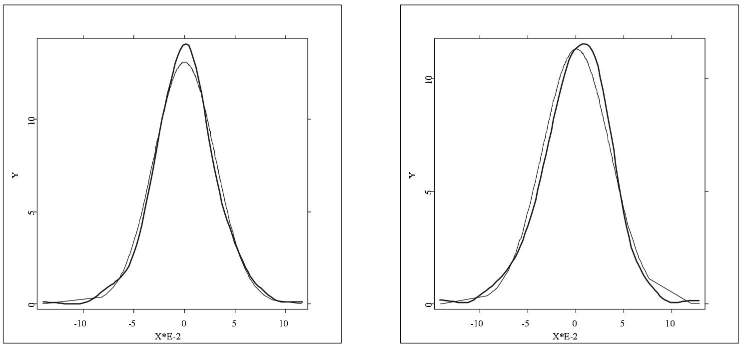
[31]The Jarque-Bera test is a test for skewness and kurtosis. Under the null hypothesis, skewness=0 and kurtosis=3, which are the values of normal distributions. Thus the test is sometimes called a normality test. For details, see for example Greene(2000).

[32]For details on non-parametric density estimation, see for example, Haerdle et al. (1990).

[33]The empirical density functions were estimated using XploRe 4.3.

further statistical tests and estimation.

Figure 2.1b. *Kernel Density Estimates*



PX-50 DAX

As we noted in Section 1.4, the fat tails of the unconditional empirical distributions can be attributed to conditional heteroskedasticity and/or conditional non-normality.

1.6 Empirical Results

This section reports the results of the analysis of stock return predictability performed on the Central-European stock market indexes. We start with the heteroskedasticity-consistent random walk test described in Section 1.2. Next come the results from ARIMA estimation (when applicable) followed by the ARCH-LM test results for conditional heteroskedasticity. Then we present the estimated GARCH-type models and the likelihood ratio tests for the null hypotheses that a) the residuals from the ARIMA models are conditionally Gaussian, and b) the residuals are homoskedastic and unconditionally Student's t-distributed. We also examine the risk aversion of investors using the likelihood ratio test for the presence of the conditional volatility term in the mean equation for returns (ARCH-in-mean) and test for the asymmetric impact of shock to volatility (TARCH). Finally, we asses the ability of the proposed GARCH models to remove all non-linearity in the standardized residuals using the Ljung-Box Q test and the BDS test.

Table 2.2 summarizes the results of the variance ratio test applied to the time series of weekly closing log-values of WIG, BUX, PX-50 and DAX indexes. Following Lo and MacKinlay (1988), we compute the variance ratios $(1 + \hat{M}_r(q))$ and the corresponding robust test statistics for $q = 2, 4, 8, 16$. The null hypothesis of random walk is decisively rejected for the Czech index PX-50. The variance ratio for PX-50 is larger than 1 for $q = 2$ implying positive first-order autocorrelation in weekly returns on this index (see equation (1.9)). The first-order autocorrelation coefficient is approximately equal to 11%. In case of the Hungarian BUX index, the variance ratio for $q = 8$ is statistically different from one at the 5% significance level, thereby rejecting the random walk null hypothesis. For WIG and

DAX indices, the null hypothesis of random walk cannot be rejected.

Table 2.2. *Random Walk Test Results*

		$q = 2$	$q = 4$	$q = 8$	$q = 16$
WIG	$1 + \hat{M}_r(q)$	0.997	1.032	1.106	1.003
	$z^*(q)$	-0.057	0.384	0.931	0.023
	Prob.	0.954	0.701	0.352	0.982
BUX	$1 + \hat{M}_r(q)$	0.964	1.162	1.299	1.202
	$z^*(q)$	-0.532	1.445	2.209	1.117
	Prob.	0.595	0.149	0.027	0.264
PX-50	$1 + \hat{M}_r(q)$	1.108	1.306	1.430	1.266
	$z^*(q)$	1.902	3.722	3.901	1.622
	Prob.	0.057	0.000	0.000	0.105
DAX	$1 + \hat{M}_r(q)$	1.097	1.131	1.158	1.295
	$z^*(q)$	1.392	1.207	1.134	1.578
	Prob.	0.164	0.227	0.257	0.114

Before we turn to the discussion of the results of the random walk test we should briefly comment on the power of the test against various alternatives. In particular, we may be interested in the power of the test against the ARIMA(1,1,0) model. The ARIMA(1,1,1) model of market 'fads' due to Summers (1986) does not seem to be an appropriate alternative. Recall that this model predicts that the first-order autocorrelation coefficient of returns should be negative for all holding periods (see equation (1.17)). But the first-order autocorrelation coefficient for daily returns is positive and statistically significant for all indexes used in our analysis[34], which contradicts the predictions of the Summer's model. To investigate the power of the variance ratio test against the ARIMA(1,1,0) alternative, Lo and MacKinlay (1989) perform a Monte Carlo experiment under the assumptions that the disturbances are *iid* $N(0{,}1)$ and the autoregressive parameter is set equal to 0.2. Based on 20,000 replications, they find that for 256 observations the power of the two-sided 5%-test is equal to 0.887, 0.744, 0.498 and 0.298 for $q = 2, 4, 8$ and 16, respectively. Thus their test has considerable power against the ARIMA(1,1,0) model even for moderate sample sizes. As the sample size decreases, however, the power of the test declines substantially[35].

The results of the random walk test indicate that the Czech and Hungarian stock contain predictable components. These findings are in contradiction to an earlier study due to Hanousek and Filer (1996) who were not able to reject the null hypothesis of random walk for PX-50, BUX and WIG using weekly data for the period starting at the data each index was first calculated and ending in June, 1996. Since our sample and that used by Hanousek and Filer(1996) are almost non-overlapping (they have only 24 common observations) it appears that predictability in the Czech and Hungarian stock returns has increased over the last decade. It may be interesting to divide our sample into two non-overlapping subsamples and run the random walk test independently in the subsamples. Our sample, however, contains only 365 weekly observations and thus the size of the subsamples would be small resulting into the variance ratio test having low power.

[34] We do not report the estimated first-order autocorrelation coefficients for daily returns here for the sake of brewity. The estimates are available from the author upon request.

[35] See Lo and MacKinlay (1989).

Since the random-walk model was rejected for the Czech and Hungarian stocks we will now search for a parametric model that would give us more insight into the nature and degree of predictability. We do this by identifying and estimating an ARIMA(p,1, q) model discussed in Section 1.3 using the Box-Jenkins methodology. For the PX-50 index we find an ARIMA(1,1,1) model to best describe the data generating process and for BUX index, ARIMA(2,1,0) appears to be the best alternative. We do not report the estimation output here since both models will be re-estimated allowing for ARCH effects in the residuals in the subsequent paragraph[36]. Next we run the ARCH-LM test for the null hypothesis of no conditional heteroskedasticity in the residuals from the ARIMA models. For WIG and DAX indexes, we apply the test to the residuals from a regression of returns on a constant only, since for these indexes the null hypothesis of random walk could not be rejected (i.e. the residuals should approximate a white noise sequence). We set the order of the test, $p = 8$. Also we run the Ljung-Box Q-test of order eight on the estimated squared residuals. Table 2.3 summarizes the results.

Table 2.3. ARCH-LM *Test Results*

	WIG	BUX	PX-50	DAX
ARCH-LM Test Stat.	39.669	12.560	43.843	73.030
Prob.	0.0000	0.1279	0.0000	0.0000
Ljung-Box Q-stat.	45.557	13.791	45.511	107.76
Prob.	0.0000	0.0870	0.0000	0.0000

Except for the Hungarian BUX index, both test clearly indicate presence of conditional heteroskedasticity in the estimated residuals. In case of the BUX index the null hypothesis of homoskedasticity is rejected at the 10% significance level using the Ljung-Box Q-test but the ARCH-LM test fails to reject the null at the conventional significance levels. Hence the presence of conditional heteroskedasticity in the residuals from the ARIMA(2,1,0) model for BUX index has to be further investigated.

We now turn to modelling conditional heteroskedasticity in the index returns. In particular, we assume the GARCH(1,1) model for the residuals from the corresponding mean regression equations. It turns out that this model has superior performance in comparison to other GARCH specifications. Thus for the Czech PX-50 index we estimate an ARIMA(1,1,1)-GARCH(1,1) model, for the BUX index an ARIMA(2,1,0)-GARCH(1,1) model and for WIG and DAX indexes we assume a random walk model whose increments obey a GARCH(1,1) process. Recall that an ARIMA(p,1,q)-GARCH(1,1) model is given by

$$\begin{aligned} r_t &= \mu + \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots \theta_q \varepsilon_{t-q}, \\ \varepsilon_t &= u_t \sqrt{h_t}, \\ h_t &= \omega + \delta_1 h_{t-1} + \alpha_1 \varepsilon_t^2. \end{aligned}$$

We estimate these models by the method of maximum likelihood under the assumption that the residuals are conditionally Student-t distributed with v degrees of freedom[37]. The

[36] The estimation outputs are available from the author upon request.

[37] The estimation is carried out using RATS 5.02. The BHHH algorithm is employed to maximize the likelihood function.

estimated coefficients are asymptotically normally distributed. Table 2.4a reports the estimated random walk model with conditionally heteroskedastic increments for the Polish WIG index.

Table 2.4a. GARCH(1,1)-t *Model for* WIG

	Coefficient	Std. Error	t-stat.	Probability
μ	$7.66{\cdot}10^{-4}$	$1.90{\cdot}10^{-3}$	0.403	0.687
ω	$1.04{\cdot}10^{-4}$	$0.92{\cdot}10^{-5}$	1.125	0.260
δ_1	0.854	0.089	9.631	0.000
α_1	0.082	0.045	1.823	0.068
v	6.521	2.627	n.m.	n.m.

The estimated GARCH coefficients δ_1 and α_1 are significant at the 1% and 10% level, respectively and their sum is less then one implying that the GARCH model is stationary, thought the volatility is fairly persistent since $(\delta_1+\alpha_1)$ is close to one. The estimated number of degrees of freedom of the conditional t-distribution is equal to 6.5 which suggests that the returns on the WIG index are conditionally non-normally distributed.

Table 2.4b. GARCH(1,1)-t *Model for* BUX

	Coefficient	Std. Error	t-stat.	Probability
μ	0.005	0.002	2.453	0.014
ϕ_2	0.133	0.049	2.684	0.007
ω	$1.34{\cdot}10^{-4}$	$8.18{\cdot}10^{-5}$	1.643	0.100
δ_1	0.765	0.108	7.101	0.000
α_1	0.060	0.032	1.863	0.062
v	4.218	0.964	n.m.	n.m.

$R^2 = 0.034$

To summarize, both heteroskedasticity and non-normality account for the fat tails in the empirical distribution of the returns on WIG.

Recall that both the Ljung-Box Q test for serial correlation in the squared residuals and the ARCH-LM test failed to reject the null hypothesis of no ARCH effects in the residuals from the ARIMA(2,1,0) model for the Hungarian BUX index. From the estimation output, reported in Table 2.4b we see, however, that the GARCH coefficients δ_1 and α_1 are statistically significant at the 1% and 10% level, respectively. The estimated number of degrees of freedom is equal to 4.2 which is a very small value given the fact that for $v \leq 4$ the kurtosis of a Student-t random variable with v degrees of freedom is infinite. But the low value of $\hat{v}$ is not surprising since the empirical kurtosis of the BUX returns is quite high, equal to 11.67. Thus, as with the WIG index, the leptokurtosis in the empirical distribution function of the returns on the BUX index is induced by both heteroskedasticity and non-normality.

Table 2.4c. GARCH(1,1)-t *Model for* PX-50

	Coefficient	Std. Error	t-stat.	Probability
μ	$1.27 \cdot 10^{-4}$	$7.23 \cdot 10^{-4}$	0.176	0.860
ϕ_1	0.633	0.224	2.823	0.005
θ_1	0.508	0.254	1.996	0.046
ω	$1.95 \cdot 10^{-5}$	$2.19 \cdot 10^{-5}$	0.887	0.375
δ_1	0.928	0.045	20.36	0.000
α_1	0.055	0.030	1.812	0.069
v	7.238	3.475	n.m.	n.m.

$R^2 = 0.021$

In Table 2.4c we summarize the estimated ARIMA(1,1,1)-GARCH(1,1) model for the Czech index, PX-50. We obtain very similar results to those for WIG and BUX in that the GARCH parameters δ_1 and α_1 are statistically significant at the 1% and 10% level, respectively, and $\hat{v}$ is much less than ten. Note that although the GARCH(1,1) model is stationary, $\delta_1 + \alpha_1 = 0.983$, i.e. the volatility of the returns on PX-50 is very persistent.

Table 2.4d. GARCH(1,1)-t *Model for* DAX

	Coefficient	Std. Error	t-stat.	Probability
μ	0.004	0.002	2.319	0.020
ω	$3.14 \cdot 10^{-5}$	$2.58 \cdot 10^{-5}$	1.219	0.223
δ_1	0.807	0.053	15.30	0.000
α_1	0.186	0.053	3.531	0.001
v	18.62	12.74	n.m.	n.m.

Finally, we examine the German index DAX (Table 2.4d.). Surprisingly, the estimated number of degrees of freedom of the conditional Student-t distribution for returns is quite large compared to the those for the market indices of the Central-European stock markets. It falls into the inconclusive region defined in Connolly(1992) and hence it is difficult to comment on the distributional properties of returns without further testing. Intuitively, since the GARCH parameters are highly significant and the volatility of returns is very persistent ($\delta_1 + \alpha_1 = 0.993$) it may be the case that the excess kurtosis in the DAX returns is induced primarily by conditional heteroskedasticity[38]. To test this hypothesis rigorously, we employ the likelihood ratio tests discussed in Section 1.4.2.

Table 2.5 presents the likelihood ratios for a) the null hypothesis that the returns on the corresponding market indexes are conditionally normally distributed ($1/v = 0$), and b) the null hypothesis of unconditionally Student-t distributed returns with v degrees of freedom

[38]A GARCH(1,1) model with the restriction that $\alpha_1 + \delta_1$ is called an *integrated* GARCH(1,1) model, denoted IGARCH(1,1) (see Bollerslev and Engle, 1986). The volatility is not covariance-stationary in this model, i.e. a shock to volatility has a permanent effect on the level of volatility. Moreover, the unconditional variance for an IGARCH model does not exist. Since $\alpha_1 + \delta_1$ is close to one for PX-50 and DAX, we might be interested in testing the restriction that $\alpha_1 + \delta_1 = 1$. But since the asymptotic properties of an IGARCH model in moderate samples are not yet well-understood (see Bollerslev et al., 1992), we leave this test to further research.

($\alpha_1 = \delta_1 = 0$). The likelihood ratios are asymptotically χ^2-distributed with a) one, and b) two degrees of freedom. The 5%-critical values are thus 3.84 and 5.99, respectively. Note, however, that according to Bollerslev (1987) the distribution of the test statistic $LR_{1/v=0}$ in samples of moderate sizes tends to be more concentrated towards the origin than the χ^2_1 distribution and hence a test based on the 3.84 critical value has lower size (is more conservative). Bollerslev(1987) provides the appropriate 5%-critical value, 2.71.

Table 2.5. LR *Test Results*

	$\hat{v}$	$LR_{1/v=0}$	$LR_{\alpha_1=\delta_1=0}$
WIG	6.521	11.50*	16.04*
BUX	4.218	52.97*	16.52*
PX-50	7.238	7.82*	13.76*
DAX	18.62	2.35	61.37*

*significant at the 1%-level

The null hypothesis of unconditionally Student-t distributed is decisively rejected for all four market indices. These findings are in line with our estimates of the GARCH parameters: they all were statistically significant. Hence, conditional heteroskedasticity is present in stock returns across all markets. The null hypothesis of conditional normality is rejected for the Central-European stock market indices. The likelihood ratio test fails to reject the null for the German DAX. Again, these results are in accordance with our estimates of v: the smaller the $\hat{v}$, the larger the likelihood ratio for the null hypothesis of conditional normality. To summarize, on the basis of the GARCH-t model and the likelihood ratios we have found that in case of the returns on the Central-European stock market indices, both conditional heteroskedasticity and non-normality account for excess kurtosis in the empirical distribution functions of index returns. For the German index DAX, the picture is less clear cut, though it appears that the primary source of the 'fat tails' in the empirical distribution is time-varying volatility.

Table 2.6. LR *Test Results*

	$LR_{\lambda=0}$	$LR_{\pi=0}$
WIG	1.30	6.01*
BUX	n.m.	2.21
PX-50	3.10	0.07
DAX	n.m.	0.01

*significant at the 5%-level

In Section 1.4.4. we introduced two extensions of a simple GARCH(1,1) model specification. First, we considered an asymmetric impact of positive and negative shocks to returns on their expected volatility (TARCH). Second, we included the expected volatility term, h_t, into the mean equation for returns (ARCH-in-mean). Both alternative specifications can be tested using the usual likelihood ratio. Table 2.6 reports the likelihood ratios for a) the null hypothesis that the shocks to returns have a symmetric effect on future volatility ($\lambda = 0$), and b) the null hypothesis of risk-neutrality of investors ($\pi = 0$). Under both null hypotheses the likelihood ratios are asymptotically χ^2distributed with one degree of freedom, hence the 5%-critical value is equal to 3.84. Note that we do not report the values of $LR_{\lambda=0}$ for BUX and DAX indices because the estimates of the constant term ω in the TARCH variance equation are negative which violates the assumption that all GARCH coefficients be positive (a negative ω may cause h_t to be negative for some t which is not consistent with the non-negativity property of variance). For WIG and PX-50 the null hypothesis that the shocks to returns have symmetric impact on volatility cannot be rejected on the conventional significance levels. Turning to the null hypothesis of risk-neutrality, the likelihood ratio test fails to reject the null for BUX, PX-50 and DAX, but does reject it at the 1% significance level for the Polish WIG index.

Table 2.7. GARCH-M-t *Model for* WIG

	Coefficient	Std. Error	t-stat.	Probability
μ	-0.030	0.0144	-2.077	0.038
π	0.817	0.377	2.167	0.030
ω	$8.53\cdot10^{-5}$	$6.42\cdot10^{-5}$	1.328	0.184
δ_1	0.875	0.060	14.49	0.000
α_1	0.071	0.033	2.136	0.033
v	6.406	2.457	n.m.	n.m.

Since the null hypothesis of risk neutrality was rejected in case of the WIG index, we now estimate the ARCH-in-mean model assuming conditional Student-t distribution of the error terms. Table 2.7 summarizes the estimation output. The estimate of π, which measures the degree of risk aversion, is positive and significant at the 5% level. It implies that the investors trading on the Polish stock market are risk-averse: their require higher returns in periods when they expect higher volatility (risk)[39]. The fact that expected volatility, h_t, is significant in the mean equation has also important implications for return predictability. To see this, recall that h_t is a function of past squared residuals from the mean equation. The expected return on the WIG index for period $t+1$ based on the information available in period t, can be thus written as

$$\begin{aligned} E\left[r_{t+1} \mid \Phi_t\right] &= \mu + \pi h_{t+1}, & (1.43)\\ &= \mu + \pi\left(\omega + \delta_1 h_t + \alpha_1 \varepsilon_t^2\right), & (1.44)\\ &= \mu + \pi\left(\zeta + \psi(L)\varepsilon_t^2\right), & (1.45)\end{aligned}$$

where $\psi(L)$ is an infinite-order polynomial in the lag operator[40]. Now because both h_t and ε_t^2 are known at time t the return for period $t+1$ is predictable from the time series of historical volatility and past squared residuals. The conclusion that WIG's returns are predictable is in contradiction to the result of the robust variance ratio test for the null

[39]There has been evidence in the literature on the ARCH-in-mean models of the sensitivity of the parameter estimates to the distributional assumptions made to obtain the estimates by MLE (see Bollerslev et al., 1992, for an overview). It has been reported, for example, that by changing the conditional distribution from normal to Student-t, the estimated π dropped substantially and was no-longer significant. When we use normal distribution to estimate the ARCH-in-mean model for the WIG index, we obtain $\hat{\pi} = 7.73$ with (Bollerslev and Woodridge, 1991) robust p-value equal to 0.087. Thus the sensitivity observed on developed markets carries over to the Polish stock market.

[40]We assume that $\delta_1 < 1$ and hence the GARCH(1,1) process can be expressed as an infinite order ARCH.

hypothesis of random walk. But this should not be surprising. From equation (1.9) we see that the variance ratio test, by construction, has power only against *linear* dependencies in the time series of returns. Equation (1.43), on the other hand, implies that in the ARCH-in-mean model there is a correlation between current return and lagged *squared* disturbances. It follows that the variance ratio test cannot detect such non-linearities and hence fails to reject the null hypothesis of random walk[41]. The ARCH-in-mean model therefore provides us with an important insight into the time-series predictability of stock returns.

Table 2.8. *Diagnostic Tests*

	WIG	BUX	PX-50	DAX
$Q(4)$	2.27	2.46	3.92	6.39
$Q(8)$	3.95	5.11	6.99	11.91
$Q(12)$	12.53	7.72	9.47	19.05
$Q^2(4)$	2.39	2.54	2.07	1.60
$Q^2(8)$	11.73	4.26	16.57*	4.67
$Q^2(12)$	15.23	5.62	20.64*	4.93
BDS(2)	-0.97	-0.51	0.35	0.58
BDS(4)	0.45	-0.08	-1.14	0.09
BDS(6)	1.08	0.83	-1.97*	0.57

*significant at the 5%-level

Finally, having estimated the corresponding ARIMA-GARCH models for the studied stock market indices, we should perform diagnostic testing to asses whether the models are correctly specified. We start by testing the standardized residuals for serial correlation. This may seem superfluous, for the ARIMA models were specified in a way removing possibly all autocorrelation in the residuals, but because we did not report the estimates and diagnostic checks before for the sake of brevity, we should do it now. The first three rows of Table 2.8 therefore report the Ljung-Box Q statistics applied to the standardize residuals from the models of Tables 2.4b,c,d and 2.7. Clearly, the null hypothesis of no serial correlation cannot be rejected on the conventional significance levels and hence the mean equations appear to be correctly specified. Next we run the Ljung-Box Q test on the squared standardized residuals (we label the test statistic Q^2 to emphasize that the test is run on squared standardized residuals). The null hypothesis of no serial correlation in squared standardized residuals can be rejected only for the PX-50 index. Hence the GARCH(1,1) model in this case fails to appropriately fit the volatility process. No reasonable alternative specifications, however, have been found to perform better[42]. Finally, we apply the BDS test on the standardized residuals to determine, whether the GARCH model removes all non-linearities in the time-series of returns. We set the proximity parameter to 0.5 and run

[41] To detect non-linearities in stock returns, we could also apply the BDS test to the returns. But since we have seen that the stock returns are heteroskedastic the result of the BDS test would be necessarily ambiguous: we would not know whether we should attribute the rejection of the *iid* null hypothesis to non-linearities in the mean or non-linearities in the variance.

[42] The correlogram of the squared standardized residuals reveals that there is a significant spike in both the autocorrelation and partial autocorrelation functions at lag 7. This indicates that a GARCH(1,8) may better fit the data. Some of the estimated coefficients in this model are, however, negative, which is rulled out by assumption. The estimates are available from the author upon request.

the test for embedding dimensions 2,4, and 6. The null hypothesis that the standardized residuals are *iid* can be rejected at 5% for PX-50 index only, which is consistent with the findings of the Ljung-Box Q test. To summarize, the diagnostic checks performed on the standardized residuals provide convincing evidence that, except for the PX-50 index, the ARIMA-GARCH models were appropriately specified and can thus be used for rational predictions of both index returns and their volatility.

1.7 Concluding Remarks

The analysis of stock return predictability performed in the previous section was aimed at detecting regular patterns in stock prices. We investigated time-series predictability of three Central-European stock markets - Czech, Hungarian and Polish, using weekly data on respective major value-weighted market indices. For comparison, we also considered the predictability of the German index DAX. We extended the previous empirical work by explicitly modeling the 'fat tails' in the unconditional distributions of stock returns using a GARCH-type model with Student-t distributed disturbances. To briefly summarize our analysis, we found that:

1. The Central-European stock prices do not follow random walk,
2. Stock returns are predictable from the time-series of historical returns using a linear (ARMA) or non-linear (ARCH-in-mean) parametric model,
3. Stock return volatility changes over time and can be predicted using a GARCH-type model,
4. Empirical density functions of stock returns exhibit excess kurtosis which can be attributed to both conditional heteroskedasticity and non-normality.

We hasten to emphasize that the predictability of stock returns need not be a symptom of market inefficiency (recall equation (1.3)). Even in an efficient market stock returns may be correlated simply because the expected return systematically changes over time. But our analysis did not study the way expected return is determined at all and hence any conclusion regarding the EMH is impossible.

Moreover, the fact that stock prices contain predictable components does not necessarily imply that the predictability is economically significant. We have mentioned earlier that some of the empirically observed predictability may be indeed spurious, induced by non-synchronous trading and/or the bid-ask spread. We will examine this possibility in greater detail in Chapter 3. But even if the predictability in the Central-European stocks is real, it is quite unclear whether is can be exploited to earn superior returns. The ARIMA models for PX-50 and BUX are poorly determined (the R^2 is equal to 0.021 and 0.034, respectively) and we have not developed and followed any out-of-sample trading strategy based on our estimated models that could be compared to a passive trading strategy to see whether our active trading approach provides higher return. Nor have we taken into account transaction costs that may well wipe out any incremental return yielded by the active trading strategy. These questions are definitely worth further research and we will try to address them in

Chapter 4 where we develop a dynamic trading strategy based on a maximally predictable portfolio.

Our empirical findings have also important implications for asset pricing. Consider, for simplicity, the well-known Black-Scholes model for pricing European-style call options on a non-dividend paying stock[43]. This model is derived under three crucial assumptions: a) log-stock prices follow random walk, b) returns are normally distributed, and c) volatility is constant through time. But non of these assumptions are satisfied empirically, so the Black-Scholes cannot be used for rational pricing of options on the Central-European stock market indices.

The reader may now wonder why we have undertaken the lengthy endeavour of studying time-series predictability when we now conclude that the results are far from being clear-cut. We hope to have provided a natural starting point to further investigation into the sources and patterns of stock return predictability. Only when we learn more about the time-series behavior of stock returns from a purely econometric point of view can we think of exploiting the forecastability of returns by means of a particular trading strategy. Knowing the empirical autocorrelation function can guide our selection of a candidate theory for studying the extent of spurious autocorrelation in stock returns. And as we just discussed, the time-series and distributional properties can help us identify an appropriate asset pricing theory (and, of course, reject a misspecified one).

[43]Black and Scholes (1973).

Chapter 2

Cross-Country Predictability and Cointegration

In the previous chapter, univariate time-series techniques were applied to study the predictability of selected Central-European stock market index returns. The individual time series were treated as if they evolved independently from one another. The advent of new technology and the ongoing market integration around the world, however, imply that national stock markets may indeed exhibit some degree of co-movement and in that case it is necessary to study the predictability of stock returns in a multivariate context. The concept of cointegration will be invoked in this chapter to investigate whether the Czech, Hungarian, Polish and German stock markets are linked through common stochastic trend(s). We will apply the multivariate cointegration analysis developed in Johansen (1988, 1991) to stock index prices expressed in (i) local currencies, and (ii) in Euros. We consider both alternatives for the following reason. Internationally trading investors may or need not hedge their positions in foreign stocks against exchange rate movements. Those investors who immunize their portfolios against foreign exchange risk are primarily concerned about stock returns in local currencies. On the contrary, if an investor does not hedge her portfolio against exchange rate movements, the rate of return on her portfolio is composed of both the return in local currencies and exchange rate appreciation/depreciation. To satisfy both classes of investors we perform the cointegration analysis on stock index prices expressed in local currencies as well as in Euros.

The chapter is organized as follows. In Section 2.1, we review informally the theoretical arguments for stock market co-movements commonly used in the empirical literature. Section 2.2 is dedicated to unit root testing methodology. In Section 2.3 we present the definition of cointegration and the methodology of the Johansen cointegration test. We also show that a vector autoregression of cointegrated variables can be alternatively expressed in an error-correction form which is particularly useful for studying the long-term and short-term relationships among cointegrated variables. Section 2.4 contains a brief data description. The emprical results of cointegration tests and error-correction model estimation are reported in Section 2.5. We conclude with a short discussion in Section 2.6.

2.1 Stock Market Co-movements

There are various hypotheses explaining national stock market integration. First, the ongoing economic globalization and market liberalization increase economic interdependence among countries. As a result, the macroeconomic fundamentals of national economies become more interrelated and in some groups of countries even converge. And since stock market returns are generally believed to be primarily driven by economic fundamentals, it follows that countries with co-moving macroeconomic variables will also tend to have integrated stock markets. A prominent example is provided in the seminal paper by Yang et al. (2003). They investigate stock market integration within the European Monetary Union[1]. Yang et al. (2003) argue that the introduction of a new common currency and a single monetary policy among eleven EU member countries should result into increased integration of their respective stock markets. The empirical results support this hypothesis. Yang et al. (2003) note, however, that the increased degree of integration can also be attributed to "faster information transmission and processing due to technological advances, recent consolidation and merger of stock exchanges in Europe" (Yang et al., 2003, p.17).

Forbes and Rigobon (1999) provide a similar explanation for stock market co-movements. They distinguish among three different types of shocks affecting economic fundamentals and thereby stock returns: (i) aggregate shocks affecting the fundamentals of more than one country, (ii) country-specific shocks affecting the economic fundamentals of other countries, and (iii) shocks which are not explained by fundamentals and are thus considered as pure contagion. The reader can easily come up with examples for the three types of shocks so we will not discuss them in detail here[2]. The conclusion of Forbes and Rigobon (1999) is that there exist strong stock market linkages due to shocks other than contagion.

The fact that stock markets are cointegrated has important implications for joint market efficiency. According to Granger (1986), if two stock markets are jointly efficient they cannot be cointegrated. If two markets were indeed cointegrated then the prices from one market could be used to forecast the prices from the other, which violates the market efficiency hypothesis. This fact will become apparent later as we define cointegration formally. It is important to note, however, that market co-movements in general do not necessarily imply inefficiency. Even if two markets are not cointegrated, their returns can be contemporaneously correlated and hence the markets "move together" and are jointly efficient.

Finally, cointegration of national stock markets implies that it is not possible to diversify internationally[3]. This follows immediately from the fact that cointegrated markets exhibit co-movements and hence a steady decline in one market is followed by a steady decline in other markets. The systematic risk cannot be therefore diversified away and it is not in interest of investors who seek diversified portfolios to invest in cointegrated stock markets.

Although the literature on cointegration among developed stock markets is fairly voluminous[4], this is not the case for the Central-European emerging markets. In a pioneering

[1] Yang et. al. (2003) also provide an extensive overview of the empirical literature on cointegration among European stock markets.

[2] See Forbes and Rigobon (1999, p.6-7).

[3] See, for example, Phylaktis and Ravazzolo (2000).

[4] See, for example, Chan et al. (1997) and the references therein.

work, Neubauer (2001) tests for cointegration among the Czech, Hungarian, Polish and German equity markets in the period 1994:1 to 1998:2 with weekly index prices expressed in terms of Deutschemark. When using the PX-50, BUX, WIG and DAX indices, respectively, he finds evidence of a significant cointegrating relationship among the markets under study. Neubauer (2001) also investigates the behavior of the cointegrating relationship over time by running the test in two consecutive non-overlapping subsamples and concludes that cointegration found in 1993:1-1996:12 disappears in 1997:1-1998:2. In our analysis, we extend and update the work by Neubauer (2001) by considering more recent data set (1999:1-2002:12) and perform the analysis on stock index prices expressed in local currencies as well as in terms of Euro.

2.2 Unit Root Test

This section provides a brief overview of the methodology of the Augmented Dickey-Fuller unit root test developed by Dickey and Fuller (1979). Let y_t denote a log-stock price at time t. Suppose that the log-stock price process $\{y_t\}$ can be modeled as an autoregressive process of order p, i.e. $y_t \sim$AR(p):

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \tag{2.1}$$

where $\{\varepsilon_t\}$ is a white-noise sequence. The process in (2.1) is stationary provided that all roots of the polynomial

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^P = 0 \tag{2.2}$$

lie outside the unit circle. If one of the roots of (2.2) is unity and the other roots are outside the unit circle, the process that generated y_t is said to contain a unit root. A process y_t that has a single unit root is non-stationary but stationarity can be achieved by first-differencing the series y_t. For the purpose of testing for a unit root in a time-series it is more convenient to rewrite equation (2.1) as follows[5]:

$$y_t = \mu + \delta t + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots \zeta_{p-1} \Delta y_{t-p+1} + \varepsilon_t, \tag{2.3}$$

where $\rho = \sum_{i=1}^{p} \phi_i$ and $\zeta_j = -\sum_{i=j+1}^{p} \phi_i, j = 1, 2, \ldots, p-1$. Under the null hypothesis, y_t obeys a unit root process with drift, i.e. $\delta = 0$ and $\rho = 1$ in equation (2.3). Under the alternative hypothesis, y_t is trend-stationary and $\delta \neq 0$ and $\rho < 1$. If the time-series under study does not appear to be trending but does appear to have a non-zero mean, the test equation (2.3) is estimated without the deterministic trend term δt. The null hypothesis is then that y_t follows a unit root process without drift, i.e. $\mu = \delta = 0$ and $\rho = 1$, and under the alternative hypothesis, y_t is a stationary process, i.e. $\mu \neq 0$, $\delta = 0$ and $\rho < 1$. In both cases, equation (2.3) is estimated by OLS and the null hypothesis of a unit root ($\rho = 1$) is tested by referring the t-statistics corresponding to the coefficient ρ to the appropriate critical values tabulated by Dickey and Fuller (1979)[6].

---

[5] See Hamilton (1994, p. 517) for derivation.

[6] Under the null hypothesis of a unit root, the test statistic does not have a standard distribution. Dickey and Fuller (1979) therefore tabulate critical values calculated by Monte Carlo simulation. The crutial

There are several methods which can be used to determine the optimal number of lags in the test equation (2.3). One approach is to assume some maximum number of lags, $p^{\max}$, estimate equation (2.3) with $p^{\max}-1$ lags and test the null hypothesis that $\zeta_{p^{\max}-1} = 0$ using the usual t-test. If the null hypothesis cannot be rejected then the regression is reestimated with $p^{\max} - 2$ lags and the process is repeated until the null hypothesis $\zeta_{p^{\max}-i} = 0$ cannot be rejected for some i. The optimal number of lags used for the Augmented Dickey-Fuller test is then set equal to $p^{\max} - i$. Alternatively, one can use some information criteria to determine the number of lags in (2.3). The two widely used are the Akaike (AIC) and Schwartz (SIC) information criteria. As a diagnostic check for the appropriateness of the selected number of lags, the Ljung-Box Q-test[7] for the presence of autocorrelation in residuals should be used: if the regression is correctly specified, the should be no significant autocorrelation among the residuals.

2.3 Cointegration and the Error-Correction Model

In this section we define cointegration and describe the methodology of the Johansen (1988,1991) maximum likelihood analysis of cointegrated systems.

Definition 1 *Let $\mathbf{y}_t$ denote an $(n \times 1)$ vector time series. The vector time series $\mathbf{y}_t$ is said to be cointegrated if each of the series is individually integrated of order one, i.e. $I(1)$, and there exists some nonzero $(n \times 1)$ vector $\boldsymbol{\gamma}$ such that the linear combination of the series $\boldsymbol{\gamma}^T\mathbf{y}_t$ is stationary. A vector satisfying this condition is called a cointegrating vector.*

The cointegrating vector is usually perceived as defining an equilibrium relationship among the individual nonstationary variables comprising $\mathbf{y}_t$. In particular, the equilibrium is given by $\boldsymbol{\gamma}^T\mathbf{y}_t = 0$. The intuition behind this interpretation is that if $\boldsymbol{\gamma}$ is a cointegrating vector, the linear combination $\boldsymbol{\gamma}^T\mathbf{y}_t$ is stationary and thus any deviation from the equilibrium relationship at time t, defined by $\boldsymbol{\gamma}^T\mathbf{y}_t = 0$, is transitory. Hence if the system is shocked in period t, the shock dies out over time and the system returns back to its equilibrium. It has to be noted, however, that the term "equilibrium" used in the context of cointegration does not necessarily imply anything about the *economic* equilibrium. Rather it describes the tendency of a system of economic variables to move toward particular outcomes (Granger, 1986).

The interpretation is, however, less clear-cut in cases when there are more than one cointegrating vectors. It can be shown that in a system with n variables there can be up to $n-1$ linearly independent cointegrating vectors (up to a normalizing constant)[8]. The number of cointegrating vectors is called the cointegrating rank of the vector time series $\mathbf{y}_t$. Another important insight, due to Stock and Watson (1988), is that in a system with

assumption made in this simulation is that the innovations are homoskedastic and normally distributed. We have seen in the previous chapter that this assumption is not valid for stock returns. When normality of homoskedasticity assumptions are questionable, Davidson and McKinnon (1999) suggest using asymptotic critical values for the ADF test, which are independent of the these assumptions (provided, of course, that the sample is reasonably large).

[7]Ljung and Box (1978).

[8]For an informal but intuitive proof of this result, see Greene (2000, pp. 791).

n variables and h cointegrating vectors the variables y_{it} comprising the vector $\mathbf{y}_t$ share exactly $g = n - h$ common deterministic time trends and g common random walk variables. This property and the fact that the equilibrium errors $\boldsymbol{\gamma}^T\mathbf{y}_t$ are transitory induce the co-movement of the y_{it} ´s.

To test for the number of significant cointegrating vectors of $\mathbf{y}_t$, Johansen (1988, 1991) develops methodology based on full-information maximum likelihood estimation. Following Johansen (1988,1991) we assume throughout this Section that $\mathbf{y}_t$ can be modelled as a $VAR(p)$ process, i.e.

$$\mathbf{y}_t = \mathbf{A}_0 + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t. \tag{2.4}$$

As shown in Hamilton (1994, pp. 549), any $VAR(p)$ process in the form (2.4) can be alternatively represented as follows:

$$\Delta\mathbf{y}_t = \boldsymbol{\zeta}_1\Delta\mathbf{y}_{t-1} + \boldsymbol{\zeta}_2\Delta\mathbf{y}_{t-2} + \cdots + \boldsymbol{\zeta}_{p-1}\Delta\mathbf{y}_{t-p+1} + \mathbf{A}_0 + \boldsymbol{\zeta}_0\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{2.5}$$

where $\boldsymbol{\zeta}_0 \equiv \sum_{i=1}^{p}\mathbf{A}_i - \mathbf{I}_n$ and $\boldsymbol{\zeta}_s = -\sum_{i=s+1}^{p}\mathbf{A}_i$ for $s = 1, 2, \ldots, p-1$. Expression (2.5) is the *error-correction* form of a cointegrated system. If there are exactly h cointegrating vectors, then the matrix $\boldsymbol{\zeta}_0$ can be written as a product of two $(n \times h)$ matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$:

$$\boldsymbol{\zeta}_0 = -\boldsymbol{\alpha}\boldsymbol{\gamma}^T. \tag{2.6}$$

The matrix $\boldsymbol{\gamma}$ is composed of the h cointegrating vectors, and the matrix $\boldsymbol{\alpha}$ is the matrix of weights with which each cointegrating vector enters the n equations. The decomposition of $\boldsymbol{\zeta}_0$ in (2.6) illuminates, why (2.5) is called the *error-correction* model for a vector time series $\mathbf{y}_t$: an equilibrium *error* in period $t-1$, given by $\boldsymbol{\gamma}^T\mathbf{y}_{t-1}$, tends to be *corrected* through the parameters of the matrix $\boldsymbol{\alpha}$ in period t.

Since the $(n \times n)$ matrix $\boldsymbol{\zeta}_0$ is a product of two $(n \times h)$ matrices, it has rank h, which is the number of cointegrating vectors. Johansen (1988,1991) exploits this property to build a test for the number of cointegrating vectors: he first estimates equation (2.6) and then examines the rank of $\hat{\boldsymbol{\zeta}}_0$ by testing for the number of characteristic roots of $\hat{\boldsymbol{\zeta}}_0$ that insignificantly differ from zero. In a system of n variables with exactly h cointegrating vectors, the matrix $\boldsymbol{\zeta}_0$ has $n-h$ characteristic roots equal to zero. Hence the test statistic for the null hypothesis that there are h or less distinct cointegrating vectors against a general alternative is given by:

$$\lambda_{\text{trace}}(h) = -T\sum_{i=h+1}^{n}\log(1-\hat{\lambda}_i) \tag{2.7}$$

where the $\hat{\lambda}_i$ ́s are the characteristic roots of $\hat{\boldsymbol{\zeta}}_0$ ordered such that $\lambda_1 > \lambda_2 > \cdots > \lambda_n$. To test the null hypothesis of h cointegrating vectors against the alternative of $h+1$ cointegrating vectors, we use the test statistic

$$\lambda_{\max}(h, h+1) = -T\log(1-\hat{\lambda}_{h+1}). \tag{2.8}$$

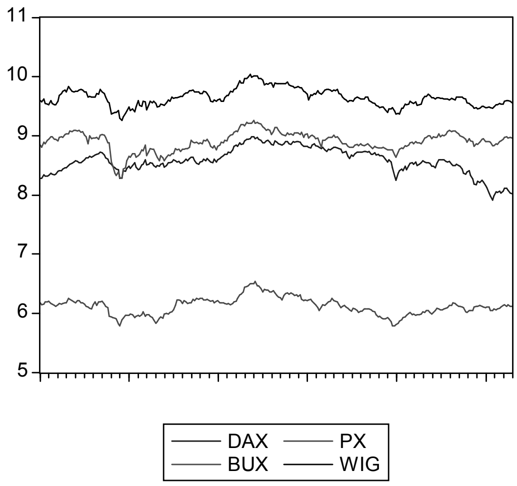
The likelihood ratio test statistics (2.7) and (2.8) do not have the usual χ^2 distribution. The appropriate critical values for (2.7) and (2.8), obtained by Monte Carlo simulations, are provided by Johansen and Juselius (1990).

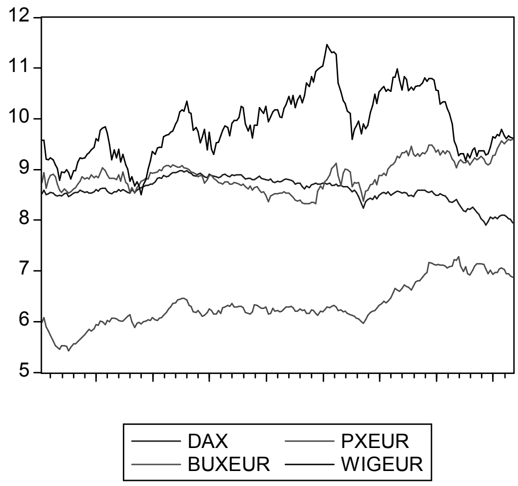
Cointegration has important implications for investigating stock markets co-movements. In some studies, the stock market interlinkages and short-term dynamics are analyzed using a VAR model for stock returns (i.e. first differences of log-stock prices)[9]. Another approach is to test for Granger causality between two markets. But from expression (2.5) follows, that none of these approaches is valid if the stock markets under study are cointegrated: if a cointegrated system is modelled as a VAR in differences, we have a misspecification error since the error-correction term is not included as a regressor in (2.5). And the same holds for the usual Granger-causality test based on the equation[10]

$$\Delta y_t = \mu + \alpha_1 \Delta y_{t-1} + \alpha_2 \Delta y_{t-2} + \cdots + \beta_1 \Delta x_{t-1} + \beta_2 \Delta x_{t-2} + \cdots + \varepsilon_t. \tag{2.9}$$

Hence, it is of crucial importance to test for cointegration among the variables under study before we select a method for describing the dynamics of the system. When the null hypothesis of no cointegration cannot be rejected, VAR in first-differences or Granger causality analysis are valid tools. If, on the contrary, there is at least one cointegrating vector, the error-correction representation in equation (2.5) is the only appropriate model for studying the short-term and long-term dynamics of the system.

Figure 2.1. *Log-Index Prices over Time*





2.4 Data Description

To perform the cointegration analysis described above we focus on the four-year period from January, 1999 through December, 2002. We use weekly closing nominal prices of the following indices: DAX (Germany), BUX (Hungary), PX-50 (Czech Rep.) and WIG (Poland). These value-weighted indices are widely used in empirical studies on emerging markets and are believed to well describe the evolution of the corresponding equity markets.

[9] See, for example, Koch and Koch (1993).

[10] The variable Δx_t is said to Granger-cause Δy_t, if lagged values of Δx_t help to forecast Δy_t . The Granger causality test is an F-test for the joint significance of the $\Delta x_{t-1}, \ldots, \Delta x_{t-p}$ terms in regression (2.9).

The prices are expressed (i) in local currencies, and (ii) in Euros and were converted to natural logarithms for the purpose of this study. The time-series of exchange rates were standardized to equal unity in the beginning of the sample. The data were downloaded from Bloomberg. Figure 2.1 depicts the evolution of the indices over the four-year period under study. The left-hand side graph shows the time-series expressed in local currencies, whereas in the right-hand side graph, the indices are expressed in Euros. It is hard to comment on the stationarity of the indices by simply visually inspecting the data. But it appears at the first glance that the time-series expressed in local currencies do, to some extent, move together, whereas those expressed in Euros seem to wander arbitrarily far apart from one another.

2.5 Empirical Results

This section reports the results of the cointegration analysis described above[11]. We start by testing the individual time-series of stock index prices for the presence of a unit root. We run the Augmented Dickey-Fuller test on levels of the series using test equation (2.3) with and without trend and then on first differences using test equation (2.3) without trend. The choice of optimal number of lags in the test equation is given by considerations of minimizing the Schwarz information criterion and presence of no serial correlation in the residuals. Table 2.1 summarizes the results. For each index, the inputs in the first row correspond to the test applied to prices in local currencies whereas the second row corresponds to prices expressed in Euros.

Table 2.1. ADF *Test Results*

	Levels with trend	Levels without trend	Differences without trend
DAX	-1.39	-0.01	-13.02**
BUX	-2.35	-2.42	-16.49**
	-2.09	-1.36	-16.51**
PX-50	-1.95	-1.68	-12.48**
	-2.16	-0.72	-12.73**
WIG	-2.09	-1.71	-15.63**
	-1.99	-2.08	-8.87**

*,**significant at 5% and 1% level, respectively

The null hypothesis of unit root is clearly rejected for all time series of log-stock index prices. When applied to first-differences, i.e. stock returns, the ADF test fails to reject the null hypothesis. We conclude that the log-stock index prices are integrated of order one and proceed with the Johansen cointegration analysis.

The four stock market index prices are assumed to follow a $VAR(p)$ process. The optimal number of lags for equation (2.5) was selected using the Hannah-Quinn information criterion[12]. One lag is chosen for the time series expressed in local currencies and no lags for

[11] The cointegration analysis was performed using EViews 4.1.

[12] The Schwartz information criterion yields the same optimal number of lags.

those expressed in Euros. Table 2.2 reports the trace test statistics for cointegration. For each h, the inputs in the first row correspond to the test applied to prices in local currencies whereas the second row corresponds to prices expressed in Euros. The null hypothesis of no cointegration is rejected at the 5% significance level for the index prices expressed in local currencies. The trace test indicates a single significant cointegrating vector. On the contrary, the null hypothesis of no cointegration cannot be rejected for index prices in Euros.

Table 2.2. *Trace Test for Cointegration*

h	Char. Root	Trace Stat.	5% Critical Value
zero	0.155471 0.076653	49.02144 30.61035	47.21 47.21
at most 1	0.040182 0.043006	13.53652 14.02228	29.68 29.68
at most 2	0.023110 0.022846	4.924027 4.878938	15.41 15.41
at most 3	6.65E-05 0.000345	0.013962 0.071797	3.76 3.76

The $\lambda_{\max}$ test results reported in Table 2.3 are consistent with the findings of the trace test: the time series of stock index prices expressed in local currencies appear to be cointegrated with one cointegrating vector, but there is no cointegration relationship among the index prices in terms of Euro.

Table 2.3. $\lambda_{\max}$*-test for Cointegration*

h	Char. Root	Trace Stat.	5% Critical Value
zero	0.155471 0.076653	35.48492 16.58807	27.07 27.07
at most 1	0.040182 0.043006	8.612493 9.143346	20.97 20.97
at most 2	0.023110 0.022846	4.910064 4.807141	14.07 14.07
at most 3	6.65E-05 0.000345	0.013962 0.071797	3.76 3.76

From the cointegration analysis follows that the appropriate models for studying the stock market co-movements and predictability of stock returns are the error-correction model for prices in local currencies and Granger causality tests for stock returns in terms of Euro. We start by estimating the former and then turn to the analysis of the latter.

The estimated error-correction model is summarized in Tables 2.4a and 2.4b.

Table 2.4a. *Cointegrating Vector*

	const.	DAX	BUX	PX-50	WIG
Coeff.(γ)	10.78	1.00	0.44	2.60**	-4.07**

**significant at 1%

The long-run relationship among the German, Hungarian, Czech and Polish stock markets is given by the cointegrating vector in Table 2.4a. Without loss of generality, the coefficient for the DAX index was normalized to unity. From the estimated error-correction model of Table 2.4b follows that only the Czech and Polish stock markets respond to a deviation from the long-run relationship: the estimated α's are significant at the 1% level for PX-50 and WIG but are insignificant at the conventional levels for DAX and BUX indices. The short-term dynamics, given by the estimated $\boldsymbol{\zeta}_1$ matrix, reveals that each and every individual stock market can be predicted using lagged returns from other markets. Surprisingly, lagged returns on the Czech index significantly predict the returns on the German DAX. We have no reasonable explanation for this finding. The degree of predictability of stock index returns as measured by the coefficient of determination is fairly high compared to the estimated ARIMA models of Chapter 1. For example, lagged returns on the four indices explain about 32.5% of the variance of weekly returns on the Hungarian BUX in the error-correction equation as compared to the 3.4% in the ARIMA(2,1,0) model. Similarly, the R^2 for the PX-50 is about seven times higher that in the ARIMA(1,1,1). Hence the long-term and short-term interactions among the German, Hungarian, Czech and Polish stock markets can be exploited to substantially increase the predictability of individual stock index returns.

Table 2.4b. *The Error-Correction Model*

	Δ(DAX)	Δ(BUX)	Δ(PX-50)	Δ(WIG)
Coint.Eq.($\boldsymbol{\alpha}$)	-0.000677	0.009596	-0.051192**	0.034923**
$\Delta(\text{DAX})_{-1}$	0.145368	0.144581*	0.141677*	0.083333
$\Delta(\text{BUX})_{-1}$	-0.135455	-0.313462**	0.009323	-0.063152
$\Delta(\text{PX-50})_{-1}$	0.444787**	0.640493**	0.194253**	0.484016**
$\Delta(\text{WIG})_{-1}$	0.013138	0.170601*	-0.267150**	-0.022882
const.	-0.001772	0.001597	0.000983	-0.000297
R^2	0.151573	0.325874	0.155483	0.260624

*,**significant at 5% and 1% level, respectively

We now turn to the analysis of the time series of stock index prices expressed in Euros. Since the null hypothesis of no cointegration cannot be rejected in this case we examine the cross-country predictability of stock returns by pairwise Granger-causality tests. The test is based on equation (2.9). Under the null hypothesis that Δx_t does not Granger-cause Δy_t, the terms $\Delta x_{t-1}, \ldots, \Delta x_{t-p}$ should be insignificantly different from zero. We use a Wald-type test for the null hypothesis of no Granger-causality[13]. The test statistic is asymptotically χ^2 distributed with p degrees of freedom. To select the optimal number of lags in regression (2.9) we consider the Schwartz information criterion and check for the presence of autocorrelation in residuals. Table 2.5 summarizes the Granger-causality

[13]The usual F-test is not appropriate in our setting, since it rests on the hypotheses of normality of residuals and fixed regressors. But we saw in the previous chapter that this assumption is likely to be violated for stock returns.

analysis applied to stock index returns in terms of Euro.

Table 2.5. *Granger-causality Tests*

	Dep.var.	Test-stat.	Granger-causality	R^2
DAX and BUX	DAX	2.94	BUX→DAX: no	0.046
	BUX	3.73*	DAX→BUX: yes	0.049
DAX and PX-50	DAX	2.95	PX-50→DAX: no	0.046
	PX-50	0.44	DAX→PX-50: no	0.029
DAX and WIG	DAX	3.40*	WIG→DAX: yes	0.050
	WIG	1.85	DAX→WIG: no	0.069
BUX and PX-50	BUX	2.33	PX-50→BUX: no	0.069
	PX-50	0.63	BUX→PX-50: no	0.033
BUX and WIG	BUX	0.53	WIG→BUX: no	0.018
	WIG	0.88	BUX→WIG: no	0.060
PX-50 and WIG	PX-50	0.00	WIG→PX-50: no	0.015
	WIG	4.19*	PX-50→WIG: yes	0.033

*,**significant at 5% and 1% level, respectively

Surprisingly, the Polish WIG index appears to Granger-cause the German DAX. We would anticipate the opposite to hold since the German stock market is much larger and more mature than the Polish one. Similarly, the result of the test that the Czech index Granger-causes the Polish WIG seems to be quite intuitively implausible. To summarize the Granger-causality tests, some of the four studied stock markets appear to be interlinked and lagged returns from one market can have some forecast power in other markets. The predictability of individual stock index returns as measured by the coefficient of determination can in some cases increase when lagged returns from other stock markets are included in the regression. This conclusion is similar to that obtained for the time-series expressed in local currencies.

2.6 Concluding Remarks

The purpose of this chapter has been to investigate cross-country predictability and cointegration among the German, Hungarian, Czech and Polish stock markets. We have applied the multivariate cointegration analysis developed by Johansen (1988,1991) to the time-series of weekly prices expressed in both local currencies and in Euros.

First, the stock market index prices expressed in local currencies were found to be cointegrated with a single significant cointegrating vector. The estimated error-correction model implies that only the Czech and Polish stock markets respond to a deviation from the long-run equilibrium. All stock markets are, however, influenced by lagged returns from at least one other stock market. Hence there exists significant cross-country predictability among the four markets under study and the forecast of future returns on one market can be substantially improved by including past returns from other markets as well.

Second, there appears to be no cointegration relationship among the stock markets when the index prices are expressed in terms of Euro. This finding is consistent with the conclusion of Neubauer (2001) that the cointegration found in the period 1993:1-1997:12 disappears over time. The bivariate Granger-causality test results indicate Granger-causality running from the German DAX to the Hungarian BUX, from the Czech PX-50 to the Polish WIG, and surprisingly, from the Polish WIG to the German DAX. The predictive power of the foreign indices is usually quite low as measured by the coefficient of determination. It is thus questionable, whether the empirically observed Granger-causality can be exploited to earn above average returns, especially when adjusted for transaction costs and the costs of hedging against exchange rate movements. Further research is therefore required to clarify this question.

Chapter 3

Nonsynchronous Trading and Predictability

It is the common practice in the vast majority of stock market empirical studies to implicitly assume that the data used for the study are sampled at equidistant points in time. Usually, a researcher chooses a particular data frequency (daily, weekly, monthly, etc.) best suiting the purpose of his/her analysis and takes for granted that the data were really sampled at the times implied by the data frequency. To give an example, it is fairly common to use daily data on stock prices for testing various hypothesis about emerging stock markets because the history of these markets is short, which rules out the possibility of obtaining a sample of plausible size with lower data frequency. The daily stock price is defined as the closing security price on that day. Hence the researcher uses stock prices prevailing at the time of the stock exchange closing. Note, however, that these prices were not necessarily sampled at time of closing of the exchange. If the last trade in the security occurred an hour before the close then the price for that day was actually sampled an hour before the close and not at the close. Clearly, if trading in a security does not take place continuously, daily data are not sampled equidistantly. Furthermore, if a security does not trade at all on a particular day, the return for that day is mistakenly assumed to be zero, even though the theoretical return for that day may be non-zero but was not realized because the stock simply did not trade (due to low liquidity, fixed transaction costs, etc.). It is the purpose of this chapter to show that nontrading induces spurious autocorrelation into the time-series of observed stock returns. But when the researcher ignores the effect of nontrading (i.e. if he/she ignores the fact that the data were sampled at non-equidistant points in time) he/she must necessarily confuse spurious autocorrelation with the autocorrelation inherent to the returns. As a result, completely false inferences about the hypothesis under study may be drawn[1].

In a similar way, the nonsynchronicity of data sampling has important implications for the time-series properties of portfolio returns. To develope some intuition for this claim, consider for simplicity a portfolio comprised of two stocks only: a stock of a large corporation

[1] As an example, a researcher may find a stock market to be weak-form inefficient due to high autocorrelation in stock returns, although the market may be fairly efficient once the effect of non-synchronous trading (and hence spurious autocorrelation) is accounted for.

trading almost continuously[2] and a stock of a small company trading less frequently. If news affecting the aggregate stock market arrives shortly before the close of the stock exchange it is likely to be incorporated into the closing prices of the frequently traded stocks. But the stocks of small companies need not trade immediately after the arrival of the news and hence their prices may reflect the news with a lag of a day. As a result, our two-stock portfolio will exhibit spurious autocorrelation induced by nonsynchronous trading, even though the individual stocks have the same underlying stochastic processes driving their returns.

The Lo-MacKinlay model of nonsynchronous trading is discussed in this chapter (Lo and MacKinlay (1990)) to address the problem considered above. In the first section, the model is derived under the hypothesis that the common factor generating "virtual" stock returns follows a stationary first-order autoregressive process. This generalization is proposed in Lo and MacKinlay (1990) although they do not derive the model explicitly under this hypothesis. We present the formulas for the mean, variance, autocovariance and cross-autocovariance of individual as well as portfolio returns under the more general hypothesis. The implications of nonsynchronous trading on the predictability of the Central-European stock index returns are discussed in the concluding Section 3.2.

3.1 An Econometric Model of Nonsynchronous Trading

Following Lo and MacKinlay (1990) we will assume throughout this chapter that there are N securities whose unobservable "virtual" returns r_{it} at time t, $i = 1, .., N$ are generated by the following stochastic process:

$$r_{it} = \mu_i + \beta_i \Lambda_t + \varepsilon_{it}, \tag{3.1}$$

where Λ_t is a zero-mean common factor and ε_{it} is a zero-mean idiosyncratic noise satisfying $E[\varepsilon_{it}\varepsilon_{jt-n}] = 0$ for all i, j, n and t. Unlike Lo and MacKinlay (1990), however, we allow the common factor Λ_t to follow a stationary first-order autoregressive process:

$$\Lambda_t = \alpha \Lambda_{t-1} + \eta_t, \tag{3.2}$$

where $\{\eta_t\}$ is a white noise process and $\alpha \in (-1, 1)$. We further assume that Λ_t is independent of ε_{it-n} for all i, n and t. Under these conditions the moments of the unobserved virtual returns are given by:

$$E[r_{it}] = \mu_i, \tag{3.3}$$

$$Var[r_{it}] = \beta_i^2 \frac{\sigma_\eta^2}{1-\alpha^2} + \sigma_{\varepsilon_i}^2, \tag{3.4}$$

$$Cov[r_{it}, r_{it-n}] = \beta_i^2 \frac{\alpha^n \sigma_\eta^2}{1-\alpha^2}, n > 0, \tag{3.5}$$

where $\sigma_\eta^2 \equiv Var[\eta_t]$ and $\sigma_{\varepsilon_i}^2 \equiv Var[\varepsilon_{it}]$.

[2]By 'almost continuously' we mean within intervals of tens of seconds.

In each period there is some positive probability p_i that security i does not trade. The observed return r_{it}^0 in such period is simply equal to zero, although the virtual return r_{it} is given by (3.1). In the following period security i again does not trade with probability p_i and the process continues. If the security trades in period t and did not trade for the past n consecutive periods then the observed return for period t is simply defined as the sum of the virtual returns for the n periods that the security did not trade and the virtual return for period t. To determine the impact of nontrading on the observed returns r_{it}^0 we need to define a binary stochastic process that will govern the trading/nontrading process. This process is identical to that in Lo and MacKinlay (1990).

Definition 2 *Let δ_{it} and $X_{it}(k)$ be the following Bernoulli random variables:*

$$\begin{aligned}
\delta_{it} &= \begin{cases} 1 & \text{with probability } p_i, \\ 0 & \text{with probability } (1-p_i), \end{cases} \\
X_{it}(k) &\equiv (1-\delta_{it})\delta_{it-1}\delta_{it-2}\cdots\delta_{it-k} \\
&= \begin{cases} 1 & \text{with probability } (1-p_i)p_i^k, \\ 0 & \text{with probability } 1-(1-p_i)p_i^k, \end{cases} \\
X_{it}(0) &\equiv 1-\delta_{it}.
\end{aligned}$$

where we assume that δ_{it} is an independently identically distributed random variable for all i.

Now having defined $X_{it}(k)$ the reader can easily verify that the process generating the observed returns r_{it}^0 described in words above is equivalent to the following definition:

Definition 3 *The observed return process r_{it}^0 is given by the following stochastic process:*

$$r_{it}^0 = \sum_{k=0}^{\infty} X_{it}(k) r_{it-k},$$

for all i.

If security i does not trade at time t, then $\delta_{it} = 1$, thus $X_{it}(k) = 0$ for all k and the observed return at time is $r_{it}^0 = 0$. If security i does trade at t then the observed return at t is equal to the sum of the virtual return at t, r_{it} and its past $\tilde{k}_{it}$ virtual returns, where the random variable $\tilde{k}_{it}$ denotes the number of past consecutive periods for which security i did not trade. We call $\tilde{k}_{it}$ the duration of non-trading and define it as

$$\tilde{k}_{it} \equiv \sum_{k=0}^{\infty} \left\{ \prod_{j=1}^{k} \delta_{it-j} \right\}. \tag{3.6}$$

Using the duration of nontrading the observed return can be equivalently defined as

$$r_{it}^0 = \sum_{k=0}^{\tilde{k}_{it}} r_{it-k}. \tag{3.7}$$

Although this definition may seem to be a more intuitive one, Definition 1 will prove to be more useful in the subsequent derivation of the moments of r_{it}^0. This is because it is quite difficult to work with a random sum of random variables like that in (3.7).

To understand how the probability of non-trading effects the duration of nontrading consider the expected value and variance of $\tilde{k}_{it}$:

$$E\left[\tilde{k}_{it}\right] = \frac{p_i}{1-p_i}, \tag{3.8}$$

$$Var\left[\tilde{k}_{it}\right] = \frac{p_i}{(1-p_i)^2}. \tag{3.9}$$

Since the derivation of (3.8) and (3.9) is trivial, we leave it to the reader for the sake of brevity. If, for instance, $p_i = \frac{1}{2}$ then the average number of consecutive periods for which security i does not trade is equal to one. If $p_i = \frac{4}{5}$ then the expected duration of nontrading is equal to 4. Clearly, if the security trades at all times (i.e. $p_i = 0$) then $\tilde{k}_{it}$ is deterministic and equal to zero.

To see how nonsynchronous trading affects the time-series properties of returns we need to derive the moments of individual and portfolio returns. This is done in a way that is very similar to the case when Λ_t is an *iid* process (i.e. when $\alpha = 0$) and the reader who is not interested in studying the lengthy proofs of the theorems below may wish to simply check, that setting $\alpha = 0$ throughout the theorems reproduces exactly the results obtained in Lo and MacKinlay (1990).

3.1.1 Individual Returns

In this paragraph we derive the mean, variance, autocovariance and cross-autocovariance of observed individual stock returns. For expositional convenience, we summarize the results in a theorem and present the proof in the Appendix.

Theorem 4 *Under Definition 2 the observed return processes* $\{r_{it}^0\}$, $i = 1, .., N$, *are covariance-stationary with the following moments:*

$$E\left[r_{it}^0\right] = \mu_i,$$

$$Var\left[r_{it}^0\right] = \sigma_{\varepsilon_i}^2 + \beta_i^2\left(1 + \frac{2\alpha p_i}{1-\alpha p_i}\right)\frac{\sigma_\eta^2}{1-\alpha^2} + \frac{2p_i}{1-p_i}\mu_i^2,$$

$$Cov\left[r_{it}^0, r_{it+n}^0\right] = \begin{cases} -\mu_i^2 p_i^n + \left(\frac{\alpha\beta_i^2\sigma_\eta^2(1-p_i)^2}{(1-\alpha^2)(1-\alpha p_i)}\right)\frac{\alpha_i^n - p_i^n}{\alpha - p_i}, \; n > 0, \; \alpha \neq p_i, \\ \left(\frac{n\beta_i^2\sigma_\eta^2}{(1+p_i)^2} - \mu_i^2\right)p_i^n, \; n > 0, \; \alpha = p_i, \end{cases},$$

$$Cov\left[r_{it}^0, r_{jt+n}^0\right] = \begin{cases} \beta_i\beta_j\frac{\sigma_\eta^2(1-p_i)(1-p_j)}{(1-p_ip_j)(1-\alpha^2)}\left(\alpha^n + p_j^n + \frac{\alpha^{n+1}p_i}{1-\alpha p_i} + \frac{\alpha p_j^{n+1}}{1-\alpha p_j} + \frac{\alpha^n p_j - \alpha p_j^n}{\alpha - p_i}\right), \; \alpha \neq p_j, \\ \beta_i\beta_j\frac{\sigma_\eta^2(1-p_i)}{(1-p_ip_j)(1+p_j)}\left(1 + n + \frac{p_j^2}{1-p_j^2} + \frac{p_ip_j}{1-p_ip_j}\right)p_j^n, \; \alpha = p_j, \; n \geq 0. \end{cases}$$

From Theorem 4 we see that nontrading does not affect the mean of observed returns. The effect on the variance of observed returns is ambiguous. Assume, for simplicity, that

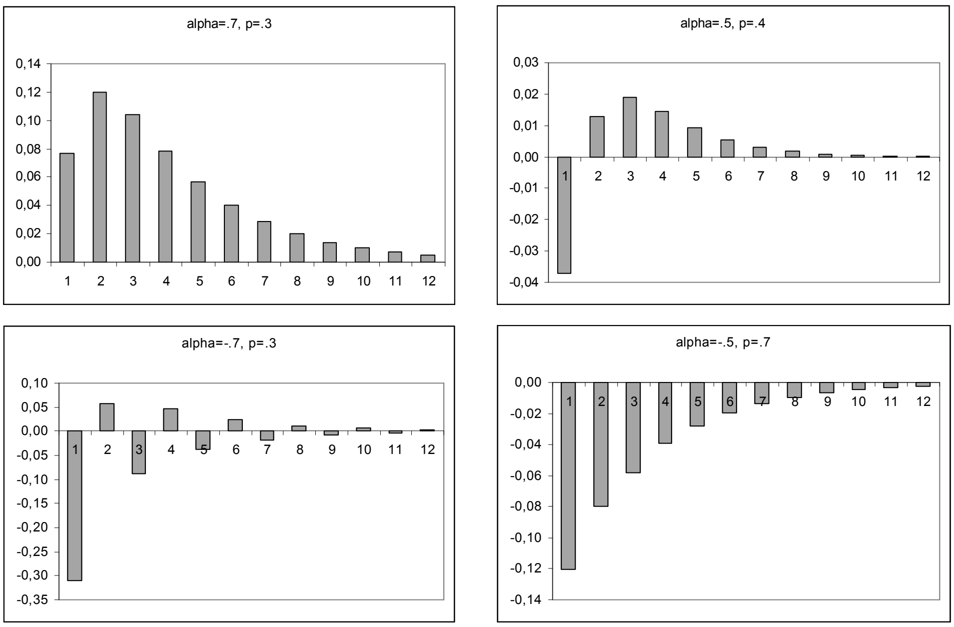
$\mu_i = 0$. Then $Var\left[r_{it}^0\right]$ is larger then $Var\left[r_{it}\right]$ if

$$\left(1 + \frac{2\alpha p_i}{1 - \alpha p_i}\right) = \frac{1 + \alpha p_i}{1 - \alpha p_i} > 1. \tag{3.10}$$

Since $\alpha \in (-1, 1)$ and $p \in \langle 0, 1)$ it follows that (3.10) holds for α positive and hence nonsynchronous trading increases the variance of observed returns. The opposite is true for a negative α. When the mean of returns μ is non-zero, the analysis is more complicated but similarly to our simple example, $Var\left[r_{it}^0\right]$ can be either higher of lower then $Var\left[r_{it}\right]$ depending upon the values of the parameters of the model.

Unlike in Lo and MacKinlay (1990) our specification of the common factor allows for a more general pattern of autocorrelation. Depending upon the values of the parameters of the model, the autocorrelation of observed returns can be either positive or negative. In both cases it decays as $n \to \infty$ but the pattern of the decay can be very different for different parameter values. The autocorrelation function can take virtually any form. To give a few examples, we set $\mu_i = 0.1, \beta_i = 1, \sigma_{\varepsilon i}^2 = 0.1$, and $\sigma_\eta^2 = 0.15$ and plot the implied autocorrelation functions of observed returns for different values of α and p_i. As shown in Figure 3.1, the serial correlation can decay geometrically, it can also first rise and then gradually fall, or it can exhibit oscillatory decay. It is apparent from the formula for $Cov\left[r_{it}^0, r_{jt+n}^0\right]$ that we would obtain similar results when we plotted the cross-autocorrelation function of observed returns on two assets.

Figure 3.1. *Autocorrelation Function of* r_{it}^0.



3.1.2 Portfolio Returns

We have seen in the previous paragraph that the moments of the individual observed returns are complicated functions of the parameters of the underlying stochastic processes. As such, they are hardly to be a subject of an empirical application because the number of parameters to be estimated is too large. Moreover we are interested in the effect of nontrading under the conditions of an autocorrelated common factor and not in estimating a large number of parameters, which would only make the analysis more complicated without gaining much insight into the main subject of this chapter.

In this paragraph, we show that if a sufficient number of individual stocks with identical nontrading probabilities are grouped into a portfolio, the formulas for the moments of the portfolio observed returns simplify tremendously. For instance, the formula for the first-order autocorrelation coefficient of the portfolio observed returns is a function of the probability of non-trading, p_i, and the autoregressive parameter, α,only. This specification allows us to decompose the empirically estimated autocorrelation coefficient into the part induced by nontrading and the part is inherent to the virtual returns. We start by defining the portfolio return, then invoke the Kolmogorov Strong Law of Large Numbers to simplify the stochastic process generating observed returns of well-diversified portfolios and finally, we derive the portfolio observed return moments.

Definition 5 *Let I_a denote the set of securities with identical probabilities of non-trading, p_a, and let N_a denote the number of securities in I_a. Let $r^0_{it}, i \in I_a$,denote the observed return on security i in time t. Then the observed return on a portfolio a is given by*

$$r^0_{at} \equiv \frac{1}{N_a} \sum_{i \in I_a} r^0_{it}.$$

Under Definitions 2 and 3 the observed portfolio return can be written as

$$\begin{aligned} r^0_{at} &\equiv \frac{1}{N_a} \sum_{i \in I_a} r^0_{it}, && (3.11)\\ &= \frac{1}{N_a} \sum_{i \in I_a} \sum_{k=0}^{\infty} X_{it}(k) r_{it-k}, && (3.12)\\ &= \sum_{k=0}^{\infty} \left(\frac{1}{N_a} \sum_{i \in I_a} \mu_i X_{it}(k) + \frac{\Lambda_{t-k}}{N_a} \sum_{i \in I_a} \beta_i X_{it}(k) + \frac{1}{N_a} \sum_{i \in I_a} \varepsilon_{it-k} X_{it}(k) \right]. && (3.13) \end{aligned}$$

Now the three sums in the parentheses can be simplified by applying the Kolmogorov Strong Law of Large Numbers. Since by assumption the nontrading stochastic process $X_{it}(k)$ is independent across securities and since we also assume that $X_{it}(k)$ is independent of ε_{it-k} for all securities at all lead and lags, and ε_{it-k} is independent across securities at all leads and lags, the summands in the terms in parentheses in (3.13) are independent random variables and hence

$$\frac{1}{N_a} \sum_{i \in I_a} \mu_i X_{it}(k) - E\left[\frac{1}{N_a} \sum_{i \in I_a} \mu_i X_{it}(k) \right] \overset{a.s.}{\rightarrow} 0, \qquad (3.14)$$

$$\frac{1}{N_a} \sum_{i \in I_a} \beta_i X_{it}(k) - E\left[\frac{1}{N_a} \sum_{i \in I_a} \beta_i X_{it}(k) \right] \overset{a.s.}{\rightarrow} 0, \qquad (3.15)$$

$$\frac{1}{N_a} \sum_{i \in I_a} \varepsilon_{it-k} X_{it}(k) - E\left[\frac{1}{N_a} \sum_{i \in I_a} \varepsilon_{it-k} X_{it}(k) \right] \overset{a.s.}{\rightarrow} 0. \qquad (3.16)$$

It is straightforward to show (see Definition 1) that

$$E\left[\frac{1}{N_a}\sum_{i\in I_a}\mu_i X_{it}(k)\right] = (1-p_a)p_a^k\mu_a,$$
$$E\left[\frac{1}{N_a}\sum_{i\in I_a}\beta_i X_{it}(k)\right] = (1-p_a)p_a^k\beta_a,$$
$$E\left(\frac{1}{N_a}\sum_{i\in I_a}\varepsilon_{it-k} X_{it}(k)\right] = 0,$$

where $\mu_a = \frac{1}{N_a}\sum_{i\in I_a}\mu_i$ and $\beta_a = \frac{1}{N_a}\sum_{i\in I_a}\beta_i$. Substituting these results into (3.13) the following equality obtains almost surely as the number of securities, N_a, increases without bound:

$$r_{at}^0 \overset{a.s.}{=} \mu_a + (1-p_a)\beta_a\sum_{k=0}^{\infty} p_a^k\Lambda_{t-k}. \tag{3.17}$$

Since this expression is identical to that derived under the hypothesis of an *iid* common factor we urge the interested reader to consult Lo and MacKinlay (1990) on further details on the derivation of (3.17) and proceed with computing the asymptotic mean, variance, auto-covariance and cross-autocovariance of portfolio returns.

Theorem 6 *As the number of securities in portfolios A and B increases without bound, the first and second moments of portfolio returns are given by*

$$E\left[r_{\kappa t}^0\right] = \mu_\kappa,$$
$$Var\left[r_{\kappa t}^0\right] = \beta_\kappa^2\left(\frac{1-p_\kappa}{1+p_\kappa}\right)\left(1+\frac{2\alpha p_\kappa}{1-\alpha p_\kappa}\right)\frac{\sigma_\eta^2}{1+\alpha^2},$$
$$Cov\left[r_{\kappa t}^0, r_{\kappa t+n}^0\right] = \begin{cases}\beta_\kappa^2\frac{(1-p_\kappa)\sigma_\eta^2}{(1+p_\kappa)(1+\alpha^2)}\left(\alpha^n + p_\kappa^n + \frac{\alpha^{n+1}p_\kappa+\alpha p_\kappa^{n+1}}{1-\alpha p_\kappa} + \frac{\alpha^n p_\kappa - \alpha p_\kappa^n}{\alpha - p_\kappa}\right), & \alpha \neq p_\kappa, \\ \beta_\kappa^2\frac{\sigma_\eta^2}{(1+p_\kappa)^2}\left(1+n+\frac{2p_\kappa^2}{1-p_\kappa^2}\right)p_\kappa^n, & \alpha = p_\kappa,\end{cases}$$
$$Corr\left[r_{\kappa t}^0, r_{\kappa t+n}^0\right] = \begin{cases}\frac{\alpha^n + p_\kappa^n + \frac{\alpha^{n+1}p_\kappa+\alpha p_\kappa^{n+1}}{1-\alpha p_\kappa} + \frac{\alpha^n p_\kappa - \alpha p_\kappa^n}{\alpha - p_\kappa}}{1+\frac{2\alpha p_\kappa}{1-\alpha p_\kappa}}, & \alpha \neq p_\kappa, \\ \left(1+\frac{n(1-p_\kappa)^2}{1+p_\kappa^2}\right)p_\kappa^n, & \alpha = p_\kappa,\end{cases}$$
$$Cov\left[r_{at}^0, r_{bt+n}^0\right] = \begin{cases}\beta_a\beta_b\frac{(1-p_a)(1-p_b)\sigma_\eta^2}{(1-p_a p_b)(1+\alpha^2)}\left(\alpha^n + p_b^n + \frac{\alpha^{n+1}p_a}{1-\alpha p_a} + \frac{\alpha p_b^{n+1}}{1-\alpha p_b} + \frac{\alpha^n p_b - \alpha p_b^n}{\alpha - p_b}\right), & \alpha \neq p_b, \\ \beta_a\beta_b\frac{(1-p_a)\sigma_\eta^2}{(1-p_a p_b)(1+p_b^2)}\left(1+n+\frac{p_b^2}{1-p_b^2}+\frac{p_a p_b}{1-p_a p_b}\right)p_b^n, & \alpha = p_b,\end{cases}$$

where $n \geq 0$, $\mu_\kappa = \frac{1}{N_\kappa}\sum_{i\in I_\kappa}\mu_i$ *and* $\beta_\kappa = \frac{1}{N_\kappa}\sum_{i\in I_\kappa}\beta_i$ *and the symbol* $\overset{a}{=}$ *denotes that the equality obtains only asymptotically.*

Again, the formulas for portfolio observed returns are fairly complicated functions of the underlying parameters of the stochastic processes. Unlike for individual returns, however, the autocorrelation coefficient of observed portfolio returns is a function of the probability of non-trading p_κ, the autoregressive parameter α, and the order of autocorrelation n only.

This specification allows for a direct analysis of the effect of nonsynchronous trading on the time series properties of observed portfolio returns. In particular, we can decompose the estimated autocorrelation coefficient of observed returns into the part induced by nonsynchronous trading and the part that comes from the autoregressive common factor. To see this, let us concentrate on the first-order autocorrelation coefficient of portfolio returns:

$$\rho_\kappa(1) \equiv Corr\left[r^0_{\kappa t}, r^0_{\kappa t+1}\right] = \frac{\alpha + p_\kappa + \frac{\alpha^2 p_\kappa + \alpha p_\kappa^2}{1-\alpha p_\kappa}}{1 + \frac{2\alpha p_\kappa}{1-\alpha p_\kappa}} = \frac{\alpha + p_\kappa}{1+\alpha p_\kappa}. \tag{3.18}$$

The first-order autocorrelation coefficient can be consistently estimated from the time-series of observed portfolio returns. The probability of non-trading can be also estimated directly from the time-series of trading volumes of the stocks in the portfolio, which we denote $\{v_{it}\}_{i=1}^T, i \in I_\kappa$. This is done by constructing a binary sequence $\{\delta_{it}\}_{t=1}^T$ for each i, where

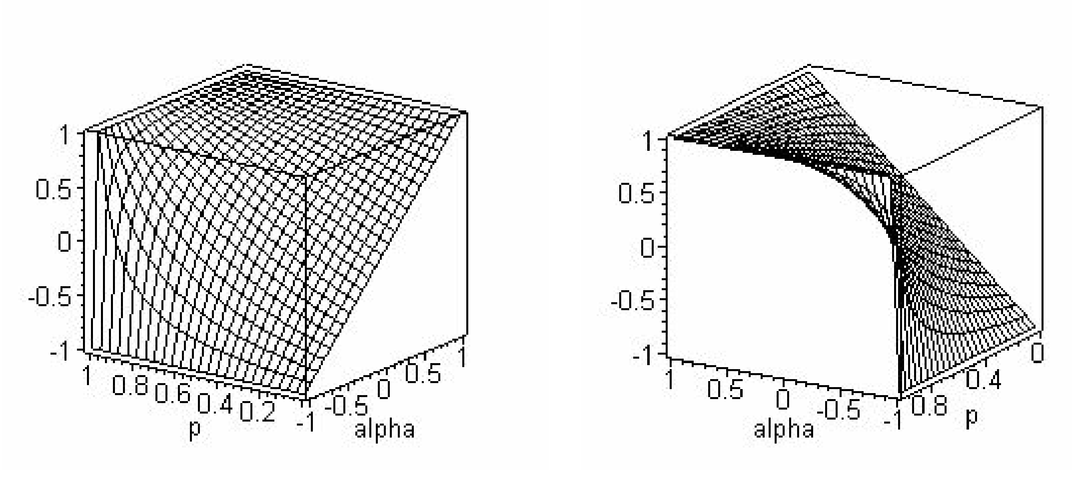
$$\delta_{it} = \begin{cases} 1 & if \quad v_{it} = 0 \\ 0 & otherwise \end{cases}$$

Clearly, δ_{it} is equal to one if security i does not trade at time t, i.e. when the trading volume is equal to zero at t, and zero otherwise. The estimator of the probability of non-trading is then constructed as

$$\hat{p}_\kappa = \frac{1}{N_{T_\kappa}} \sum_{t \in T_\kappa} \left(\frac{1}{N_\kappa} \sum_{i \in I_a} \delta_{it} \right), \tag{3.19}$$

where the subscript of the first sum indicates that we sum over particular dates that are elements of the set T_κ. For example, Lo and MacKinlay (1990) use only month-end trading days over (nonconsecutive) sixteen years to estimate p_κ. Hence in their case, T_κ includes only month-end trading days. The term in the parentheses is simply the fraction of securities in portfolio κ that do not trade at time t. Thus, the estimator $\hat{p}_\kappa$ is an average of the fraction of securities in the portfolio that do not trade at dates $t \in T_\kappa$. Note, that since we assume

Figure 3.2. *Behavior of* $\rho_\kappa(1)$



that the portfolio κ is well-diversified, i.e. N_κ is large, not much is gained in efficiency by choosing a large set T_κ. Asymptotically (as $N_\kappa \to \infty$) the term in the parentheses approaches the true value of the parameter p_κ and summing over the elements of T_κ and dividing by N_{T_κ} only reproduces the same result. Hence, when the portfolio is sufficiently large, month-end trading days (e.g. approximately 1/21st of all daily observations in the sample) are sufficient to estimate p_κ efficiently.

Now having consistently estimated $\rho_\kappa(1)$ and p_κ, we can compute the estimate of the unobserved parameter α by inverting the function $\rho_\kappa(1)$ and substituting $\hat{\rho}_\kappa(1)$ and $\hat{p}_\kappa$ for $\rho_\kappa(1)$ and p_κ respectively. This yields a consistent estimator $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\hat{p}_\kappa - \hat{\rho}_\kappa(1)}{\hat{p}_\kappa \hat{\rho}_\kappa(1) - 1}. \tag{3.20}$$

of the autoregressive parameter α. When trading takes place at all times ($p_\kappa = 0$) then the only source of autocorrelation in observed returns is the common factor. In such a case, the first-order autocorrelation coefficient of observed returns is equal to α, and we denote it $\tilde{\rho}(1)$. Whenever $p_\kappa > 0$, however, the coefficient is given by (3.18). This implies a straightforward decomposition of the estimated first-order autocorrelation coefficient:

$$\begin{aligned} \hat{\rho}_\kappa(1) &= \tilde{\rho}(1) + \left[\hat{\rho}_\kappa(1) - \tilde{\rho}(1)\right], \\ &= \hat{\alpha} + \left[\frac{\hat{\alpha} + \hat{p}_\kappa}{1 - \hat{\alpha}\hat{p}_\kappa} - \hat{\alpha}\right] \end{aligned}$$

When the probability of nontrading is equal to zero, the only source of serial correlation in observed returns is the autocorrelated common factor. For $p_\kappa \neq 0$, the first-order autocorrelation coefficient is an increasing function of p_κ since

$$\frac{\partial \rho_\kappa(1)}{\partial p_\kappa} = \frac{1 - \alpha^2}{(1 + \alpha p_\kappa)^2} > 0$$

for all admissible α and p_κ. See Figure 3.2 for a graph of $\rho_\kappa(1)$ as a function of α and p_κ.

In their original model, Lo and MacKinlay (1990) also investigate the impact of daily nontrading probability on the time-series properties of weekly, monthly and lower-frequency returns. Although it would be computationally straightforward to do such analysis in our framework, we refrain from that, because the formulas for the moments of observed returns are already quite complicated functions of the underlying parameters. We anticipate that the results would be qualitatively fairly similar to those of Lo and MacKinlay (1990), e.g. that the spurious autocorrelation induced by nontrading decreases as the data frequency decreases.

3.2 Concluding Remarks

The Central-European stock markets are known for a lower degree of liquidity than the developed equity markets. A prominent example is the Prague Stock Exchange, where out of a total of 76 listed securities, only 7 stocks trade on a daily basis in recent years. The liquidity in the rest of the market is nil. Unfortunately, the econometric model of nontrading derived above cannot be directly applied to the Czech, Hungarian and Polish stock returns. The number of infrequently traded stocks with similar nontrading probabilities is to low for the asymptotic portfolio moments to reasonably approximate their finite-sample counterparts. But the model still sheds some light on the possible sources of the predictability of the Czech and Hungarian stock index returns found in Chapter 1. Both the Czech PX-50 and the Hungarian BUX indices contain infrequently traded securities. Although we cannot quantify the effect of nontrading on the time-series properties of the indices directly, it is highly suspicious that the autocorrelation of their returns is entirely real. The problem of spurious autocorrelation is necessarily present and hence the predictability of the Czech and Hungarian stocks should be evaluated using indices composed of daily traded securities only. But this fact is often ignored in the empirical literature, and the PX-50 and BUX indices are widely used in various empirical studies[3] because it is believed that these indices best mirror the evolution of the respective equity markets. As a result, spurious a real autocorrelation may be confused, leading to false inferences about the hypotheses under study.

[3] See, for example, Filáček et al. (1998), Hanousek and Filer (2000), and Gilmore and McManus (2001).

Chapter 4

Maximizing Predictability of Asset Returns

We have seen in the first two chapters of this thesis that the Central-European stock market index returns contain predictable components. In the previous chapter, we have shown that the predictability of individual or portfolio returns as measured by the first-order autocorrelation coefficient may be in part spurious, induced by infrequent and/or nonsynchronous trading. Thus the remaining question which needs to be addressed is whether our findings of return predictability can be useful in the real-world investment process.

The purpose of this chapter is to investigate whether the predictability of stock returns is economically significant, i.e. if it can be exploited in practice to earn abnormal returns when adjusted for transactions costs. To do this, we first present the methodology of maximizing predictability of stock returns. Although individual security prices can be barely predictable, optimally formed portfolios may exhibit much higher degree of predictability. We then develop a simple dynamic trading strategy based on the optimal portfolios and asses its performance using various measures of market timing. To avoid the problem of nonsynchronous trading we apply the analysis to the most liquid Central-European stocks only, i.e. to securities trading on a daily basis.

The chapter is organized as follows. In Section 4.1, we briefly discuss the principal component analysis (PC) and the way it can be used to forecast stock returns from the time-series of historical returns only. In Section 4.2 we consider the predictive power of various macroeconomic factors and term-structure variables and present the methodology of the Maximally Predictable Portfolio (MPP) due to Lo and MacKinlay (1997). Various measures of market timing and investment performance are described in Section 4.3. Finally, in Section 4.4. we apply the principal component analysis to selected Czech stocks and the MPP analysis to selected Polish stocks to forecast their respective returns[1]. We follow a simple trading rule, in which we are invested all in a stock portfolio if the expected return

[1] The choice of method for maximizing predictability is motivated by previous research. In Hanousek and Filer (1999), lagged macroeconomic variables were found to Granger-cause stock returns on Polish stock market, whereas no Granger-causality was found for the Czech stock market. We thus use a multifactor model for forecasting Polish stock returns with macroeconomic and term structure variables as factors and use principle component analysis to extract the most important factors that drive the Czech stock returns.

on this portfolio exceeds that of a risk-free asset and all in the risk-free asset if the opposite holds. We then evaluate the performance of this trading strategy using the market timing measures of Section 4.3. We conclude with a short discussion in Section 4.5.

4.1 Principal Component Analysis

A popular approach to studying predictability is to assume a multifactor asset pricing model and then investigate the predictability of the individual factors[2]. This is a two-step procedure. First, a linear cross-sectional model for stock returns is estimated:

$$\mathbf{r}_t = \boldsymbol{\alpha} + \boldsymbol{\delta}^T \mathbf{F}_t + \boldsymbol{\varepsilon}_t, \tag{4.1}$$

where $\mathbf{r}_t \equiv [r_{1t}\ r_{2t}\ \cdots\ r_{nt}]^T$ is a vector of n stock returns in period t, $\mathbf{F}_t \equiv [F_{1t}\ F_{2t}\ \cdots\ F_{pt}]^T$ is a vector of p factors in period t, and $\boldsymbol{\varepsilon}_t \equiv [\varepsilon_{1t}\ \varepsilon_{2t}\ \cdots\ \varepsilon_{nt}]$ is a vector white noise with covariance matrix $\boldsymbol{\Sigma}$. $\boldsymbol{\alpha}$ denotes an $(n \times 1)$ vector of intercepts and $\boldsymbol{\delta}^T$ is an $(n \times p)$ matrix of coefficients. When the p factors are known, the system of regression equations (4.1) can be estimated equation-by-equation by OLS[3]. If $\mathbf{F}_t$ is unknown, the principal component analysis can be applied to extract the most important factors driving the stock returns. Let $\boldsymbol{\Gamma}_r \equiv Var\,[\mathbf{r}_t]$ denote the covariance matrix of $\mathbf{r}_t$. The first principal component is a linear combination $\boldsymbol{\gamma}_{PC1}^T \mathbf{r}_t$ that has the maximum possible variance among all linear combinations satisfying $\sum_{i=1}^{p} \gamma_i^2 = 1$, i.e. $\boldsymbol{\gamma}_{PC1}^T$ it is a solution to the following optimization problem:

$$\arg\max_{\gamma \in R^n} Var\left[\boldsymbol{\gamma}^T \mathbf{r}_t\right] \quad \text{s.t.} \quad \boldsymbol{\gamma}^T \boldsymbol{\gamma} = 1.$$

It can be shown[4] that $\boldsymbol{\gamma}_{PC1}$ is given by the eigenvector corresponding to the largest eigenvalue λ_1 of the covariance matrix $\boldsymbol{\Gamma}_r$. The first principal component is therefore a portfolio that captures most of the variance of the stock returns. Since stock portfolio weights ω_i sum up to one, we further standardize the vector $\boldsymbol{\gamma}_{PC1}$ such that $\boldsymbol{\omega}_{PC1} = \boldsymbol{\gamma}_{PC1}/(\boldsymbol{\gamma}_{PC1}^T \iota)$, where ι is an $(n \times 1)$ vector of ones. Consequently, the rate of return on the first principal component portfolio is equal to $r_t^{PC1} \equiv \boldsymbol{\omega}_{PC1}^T \mathbf{r}_t$.

Second, we analyze the predictability of the most important factor, i.e. the portfolio given by the first principal component weights $\boldsymbol{\omega}_{PC1}^T$. A straightforward measure is the squared first-order autocorrelation coefficient of r_t^{PC1}, which corresponds to the R^2 from the regression

$$r_t^{PC1} = a + \rho_1 r_{t-1}^{PC1} + \epsilon_t, \tag{4.2}$$

where $\{\epsilon_t\}_{t=1}^T$ is a white-noise sequence with $Var\,[\epsilon_t] = \sigma_\epsilon^2$. Note, however, that the pattern of autocorrelation of r_t^{PC1} may be more general than that implied by an $AR(1)$ process. In such a case, a relevant measure of the predictability of r_t^{PC1} is the R^2 from 'the appropriate' $ARMA$ model[5]. We will apply this approach to selected Czech stocks in Section 4.4 and asses the out-of-sample predictability of r_t^{PC1}.

[2] See Lo and MacKinlay (1997), and the references therein.

[3] Since each equations contains the same regressors, there is no gain by running SUR (Seemingly Unrelated Regressions).

[4] See Haerdle and Simar (2001), Theorem 2.5.

[5] By 'appropriate' we mean appropriate in the Box and Jenkins (1984) sense.

4.2 The Maximally Predictable Portfolio

Lo and MacKinlay (1997) propose an alternative to cross-sectional factors derived by the principal component analysis. They show that although r_t^{PC1} captures most of the variance of the n stock returns, it need not reflect much of the inherent stock return predictability. Instead of maximizing variance, they construct a portfolio by maximizing predictability as measured by the well-known coefficient of determination.

Assume we regress the stock returns on lagged rather than contemporaneous values of factors driving stock returns:

$$\mathbf{r}_t = \boldsymbol{\alpha} + \boldsymbol{\delta}^T \mathbf{F}_{t-1} + \boldsymbol{\varepsilon}_t. \tag{4.3}$$

Each of the n coefficients of determination, $R_i^2, i = 1 \ldots n$, measures the predictability of corresponding individual stock returns included in $\mathbf{r}_t$. Now suppose we form a linear combination $\boldsymbol{\gamma}^T \mathbf{r}_t$ of the n stocks, and consider the coefficient of determination when we regress $\boldsymbol{\gamma}^T \mathbf{r}_t$ on a constant and $\mathbf{F}_{t-1}$:

$$R^2(\boldsymbol{\gamma}) = \frac{Var\left[\boldsymbol{\gamma}^T(\boldsymbol{\delta}^T \mathbf{F}_{t-1})\right]}{Var\left[\boldsymbol{\gamma}^T \mathbf{r}_t\right]} = \frac{\boldsymbol{\gamma}^T Var\left[\boldsymbol{\delta}^T \mathbf{F}_{t-1}\right] \boldsymbol{\gamma}}{\boldsymbol{\gamma}^T Var\left[\mathbf{r}_t\right] \boldsymbol{\gamma}} = \frac{\boldsymbol{\gamma}^T \Gamma_F \boldsymbol{\gamma}}{\boldsymbol{\gamma}^T \Gamma_r \boldsymbol{\gamma}}. \tag{4.4}$$

where $Var\left[\boldsymbol{\delta}^T \mathbf{F}_{t-1}\right] \equiv \Gamma_F$ and $Var\left[\mathbf{r}_t\right] \equiv \Gamma_r$. To maximize the predictability of a portfolio is then equivalent to choosing $\boldsymbol{\gamma}$ such that (4.4) is maximized and $\boldsymbol{\gamma}$ is a portfolio, i.e. $\boldsymbol{\gamma}^T \iota = 1$. But because $R^2(\boldsymbol{\gamma}) = R^2(k\boldsymbol{\gamma})$ for any constant k, we can maximize (4.4) without imposing the constraint and then rescale the solution $\boldsymbol{\gamma}$ such that its elements sum up to unity. The maximum $R^2(\boldsymbol{\gamma})$ is given by the largest eigenvalue λ_1 of the matrix $\mathbf{B} \equiv \boldsymbol{\Gamma}_r^{-1} \Gamma_F$ and is attained by the eigenvector γ_{MPP} corresponding to the largest eigenvalue of $\mathbf{B}$[6]. Properly normalizing γ_{MPP} then yields the maximally predictable portfolio ω_{MPP}.

It is important to realize that the probability distribution of the maximum R^2 is not the same as the distribution of the individual maximands. Lo and MacKinlay (1997) run a Monte Carlo experiment to compute the critical values for the maximal R^2 under the null hypothesis of no predictability. They run their experiment for different number of assets (n), factors (p) and observations. To give an example of their results, for five assets, six conditional factors and 47 observations, the appropriate 5% critical value for the maximal R^2 is 45.2%, whereas without maximization, the appropriate 5% critical value is only 22.4%. These results imply the need to interpret the maximum R^2 with caution. A high value of the maximum R^2 may be a symptom of data-mining and not genuine predictability.

The vector of factors $\mathbf{F}_{t-1} = \left[F_{1t}\ F_{2t}\ \cdots\ F_{pt}\right]^T$ may include various economic variables such as dividend yields, term-structure variables and leading economic indicators believed to influence stock returns. The selection of factors is usually motivated by theoretical considerations, but intuition and the results of previous research also play an important role. The literature on the cross-section of expected returns is extremely voluminous and an interested reader is urged to consult Lo and MacKinlay (1997), Fama (1991) and the references therein.

Before we turn to empirical applications, two econometric issues regarding the model in equation (4.3) are worth noting. The selection process of factors almost always involves some sort of data mining, which biases the usual statistical inference based on the

[6] For proof of this result, see Lo and MacKinlay (1997) and the references therein.

t-statistics in regression (4.3). Moreover, as shown in Ferson et al. (2003), if the true (but unobserved) expected return is highly persistent, the data mining problem interacts with spurious regression bias even though realized stock returns are not very persistent themselves[7]. Consequently, Ferson et al. (2003) find that many of the regressions in financial economics literature may be indeed spurious. Note that in our case, the problem of spurious regression may make it impossible to evaluate the economic significance of stock return predictability. If the regression in (4.3) is spurious due to highly persistent expected return, then the model has no forecast power, and hence any dynamic trading strategy based on the model has no incremental value. We (falsely) conclude that the predictability of stock returns is economically insignificant, although the true expected return is highly autocorrelated and thus forecastable! Since it is difficult to correctly adjust the critical values for spurious regression bias in practice, we use standard critical values in the subsequent empirical application, but bare in mind that spurious regression bias may be present when interpreting the results.

4.3 Measures of Market Timing and Investment Performance

In the last two sections we presented the methodology for maximizing predictability of asset returns. We now develop a simple trading rule based on the PC1 or MPP portfolios and present three measures of market-timing[8] aimed at determining whether the predictability in stock returns is genuine and economically significant.

Let $z^{\kappa}_{t+1} \equiv E\left[r^{\kappa}_{t+1} \mid \Phi_t\right], \kappa = PC1, MPP$, denote the expected return on an optimal portfolio κ for the period $t+1$ on the basis of the information set Φ_t available at time t, and let r_{t+1} denote the rate of return on a risk-free asset for the period $t+1$, whose value is know at time t. Our simple asset-allocation strategy takes the following form[9]:

$$\begin{aligned} \theta_{t+1} &= 1 \text{ if } z^{\kappa}_{t+1} > r_{t+1}, \\ \theta_{t+1} &= 0 \text{ otherwise,} \end{aligned} \tag{4.5}$$

where θ_{t+1} is the fraction of assets invested in the portfolio κ during the period $t+1$. Hence, we are invested all in the optimal stock portfolio if the expected return on this portfolio exceeds the return on a risk-free asset, and all in the risk-free asset if the opposite holds. Let x_{t+1} denote the realized return on this trading strategy in period $t+1$. Clearly, x_{t+1} is given by

$$x_{t+1} = \theta_{t+1} r^{\kappa}_{t+1} + (1 - \theta_{t+1}) r_t.$$

[7]We can write the rate of return on a stock as composed of two parts: $r_t = E\left[r_t \mid \Phi_{t-1}\right] + \varepsilon_t$, where $E\left[r_t \mid \Phi_{t-1}\right]$ denotes expected return for period t on the basis of the information set Φ_{t-1} and ε_t is a zero-mean noise component. Since ε_t usually accounts for a substantial portion of the variation in r_t, the persistence of r_t may be small although $E\left[r_t \mid \Phi_{t-1}\right]$ is highly autocorrelated.

[8]We use the term 'market timing' although we are not timing the whole market portfolio but rather our optimal portfolio. No confusion should arise.

[9]This simple trading strategy is fairly common in the literature. See, for example, Merton (1981), Henriksson and Merton (1981), Breen, Glosten and Jagannathan (1989) and Lo and MacKinlay (1997).

We assume that the equity market is sufficiently large and all investors behave as price-takers, i.e. no single investor alone can influence either the stock market prices or the risk-free rate of return in any period.

There are various approaches to evaluating performance of an active asset trading strategy. The simplest and probably most frequently used is the Sharpe ratio (Sharpe, 1966) defined as

$$S_A \equiv \frac{\bar{r}_A - \bar{r}}{\sigma_A},$$

where $\bar{r}_A$ is the average rate of return on an actively managed portfolio A (in our case $\bar{r}_A = \bar{x}$), $\bar{r}$ is the average risk-free rate of return and σ_A is the standard deviation of r_A. Higher value of S_A implies better performance. In a similar way, Treynor (1965) proposes to asses a portfolio's performance using the following ratio:

$$T_A \equiv \frac{\bar{r}_A - \bar{r}}{\beta_A},$$

where β_A is the portfolio's beta obtained from the usual CAPM regression

$$(r_{At} - r_t) = \alpha + \beta_A(r_{Mt} - r_t) + \varepsilon_{At}. \tag{4.6}$$

Again, higher value of T_A implies better performance. The only difference between the Sharpe and Treynor ratios is that the former uses total portfolio risk (σ_A) to standardize the excess return ($\bar{r}_A - \bar{r}$), whereas the latter uses systematic risk only as measured by the beta coefficient. A common drawback of evaluating performance by S_A and/or T_A is the absence of a statistical theory that would allow for testing hypotheses about the difference between two Sharpe or Treynor ratios: although $S_A > S_B$ ($T_A > T_B$) for some portfolios A and B, we do not know if the difference is statistically significant. More sophisticated methods for performance measurement have been therefore suggested in the literature.

4.3.1 The Henriksson-Merton Approach

Henriksson and Merton (1981) propose two measures to evaluate the performance of the simple trading strategy described above. The first measure is a non-parametric one whereas the second is parametric. We start with the former.

Let p_1 denote the probability of a correct forecast in an "down" market and let p_2 denote the probability of a correct forecast in an "up" market. Formally,

$$\begin{aligned} p_1 &= \text{Prob}\left[\theta_t = 0 \mid r_t^\kappa \leq r_t\right], \\ p_2 &= \text{Prob}\left[\theta_t = 1 \mid r_t^\kappa > r_t\right]. \end{aligned}$$

It is shown in Merton (1981) that $p_1 + p_2$ is a sufficient statistic for assessing of forecasting skills. The forecast has no value if $p_1 + p_2 = 1$. A sufficient condition for the forecast to have a positive value is that $p_1 + p_2 > 1$. To test the null hypothesis of no predictability, i.e. $H_0 : p_1 + p_2 = 1$, against the alternative that $p_1 + p_2 > 1$, it is unnecessary to estimate either of the conditional probabilities[10]. Define the following variables: $N_1 \equiv$ number of

[10] See Merton and Henriksson (1981, pp. 517-520) for details.

observations where $r_t^\kappa \leq r_t$, $N_2 \equiv$ number of observations where $r_t^\kappa > r_t$, $N \equiv N_1 + N_2 =$ total number of observations, $n_1 \equiv$ number of successful predictions, given $r_t^\kappa \leq r_t$, $n_2 \equiv$ number of unsuccessful predictions, given $r_t^\kappa > r_t$, $n \equiv n_1 + n_2 =$ total number of predictions that $r_t^\kappa \leq r_t$. Under H_0, n_1 has a hypergeometric distribution that can be asymptotically approximated by normal distribution:

$$n_1 \overset{a}{\sim} N\left(\frac{nN_1}{N}, \frac{n_1 N_1 N_2 (N-n)}{N^2(N-1)}\right)$$

Thus $H_0 : p_1 + p_2 = 1$ can be tested by referring n_1 to the critical values of normal distribution. A disadvantage of the nonparametric approach to testing market timing skills is that it requires the knowledge of the time-series of forecasts $\{\theta_t\}_{t+1}^T$, which need not be publicly observable.

A parametric alternative for testing market timing skills proposed by Henriksson and Merton (1981) allows for analyzing the incremental performance from macroforecasting and microforecasting separately. Macroforecasting is forecasting the market as a whole, i.e. performing market timing. Microforecasting corresponds to forecasting the returns on individual securities and investing into the securities with positive expected excess returns. The test is based on the following regression model:

$$x_t - r_t = \alpha + \beta_1 (r_{Mt} - r_t) + \beta_2 y_t + \varepsilon_t, \tag{4.7}$$

where r_{Mt} denotes the return on the market portfolio in period t, $y_t \equiv \max[0, r_t - r_{Mt}]$, and ε_t is a random shock satisfying the usual linear regression assumptions. The intuition behind equation (4.7) is simple. If the market-timer correctly predicts a "down" market, the return he achieves for period $t+1$ given by (4.7) is

$$x_t - r_t = \alpha + (\beta_1 - \beta_2)(r_{Mt} - r_t) + \varepsilon_t, \tag{4.8}$$

whereas if he correctly predicts an "up" market he obtains

$$x_t - r_t = \alpha + \beta_1 (r_{Mt} - r_t) + \varepsilon_t. \tag{4.9}$$

Clearly, if macroforecasting has any value (i.e. if the market-timer is successful in forecasting the market as a whole), the coefficient β_2 in (4.7) must be positive and statistically significant. Indeed, if the market timer has perfect foresight, then $\beta_1 = \beta_2$ and from (4.8) the market timer always obtains $\alpha + r_t$ in a "down" market, up to an additive noise term ε_t . Hence he is perfectly immunized against down-side risk. If, on the other hand, the market timer does not produce any value-added, the coefficient β_2 is insignificant in (4.7) and the return is given by (4.9). But this return can be achieved by a passive buy-and-hold strategy of investing into a portfolio with beta equal to β_1, and thus there is no gain from active market-timing. The value to microforecasting is measured by the coefficient α. If the estimated α is positive and statistically significant then the manager achieves higher return than what is justified by the systematic risk as measured by the beta coefficient. He therefore exhibits superior stock selection skills.

To quantify the value of market-timing in dollar terms, Merton (1981) observes that y_t in regression (4.7) is the return on a one-period put option on the market portfolio with a current value of one dollar and a strike price equal to one plus the risk-free rate. Let η_1 denote the target beta of the market timer's portfolio when he predicts a "down" market and let η_2 be the target beta when the manager predicts an "up" market. Merton (1981) shows, that up to an additive noise term the returns generated by the market-timing strategy are the same as those that would be generated by a protective put option investment strategy where for each dollar invested, the fraction $[p_2\eta_2 + (1 - p_2)\eta_1]$ dollars is invested in the market portfolio and $(p_1 + p_2 - 1)(\eta_2 - \eta_1)$ put options (per dollar invested) with the return y_t are purchased. The remaining balance is invested into the risk-free asset. The value of market-timing (per dollar invested) comes from the fact, that the $(p_1 + p_2 - 1)(\eta_2 - \eta_1)$ put options are purchased at no cost by investing into the actively managed portfolio. In equilibrium, the value of market-timing must be the same as the value of the put options purchased, otherwise arbitrage opportunities would exist. In case of our asset-allocation rule, $\eta_2 = \beta_\kappa$ and $\eta_1 = 0$, since we are invested all in portfolio κ when we anticipate an "up" market and all in the risk-free asset if we expect a "down" market (the beta of a risk-free asset is zero by definition). If f denotes the price of a one-period put option on the market portfolio with a current price of one dollar and a strike price equal to one plus the risk-free rate, then our trading strategy is worth $(p_1 + p_2 - 1)\beta_\kappa f$ per period per dollar invested. When we use monthly data, this implies an annual management fee of $(100)(12)(p_1 + p_2 - 1)\beta_\kappa f$ percent. Note that market-timing has no value if $p_1 + p_2 = 1$, which is an identical condition to that in the nonparametric test.

To compute the value of the market-timing strategy it is unnecessary to estimate either β_κ or $(p_1 + p_2 - 1)$. Henriksson and Merton (1981) show that plim $\hat{\beta}_2 = (p_1 + p_2 - 1)\beta_\kappa$, where $\hat{\beta}_2$ is an ordinary least squares estimate of β_2 from regression (4.7). When the sample is sufficiently large, the parametric alternative does not require the knowledge of $\{\theta_t\}_{t+1}^T$ and $(p_1 + p_2 - 1)\beta_\kappa$ can be estimated using OLS. The value of the put option can be calculated using the well-known Black-Scholes formula[11].

4.3.2 The Break-Even Transaction Costs

A direct measure of the economic significance of stock return predictability are the break-even transaction costs equating the total return on an active market-timing trading strategy with the total return on a passive investment. Following Lo and MacKinlay (1997) we define

[11] Black and Scholes (1973). The value of the put is given by: $f_t = N(-d_2) - N(-d_1)$, where

$$d_1 \equiv \frac{\left(r_t + \frac{\sigma_m^2}{2}\right)\frac{1}{12}}{\sigma_m\left(\sqrt{\frac{1}{12}}\right.}, \quad d_2 \equiv d_1 - \sigma_m\sqrt{\frac{1}{12}},$$

$N(\bullet)$ denotes cumulative standard normal distribution and σ_m^2 is the variance of the market portfolio. Note that the value of market timing can be alternatively expressed in terms of call options, since the put-call parity implies that the put and call prices are equal in this particular case.

the end-of-period value of a dollar investment over the entire period as

$$W_T^P \equiv \prod_{t=1}^{T} (1 + r_t^\kappa),$$
$$W_T^A \equiv \prod_{t=1}^{T} \left[\theta_t(1 + r_t^\kappa) + (1 - \theta_t)(1 + r_t)\right],$$

where A, P stand for active and passive, respectively. If the active strategy requires k switches into or out of the portfolio κ over the entire investment period, then the one-way break-even transaction costs $100 \times c$ are a solution to the equation

$$W_T^P = W_T^A \times (1 - c)^k.$$

Thus

$$c = 1 - \left(\frac{W_T^P}{W_T^A}\right)^{1/k}. \tag{4.10}$$

Comparing the implied transaction costs $100 \times c$ with the real-world transaction costs provides a straightforward measure of the economic significance of stock return predictability.

4.4 Empirical Application

4.4.1 Czech Stocks

The Czech stock market index, PX-50, was found significantly autocorrelated in the period from 1996:1 to 2002:12 (see Section 1.6). Since the index contains only about 18 stocks, both liquid and infrequently traded, it is difficult to asses whether the estimated serial correlation is spurious or not[12]. Nevertheless, we can evaluate the predictability of Czech stocks by focusing solely on the most liquid ones, i.e. those trading on a daily basis. In particular, we use the time-series of weekly returns on the following stocks trading on the Prague Stock Exchange: CEZ (CEZ), Ceske Radiokomunikace (CR), Cesky Telecom (CT), Komercni banka (KB), and Phillip Morris CR (PM). The time period under study is the same as in Section 1.6, i.e. 1996:1 - 2002:12. The data were downloaded from Bloomberg.

We use the principal component analysis procedure presented in Section 4.1 to evaluate the economic significance of the Czech stock return predictability. We do not apply the MPP analysis based on a multifactor model with various economic variables as factors. This is due to the empirical findings of Hanousek and Filer (2000) that neither lagged nor conteporaneous macroeconomic variables affect Czech stock returns in the period 1993 through mid-1999. As Hanousek and Filer (2000, pp. 629) write, "the Czech stock market appears to have become increasingly divorced from reality". We are therefore left with

[12]Recall that the moments of observed returns on portfolios of infrequently traded stocks derived in Chapter 3 hold only asymptotically, e.g. as the number of securities in the portfolio increases without bound. Clearly, the analysis does not apply to a portfolio of only 18 stocks.

time-series techniques only to measure the predictability of Czech stocks.

Table 4.1. *First Principal Component Weights*

	96:1-99:12	97:1-00:12	98:1-01:12	99:1-02:12
CEZ	0.226	0.236	0.223	0.176
CR	0.131	0.099	0.122	0.228
CT	0.156	0.160	0.172	0.244
KB	0.404	0.446	0.431	0.311
PM	0.083	0.059	0.052	0.041
Var. Explained	56.03%	57.62%	54.66%	46.16%

Before we turn to the dynamic asset allocation rule given in (4.5), we make some preliminary checks as to whether the first principal component portfolio $\boldsymbol{\omega}_{PC1}$ is predictable at all in various subsamples. It may well be that the autocorrelation in weekly returns on the PX-50 index is spurious, induced by infrequent trading of the less liquid stocks contained in the index. In such a case, it would make little sense to run a dynamic strategy using the liquid stocks, since these can be unforecastable.

Table 4.1 reports the first principal component portfolio weights in four overlapping subsamples, along with the proportion of the variation in the stock returns explained by the first principal component. In Table 4.2 are presented the autocorrelation functions of $\boldsymbol{\omega}_{PC1}$ in the same four subsamples. The null hypothesis of no autocorrelation up to the lag given in the parentheses of $\hat{\rho}(\bullet)$ is tested using the Ljung-Box Q-statistic[13]. The time-series predictability of the first principal component portfolio is nil in the first three subsamples. On the contrary, there is significant positive first-order autocorrelation in the $\boldsymbol{\omega}_{PC1}$ in the last subsample. The R^2 corresponding to the regression in (4.2) is, however, only about 6%.

Table 4.2. *Autocorrelation Function of* $\boldsymbol{\omega}_{PC1}$

	96:1-99:12	97:1-00:12	98:1-01:12	99:1-02:12
$\hat{\rho}(1)$	-0.020	0.044	0.077	0.248**
$\hat{\rho}(3)$	0.069	0.030	0.054	-0.010**
$\hat{\rho}(5)$	0.117	0.088	0.076	0.001*
$\hat{\rho}(7)$	-0.109	-0.105	-0.109	0.012

*,** significant at 5% and 1% level, respectively

The main result from Table 4.2 is the impossibility of attaining abnormal return by running a dynamic trading strategy on the first principal component portfolio in the period 1996 to 2002. To evaluate the predictability of the returns on $\boldsymbol{\omega}_{PC1}$, it is necessary to compare the performance of a trading rule based on the out-of-sample forecasts of r_t^{PC1} with an unmanaged portfolio over a sufficiently long period. But from Table 4.2 follows that the r_t^{PC1} is unpredictable from the time-series of historical returns in the subsamples 96:1-99:12, 97:1-00:12, and 98:1-01:12. In these time periods, the best one-step-ahead forecast of the return on $\boldsymbol{\omega}_{PC1}$ is therefore the unconditional mean of r_t^{PC1}. As a result, there can be no gain from market-timing, and the trading strategy has no value. Of course, the

[13] Ljung and Box (1979).

autocorrelation of the returns on $\boldsymbol{\omega}_{PC1}$ in the subsample 99:1-02:12 can be used to forecast returns in year 2003, but our sample does not contain the data on stock returns in 2003 to evaluate the economic significance of predictability in this time period.

4.4.2 Polish Stocks

Since Hanousek and Filer (2000) find that several lagged macroeconomic variables significantly forecast Polish equity market returns we implement the dynamic trading strategy based on the maximally predictable portfolio to evaluate the economic significance of Polish stock return predictability. Using monthly data from 1997:1 to 2002:12, we consider 39 most liquid Polish stocks trading on the Main Market of the Warsaw Stock Exchange. We group these stocks by sectors into the following five equal-weighted portfolios:

Services (SE) - media, telecoms and IT,

Finance (FI) - banks and insurance companies,

Heavy Industry (HI) - chemical industry, machinery, wood and wooden products,

Light Industry (LI) - food and other light industry,

Construction (CO) - construction and construction materials.

To develop a suitable forecasting model for the returns of these equal-weighted sector portfolios, we draw on previous empirical results. There is substantial literature documenting the predictive power of various conditional factors on both developed and emerging markets[14]. For our analysis, we choose the following conditional factors:

IP_t - growth rate of industrial production in month t,

IF_t - inflation rate defined as a percentage change in the producer price index (PPI) in month t,

GD_t - percentage change in the budged deficit in month t,

TD_t - percentage change in the trade deficit in month t,

$\mathrm{W1M}_t$ - one-month Warsaw Interbank Offer Rate at the end of month t,

$\mathrm{M3}_t$ - growth rate of M3 in month t.

Our multifactor forecasting model is thus given by

$$\mathbf{r}_t = \alpha + \beta_1 \mathrm{IP}_{t-1} + \beta_2 \mathrm{IF}_{t-1} + \beta_3 \mathrm{GD}_{t-1} + \beta_4 \mathrm{TD}_{t-1} + \beta_5 \mathrm{W1M}_{t-1} + \beta_6 \mathrm{M3}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (4.11)$$

where $\mathbf{r}_t$ denotes the (5×1) vector of returns on the equal-weighted sector portfolios. The time-series of stock prices and macroeconomic variables were downloaded from Bloomberg.

[14]See, for example, Lo and MacKinlay (1997), Hanousek and Filer (2000), and the references therein.

To forecast returns, we first estimate (4.11) and compute the maximally predictable portfolio weights ω_{MPP} using the first 36 months of our sample, from 1997:1 to 1999:12. The one-month-ahead forecast is then generated month by month beginning in 2000:1 and ending in 2002:12 using a rolling procedure, where the earliest observation in the sample is dropped for each new added, keeping the rolling sample size constant at 36 months. The coefficients in (4.11) and the MPP weights are therefore updated monthly.

Table 4.3 reports ordinary least squares (OLS) estimates of the multifactor model (4.11) for the five asset groups, using the first 36-month sample from 1997:1 to 1999:12. Heteroskedasticity-consistent standard errors are used to asses the statistical significance of the estimates.

Table 4.3. OLS *estimates for sector-grouped portfolio returns*

	const.	IP	IF	GD	TD	W1M	M3	R^2
Services	-0.039	-0.473	4.971	-0.011	0.116	-2.621	-2.561	0.118
Finance	-0.001	-0.447	3.011	-0.123	0.086*	-1.704	0.403	0.140
Light Ind.	-0.003	-0.515**	3.632**	0.025	0.062*	-1.282	1.218**	0.187
Heavy Ind.	0.041	-0.387	5.412	0.145	0.002	-6.032	1.218	0.173
Construction	0.005	-0.531	4.147	0.043	0.049	-2.726	0.499	0.137

*,**,***significant at 10%, 5% and 1% level, respectively

The estimated coefficients for the growth rate of industrial production are uniformly negative across the asset groups, which is a rather counterintuitive result. We would expect an increase in industrial production to have a positive impact on the subsequent stock returns. The estimated coefficients for inflation are all positive as anticipated. If we assume that expected inflation for the next month is roughly equal to current inflation, then higher inflation in this month implies higher nominal return on stock in the next month. The estimated impact of the budget deficit on future stock returns is ambiguous. There is, however, no sound economic theory indicating any link between those two variables. We included GD_t to the model because it was found significant in Hanousek and Filer (2000). The negative, although insignificant, correlation between stock returns and the nominal interest rate (WIBOR) is in line with the empirical findings from developed stock markets[15]. Finally, the positive association between the lagged growth rate of money (M3) and stock returns is quite intuitive, since higher growth rate of money implies higher future inflation and thereby nominal returns.

Table 4.4a. MPP *weights*

Services	0.856
Finance	1.118
Light Industry	-4.456
Heavy Industry	2.914
Construction	0.567

Having estimated the multifactor model in Table 4.3, the maximally predictable portfolio can be readily constructed. We assume that unlimited short sales are possible, which

[15]See Breen et al. (1989) and the references therein.

allows us to maximize predictability without imposing any constraint on the MPP portfolio weights. The MPP portfolio ω_{MPP} is the properly normalized eigenvector corresponding to the largest eigenvalue of the matrix $\hat{\mathbf{B}} \equiv \hat{\mathbf{\Gamma}}_r^{-1}\hat{\Gamma}_F$. The R^2 from the regression of the MPP returns on the conditional factors is then the maximum R^2. Table 4.4a reports the MPP weights and in Table 4.4b are presented the OLS estimates of the multifactor model in (4.11) for the MPP returns. Heteroskedasticity-consistent standard errors are used to asses the statistical significance of the estimates.

Table 4.4b. OLS *estimates for the* MPP *portfolio*

	const.	IP	IF	GD	TD	W1M	M3	R^2
r_t^{MPP}	-0.283	-0.998	-3.798	-0.670**	0.372***	17.416**	0.451	0.379

*,**,***significant at 10%, 5% and 1% level, respectively

Not surprisingly, the maximal R^2 is larger than the individual R^2's. Although we do not have in hand the appropriate critical values for the maximum R^2 under the null hypothesis of no predictability, given five asset groups, six conditional factors and 36 observation, we can use as a first approximation the critical value tabulated by Lo and MacKinlay (1997) for the same number of asset and factors, but 47 observation. The 5% critical values is equal to 45.2%, which is higher than 37.9% in our model and thus the null hypothesis of no predictability cannot be rejected in the period 1997:1 to 1999:12.

We now roll the MPP procedure over the entire out-of-sample period from 2000:1 to 2002:12. We do not report the updated coefficient estimates of (4.11) and the MPP weights for the sake of brewity[16]. At the end of each month, the MPP portfolio return for the next month is forecasted and compared with the one-month risk-free rate, which we approximate by the one-month Warsaw Interbank Offer Rate. When the expected return on the MPP exceeds that of the risk-free asset, we invest all funds into the MPP. If the opposite holds, we are invested all in the risk-free asset. As in Section 4.3, we denote by x_t the realized rate of return on this active trading strategy in month t, r_t^{MPP} denotes the rate of return on the MPP in month t, and r_t denotes the one-month WIBOR rate for month t. To evaluate the performance of out active asset allocation strategy, we also follow two other strategies: (i) investing all funds into the MPP in each month regardless of its expected return, (ii) investing all funds into the risk-free asset in each month. Table 4.5 summarizes the mean monthly return and monthly volatility for the returns on the corresponding trading strategies and the terminal value of 1000 Polish Zloty in 2002:12 invested into the corresponding strategy in 2000:1.

Table 4.5. *Performance of trading strategies*

	ACTIVE	MPP	RF
Mean Return	3.71%	1.55%	1.17%
Volatility	43.18%	49.60%	0.38%
Terminal Value	3804.59	1749.13	1525.25

From Table 4.5 follows that our active trading strategy achieved superior performance in the period 2000:1 to 2002:12. Compared to the strategy of investing every month into

[16]The MPP weights are available from the author upon request.

the MPP, our strategy yielded higher mean return with lower volatility. Note, however, that the volatility of both ACTIVE and MPP strategies is extremely high, and thus these investments are very risky. To asses the predictability of the MPP, we calculate the one-way break-even transaction costs according to the formula in (4.10). The ACTIVE strategy required 10 switches into and out of the MPP, hence the one-way break-even transaction costs are given by

$$c = \left[1 - \left(\frac{W_T^P}{W_T^A}\right)^{1/k}\right] \times 100 = \left[1 - \left(\frac{1749.13}{3804.59}\right)^{1/10}\right] \times 100 \approx 7.48\%.$$

Clearly, the break-even transaction costs by far exceed those incurred in reality, which implies economically significant predictability of Polish stocks.

Finally, we asses the performance of our active asset allocation rule using the Henriksson and Merton (1981) approach. Following Lo and MacKinlay (1997) we assume that the MPP portfolio can be considered the 'market'. We apply the nonparametric test because our sample of moderate size only. Out of the total of 36 observations ($N = 36$), there are $N_1 = 17$ "down" markets, $N_2 = 19$ "up" markets, $n_1 = 11$ correct "down" forecast, and $n_2 = 8$ incorrect "down" forecasts. The estimated probability of a correct forecast in a "down" market is thus $p_1 = 0.647$, the probability of a correct forecast in an "up" market is $p_2 = 0.579$, and $p_1 + p_2 = 1.226$. Under the null hypothesis of no predictability ($p_1 + p_2 = 1$), n_1 is asymptotically normally distributed with mean 9 and variance 1.33. The null hypothesis can be tested by referring $\frac{n_1 - 9}{1.154} \approx 1.733$ to the critical values of the standard normal distribution[17]. The 5% critical value for a one-sided test is 1.65. The null hypothesis can be therefore rejected at the 5% level. Given the monthly volatility of the MPP portfolio of 49.6%, the average monthly risk-free rate of 1.17% and the fact that $\beta_\kappa = 1$ (the target beta of our portfolio in the "up" market is the beta of the MPP, which is equal to unity since we assume that MPP is the "market"), the annual management fee is estimated to be (100)(12)(0.226)(0.1974) = 60.01% of the value of the portfolio managed. The source of this extreme management fee is the relatively high value of the protective put option, which is in turn caused by the high volatility of the MPP portfolio.

To summarize the MPP analysis implemented above, the predictability of Polish stock returns is economically significant. The one-way break-even transaction costs of 7.48% imply that it was possible to earn above-average rate of return by means of out market-timing strategy in the period 2000:1 to 2002:12.

4.5 Concluding Remarks

The aim of this chapter has been to evaluate the economic significance of Czech and Polish stock return predictability. We applied the principal component analysis to the most liquid Czech stock and found no significant predictability in their weekly returns for the period from 1997:1 to 2001:1. This result contradicts our finding of significant serial correlation in weekly returns on the Czech value-weighted stock index, PX-50. It appears that the auto-correlation of weekly returns on the PX-50 index is in part spurious, induced by infrequent

[17] It is shown in Henriksson and Merton (1981), that the asymptotic approximation works well even in moderate samples, provided that $N_1 \approx N_2$, which is our case.

trading of the less liquid stock included in the index. Although we cannot quantify the effect of nontrading directly, due to the low number of infrequently traded stocks contained in the PX-50 index, the fact that the most liquid Czech stocks are unpredictable implies that the autocorrelation of returns on PX-50 is almost entirely spurious.

Contrary to the Czech stocks, the predictability of Polish stocks was found economically significant. We followed an out-of-sample rolling trading strategy based on the maximally predictable portfolio using monthly returns on five sector-sorted portfolios of the 39 most liquid Polish stocks. For the out-of-sample period from 2000:1 through 2002:12, our active asset allocation rule achieved higher average rate of return than that of a passive trading strategy. The implied one-way break-even transaction costs of 7.48% provide evidence of economically significant stock return predictability.

Conclusion

The aim of this thesis has been to investigate the predictability of Central-European common stock returns. We have applied modern econometric techniques to the Czech, Hungarian and Polish stock prices and/or returns to learn more about their time-series and distributional properties. We studied the stock return predictability in both univariate and multivariate context, focusing on the time-period from 1996:1 through 2002:12. Besides empirical results, we have also provided a brief overview of the underlying financial theory, and pointed out the difficulties that may arise when interpreting the results. In particular, we discussed the problem of spurious autocorrelation of stock returns induced by infrequent trading, and the problem of determining the economic significance of stock return predictability. In the next few paragraphs, we summarize the empirical findings and suggest directions for future research.

The Random Walk Hypothesis did not hold on the Central-European stock markets in the period 1996:1 to 2002:12. Both the mean and the variance of weekly stock returns were predictable from the time-series of historical returns. We found linear dependencies in the time-series of Czech and Hungarian stock index returns and estimated simple ARIMA-type models to learn more about the pattern of serial correlation in the stock index returns. The coefficients of determination were extremely small for both models (0.034 and 0.021 for the Hungarian and Czech stock indices, respectively) implying little economic significance of return predictability. A similar result was obtained for the Polish stock market index, where the dependencies in returns were found nonlinear (GARCH-M). The behavior of return volatility was modelled by the simple GARCH(1,1) model under the assumption that the innovations are conditionally Student t-distributed with v degrees of freedom, where v was treated as an unknown parameter to be estimated. The estimates of v along with the results of likelihood ratio tests indicate that both conditional heteroskedasticity and non-normality account for the 'fat tails' in the unconditional distributions of the stock index returns. The GARCH(1,1) model performs well in explaining the time-varying volatility of the Hungarian and Polish indices but fails in case of the Czech stock market index.

We next investigated the stock return predictability in a multivariate context. Applying the Johansen (1988,1991) multivariate cointegration test we showed that in the period 1998:1-2002:12 the Czech, Hungarian, Polish and German stock market indices were cointegrated when the index prices were expressed in local currencies. This result implies joint market inefficiency and the impossibility to diversify internationally in the Central-European region for investors who hedge their stock positions against exchange rate risk. On the contrary, no cointegration was found for the stock index prices expressed in terms of Euro. Thus investors from the European Monetary Union member states not hedging their foreign exchange risk exposures can obtain benefits from international diversification of their portfolios. Our results also indicate significant cross-country predictability among the four markets under study: the forecast of future returns on one market can be substantially improved by including past returns from other markets as well.

Since all emerging markets suffer from lower liquidity, we dedicated an entire chapter to studying the effect of infrequent trading on the time-series properties of individual stock an portfolio returns. We generalized the econometric model of nonsynchronous trading developed by Lo and MacKinlay (1990) by allowing the common factor generating the 'virtual' returns follow a stationary first-order autoregressive process. We derived the moments of individual stock a portfolio returns and investigated the behavior and pattern of autocorrelation of observed stock returns induced by the interaction of the autoregressive parameter of the process for the common factor and the probability of nontrading. Our results imply that the sign and pattern of serial correlation of infrequently traded stocks (and portfolios of these stocks) can be fairly general in our specification of the model. Also, our specification allows for a direct decomposition of the estimated first-order autocorrelation coefficient of portfolio returns into the (spurious) part induced by nontrading and the (real) part inherent to the portfolio returns. The main conclusion the model of nonsynchronous trading is that the empirically found predictability of stock returns may be in part spurious due to infrequent trading. As a result, it cannot be exploited to earn abnormal returns.

Unfortunately, the model could not be applied to study the effect of nontrading on the time-series properties of Central-European stock portfolios directly. The number of illiquid stocks traded on these markets is too low for the estimated moments of observed returns to reasonably approximate their asymptotic counterparts derived in our model. Nevertheless, some informal arguments are still possible to make. We showed that the weekly returns on the Czech stock market index PX-50 are significantly autocorrelated, whereas the returns on the most liquid Czech stocks (i.e. stock trading on a daily basis) approximate white noise. Since the PX-50 index also includes securities with extremely high probabilities of nontrading, it follows that much of the serial correlation in the returns on PX-50 is spurious due to the effect on nontrading. Of course, a more rigorous analysis is needed to gain a better insight into this matter but this is likely to remain impossible within the framework of the Lo and MacKinlay (1990) model of nonsynchronous trading in the near future, given the evolution and prospects of the Central-European stock markets.

Finally, in the last chapter of this thesis we evaluated the economic significance of the predictability of Polish stock returns. We argued that stock return predictability is economically meaningful if and only if it can be exploited to earn statistically significant abnormal return. To test this hypothesis, we focused on 39 most liquid Polish stocks traded on the Warsaw Stock Exchange and considered the out-of-sample period from 2000:1 to 2002:12. For sector-grouped portfolios, we estimated a multifactor forecasting model with macroeconomic and term-structure variables as factors in each month and formed the maximally predictable portfolio (MPP). The expected return on the MPP for the next month was then compared with the risk-free rate and all (virtual) funds were invested into the asset with higher expected return. This trading strategy was rolled over the entire out-of-sample period. Comparing the ex-post return on our active trading strategy with the return on a passively managed portfolio yields a direct measure of economic significance of stock return predictability. The implied break-even transaction costs of 7.48% by far exceed those incurred in practice, providing evidence on the existence of profitable trading strategies based on return predictability.

We have to emphasize once again that significant stock return predictability need not be a symptom of market inefficiency. To test market efficiency rigorously we would have to assume a particular asset pricing model and take into account the (dynamic) risk preferences of investors. But as we argued above, this necessarily leads to the joint hypothesis problem making it virtually impossible to test market efficiency in practice.

The analysis of stock return predictability on Central-European stock markets is far from being complete. We characterized the time-series and distributional properties of stock index returns, while pointing out to the problem of spurious autocorrelation induced by infrequent trading. We also stressed the importance of evaluating economic significance of return predictability and presented an example from Polish stock market. It is the task of future research to asses the predictability of the Czech and Hungarian stock returns in greater detail as well. Also, a formal analysis of the nonsynchronous trading effect is required to learn more about the pattern of autocorrelation of stock returns. Besides time-series predictability, it is also important to study cross-sectional predictability of stock returns. Numerous asset pricing models await to be put to test and can help us better characterize the stock return generating process. To summarize, this thesis is just a starting point to the analysis of the nature and sources of stock return predictability on the Central-European stock markets.

Appendix - Proof of Theorems

Proof of Theorem 4.

The proof of part (i) of the theorem is identical to Lo and MacKinlay (1990) since Λ_t has a zero mean and the fact that Λ_t is an AR(1) process has no effect on the mean of r_{it}^0.

To prove (ii) we first derive the second uncentered moment of r_{it}^0:

$$\begin{aligned}
E\left[\left(r_{it}^0\right)^2\right] &= E\left[\sum_{k=0}^{\infty} X_{it}(k) r_{it-k} \cdot \sum_{l=0}^{\infty} X_{it}(l) r_{it-l}\right], \\
&= \sum_{k=0}^{\infty} E\left[X_{it-k}^2 r_{it-k}^2\right] + 2\sum\sum\nolimits_{k<l} E\left[X_{it-k}X_{it-l}\right] E\left[r_{it-k} r_{it-l}\right], \\
&= \left(\mu_i^2 + \beta_i^2 \frac{\sigma_\eta^2}{1-\alpha^2} + \sigma_{\varepsilon_i}^2\right) \sum_{k=0}^{\infty} (1-p_i) p_i^k \\
&\quad +2\sum_{k=0}^{\infty} \sum_{l=k+1}^{\infty} (1-p_i) p_i^l \left(\mu_i^2 + \theta(k-l)\sigma_{\varepsilon_i}^2\right) \\
&\quad +2\sum_{k=0}^{\infty} \sum_{l=k+1}^{\infty} (1-p_i) p_i^l \left(\beta_i^2 \frac{\theta(k-l)\sigma_\eta^2 + (1-\theta(k-l))\alpha^{|k-l|}\sigma_\eta^2}{1-\alpha^2}\right) \\
&\quad \text{where } \theta(x) \text{ is 1 if } x \text{ is equal to zero and 0 otherwise,} \\
&= \mu_i^2 + \beta_i^2 \frac{\sigma_\eta^2}{1-\alpha^2} + \sigma_{\varepsilon_i}^2 \\
&\quad +2(1-p_i)\left(\frac{\mu_i^2 p_i}{1-p_i} \sum_{k=0}^{\infty} p^k + \beta_i^2 \frac{\sigma_\eta^2}{1-\alpha^2} \frac{\alpha p_i}{1-p_i} \sum_{k=0}^{\infty} (\alpha p)^k\right), \\
&= \mu_i^2 + \beta_i^2 \frac{\sigma_\eta^2}{1-\alpha^2} + \sigma_{\varepsilon_i}^2 + \frac{2\mu_i^2 p_i}{1-p_i} + \frac{2\alpha\beta_i^2\sigma_\eta^2 p_i}{(1-\alpha^2)(1-\alpha p_i)}.
\end{aligned}$$

Now the variance of r_{it}^0 follows directly from the well-known results that $Var\left[x\right] = E\left[x^2\right] - \left(E\left[x\right]\right)^2$.

Similarly, to derive the autocovariance, we first compute $E\left[r_{it}^0 r_{it+n}^0\right]$:

$$\begin{aligned}
E\left[r_{it}^0 r_{it+n}^0\right] &= E\left[\sum_{k=0}^{\infty} X_{it}(k) r_{it-k} \cdot \sum_{l=0}^{\infty} X_{it+n}(l) r_{it-l+n}\right], & (12)\\
&= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} E\left[X_{it}(k)\right] E\left[X_{it+n}(l)\right] E\left[r_{it-k} r_{it-l+n}\right],\\
&= \sum_{k=0}^{\infty}\sum_{l=0}^{n-1}(1-p_i)^2 p_i^{k+l}\left(\mu_i^2 + \beta_i^2 \frac{\alpha^{k+n-l}\sigma_v^2}{1-\alpha^2}\right),\\
&= (1-p_i)^2\mu_i^2 \sum_{k=0}^{\infty}\sum_{l=0}^{n-1} p_i^{k+l},\\
&\quad +(1-p_i)^2 \frac{\beta_i^2\sigma_v^2}{1-\alpha^2}\sum_{k=0}^{\infty}\sum_{l=0}^{n-1}\alpha^{k+n-l} p_i^{k+l},\\
&= \mu_i^2(1-p_i^n) + \frac{\alpha\beta_i^2\sigma_\eta^2\left(1-p_i\right)^2}{\left(1-\alpha^2\right)\left(1-\alpha p_i\right)}\left[\frac{\alpha^n - p_i^n}{\alpha - p_i}\right]. & (13)
\end{aligned}$$

when $\alpha \neq p_i$ and $n > 0$. When $\alpha = p_i$ the last term in the product is not defined since the denominator is equal to zero. It can be, however, continuously defined by the its limit as $\alpha \to p_i$. To see this, define $f : (-1, 1) \to R$ as follows:

$$f(\alpha) \equiv \frac{\alpha^n - p_i^n}{\alpha - p_i}, \tag{14}$$

where $p_i \in \langle 0, 1)$. Now the limit of $f(\alpha)$ as $\alpha \to p_i$ is given by

$$\lim_{\alpha\to p_i} f(\alpha) = \lim_{\alpha\to p_i}\frac{\alpha^n - p_i^n}{\alpha - p_i} = \lim_{\alpha\to p_i} n\alpha^{n-1} = np_i^{n-1} < \infty, \tag{15}$$

where the second equality follows from the L'hospital rule. Substituting into (12) and setting $\alpha = p_i$ throughout the expression produces

$$E\left[r_{it}^0 r_{it+n}^0\right] = \mu_i\mu_j + \left[\frac{n\beta_i^2\sigma_\eta^2}{(1+p_i)^2} - \mu_i^2\right] p_i^n.$$

The formula for autocovariance then follows from the result that $Cov\left[x, y\right] = E\left[xy\right] - E\left[x\right] E\left[y\right].$

Finally, we derive the cross-autocovariance of individual returns. The computation in this case is a bit more involved. As before, we start by computing $E\left[r^0_{it}r^0_{jt+n}\right]$:

$$\begin{aligned}
E\left[r^0_{it}r^0_{jt+n}\right] &= E\left[\sum_{k=0}^{\infty} X_{it}(k)r_{it-k}\cdot\sum_{l=0}^{\infty} X_{it}(l)r_{jt-l+n}\right],\\
&= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} E\left[X_{it}(k)\right]E\left[X_{jt+n}(l)\right]E\left[r_{it-k}r_{jt-l+n}\right],\\
&= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(1-p_i)p_i^k(1-p_j)p_j^l\left[\mu_i\mu_j+\beta_i\beta_j\frac{\theta(k-l+n)\sigma^2_\eta}{1-\alpha^2}\right]\\
&\quad+\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(1-p_i)p_i^k(1-p_j)p_j^l\left[\beta_i\beta_j\frac{\alpha^{|k-l+n|}\sigma^2_\eta}{1-\alpha^2}\right]\\
&\quad-\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(1-p_i)p_i^k(1-p_j)p_j^l\left[\beta_i\beta_j\frac{\theta(k-l+n)\alpha^{|k-l+n|}\sigma^2_\eta}{1-\alpha^2}\right]\\
&= \mu_i\mu_j+\beta_i\beta_j(1-p_i)(1-p_j)\frac{\sigma^2_\eta}{1-\alpha^2}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}p_i^kp_j^l\alpha^{|k-l+n|}.
\end{aligned}$$

Now to remove the absolute value operator, we consider five mutually exclusive cases. First, let $k>l$. Since $k\geq 0, l\geq 0$ and $n\geq 0$ we have $|k-l+n|=k-l+n$. Hence

$$\begin{aligned}
\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}p_i^kp_j^l\alpha^{|k-l+n|} &= \alpha^n\sum_{k=l+1}^{\infty}\sum_{l=0}^{\infty}p_i^kp_j^l\alpha^{k-l},\\
&= \frac{\alpha^{n+1}p_i}{(1-\alpha p_i)(1-p_ip_j)}.
\end{aligned}$$

Second, let $k=l$. Then

$$\begin{aligned}
\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}p_i^kp_j^l\alpha^{|k-l+n|} &= \sum_{k=0}^{\infty}(p_ip_j)^k\alpha^n,\\
&= \frac{\alpha^n}{1-p_ip_j}.
\end{aligned}$$

Third, let $l-n<k<l$. It follows that

$$\begin{aligned}
\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}p_i^kp_j^l\alpha^{|k-l+n|} &= \sum_{k=0}^{\infty}\sum_{l=k+1}^{k+n-1}p_i^kp_j^l\alpha^{k-l},\\
&= \frac{\alpha^np_j-\alpha p_j^n}{\alpha-p_j}, \alpha\neq p_j.
\end{aligned}$$

Fourth, let $k=l-n$. Thus

$$\begin{aligned}
\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}p_i^kp_j^l\alpha^{|k-l+n|} &= \sum_{k=0}^{\infty}p_i^kp_j^{k+n},\\
&= \frac{p_j^n}{1-p_ip_j}.
\end{aligned}$$

Finally, let $k < l - n$. Then

$$\begin{aligned}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p_i^k p_j^l \alpha^{|k-l+n|} &= \sum_{k=0}^{\infty}\sum_{l=k+n}^{\infty} p_i^k p_j^l \alpha^{l-k-n}, \\ &= \frac{\alpha p_j^{n+1}}{(1-\alpha p_j)(1-p_i p_j)}.\end{aligned}$$

Since the five cases were mutually exclusive, we have

$$\begin{aligned}E\left[r_{it}^0 r_{jt+n}^0\right] &= \mu_i\mu_j + \beta_i\beta_j \frac{\sigma_\eta^2}{1-\alpha^2}\left[\frac{(1-p_i)(1-p_j))}{1-p_i p_j}\right. \\ &\times \left[\alpha^n + p_j^n + \frac{\alpha^{n+1}p_i}{1-\alpha p_i} + \frac{\alpha p_j^{n+1}}{1-\alpha p_j} + \frac{\alpha^n p_j - \alpha p_j^n}{\alpha - p_j}\right],\end{aligned}$$

when $\alpha \neq p_j$. If $\alpha = p_j$ the last term in the parentheses is not defined, but it can be defined continuously by the limit as $a \to p_j$. To see this define $g : (-1, 1) \to R$ as follows

$$g(\alpha) \equiv \frac{\alpha^n p_j - \alpha p_j^n}{\alpha - p_j}, \tag{16}$$

where $p_j \in \langle 0, 1)$. Then

$$\lim_{\alpha\to p_j} g(\alpha) = \lim_{\alpha\to p_j} \frac{\alpha^n p_j - \alpha p_j^n}{\alpha - p_j} = \lim_{\alpha\to p_j} n\alpha^{n-1}p_j - p_j^n = (n-1)p_j^n < \infty. \tag{17}$$

where the second equality follows from the L'hospital rule. Substituting this result into $E\left[r_{it}^0 r_{jt+n}^0\right]$ and setting $\alpha = p_j$ produces the result given in the theorem. ■

Proof of Theorem 6.

Deriving the mean of $r_{\kappa t}^0$ is straightforward. Taking expectations of both sides of (3.17) we have

$$\begin{aligned}E\left[r_{\kappa t}^0\right] &= \mu_\kappa + (1-p_\kappa)\beta_\kappa \sum_{k=0}^{\infty} p_\kappa^k E\left[\Lambda_{t-k}\right], \\ &= \mu_\kappa,\end{aligned}$$

since Λ_{t-k} has a zero unconditional mean.

To derive the variance of $r^0_{\kappa t}$ we first compute the second uncentered moment:

$$
\begin{aligned}
E\left[\left(r^0_{\kappa t}\right)^2\right] &= \mu_\kappa^2 + (1-p_\kappa)^2\beta_\kappa^2 E\left[\left(\sum_{k=0}^{\infty} p_\kappa^k \Lambda_{t-k}\right)^2\right], \\
&= \mu_\kappa^2 + (1-p_\kappa)^2\beta_\kappa^2 E\left[\sum_{k=0}^{\infty} p_\kappa^{2k} \Lambda_{t-k}^2\right] + \\
&\quad +2(1-p_\kappa)^2\beta_\kappa^2 E\left[\sum\sum\nolimits_{k<l} p_\kappa^{k+l}\Lambda_{t-k}\Lambda_{t-l}\right], \\
&= \mu_\kappa^2 + (1-p_\kappa)^2\beta_\kappa^2 \sum_{k=0}^{\infty} p_\kappa^{2k} E\left[\Lambda_{t-k}^2\right] + \\
&\quad +2(1-p_\kappa)^2\beta_\kappa^2 \sum\sum\nolimits_{k<l} p_\kappa^{k+l} E\left[\Lambda_{t-k}\Lambda_{t-l}\right], \\
&= \mu_\kappa^2 + (1-p_\kappa)^2\beta_\kappa^2 \sum_{k=0}^{\infty} p_\kappa^{2k} \frac{\sigma_\eta^2}{(1-\alpha^2)} + \\
&\quad +2(1-p_\kappa)^2\beta_\kappa^2 \sum_{k=0}^{\infty}\sum_{l=k+1}^{\infty} p_\kappa^{k+l} \frac{\sigma_\eta^2(1-\theta(k-l))\alpha^{l-k}}{1-\alpha^2},
\end{aligned}
$$

where $\theta(x)$ is 1 if x is equal to zero and 0 otherwise. Thus we have

$$
\begin{aligned}
E\left[\left(r^0_{\kappa t}\right)^2\right] &= \mu_\kappa^2 + \beta_\kappa^2\left(\frac{1-p_\kappa}{1+p_\kappa}\right)\left(\frac{\sigma_\eta^2}{1-\alpha^2}\right) + \\
&\quad +2(1-p_\kappa)^2\beta_\kappa^2 \sum_{k=0}^{\infty} p_\kappa^{2k}\left(\sum_{l=1}^{\infty}(\alpha p_\kappa)^l\right), \\
&= \mu_\kappa^2 + \beta_\kappa^2\left(\frac{1-p_\kappa}{1+p_\kappa}\right)\left(\frac{\sigma_\eta^2}{1-\alpha^2}\right) + \\
&\quad +2\beta_\kappa^2(1-p_\kappa)^2\left(\frac{\alpha p_\kappa}{(1-\alpha p_\kappa)(1-p_\kappa^2)}\right)\left(\frac{\sigma_\eta^2}{1-\alpha^2}\right), \\
&= \mu_\kappa^2 + \beta_\kappa^2\left(\frac{1-p_\kappa}{1+p_\kappa}\right)\left(1+\frac{2\alpha p_\kappa}{1-\alpha p_\kappa}\right)\frac{\sigma_\eta^2}{1+\alpha^2},
\end{aligned}
$$

from which $Var\left[r^0_{\kappa t}\right]$ follows since $Var\left[x\right] = E\left[x^2\right] - \left(E\left[x\right]\right)^2$.

The proofs of the formulas for autocovariance and cross-autocovariance are very similar to those for individual returns and thus we leave them to the reader for the sake of brevity. ■

References

Bera, A. and C. Jarque (1981): Efficient Test for Normality, Heteroskedasticity, and Serial Correlation of Regression Residuals: Monte Carlo Evidence, *Economics Letters*, **7**, pp. 313-318.

Berndt, E., Hall, B., Hall, R. and J. Hausman (1974): Estimation and Inference in Nonlinear Structural Models, *Annals of Economic and Social Measurement*, **3/4**, pp. 653-665.

Black, F. and M. Scholes (1973): The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, **81**, pp. 637-659.

Bollerslev, T. (1986): Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, **31**, pp. 307-327.

Bollerslev, T. (1987): A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, *The Review of Economics and Statistics*, **69**, pp. 498-505.

Bollerslev, T., Chou, R. and K. Kroner (1992): ARCH Modelling in Finance, *Journal of Econometrics*, **52**, pp. 5-59.

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.

Box, G. and G. Jenkins (1984): *Time Series Analysis: Forecasting and Control*, 2nd ed., Holden Day, San Francisco.

Breen, W., Glosten, L.B. and R. Jagannathan (1989): Economic Significance of Predictable Variations in Stock Index Returns, *Journal of Finance*, **44**, pp. 1177-1189.

Brock, W., Dechert, W. and J. Scheinkman (1987): A Test for Independence Based on the Correlation Dimension, Working Paper, University of Wisconsin at Madison, Department of Economics.

Brock, W.A., Hsieh, D.A. and B. LeBaron (1991): *Nonlinear Dynamics, Chaos and Instability*, MIT Press, Cambridge.

Campbell, J., Lo A.W. and A.C. MacKinlay (1997): *The Econometrics of Financial Markets*, Princeton University Press, Princeton.

Chan, K.C., Gup, B.E. and M. Pan (1997): International Stock Market Efficiency and Intergration: A Study of Eighteen Nations, *Journal of Business Finance and Accounting*, **26**, pp. 803-813.

Chen, N., Roll, R. and S.A. Ross (1986): Economic Forces and the Stock Market, *Journal of Business*, **59**, No. 3, pp. 383-403.

Connolly, R.A. (1989): An Examination of the Robustness of the Weekend Effect, *Journal of Financial and Quantitative Analysis*, **24**, No. 2, pp. 133-169.

Conrad, J. and G. Kaul (1988): Time-Variation in Expected Returns, *Journal of Business*, **61**, No. 4, pp. 409-425.

Davidson R. and J.G. MacKinnon (1993): *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.

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

Elton, E.J. and M.J. Gruber (1995): *Modern Portfolio Theory and Investment Analysis*, 5th ed., John Wiley and Sons, New York.

Enders, W. (1995): *Applied Econometric Time Series*, John Wiley and Sons, New York.

Engle, R. (1982): Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflations, *Econometrica*, **50**, pp. 987-1008.

Engle, R., Lilien, D.M. and R.P. Robins (1987): Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model, *Econometrica*, **55**, pp. 391-407.

Fama, E.F. (1965): The Behavior of Stock Market Prices, *Journal of Business*, **38**, pp. 34-105.

Fama, E.F. (1970): Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance*, **XXV**, No. 2, pp. 383-417.

Fama, E.F. (1991): Efficient Capital Markets II, *Journal of Finance*, **26**, No. 5, pp. 1575-1617.

Fama, E.F. and K.R. French (1988): Permanent and Temporary Components of Stock Prices, *Journal of Political Economy*, **96**, pp. 246-273.

Ferson, W.E., Sarkissian, S. and T.T. Simin (2003): Spurious Regression in Financial Economics?, *Journal of Finance* (forthcomming).

Filáček, J., Kapička, M. and M. Vošvrda (1998): Testování hypotézy efektivního trhu na BCPP (in Czech), *Finance a Úvěr*, **48**, pp. 554-566.

Filer, R.K. and J. Hanousek (1996): The Extent of Efficiency in Central-European Equity Markets, Working Paper, Center for Economic Research and Graduate Education, Charles University, Prague, Czech Republic.

Gilmore, C.G. and G.M. McManus (2001): Random-Walk and Efficiency Tests of Central European Equity Markets, European Financial Management Association Conference, Lugano, Switzerland.

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

Greene, W.H. (2000): *Econometric Analysis*, 4th ed., Prentice Hall, New Jersey.

Glosten, L.R., Jagannathan, R. and D. Runkle (1993): Relationship between the Expected Value and the Volatility of the Nominal Excess Return on Stock, *Journal of Finance*, **48**, pp. 1779-1801.

Grossman, S. and J. Stiglitz (1980): On the Impossibility of Informationally Efficient Markets, *American Economic Review*, **70**, pp. 393-408.

Forbes K. and R. Rigobon (1999): No Contagion, only Interdependence: Measuring Stock Market Co-Movements, Working Paper 7267, National Bureau of Economic Research, Cambridge.

Hanousek, J. and R.K. Filer (2000): The Relationship Between the Real Economy and Equity Markets in Central and Eastern Europe, *Economics of Transition*, **8**(3), pp. 623-638.

Hamilton, J.D. (1994): *Time Series Analysis*, Princeton University Press, Princeton.

Haerdle, W. (1990): *Applied Nonparametric Regression*, Cambridge University Press, New York.

Haerdle, W. and L. Simar (2001): *Applied Multivariate Statistical Analysis*, E-book, Humbolt University, Berlin.

Henriksson, R.D. and R.C. Merton (1981): On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills, *Journal of Business*, **54**, pp. 513-533.

Hull, J.C. (2000): *Options, Futures and Other Derivatives*, Prentice Hall, New Jersey.

Johansen, S. (1988): Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, **12**, pp. 231-254.

Johansen, S. (1991): Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, **59**, pp. 1551-80.

Johansen, S. and K. Juselius (1990): Maximum Likelihood Estimation and Inference on Cointegration - with Applications to the Demand for Money, *Oxford Bulletin Of Economics And Statistics*, **52**, pp. 169-210.

Lo, A.W. and A.C. MacKinlay (1988): Stock 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 (1989): The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation, *Journal of Econometrics*, **40**, pp. 203-238.

Lo, A.W. and A.C. MacKinlay (1990): An Econometric Analysis of Nonsynchronous Trading, *Journal of Econometrics*, **45**, pp. 181-211.

Lo, A.W. and A.C. MacKinlay (1997): Maximizing Predictability in the Stock and Bond Markets, *Macroeconomic Dynamics*, **1**, pp. 309-325.

Lo, A.W. and A.C. MacKinlay (1999): *A Non-Random Walk Down Wall Street*, Princeton University Press, Princeton.

Ljung, G.and G. Box (1979): On a Measure of Lack of Fit in Time Series Models, *Biometrika*, **66**, pp. 265-270.

Koch, P.D. and T.W. Koch (1993): Dynamic Relationships among the Daily Levels of National Stock Indexes, In: S.R. Stansell, ed., International Financial Market Integration, Blackwell, pp. 299-328.

Mandelbrot, B. (1963): The Variation of Certain Speculative Prices, *Journal of Business*, **36**, pp. 394-419.

Merton, R.C. (1981): On Market Timing and Investment Performance. I. An Equilibrium Theory of Value for Market Forecasts, *Journal of Business*, **54**, pp. 363-406.

Neubauer, M. (2001): International Portfolio Diversification in the Central European Region, *Prague Economic Papers*, **1**, pp. 74-86.

Phylaktis, K. and F. Ravazzolo (2000): Stock Market Linkages in Emerging Markets: Implications for International Portfolio Diversification, Working Paper, City University Business School, London.

Poterba, J.M. and L.H. Summers (1988): Mean Reversion in Stock Prices: Evidence and Implications, *Journal of Financial Economics*, **22**, No. 1., pp. 27-59.

Roll, R. (1984): A Simple Implicit Measure of the Bid-Ask Spread in an Efficient Market, *Journal of Finance*, **39**, pp. 1127-1139.

Samuelson, P.A. (1965): Proof that Properly Anticipated Prices Fluctuate Randomly, *Industrial Management Review*, **6**, pp. 41-49.

Sharpe, W.A. (1966): Mutual Fund Performance, *Journal of Business*, **39**, pp. 119-138.

Silverman, B.W. (1986): *Density Estimation*, Chapman and Hall, London.

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

Summers, L.H. (1986): Does the Stock Market Rationally Reflect Fundamental Values?, *Journal of Finance*, **41**, No. 3., pp. 591-601.

Treynor, J. (1965): How to Rate Management of Investment Funds, *Harvard Business Review*, **43,** pp. 63-75.

Yang, J., Min, I. and Q. Li (2003): European Stock Market Integration: Does EMU Matter ?, *Journal of Business Finance and Accounting*, **30**, pp. 209-224.