Stock return predictability

- Empirical testing of various financial variables' predictive ability on stock returns

LOUISE AMB DYRDAL
MARTE MØLLER GARTHUS

Master thesis
Cand. Merc. Applied Economics and Finance
Copenhagen Business School 2014

Date of submission: 18. August 2014
Supervisor: Bent Jesper Christensen
Department of Economics and Business, Aarhus University
Number of pages (characters): 120 (205 357)

Marte M. Garthus
Marte Møller Garthus

Louise Amb Dyrdal
Louise Amb Dyrdal

ACKNOWLEDGEMENTS

This master thesis is a final assignment in the MSc. Applied Economics and Finance program at Copenhagen Business School (hereafter called CBS). We have mainly applied knowledge from the Corporate Finance, Capital Market Theory, and Applied Econometrics courses studied during our master education at CBS. The authors of this thesis are highly interested in the subjects introduced in these courses which lead to the motivation behind the choice of master thesis topic. Our desire was to challenge ourselves with this master thesis and gain additional knowledge on this specific topic.

We want to thank our supervisor, Bent Jesper Christensen, who has provided us with guidance, useful suggestions, and detailed answers to our questions throughout our work with this thesis. In addition we are thankful for him taking time to meet us when traveling to Copenhagen.

From writing this master thesis we have gained new and useful information on a topic we knew little about beforehand. Especially econometrics issues have made us rise to the challenge when facing problems along the way. It has been an interesting process writing this thesis and we consider it a worthy ending to our time at CBS.

Marte Møller Garthus

Louise Amb Dyrdal

EXECUTIVE SUMMARY

The existing literature on stock return predictability includes several different studies that over the years have gathered a significant amount of information regarding a variable's ability to predict future stock return. A variety of parameters and variables have been tested and the conclusion of different scholars indicates that several of these contain predictive power.

In this master thesis we investigate whether dividend yield, price-earnings, cay, interest rate, momentum, and Sharpe ratio can predict future stock returns over a short horizon. We have included the most recent data available and attempted to test for the effect of the financial crisis of 2008. To examine whether there exists similar patterns across different international markets we have included the U.S., the Norwegian, and the Swedish market in our analysis.

Our results indicate a presence of predictability in our explanatory variables, and especially cay, interest rate, and Sharpe ratio are significantly strong predictors. Predictability in interest rate and Sharpe ratio is robust over several markets since they show strong prediction in the Scandinavian data as well.

According to our findings, the occurrences during and after the recent financial crisis have affected the predictive ability of our explanatory variables. By excluding the observations after 2007 from our sample, several of the variables become significant predictors. Overall, we find various cases of predictability in our variables depending on the method applied, sample period, and market.

ABBREVIATIONS

ACF:	Autocorrelation function
BIC:	Bayesian Information Criterion
BG:	Breusch-Godfrey
BLUE:	Best Linear Unbiased Estimators
B/M:	Book-to-market
Cay:	Consumption wealth ratio
CRSP:	Center for Research in Security Prices
DF:	Dickey-Fuller
D/P:	Dividend-price
DY:	Dividend yield
DW:	Durbin-Watson
EMH:	Efficient market hypothesis
E/P:	Earnings-price
GLS:	Generalized Least Squares
IS:	In-sample
JB:	Jarque-Bera
NYSE:	New York Stock Exchange
OLS:	Ordinary Least Squares
OMXS:	Stockholm All Share Index
OOS:	Out-of-sample
OSEAX:	Oslo Børs All Share Index
PACF:	Partial autocorrelation function
PE:	Price-earnings
S&P 500:	Standard & Poor's 500 index
VIF:	Variance inflation factor

TABLE OF CONTENT

1. INTRODUCTION

The subject of stock return predictability has been researched by several academics after Fama & French and Campbell & Shiller made the groundbreaking discovery that dividend yield (DY) and the dividend-price (D/P) ratio can predict stock returns in 1988. The purpose of this master thesis is to test for whether the same discoveries can be found in more recent data samples. The factors we have tested for are divided into three categories; fundamental valuation factors, macroeconomic factors, and technical factors. In addition to these factors, which are well known in the previous literature, we look at whether Sharpe ratio perhaps can have predictive ability as well, a variable that has not received so much attention.

Our research is based on U.S. data due to the fact that most of the previous research is done with U.S. data, and will make our result more comparable to the previous findings. We have chosen to look at the Standard & Poor's 500 index (S&P 500) which has developed to be a leading indicator of U.S. equities, and includes 500 large cap companies across the major industries.

In addition to U.S. data we expand our testing on stock return predictability with parts of the Scandinavian market. Research on the Scandinavian market is limited, but the results will be used to check the presence of similarities across the U.S. and Scandinavian market. Based on access to data we have investigated the Norwegian and Swedish stock market. For both countries the all-share index are used, respectively Oslo Børs All Share Index (OSEAX) and Stockholm All Share Index (OMXS). We chose the all share index based on the number of shares included in the index, because a benchmark index not consists of enough shares to be comparable to the S&P 500. Scandinavian data hereafter is related to Norwegian and Swedish data.

1.1 PROBLEM STATEMENT

For the purpose of this master thesis, the following problem statement has been formulated:

"To what degree do different fundamental valuation, macroeconomic, and technical factors have predictive ability on stock returns?"

Four sub-questions have been made to reflect the problem statement:

- Can Sharpe ratio have predictability on stock returns?
- How has the financial crisis of 2008 affected the different factors' ability to predict stock returns?
- Do the explanatory variables in Scandinavian and U.S. data indicate the same ability to predict stock returns?
- How does our results compare to previous research on stock return predictability?

1.2 DELIMITATIONS

The variables included as our explanatory variables in the regressions are chosen on the basis of previous studies and a desired diversity between fundamental valuation, macroeconomic, and technical factors. There are several variables that have been tested for predictability up through the years, but we chose to focus on the ones that we have read about in different courses at CBS and that we find interesting.

The choice to focus on U.S. data was taken since the previous research we found had applied U.S. data. We use the S&P 500 index for calculations of stock returns, and we could have included several indices, but including one more index would double the amount of regressions we run. Due to the limitation of space in our thesis, we choose to only compute the stock returns from one index. In addition we wanted to include Scandinavian data since we are Norwegians who study in Denmark, and because Scandinavian data has received little attention on this area. But the Danish index was too small since it only included 20 companies, so it would be difficult to compare the results with the rest of the indices. It would be interesting to include data from other European countries, but we chose to only focus on U.S., Swedish, and Norwegian data.

In the calculations of return we have computed the log of nominal return and applied it as the dependent variable. Several previous studies have also included excess return, return subtracted by the risk-free rate, as a dependent variable. If we were to apply both versions of return we would have to double the number of our regressions, hence we chose to test for one due to the limited space in our thesis.

The data samples collected for the different variables are selected on the basis of availability. We wish to test for predictability over the longest period that is available to us. The individual sample later applied for each variable is the longest sample period that we could find for the respective variable. We also make three subsamples based on a sample period where information was available for all of our variables.

We test for predictability in a short horizon, one month, since the previous studies have mainly done the same. The cay variable is the only variable that is tested on a quarterly horizon since the information that is required to calculate this variable is only available on a quarterly basis. We chose to focus on short horizon over long horizon. We believe it would be more interesting for an investor to know the predictability of various variables in the short horizon to be able to make abnormal returns in the short run. The regression could have been extended for testing of a longer horizon by including several lags, but due to the page limit in our thesis it has not been included.

We will mainly focus on the variables' individual ability to predict stock returns, but we will also run a multiple regression including all the explanatory variables to see if they can predict stock returns together. We only focus on Ordinary Least Squares (OLS) as our model, even though there are several alternative methods that could be applied. Avramov (2002) for instance applies a Bayesian methodology, which is an alternative to OLS, but we chose to apply the method most acknowledged.

1.3 METHOD AND TOOLS

Our model regressions and econometrics tests are run using SAS Enterprise Guide 5.1. SAS was chosen as a tool since it is available to the practitioners at CBS and we learned how to apply the program in the "Applied Econometrics" course we had during the first semester of our master's degree. SAS is a software suite that can mine, alter, manage, and retrieve data from a variety of sources and perform statistical analysis on it. SAS Enterprise Guide allows us to apply tests such as the Jarque-Bera test, the Durbin-Watson test, and the Breusch-Godfrey test among others, which is essential in the search of valid OLS statistics.

Our main source when collecting data is Thomson Reuters Datastream. Datastream is available for CBS practitioners through Microsoft Excel in the Microsoft Office package and is therefore a natural tool for us to use when collecting data. The data presented by

Datastream is annualized, which has been confirmed by Vikram Srinivasa, who works at Thomson Reuters, so the data can easily be compared and applied. In addition, Excel has been a commonly used program in our gathering and calculations of variables, especially when computing the momentum effect variable. The national banks' homepage for the respective countries are also excessed for different purposes. Theoretical sources have mostly consisted of previous studies in articles and textbooks on econometrics.

1.4 THE STRUCTURE OF THE MASTER THESIS

To be able to answer our problem statement and the sub-question in the most optimal way the master thesis is divided into 6 parts and is structured in the following way. Part 2 contains of the relevant theory chosen to answer our problem statement optimally. This includes theory on stock return, the efficient market hypothesis, and the variables we apply in our regression. In addition the econometric framework used in our testing and possible econometrics issues are described. The data collecting process and the computations of the variables are presented in part 3. Our econometric analysis and the results for each of the tested variables is explained and presented in part 4. These results are further discussed in part 5, where our findings are compared with the already existing research on the topic. In addition the results are discussed in the light of the financial crises in 2008 and similarities/differences across the U.S. and Scandinavian market. The final answers to our question are summarized in part 6, where the conclusion of this master thesis is presented.

2. THEORY

2.1 STOCK RETURNS

When investors purchase stocks, they expect that their investment will secure them return. Stock returns are connected to the riskiness of the stock. The risk premium of a security represents the additional return that investors expect to earn to compensate them for the security's risk (Berk & DeMarzo, 2014, p.86). The general rule is that the more risk you take, the greater the potential for higher return – and loss.

The simple net return, R_t, on the asset between dates t-1 and t is defined as (Campbell, Lo, & MacKinlay, 1997):

(2.1) $$R_t = \frac{P_t}{P_{t-1}} - 1$$

Where P_t is the price of the asset at time t and P_{t-1} is the price of the asset at time t-1.

The risk of a security must be evaluated in relation to the fluctuations of other investments in the economy. A security's risk premium will be higher the more its returns tend to vary with the overall economy and the market index. If the security's returns vary in the opposite direction of the market index, it offers insurance and will have a negative risk premium (Berk & DeMarzo, 2014, p. 88). If an investor is risk-averse, he or she will be more reluctant to buy stocks with high risk compared to if the investor was risk-loving.

2.2 THE EFFICIENT MARKET HYOPTHESIS AND RANDOM WALK

Fama introduced the Efficient Market Hypothesis (EMH) back in 1970. The theory describes the market as efficient when the stock prices "fully reflect" all available information in the market (Fama, 1970). This theory indicates that in a world with efficient markets, movement in stock prices is unpredictable. There are three different degrees of the efficient market hypothesis. They are different in the way that the stock prices reflect different degree of information.

2.2.1 WEAK-FORM

The first version of the hypothesis is the weak-form, the hypothesis is said to be in the weak-form if the stock prices reflect all existing market trading data. This is data such as history of past price, trading volume, and short interest rate (Zvi, 2011, p. 375).

2.2.2 SEMISTRONG-FORM

With the semistrong-form hypothesis all available public information is reflected in the stock prices. In addition to the market trading data reflected in the weak-form, the semistrong-form also reflects information regarding the prospect of a firm. Fundamental data on the firm's product line, balance sheet composition, and financial statement is example of that kind of information (Zvi, 2011, p. 376).

2.2.3 STRONG-FORM

The last version of the EMH is the strong-form hypothesis. This version states that all information, both private and public, is reflected in the stock prices (Zvi, 2011, p. 376). Compared with the semistrong-form the strong-form in addition includes private information, such as insider and market maker information (Tsai & Hsiao, 2010).

2.2.4 RANDOM WALK

Another well-known theory from the history that also states that stock prices are unpredictable is the random walk theory. One of the more well-known contributors to this topic is Maurice Kendall (1953) for his analysis of economic time series and price movements. One of the findings from this empirical analysis showed that there was little serial correlation in stock-exchange movements. The random walk theory states that historical prices cannot be used to predict future prices. In light of the name of the theory the stock prices follow a random walk, the behavior of prices is random and not predictable. A market where the successive price changes are independent is called a random walk market (Fama, 1965).

2.3 BREAKTHROUGH IN THE 1980S

Turning away from the EMH and random walk, Fama and French (1988) argued against the general opinion about stock returns with their article published in 1988. In the article they claimed that prediction of stock returns using aggregated financial variables has been a financial discipline during the last two decades. They find that DY and the default spread capture similar variation in expected bond and stock returns. The DY and the default spread forecast high returns when business conditions are persistently weak and low returns when conditions are strong (Fama & French, 1989). Campbell and Shiller (1989) supported Fama and French's view in their article, where they find that earnings can be used to forecast future dividends.

Since this breakthrough several other authors have tested stock return predictability of different types of variables, fundamental valuation variables, macroeconomics variables, and technical variables.

2.4 FUNDAMENTAL VALUATION FACTORS

2.4.1 DIVIDEND YIELD

DY is calculated as:

(2.2)
$$Dividend\ yield = \frac{Annual\ dividends\ per\ share}{Price\ per\ share}$$

Which is the expected annual dividend of the stock divided by its current price. The DY is the percentage return the investor expects to earn from the dividend paid by the stock (Berk & DeMarzo, 2014, p.273). DY and capital gain rate together makes the total return of a stock.

DY is a variable that has been given much attention in the prediction of stock return. Several researchers have used DY or D/P ratio as one of their explanatory variables when trying to predict stock returns.

Avramov and Chordia (2006) seek to understand whether considering business cycle variables will benefit a real time investor, and they include DY as one of their explanatory variables. By implementing firm-level analysis, they provide new evidence about stock return predictability, and they find that returns are predictable out-of-sample (OOS) by the DY.

Lewellen (2004) find strong evidence that DY predicts both equal- and value-weighted NYSE returns from 1946 to 2000, as well as in various subsamples.

Rapach and Wohar (2006) show that a number of financial variables appearing in the literature display both in-sample (IS) and OOS predictive ability with respect to stock returns in annual data covering most of the twentieth century. They test for return predictability using a bootstrap procedure that explicitly accounts for data mining. They do tests over 1-year, 5-year, and 10-year horizons. The D/P ratio shows predictive ability in the 5-year horizon when IS CRSP data is applied, and in the 10-year horizon when OOS data is applied.

Pesaran and Timmermann (1995) examine the robustness of the evidence on predictability of U.S. stock returns, and address the issue of whether this predictability should have been historically exploited by investors to earn profits in excess of a buy-and-hold strategy in the market index. DY is one of the variables that they test, and their findings confirm the results of previous research which has emphasized the importance of predictable components in stock returns related to the business cycle.

Avramov (2002) proposes a Bayesian model averaging approach to analyze stock return predictability, and find that earnings-yield have higher predictive ability than the other traditional market ratios, such as DY and book-to-market (B/M). In essence, taking model uncertainty into account appears to substantially diminish the predictive power of some explanatory variables, like DY.

Ho and Hung (2012) mainly test the predictive ability of investor sentiment on the return and volatility at the aggregate market level in the U.S., the four largest European countries and three Asia-Pacific countries, but they also perform some testing on the predictive ability of more fundamental variables. Their findings show that DY and inflation rate exhibit the most prevalent ability for stock returns, among the fundamental variables.

2.4.2 PRICE-EARNINGS

Much of the real-world discussion on stock market valuation concentrates on the firm's price-earnings multiple, the ratio of price per share to earnings per share, commonly called the PE ratio (Zvi, 2011, p.781).

$$\frac{P}{E} = \frac{Market\ Capitalization}{Net\ Income} = \frac{Share\ Price}{Earnings\ per\ Share} \tag{2.3}$$

The PE ratio is a simple measure that is used to assess whether a share is over- or under-valued based on the idea that the value of a share should be proportional to the level of earnings it can generate for its shareholders (Berk & DeMarzo, 2014, p.41). The PE ratio tells us how much stock purchasers must pay per dollar of earnings that the firm generates (Zvi, 2011, p.71). Generally a high PE ratio means that investors are anticipating higher growth in the future.

Because the PE ratio considers the value of the firm's equity, it is sensitive to the firm's choice of leverage. The PE ratio is therefore of limited usefulness when comparing firms with markedly different leverage (Berk & DeMarzo, 2014, p.41). You get the most out of PE if you are able to compare it with other companies' PE in the same industry.

Some of the previous studies have used PE ratio as one of their explanatory variables when trying to predict stock return ability, others have used earnings-price (E/P) or earnings yield. There will be reason to compare our findings with all of these variables, since they include the same factor, earnings.

In the same study where Rapach and Wohar (2006) found that D/P have some predictive ability. They found that the PE ratio have predictive ability over a 10-year horizon when IS S&P data is applied, and over a 1-year horizon when OOS data is used.

As previously mentioned, Lewellen (2004) found strong evidence that DY has predictive power, but he also tested the B/M ratio and E/P ratio. The ratios all measure stock prices relative to fundamental parameters. Since each of them has price in the denominator, the ratios should be positively related to expected returns. The evidence for B/M and E/P is somewhat weaker, and overall they seem to have limited forecasting power.

Avramov (2002) proposes a Bayesian model to analyze stock return predictability, and in addition to DY, earnings yield is also tested as an explanatory variable. Using monthly observations, Avramov finds the earnings yield to have stronger predictive power than DY because it receives the highest cumulative probabilities.

E/P ratio was also included as a variable when Pesaran and Timmermann (1995) examined the robustness of the evidence on predictability of U.S. stock returns. Their findings point to, and confirm, that previous research has correctly related stock returns to the business cycle.

Even though, they also conclude that there does not seem to be a robust forecasting model in the sense that the determinants of the predictability of stock return in the U.S. seem to have undergone important changes throughout the period under consideration (monthly S&P observations between 1954 and 1992).

2.5 MACROECONOMIC FACTORS

Among the factors that are connected to stock return predictability in the already existing literature, different macroeconomic factors are well represented. It is expected that returns vary and follows variation in the business cycles, and this might be one of the reasons why macroeconomic factors are connected as a predictability factor. The results and conclusions are different in the various studies. This is not surprising considering the different methods applied, chosen time period, and data. However there exists evidence of a relationship between macroeconomic factors and returns, and those factors prove to be useful in predicting stock returns.

Typical macroeconomic factors used in the different studies are interest rate, unemployment rate, aggregate output, inflation rate, price level, GDP, exchange risk, labor income, and consumption. In this thesis we have chosen to look deeper into two of these factors, the interest rate and the consumption wealth ratio (cay).

2.5.1 INTEREST RATE

Several previous studies highlight interest rate as a factor that is able to predict returns. The interest rate is related to the discount rate applied to the future aggregate cash flows, thus it is an important factor in the economy and in asset pricing.

Based on the individuals' willingness to borrow and lend the interest rate is determined in the market. In addition the interest rate is also influenced by other different macroeconomic factors such as inflation, money supply, government policy, and expectation about future growth. So by looking at the interest rate we also indirectly look at other macroeconomic factors (Berk & DeMarzo, 2014).

Among the different macroeconomic factors, the interest rate has showed to be among the most robust. Rapach, Waphar and Rangvid (2005) looked at stock return predictability using macroeconomic variables in 12 industrialized countries. Their conclusion is that the interest rates are the most common and robust one. Andreas Shrimp (2010) draw the same conclusion based on his study where he looks at stock return predictability in a situation where investors are uncertain about the right state variable. By looking at five international stock markets, and nine financial and macroeconomic predictive variables, he found that interest rate-related variables were the most robust predictive variable.

In the study where Ang and Bekaert (2007) look at the predictability power with main focus on DY, they found evidence that the short term interest rate was the most robust variable for predicting future excess returns. This evidence was only significant at short horizons.

2.5.2 CONSUMPTION WEALTH RATIO

The consumption wealth ratio, also often referred to as cay, is a variable that consists of three key macroeconomic factors. These macroeconomic factors are consumption *c*, asset wealth *a*, and labor income *y*, in the equation below *w* is the average share of nonhuman wealth in total wealth. Cay is a macroeconomic variable that have showed to be a strong variable in predicting quarterly stock returns.

$$cay_t = c_t - wa_t - (1-w)y_t \tag{2.4}$$

Lettau and Ludvigson (2001) found that fluctuations in cay were good predictors for both stock return and excess stock returns at a short and intermediate horizon. In their study they looked at U.S. quarterly data in the period from 1952 to 1998. They also looked at other well known predicting variables, among others DY and market premium, and their result showed that cay outperforms these in being the most predictive variable at short horizon, whereas DY is a stronger predicter at the longer horizon. Based on Lettau and Ludvigsons result, Avramov (2002) also did an analysis on the predictive power of cay. The share of nonhuman wealth is predicted, and when they used observations available at the time of prediction they got poor results. However, he found predictive power when they used data observed after the prediction. The different levels in the results based on which sample they were using, could be an indicator that the good result he got in the first test, and Lettau and Ludvigson got, is due to a look-ahead bias (Avramov, 2002). The look-ahead bias is referred to when there is used information or data in a study that is not available at the actual predicting time of the study.

Two other researchers that have looked at postwar quarterly data for the U.S. are Rapach and Wohar (2006). They looked at three different time horizons, 1, 8 and 16 quarters, and they perform an IS and an OOS test. As Lettau and Ludvigson they also got result where they found predictive ability for cay. They got significant results for all the different horizons tested, which stretches from 1 quarter up to 4 years.

2.6 TECHNICAL FACTORS

When predicting stock returns, one of the commonly used theories is the technical theory and the technical analysis. The fundamental with this analysis is to investigate past data and try to see a pattern in the behavior of the data, and analyze if this pattern exists in future. Because of this way to look at previous information, technical analysis is also referred to as chartist. From the EMH view a technical analysis is meaningless, because in that case historical data would already be reflected in the actual stock prices and would not add any further value (Zvi, 2011).

2.6.1 MOMENTUM

Momentum effect is the phenomena where a period of rise is followed by a period with another rise, or a decreasing period is followed by reduction. Consequently, the following time period have the same direction as the previous one. Whether an investor will follow a momentum strategy or not has background in the individual investor behavior. If investors use momentum as a strategy they can choose to buy stock that has showed previous good returns and sell stock that has had previous bad result. This strategy is just based on previous returns, and rejects the EMH, more specific the weak-form hypothesis, which states that you cannot predict future stock prices based on previous and historical data (Berk & DeMarzo, 2014).

In the academic literature there has been a debate whether the contrarian strategies, buying loser and selling winner, or the momentum strategies gives abnormal returns. Jagadeesh and Titman (1993) looked at whether a momentum strategy could be used to predict future stock returns and get abnormal returns. By observing realized stock returns based on the previous 6 months returns, they find that by following the relative strength strategy (buy past winners and sell past losers) positive abnormal results would be generated. This is significant up to a 12 month holding period, after 12 months these abnormal returns disappears (Jegadeesh & Titman, 1993).

While Jegadeesh and Titman looked at data from the U.S. there has also been done research on this topic regarding international returns. A recent study done by Chui, Titman and Wei (2010) looked at different cultures and their influence on momentum strategies. They have a slightly different perspective since they analyze momentum strategies linked to behavioral bias, more specific individualism. Based on Hofstede's individualism index they found that average return are more than 0.6% higher in countries within the top 30% of the index, compared to the bottom 30% of the index. From that conclusion one can say that investor behaviors affect the trading strategies and the outcome (Chui et al., 2010).

2.7 SHARPE RATIO

The Sharpe ratio is a ratio developed by Nobel laureate William Sharpe, and it measures the ratio of reward-to-volatility provided by a portfolio (Berk & DeMarzo, 2014, p.373). By subtracting the risk-free rate from the expected return of the portfolio and dividing by the standard deviation of the portfolio, you get the Sharpe ratio.

(2.5) $$Sharpe\ Ratio = \frac{Portfolio\ Excess\ Return}{Portfolio\ Volatility} = \frac{E[R_P]-r_f}{SD(R_P)}$$

The optimal portfolio to combine with the risk-free asset will be the one with the highest Sharpe ratio, where the line with the risk-free investment just touches, and so is tangent to the efficient frontier of risky investments (Berk & DeMarzo, 2014).

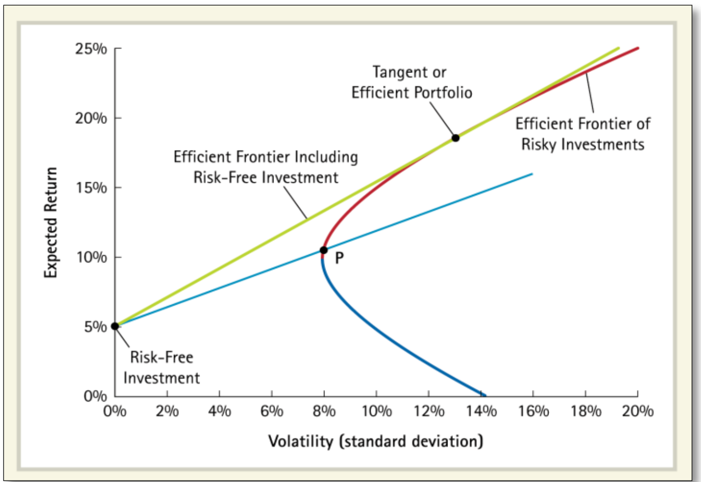


Figure 1: The Tangent or Efficient Portfolio (Berk & DeMarzo, 2014, figure 11.10)

The tangent portfolio is the portfolio with the highest Sharpe ratio. In the figure above, investments on the green line connects the risk-free investment and the tangent portfolio to provide the best risk and return trade-off available to an investor. The green line is also called the capital allocation line. The tangent portfolio is also referred to as the efficient portfolio (Berk & DeMarzo, 2014, p.374).

By computing the excess return and the standard deviation of portfolios, it is possible to predict which portfolio has the best performance. The building blocks of the Sharpe ratio, expected returns and volatilities, are unknown quantities that must be estimated statistically and are, therefore subject to estimation error (Lo, 2002). Possible factors that can affect the calculation are mean reversion, serial correlation, and aggregation methodology (Johnston, Hatem, & Scott, 2013). When calculating returns and standard deviations, and hence Sharpe ratios, the order of the returns does not matter. Therefore the simulated n-period mean returns and standard deviations will approximately equal the historical overlapping n-period historical distribution mean return and standard deviation. Hence there is no need for the simulations (Johnston et al., 2013).

2.8 ECONOMETRICS THEORY

Regression analysis is concerned with the study of the dependence of one variable, the dependent variable, on one or more other variables, the explanatory variables, with a view to estimating and/or predicting the (population) mean or average value of the former in terms of the known or fixed (in repeated sampling) values of the latter (Gujarati & Porter, 2009, p.34).

In the simple linear regression model we study the following equation:

(2.6) $$Y = \beta_0 + \beta_1 X + u$$

Where Y is defined as the dependent variable or explained variable, X is defined as the independent variable or the explanatory variable, β is defined as the coefficient, and u is the error term.

The following assumptions must be fulfilled for the OLS estimators to be Best Linear Unbiased Estimators (BLUE) (Gujarati & Porter, 2009, p.315):

1. The relationship is linear in parameters, and given by: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$
2. Values taken by the regressor X may be considered fixed in repeated samples or they may be samples along with the dependent variable Y.

3. The error term has zero expected value: $E(u_i|X_i) = 0$ for each i.

4. The error term has constant variance for all observations (homoscedasticity): $E(u_i^2|X_i) = \sigma^2$

5. The random variables u_i are statistically independent (no autocorrelation): $Cov(u_i, u_j) = 0$ for all i≠j.

6. Zero covariance between u_i and each X variable: $cov(u_i, X_{1i}) = cov(u_i, X_{2i}) = 0$

7. The number of observations (n) must be greater than the number of parameters to be estimated.

8. There must be sufficient variation in the values of the X variables.

9. No exact collinearity between the X variables.

10. The regression model is correctly specified.

11. The error terms are normally distributed.

In many text books, these assumptions are boiled down to only 6 assumptions since some of the points mentioned above are quite trivial. In particular we usually worry about four kinds of assumption violations: multicollinearity (9), heteroscedasticity (4), autocorrelation (5), and the errors being normally distributed (11), where numbers in parenthesis correspond to the numbers of the assumption above.

It is possible to apply OLS for both linear and nonlinear functions of x and y, and the parameters for x and y will remain linear. The variables can take form of the natural logarithm of x, y or both, can use quadratic forms of x, or use interactions of x variables (Gujarati & Porter, 2009, p.62).

The OLS estimated slope and intercept are:

(2.7) $$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \frac{Cov(X,Y)}{Var(X)}$$

(2.8) $$\beta_0 = \bar{Y} - \hat{\beta}_1\bar{X}$$

The slope estimate, β_1, is the sample covariance between x and y divided by the sample variance of X. If X and Y are positively correlated, the slope will be positive, and if X and Y are negatively correlated, the slope will be negative. Intuitively, OLS is a fitting line through the sample points such that the sum of squared residuals is as small as possible, hence the name least squares. The residual, $\hat{u}$, is an estimate of the error term, u, and is the difference between the sample regression function and the sample point (Gujarati & Porter, 2009, p.71).

2.8.1 T-TEST

In estimating a regression model we test the following hypothesis:

(2.9) $$H_0: \beta_1 = 0$$

(2.10) $$H_1: \beta_1 \neq 0$$

H_0 is the null hypothesis, and H_1 is the alternative hypothesis. This is a two-tailed alternative hypothesis. A one-sided alternative hypothesis can either test for:

(2.11) $$H_1: \beta_1 < 0$$

Or

(2.12) $$H_1: \beta_1 > 0$$

The t-test statistic with n-k degrees of freedom can be written as:

(2.13) $$t = \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)}$$

Our value of β_1 is drawn from the null hypothesis. Once we calculate t, we compare it to the critical value. The t-statistic is compared to the critical value from the t-student distribution which depends on the choice of significance level and number of observations in the sample. If the t-statistic is higher than the critical value, we reject the null hypothesis.

2.8.2 MULTICOLLINEARITY

Multicollinearity is correlation between the independent variables, hence it is only relevant to check for multicollinearity in regressions with more than one explanatory variable. Multicollinearity can typically be found in time series with upward/downward trend, or it can be synonymous with having too few observations. A few reasons for multicollinearity can be data collection, the constraints in the model, the model specification, or an overdetermined model.

There are five diagnostics for the presence of multicollinearity, a high R^2 but few significant t-values, high pair/wise correlation among the explanatory variables, R^2 from the auxiliary regression is larger than R^2 from the original regression, a variance inflation factor (VIF) higher than 10, or a condition index greater than 30 (Gujarati & Porter, 2009, p.340).

Multicollinearity can be solved in a few different ways, we can drop a variable or several variables from the model, acquire additional data or a new sample, rethink the model, combine cross-sectional and time series data, transform variables, or obtain a priori information (Gujarati & Porter, 2009, p.342).

In cases of near or high multicollinearity, one is likely to encounter some consequences. Although BLUE, the OLS estimators have large variances and covariances, making estimation difficult, the t ratio of one or more coefficients tends to be statistically insignificant, the R^2 can be very high, and the OLS estimators and their standard errors can be sensitive to small changes in data (Gujarati & Porter, 2009, p.327)

2.8.3 NORMALITY OF THE ERROR TERM

According to the assumptions of the regression model, the error terms must be normally distributed. The t and F tests require that the error term follows the normal distribution otherwise the testing procedure will not be valid in small samples. We can test the normality of the error term by looking at the histogram of the observations, or by looking at the Jarque-Bera (JB) statistic.

We can calculate the JB statistic manually or SAS can give the statistic in one of its outputs. To perform the JB test we first calculate the skewness and the kurtosis of the standardized residuals (Gujarati & Porter, 2009, p.815):

(2.14) $$Skewness: S = \frac{E(\hat{u}^{*}-v_u)^3}{\left[E(\hat{u}^{*}-v_u)^{2^{3/2}}\right]}$$

(2.15) $$Kurtosis: K = \frac{E(\hat{u}^{*}-v_u)^4}{\left[E(\hat{u}^{*}-v_u)^{2^{2}}\right]}$$

Then we calculate the JB test statistic as:

(2.16) $$JB = \frac{n}{6}\left[S^2 + \frac{(K-3)^2}{4}\right]$$

The JB test is asymptotically χ^2-distributed. The null hypothesis of normality is rejected if the JB test is bigger than 3.84, when there is a 5% significance level and 1 degree of freedom, or if the p-value is below your chosen significance level (normally 0.05).

2.8.4 HETEROSCEDASTICITY

In the case of homoscedasticity, the variance of the error term, $var(u_i|X_i = x)$, is constant for i=1,…,n and does not depend on x. Thus, under homoscedasticity $var(u_i) = \sigma^2$, thus implies that the data are equally spread out over our whole sample. Hence, the variance of the error term, $var(u_i|X_i = x)$, is not constant for i=1,…,n in the case of heteroscedasticity. Under heteroscedasticity, the OLS standard errors are biased, and if the standard errors are biased, we cannot use the normal t and F statistics and OLS is no longer BLUE (Gujarati & Porter, 2009, p.371).

To find out whether the assumption about constant variance is fulfilled, we plot the residuals against the explanatory variables and against the predicted value of Y. If there seems to be a pattern of variance in the residuals, there might be a case of heteroscedasticity. Other ways to check for heteroscedasticity is to perform a White's test manually (Gujarati & Porter, 2009, p.387).

(2.17) $$White`s\ test = n * R^2 \sim \chi^2_{df}$$

The degrees of freedom and R^2 is from the regression of the squared residuals on the independent variables. The n is the sample size. If the White's test statistic is higher than the critical value then we reject the hypothesis of homoscedasticity.

The problem of heteroscedasticity can basically be solved in three ways; by using heteroscedasticity corrected standard errors, by using a weighted least squares (WLS) approach, or by changing the functional form.

2.8.5 AUTOCORRELATION

A series exhibiting autocorrelation is related to its own past values. The term autocorrelation may be defined as “correlation between members of series of observations ordered in time or space”. The classical model assumes that the disturbance term relating to any other observation is not influenced by the disturbance term relating to any other observation (Gujarati & Porter, 2009, p.66).

However, if there is such a dependence, we have autocorrelation. Symbolically:

(2.18) $$E(u_i u_j) \neq 0 \quad i \neq j$$

The correlation coefficient (between x and y):

(2.19) $$\rho_{x,y} = \frac{cov(x,y)}{\sqrt{var(x)var(y)}}$$

(2.20) $$\rho_{x,y} = \frac{E(x-Ex)(y-Ey)}{\sqrt{E(x-Ex)^2 E(y-Ey)^2}}$$

For time series:

(2.21) $$\rho_j = \frac{cov(y_t, y_{t-j})}{\sqrt{var(y_t)var(y_{t-j})}}$$

To test for autocorrelation, we have to find out whether the error terms are independent over time. Therefore, we plot the residuals against time. If the error term of different observations seem to be correlated, that suggests that autocorrelation is present. If the patterns are wave-like, the autocorrelation is positive. We may also plot the lagged residuals against the residuals. If there is a linear relationship between the residuals and the lagged residuals, that is an indication of autocorrelation (Gujarati & Porter, 2009, p. 413).

If the assumption concerning autocorrelation is not fulfilled, we cannot use the t-values, the F-values, and R^2.

There are two different ways in which we can test for autocorrelation numerically, either by using the Breusch-Godfrey (BG) test or the Durbin-Watson (DW) test. In the BG test, the null hypothesis with four restrictions is:

(2.22) $$H_0: \rho_1 = \rho_2 = \rho_3 = \rho_4 = 0$$

And the BG test value ([n-p]R^2) should be compared with the critical value ($\chi^2_{0,95}(p)$). Where n is the number of observations, and p is the number of restriction (number of ρ`s) as always. We reject the null hypothesis H_0 if $(n-p)R^2 > \chi^2_{0,95}(p)$ (Gujarati & Porter, 2009, p.439).

In the DW test the H_0-hypothesis is that no autocorrelation appears.

(2.23) $$H_0: Cov(\varepsilon_i, \varepsilon_j) = 0, for\ i \neq j\ (no\ autocorrelation)$$

(2.24) $$H_1: Cov(\varepsilon_i, \varepsilon_j) \neq 0, for\ i \neq j$$

A table containing the DW d statistic is used to find the significance points of d_L and d_U at the 0.05 level of significance, by looking up n (number of observations) and k (explanatory variables) (Gujarati & Porter, 2009, p.436).

(2.25) $$d \approx 2(1 - \hat{\rho})$$

This equation can be used to find ρ. If $\rho = 0$, we assume no correlation, d=2. If $\rho = \pm 1$, we assume perfect positive or negative correlation. ρ is known as the coefficient of autocorrelation, or more accurately, the coefficient of autocorrelation at lag 1. It is critical to note that |ρ|<1, that is, the absolute value of ρ is less than 1. If ρ is 1, the variances and covariances are not defined. If |ρ|<1, we say that the AR(1) process given in equation 2.26 is stationary. If |ρ| is less than 1, then it is clear from equation 2.27 that the value of the covariance will decline as we go into the distant past.

(2.26) $$u_t = \rho u_{t-1} + \varepsilon_t \qquad -1 < \rho < -1$$

(2.27) $$cov(u_t, u_{t+s}) = E(u_t u_{t-s}) = \rho^s \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

To be able to apply the DW test, the following assumptions must be fulfilled: (Gujarati & Porter, 2009, p.434)

1. The regression model includes the intercept term. If it is not present, as in the case of the regression through the origin, it is essential to rerun the regression including the intercept term to obtain the RSS.

2. The explanatory variables, the X's, are nonstochastic, or fixed in repeated sampling.

3. The disturbances u_t are generated by the first-order autoregressive scheme:

   $u_t = pu_{t-1} + \varepsilon_{t.}$ Therefore, it cannot be used to detect higher-order autoregressive schemes.

4. The error term u_t is assumed to be normally distributed.

5. The regression model does not include the lagged value(s) of the dependent variable as one of the explanatory variables. Thus, the test is inapplicable in model of the following type: $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + \beta_k X_{kt} + \gamma Y_{t-1} + u_t$

6. There are no missing observations in the data.

In the presence of pure autocorrelation (not the result of mis-specification of the model) the method of generalized least squares (GLS) can solve the problem. When ρ is known the problem of autocorrelation can be easily solved. Consider the two-variable regression model:

(2.28) $$Y_t = \beta_1 + \beta_2 X_t + u_t$$

The error term is assumed to follow a AR(1) scheme

(2.29) $$u_t = \rho u_{t-1} + \varepsilon_t$$

If the two-variable regression holds true at time t, it also holds true at time (t-1).

(2.30) $$Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$$

By multiplying both sides of Eq.2.30 with ρ, we obtain

(2.31) $$\rho Y_{t-1} = \rho\beta_1 + \rho\beta_2 X_{t-1} + \rho u_{t-1}$$

Subtracting the equation above from Eq.2.28 gives

(2.32) $$(Y_t - \rho Y_{t-1}) = \beta_1(1-\rho) + \beta_2(X_t - \rho X_{t-1}) + \varepsilon_t$$

The error term satisfies the usual OLS assumptions, hence we can apply OLS to the transformed dependent and explanatory variable and estimators will be BLUE (Gujarati & Porter, 2009, p.442).

2.8.6 STATIONARITY

In general, a time series is stationary if its probability distribution does not change over time. The Dickey-Fuller (DF) unit root test is a test of whether or not time series are stationary. The actual procedure of implementing the DF test involves several decisions. First, we need to choose which model for the unit root test to consider, depending on the nature of the unit root process, the actual estimation procedure is as follows (Gujarati & Porter, 2009, p.755):

Model a (X_t $is\ a\ random\ walk$): $\Delta X_t = \delta X_{t-1} + \upsilon_t$
Model b (X_t $is\ a\ random\ walk\ with\ drift$): $\Delta X_t = A_1 + \delta X_{t-1} + \upsilon_t$
Model c (X_t $is\ a\ random\ walk\ with\ drift\ around\ a\ deterministic\ trend$): $\Delta X_t = A_1 + A_2 t + \partial X_{t-1} + \upsilon_t$

Secondly, in order to have an unbiased estimation of δ based on the chosen equation, we need to verify that there is no autocorrelation in the error terms. Accordingly, we run the BG test and, in case of autocorrelation, we modify the chosen model by adding as many lags as necessary to obtain a model without autocorrelation. Only when we have a model without autocorrelation we can rely on the estimation of δ and decide about stationarity. We are now able to decide which statistics of the augmented DF test we should consider to evaluate whether the time series is stationary or not.

In the DF unit root test the hypotheses are:

(2.33) $H_0: \delta = 0$ (non-stationarity in the time series)
(2.34) $H_1: \delta < 0$ (stationarity in the time series)

The estimated coefficient of X_{t-1} follows the τ (tau) statistic and not the t-student distribution (Gujarati & Porter, 2009, p.755). The critical values of the tau test are different for each of the specification of the DF test (i.e, model a, b, c). There is no way of knowing which specification is correct to begin with. Trial and error is necessary.

To assess whether a time series is stationary or not, it is necessary to perform a graphical inspection of the autocorrelation function (ACF) and the partial autocorrelation function (PACF), along with a unit root test. It indicates that a time series is non-stationary if the function declines slowly and linearly. If we have a stationary time series the autocorrelations at various lags mover around zero (in other words, the ACF could resemble the correlogram of a white noise time series or show a fast exponentially decay in the lags) (Gujarati & Porter, 2009, p.749).

If both the variables (the explanatory and the dependent) are non-stationary it should concern us, as a regression of one non-stationary series on another might be a spurious regression. A spurious regression is a regression model that is seemingly good but when taking a closer look there is no relationship between Y and X. In order to check for spuriousness we plot ΔY_t and ΔX_t to see whether there is any correlation between the differences of the variables of interest. If there is correlation between the variables in the plot, it is an indication that changes in X causes changes in Y.

Non-stationary time series can be solved by using the first difference of the dependent variable as the dependent variable and several difference of the explanatory variable as explanatory variables. Then look at the residual plots and perform a White's test to check for heteroscedasticity, and also check for normal distribution by looking at the histogram and performing a JB test. If the model is better now but some of the explanatory variables are insignificant, exclude them and make a new model. To test if the new model is a special case of the other model, test the hypothesis that says that the coefficient of the excluded variables are equal to 0 by using the F-test. If we do not reject the null hypothesis, the new model is preferred.

2.8.7 IN-SAMPLE PREDICTABILITY

The IS predictive regression model is estimated using T-1 observations in the sample when the explanatory is in first lag. The observations of the dependent and explanatory variable are the raw numbers gathered from the collected data. Eq.2.6 is regressed and the predictability of x_{t-1} is assessed by examining the t-statistic of the OLS estimate.

2.8.8 OUT-OF-SAMPLE PREDICTABILITY

OOS predictions can be generated by either applying a recursive scheme or a rolling scheme. The recursive scheme involves increasing the number of observations during the testing, while a rolling scheme keeps a constant number of observations throughout the testing.

First, the total sample, T, is divided into an IS period, R, and an OOS period, P. The following equation:

(2.35) $$Stock\ return_t = \alpha_1 + \beta_1 * Explanatory\ variable_{t-1} + u_t$$

is applied to generate the OLS estimates which will further forecast the "unrestricted" predictive regression model. The unrestricted model forecast is denoted by

(2.36) $$\hat{y}_{1,R+k} = \hat{\alpha}_{1,R} + \hat{\beta}_{1,R} * x_R$$

Where $\hat{\alpha}_{1,R}$ and $\hat{\beta}_{1,R}$ are the OLS estimates of the intercept and the explanatory variable. The first regression incudes all observations in the R period, the next regression test for the R+1 observations and this procedure is continued up to R+P observations are included in the regression. The forecasting error from the unrestricted model is denoted by

(2.37) $$\hat{u}_{1,R+k} = y_{R+k} - \hat{y}_{1,R+k}$$

Each $\hat{\alpha}_{1,R}$ and $\hat{\beta}_{1,R}$ can further be applied to calculate a new set of stock returns, and then be regressed with the explanatory variable to check for predictability. In the "restricted" model the regression is applied in the same manner, except $\beta = 0$ in Eq.2.35. Hence, the restricted model forecast is $\hat{y}_{0,R+k} = \hat{\alpha}_{0,R} where$ $\hat{\alpha}_{0,R}$ is the OLS estimate of α, and the corresponding forecast error is denoted by

(2.38) $$\hat{u}_{0,R+k} = y_{R+k} - \hat{y}_{0,R+k}$$

After running all the linear regressions there will be a total of two sets of T-R-k+1 recursive forecast errors, one for the unrestricted and one for the restricted model (Rapach & Wohar, 2006).

The MSE-F statistic can be computed to compare the OOS forecasts from the unrestricted and restricted models. The statistic is applied to test the null hypothesis that the unrestricted forecast MSE is equal to the restricted MSE against the alternative hypothesis that the unrestricted MSE is lower than the restricted MSE. First step in calculating the MSE-F statistic is to compute $\hat{d}_{t+k} = \left(\hat{u}_{0,t+k}\right)^2 - \left(\hat{u}_{1,t+k}\right)^2$ which is further applied in

(2.39) $$\bar{d} = (T - R - k + 1)^{-1} \sum_{t=R}^{T-k} \hat{d}_{t+k} = MSE_0 - MSE_1$$

Where (Clark & McCracken, 2009)

(2.40) $$MSE_i = (T - R - k + 1)^{-1} \sum_{t=R}^{T-k} \left(\hat{u}_{i,t+k}\right)^2, i = 0,1$$

Finally, it is possible to calculate the MSE-F statistic by applying the following formula

(2.41) $$MSE - F = (T - R - k + 1)\bar{d}/MSE_1$$

If the MSE-F proves to be significant it indicates that the unrestricted model forecasts are statistically greater than the restricted model forecasts. Another statistic that can be calculated in connection to OOS testing is ENC-NEW which relates to forecast encompassing. If the ENC-NEW statistic is rejected, then the financial variable contains information that is useful for predicting returns apart from a model of constant returns (Rapach & Wohar, 2006).

To compute the ENC-NEW the first step is to calculate $\hat{c}_{t+k} = \hat{u}_{0,t+k}\left(\hat{u}_{0,t+k} - \hat{u}_{1,t+k}\right)$ which is further applied in

(2.42) $$\bar{c} = (T - R - k + 1)^{-1} \sum_{t=R}^{T-k} \hat{c}_{t+k}$$

The final step is computing the ENC-NEW statistic

(2.43) $$ENC - NEW = (T - R - k + 1)\bar{c}/MSE_1$$

The asymptotic critical values for tests of equal MSE are applied when deciding whether the unrestricted model forecasts are superior to the restricted model (McCracken, 2007). The asymptotic critical values for tests of forecast encompassing are used when deciding whether the financial variables contain useful information (Clark & McCracken, 2001).

3. DATA COLLECTION AND VARIABLE COMPUTATION

When testing for the stock return predictability of different variables we need to compare data on the stock return from a given period and data on the various variables from the same period. We have decided to mainly focus on the U.S. as the market where we will check for stock return predictability. We will look at data after 1952 due to the world wars and because after the presidential election in 1952 the Fed stopped pegging interest rates and began to pursue an independent monetary policy (Campbell & Yogo, 2006). We have tried to include the most recent data and will be looking at periods up until February 2014, given that they are available for the particular variable.

As mentioned earlier we have also chosen to look at two of the countries in the Scandinavian market, Norway and Sweden, in addition to the U.S. data. For both countries we are looking at a time horizon of 15 years which is divided into two subsamples, 1987-1999 and 2000-September 2012, when we are testing for predictability. We will not include the variables momentum effect and cay for the Scandinavian countries because cay is a variable consisting of three variables that are difficult to find for Norway and Sweden, and momentum effect is a very time-consuming variable to calculate.

3.1 STOCK RETURNS

3.1.1 U.S. DATA

The S&P 500 index is used when calculating the stock returns. First we collected the price index for the S&P 500 from Datastream on a monthly basis between 1964 and today, and then we calculated the log returns using the following formula:

(3.1) $$Log\ returns = \ln\left(\frac{price_t}{price_{t-1}}\right)$$

We will try to investigate whether the other variables have predictive ability on the log returns. All the data codes applied to retrieve data from Datastream can be found in Appendix A.

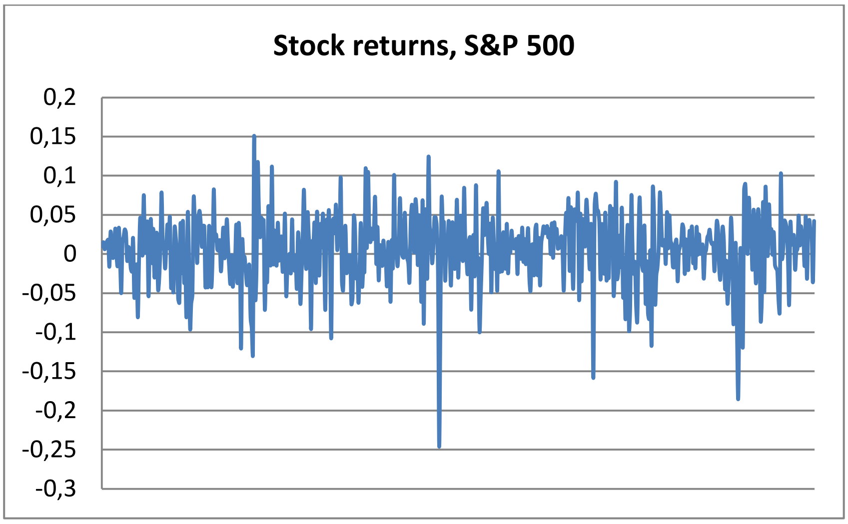


Figure 2: Stock returns S&P500 from Feb 1964 to Feb 2014-06-13

3.1.2 SCANDINAVIAN DATA

The stock returns for Norway and Sweden are calculated based on the monthly price from each of the countries all-share index and the return is calculated in log.

OSEAX is the Oslo Stock Exchange All-Share index that consists of all listed shares on Oslo Stock Exchange, in total 160. The historical data prices are collected from Oslo Stock Exchange's web page. The index goes back to December 1995 when it was developed, before this date there was a total index (TOTX). In 2001 the total index was divided into a benchmark index, mutual fund index and the all-share index. The historical data for the period before 1995 is based on the linked all-share index (Oslo Børs, 2014). This is an index where Oslo Stock Exchange have calculated an adjustment factor that is based on the ratio between the all-share index and the total index per 31 August 2001, this factor was set to be 7.0860. To get the prices for the OSEAX the prices for the total index is dividend on this adjustment factor.

In Sweden the all share index has the ticket OMXS and consist of 280 shares in total. The historical price index for the whole period is collected from Datastream.

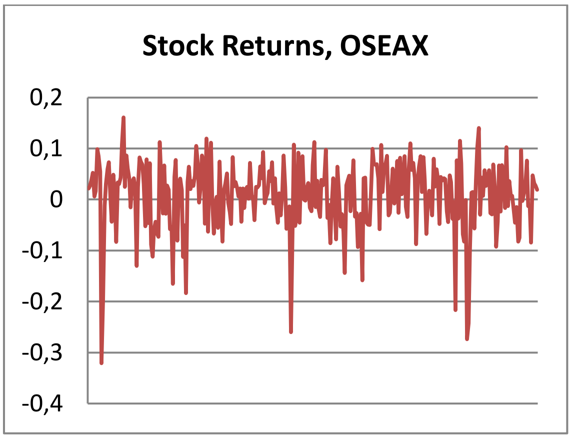


Figure 3: Stock returns OSEAX from Jan 1987 to Sept 2012

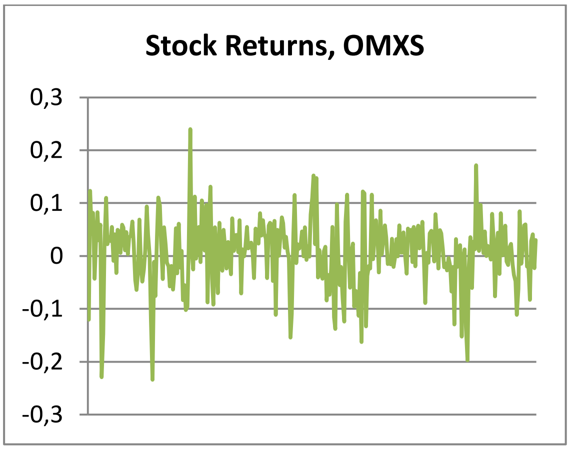


Figure 4: Stock returns OMXS from Jan 1987 to Sept 2012

3.2 DIVIDEND YIELD

3.2.1 U.S. DATA

The DY ratio is collected from Datastream on a monthly basis for a period from January 1965 to September 2012, because that was the largest sample period Datastream could provide. The DY expresses the dividend per share as a percentage of the share price, according to Datastream.

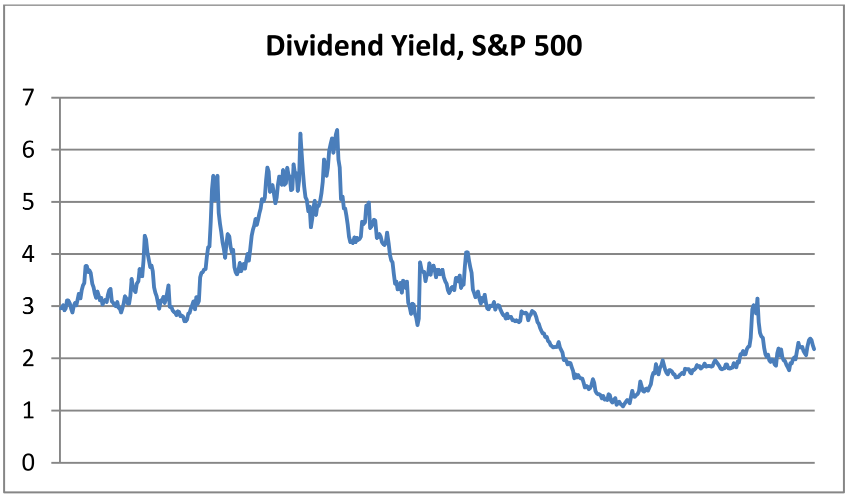


Figure 5: Dividend yield from Jan 1965 to Sept 2012

3.2.2 SCANDINAVIAN DATA

Datastream do not report DY for the index as a whole, but it reports DY for each of the shares in the index. For both OSEAX and OMXS the individual DY is collected, and we have calculated an average each month by taking all of the reported DY and divided it on the number of shares in the index, making an equal-weigthed average.

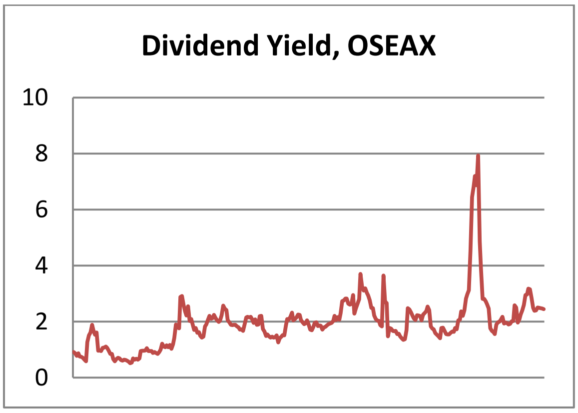


Figure 6: Dividend from Jan 1987 to Sept 2012

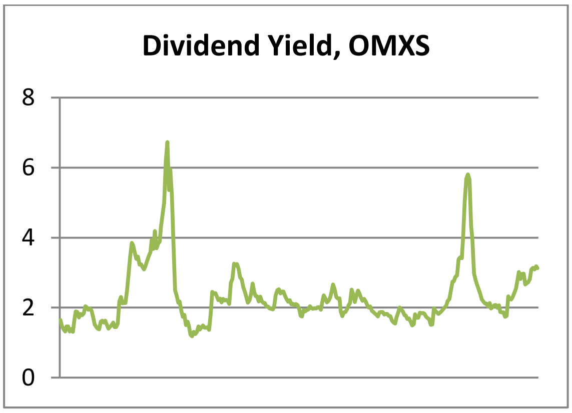


Figure 7: Dividend from Jan 1987 to Sept 2012

3.3 PRICE-EARNINGS

3.3.1 U.S. DATA

In resemblance with the DY, the PE ratio is also collected from Datastream on a monthly basis for a period from January 1968 to September 2012. Datastream defines the PE ratio as the price divided by the earnings rate per share at the required date.

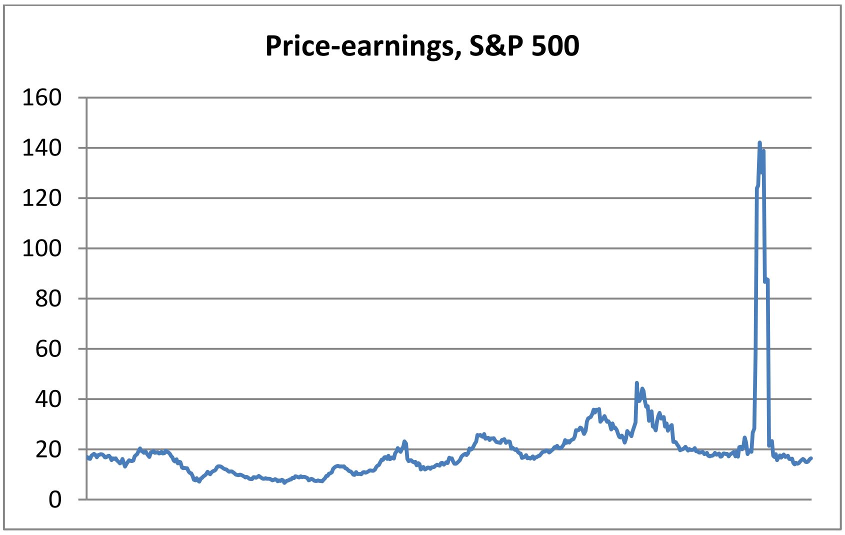


Figure 8: Price-earnings from Jan 1968 to Sept 2012

3.3.2 SCANDINAVIAN DATA

To calculate the PE for each of the indices we are following the same procedures as we did with DY. The individual shares' PE is collected from Datastream and the calculated equal-weighted average is used as the observation in each of the respective months.

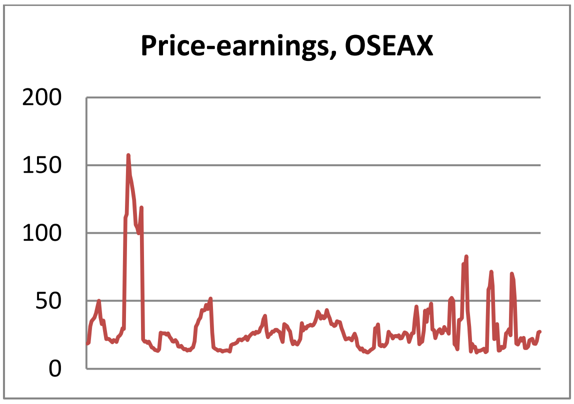


Figure 9: Price-earnings from Jan 1987 to Sept 2012

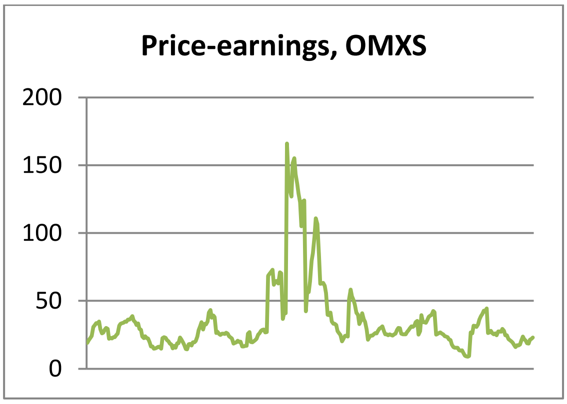


Figure 10: Price-earnings from Jan 1987 to Sept 2012

3.4 CONSUMPTION WEALTH RATIO

3.4.1 U.S. DATA

By inspiration from Gao and Huang (2004), we have collected quarterly data to calculate cay from Federal Reserve Board's web pages (Federal Reserve Board, 2014), where data from 1945 and until today are provided. The sample period we have used stretches from the first quarter in 1964 to the last quarter in 2013. Cay is a variable consisting of three different numbers, consumption, labor income, and financial wealth.

Total personal consumption expenditure is used to measure the consumption, compensation of employees is used to measure income, and net worth (assets-liabilities) from the household balance sheet is used to measure financial wealth. Cay is calculated as a log variable, so consumption, income, and wealth are also calculated in logs.

In the calculation of *w*, average share of nonhuman wealth in total wealth, the data is rather hard to collect. Avramov (2002) used the estimations made by Lettau and Ludvigson (2001) in his calculations, as will we. By using the same values for *w*, we are more capable of comparing our results to the results of their previous research. Lettau and Ludvigson (2001) estimate asset wealth (*w*) to equal 0.3054, and labor income (1-*w*) to equal 0.5891, which leads them to sum up to 0.8945 (Lettau & Ludvigson, 2001). This means that they do not sum up to unity, and that there is a part of total wealth that includes other factors than asset wealth and labor income.

Since data included in cay is observed on a quarterly basis, we have computed stock returns on a quarterly basis for this variable.

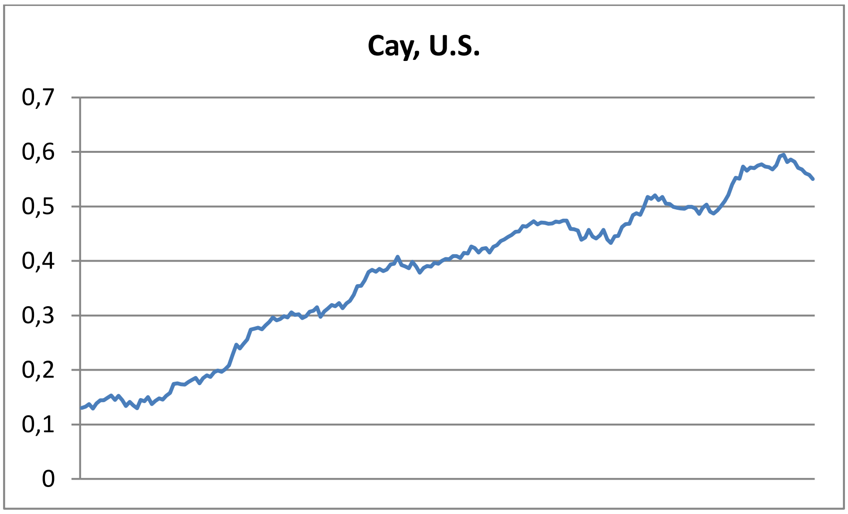


Figure 11: Cay from first quarter 1964 to last quarter 2013

3.5 INTEREST RATE

3.5.1 U.S. DATA

The 3 month Treasury bill is collected from Datastream and used as our monthly rate. We have adjusted the output from Datastream by subtracting the 12-month moving average from the output, and used the results as our rate. This is done in inspiration by an article by Rapach, Waphar, and Rangvid (2005). The sample period is from January 1973 to February 2014.

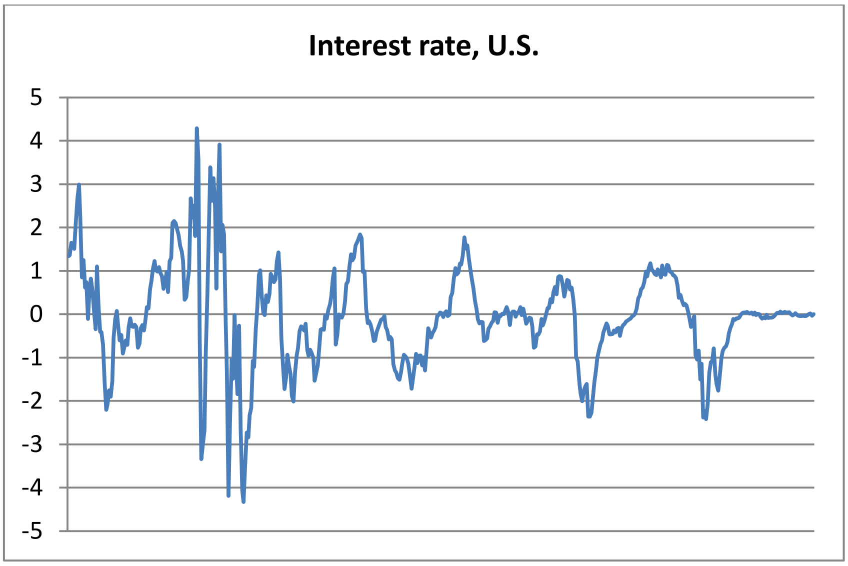


Figure 12: Interest rate from Jan 1973 to Feb 2014

3.5.2 SCANDINAVIAN DATA

As the interest rate we have applied a rate that is equivalent to the Treasury bill used on the U.S. data. The rate is called “statskasseveksler” in Norway and “statsskuldväxlar” in Sweden, and we have used monthly observations on the 3-month bill.

The Swedish rate is collected from Datastream, the code for this 3-month treasury bill is SDGBILL3. Datastream do not report the data for the Norwegian Treasury bill, so the data is collected from Norges Bank, Norway’s central bank. From the web page (Norges Bank, 2014) we downloaded historical data for the time period that was available, that was only back to February 2003. The rest of the observations needed, January 1987-January 2003, is collected from the publication Historical monetary statistics for Norway – Part II table 1.B.2 (Eitrheim, Klovland, & Qvigstad, 2007).

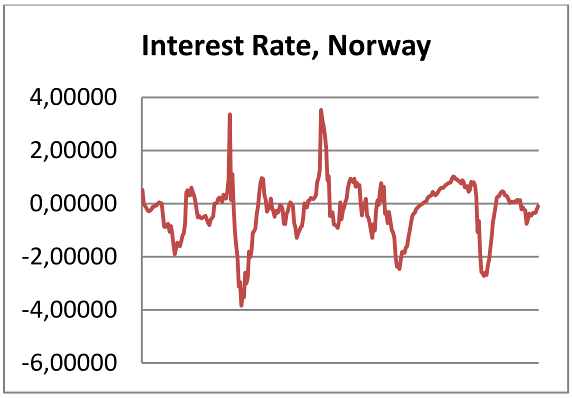


Figure 13: Interest rate from Jan 1987 to Sept 2012

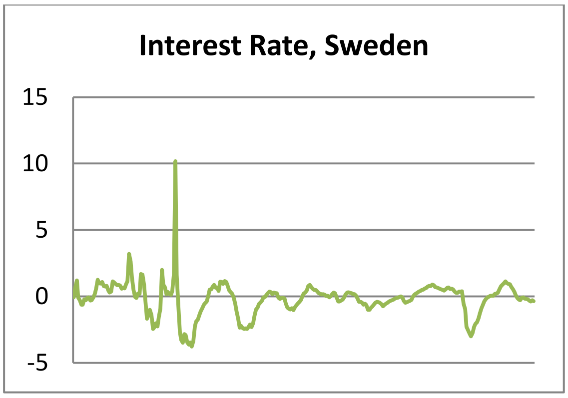


Figure 14: Interest rate from Jan 1987 to Sept 2012

3.6 MOMENTUM

3.6.1 U.S. DATA

Jegadeesh and Titman (1993) have applied different strategies in testing for momentum effect in the market, and we have decided to use one of their strategies where the portfolio is constructed based on the last 6 months return and is held for 6 months with overlapping periods. We have collected prices of the individual stocks that are included in S&P 500, but we have not included the stocks that have disappeared from the S&P 500 after 1973.

We calculated the 6 month return on every single stock in S&P 500 on a monthly basis between July 1973 and August 2013. Then, we arranged the stocks from the highest return to the lowest and found the top 10% and bottom 10% of the stocks which defines the “winner” and “loser” portfolios. Further we found the return of the top 10% and bottom 10% after 6 new months, since the strategy says the portfolio has a 6 month holding period. Then we calculated the return of the “winner” and “loser” portfolio by using equal weights and multiplying with the 6 month return. The momentum effect was then defined as the difference between the “winners” return and the “losers” return.

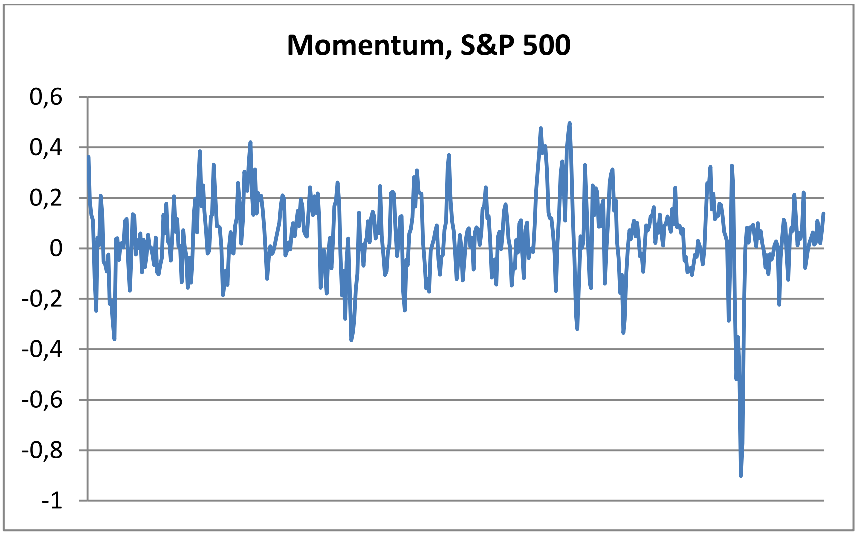


Figure 15: Momentum from July 1973 to Aug 2013

3.7 SHARPE RATIO

3.7.1 U.S. DATA

The log returns for S&P 500 are used as the expected return of the portfolio and the 3 month Treasury bill, adjusted for 12-month moving average, is used as the risk free rate. The standard deviation is calculated with the STDEV formula in Excel, using the returns on each of the stocks in S&P 500 as observations. Sharpe ratio has observations on a monthly basis between February 1973 and February 2014.

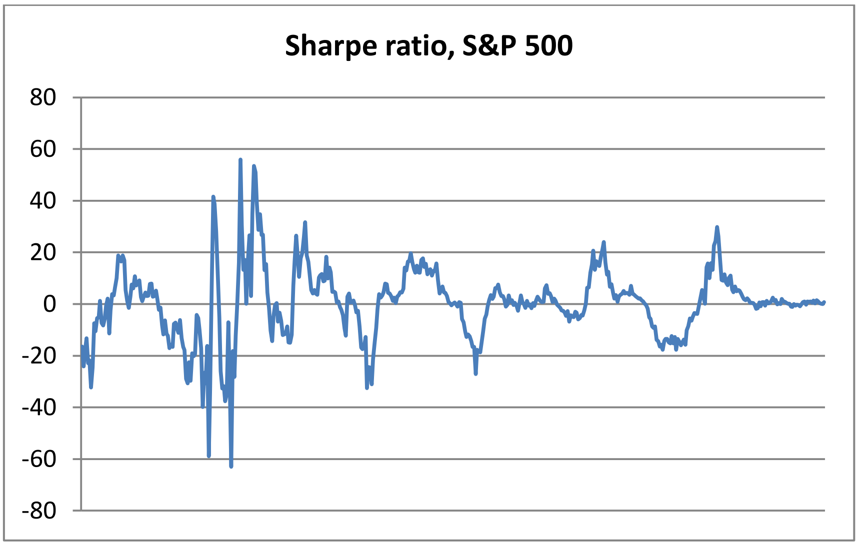


Figure 16: Sharpe ratio from Feb 1973 to Feb 2014

3.7.2 SCANDINAVIAN DATA

Sharpe ratio is calculated by using the same formula we used for the S&P 500. We used the already collected data for OSEAX and OMXS, in addition we obtain the individual prices for all the shares in the index to calculate individual return which we further used to calculate the standard deviation. The interest rate variable is used as the risk-free rate in the calculation of Sharpe ratio.

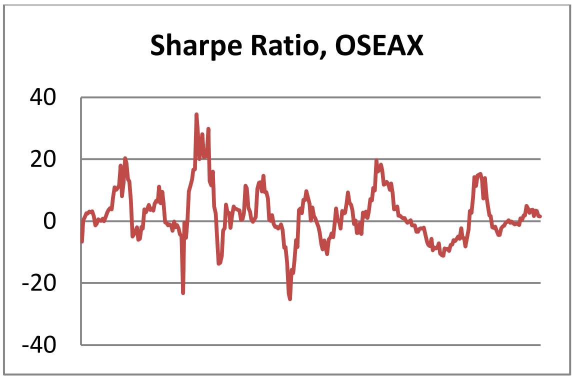


Figure 17: Sharpe ratio from Jan 1987 to Sept 2012

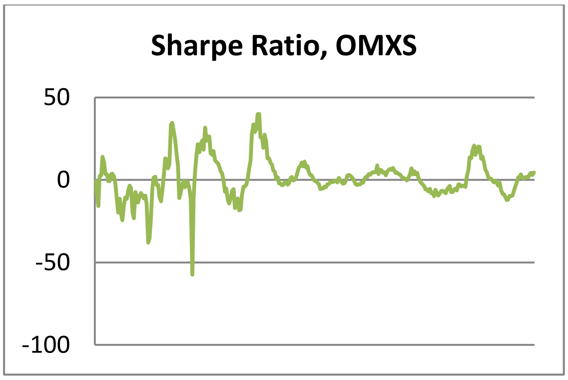


Figure 18: Sharpe ratio from Jan 1987 to Sept 2012

In the table below there is an overview of the frequency and sample periods collected for each of the variables and countries.

Variable	U.S.	Norway	Sweden	Frequency
Stock returns	02:1964 – 02:2014	01:1987 – 09:2012	01:1987 – 09:2012	Monthly
Dividend yield	01:1965 – 09:2012	01:1987 – 09:2012	01:1987 – 09:2012	Monthly
Price-earnings	01:1968 – 09:2012	01:1987 – 09:2012	01:1987 – 09:2012	Monthly
Cay	Q1:1964 – Q4:2013	-	-	Quarterly
Interest rate	01:1973 – 02:2014	01:1987 – 09:2012	01:1987 – 09:2012	Monthly
Momentum effect	07:1973 – 08:2013	-	-	Monthly
Sharpe ratio	02:1973 – 02:2014	01:1987 – 09:2012	01:1987 – 09:2012	Monthly

Table 1: Sample period and frequency for the different variables

4. ANALYSIS AND RESULTS

We have run the linear regression as most of the previous research has done:

(4.1) $$Stock\ return_t = \beta_0 + \beta_1 * Explanatory\ variable_{t-1} + u_t$$

β_0 is the coefficient of the constant or the intercept, and it measures the expected value of the stock price when the explanatory variable is zero. β_1 is the regression coefficient of the explanatory variable. The stock returns are computed in log and the explanatory variable is linear, except for the cay variable which will be computed in log. With a linear explanatory variable the coefficient value is approximately the percentage change in stock returns given a 1 unit change in the explanatory variable, since it is a log-lin model. When both the dependent variable and the explanatory variable are in log, the coefficient of the explanatory variable is the elasticity of stock returns with respect to cay. We lag the explanatory variable, so we are using the explanatory variable from the month before to test for predictive ability in the stock returns the month after.

Our main testing method will be the IS testing, which will be applied for all of our sample periods for both U.S. and Scandinavian data. Additionally, we will apply the OOS testing on one of our subsamples for the U.S. data to test for predictability differences in IS and OOS.

In the results some of the plots from SAS have been included, the remaining can found in Appendix B-E.

4.1 U.S. LONG SAMPLE

Each factor has been tested for an "individual" sample which is the largest sample period we could find for that factor, and for a "financial crisis" sample which is the individual sample stopped after 2007, excluding the data from 2008 and after.

4.1.1 DIVIDEND YIELD

Individual Sample	
Dividend Yield (DY)	
β	0.0019
Standard Error	0.0015
t-value	1.26
p-value	0.2091
R-square	0.0028
F	1.58
Jarque-Bera	175.38
Durbin Watson	1.89
Breusch-Godfrey	
- p=1	1.83
- p=3	3.82

Individual sample excl. Financial Crisis	
Dividend Yield (DY)	
β	0.0017
Standard Error	0.0015
t-value	1.14
p-value	0.2548
R-square	0.0025
F	1.30
Jarque-Bera	179.69
Durbin Watson	1.97
Breusch-Godfrey	
- p=1	0.13
- p=3	0.90

Table 2: OLS estimates for DY

First, we tested if the observations in the individual sample had predictability on stock returns. The first lag of DY was plotted in a scatter plot against stock returns to get an indication of the relationship between the two variables.

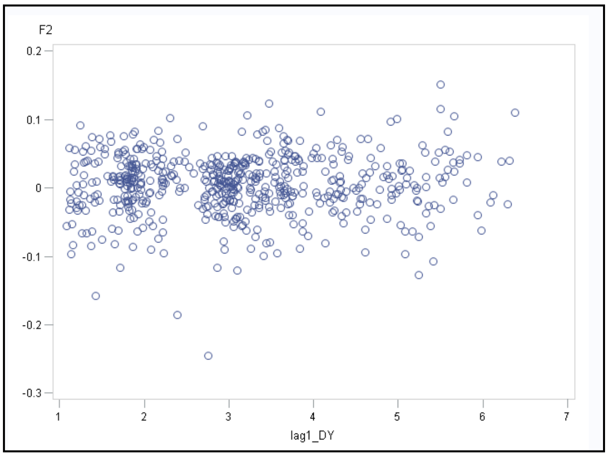


Figure 19: Relationship between stock returns and first lag of dividend yield

According to the scatter plot, the observations form no specific linear relationship, but more a random relationship where the observations are spread around. The mean looks close to zero, and we make a presumption that the coefficient of DY is very small and it is hard to say if it will be positive or negative. As table 2 shows, the coefficient is 0.0019, which is close to zero and positive. The low R^2 of 0.28%, is a suggestion that the DY explains a small fraction of stock returns.

For us to be able to rely on the t-statistic given in table 2, we have to check for the assumptions of an OLS regression. To test for normality in the error term, we run a time series regression analysis with autoregressive errors to get the output of the JB normality test. The result was a statistic of 175.38, which is significantly higher than the critical value of 3.84, so we reject the null hypothesis of normality in the error term. The nonfulfillment of the normality assumption may not be as critical as it appears because we know that the OLS point estimates still remain unbiased (Gujarati & Porter, 2009, p.318). In the case of small sample sizes, the t and F test require that the error term follows a normal distribution, otherwise they will be invalid. This will not be a problem since our sample size is rather big throughout all of the testing.

We check for heteroscedasticity by looking at the residual plot between the first lag of DY and the residuals of stock return.

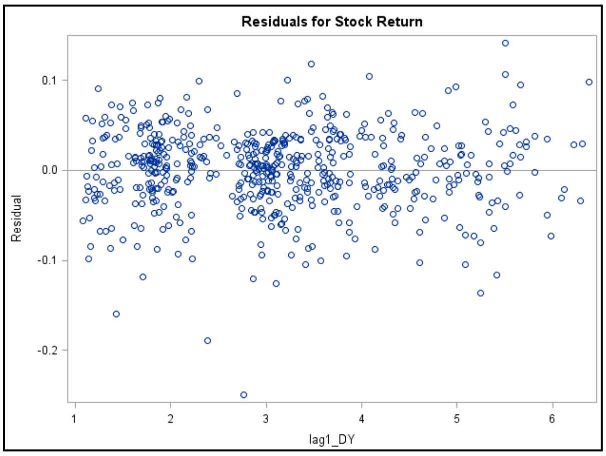


Figure 20: Residuals for stock return against the first lag of DY

If heteroscedasticity is present then there will be a pattern of variance in the plot above. As far as we can tell the observations show no sign of a pattern and accordingly there is not a case of heteroscedasticity.

The ACF and PACF graphs are plotted for both stock returns and DY to get an impression of whether the time series are stationary. For stock return, the lags in the graphs are nicely behaved and various lags move around zero. For DY, the lags are slowly and linearly declining, which is a sign that the time series are nonstationary in DY.

To test for autocorrelation, we have to find out whether the error terms are independent over time. Therefore, we plot the residuals against time. As figure 21 indicates, the observations are randomly scattered around, and the error term of the different observations seem to be uncorrelated. This suggests that autocorrelation is not present.

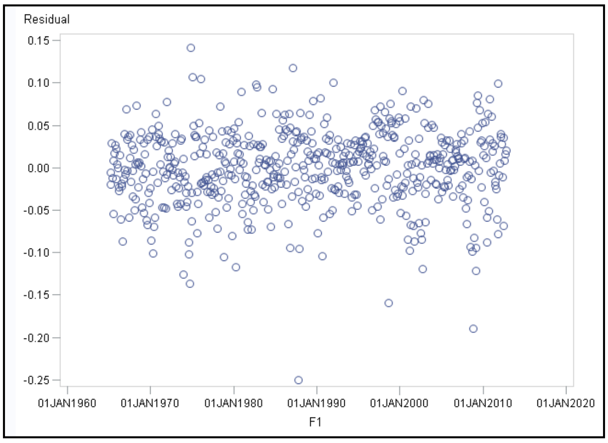


Figure 21: Residuals over time

We also run a numerical test to check for autocorrelation. We can either apply the DW test or the BG, and we choose BG since we fail to fulfill some of the assumptions of DW. Our tests of normality shows that our explanatory variable has an error term that is not normally distributed, which is in conflict with the fourth assumption. Since we are using the lagged variable in our linear regression we will always miss the first value in the first samples, according to assumption number 6, the d statistics do not allow for missing values. Regarding the order on the autoregressive scheme, we do not know exactly which order the disturbance is following and DW will not detect any potential higher-order. Since we have monthly and quarterly data we also want to test for ρ >1.

Since the variables do not fulfill assumption 3, 4, and 6 of the DW test, we choose to run the BG test which will be a more appropriate test for our sample. With the BG test we have the opportunity to test for higher-order autoregressive schemes and nonstochastic regressors, making it a more general test than DW. Both the DW and BG statistics will be reported in the results, but our conclusion about autocorrelation will be based on the results from the BG test.

In the BG test we run the test for both ρ =1 and ρ =3 since we have monthly observations for DY. The critical values of the χ^2-distributed test are:

	Significance level	
Df = ρ	0.1	0.05
1	2.70554	3.84146
3	6.25139	7.81473
4	7.77944	9.48773

Table 3: The critical values of the BG test

The BG statistic is 1.83 for ρ =1 and 3.82 for ρ =3, so both of these values are below the critical value at the 5% and 10% significance level. Hence, we keep the null hypothesis of no autocorrelation and the numerical test confirms that there is no autocorrelation.

In the t-test the null hypothesis is not rejected when the coefficient of DY is equal to 0. In a two-sided alternative hypothesis the coefficient is different from 0. In a one-sided alternative hypothesis the coefficient is either smaller or bigger than 0, depending on how the alternative hypothesis is formulated. The t-values computed in this thesis will be tested for both one-sided and two-sided alternative hypothesis at both the 5% and 10% significance level. When applying the one-sided test the coefficient will be considered significant if it rejects the null when $H_1: \beta_1 < 0$ or $H_1: \beta_1 > 0$, since the objective is to investigate any form of predictability on stock returns.

T-test					
Mark		Significance level	60 obs	120 obs	∞
**	two-sided	5%	2	1.98	1.96
*	two-sided	10%	1.671	1.658	1.645
	one-sided	5%	1.671	1.658	1.645
°	one-sided	10%	1.296	1.289	1.282

Table 4: The critical values of the t-test

In the individual sample there are 572 observations and the t-value is 1.26, so we cannot reject the null hypothesis, which means that DY does not contain predictive ability.

Next, we tested the observations from the individual sample excluding the financial crisis. For this sample period, the scatter plot of DY on stock return looks very much the same. We believe the coefficient of the first lag of DY to be small, and it is close to zero at a value of 0.0017. The R^2 is lower than in the individual sample at 0.25%. The JB statistic is 179.69, which is higher than in the individual sample, and notably higher than 3.84, so we reject the null of normality in the error term.

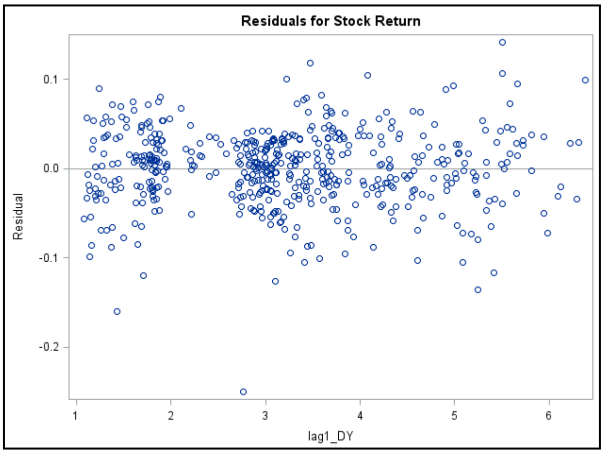


Figure 22: Residuals for stock return against the first lag of DY

The residual plot between the first lag of DY and the residuals above show no sign of pattern in the variance, and hence no sign of heteroscedasticity. The ACF and PACF of both stock return and DY show the same results as in the independent sample, which is stationary stock returns and nonstationary DY.

When we check for autocorrelation, the scatter plot between residuals and time is similar to the plot from the individual sample. Hence, there is no sign of autocorrelation. We find a BG statistics of 0.13 for ρ =1 and 0.90 for ρ =3, so we keep the null hypothesis of no autocorrelation in both the 5% and 10% significance case. The t-statistic is 1.14, which is lower than 1.282 and we cannot reject the null hypothesis of no explanatory ability.

Our results from the individual sample, with and without the financial crisis, show no signs of predictability in DY.

4.1.2 PRICE-EARNINGS

Individual Sample	
Price-Earnings (PE)	
β	0.00004
Standard Error	0.0001
t-value	0.34
p-value	0.7345
R-square	0.0002
F	0.12
Jarque-Bera	171.61
Durbin Watson	1.89
Breusch-Godfrey	
- p=1	1.64
- p=3	3.00

Individual sample excl. Financial Crisis	
Price-Earnings (PE)	
β	-0.0004
Standard Error	0.0003
t-value	-1.56°
p-value	0.1200
R-square	0.0051
F	2.43
Jarque-Bera	164.74
Durbin Watson	1.97
Breusch-Godfrey	
- p=1	0.11
- p=3	0.69

Table 5: OLS estimates for PE

We perform the same procedure with the PE ratio as we did with DY. First, we test if the individual sample can predict stock returns. We make a scatter plot with stock return along the vertical axis, and the first lag of PE along the horizontal axis.

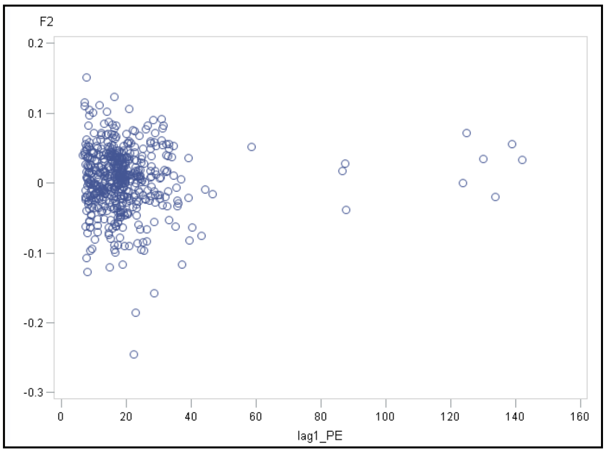


Figure 23: Relationship between stock returns and the first lag of PE

The observations are randomly scattered around, except for a few very large outliers. These outliers may have an impact on the assumptions of OLS since OLS is very sensitive to large outliers, and especially the residuals. Our hypothesis for the coefficient of PE is that it is relatively small and positive. As the table above presents, the coefficient is 0.00004 and the R^2 is 0.02%, and our prediction is correct.

We check for the normality assumption by investigating the JB statistic, and it is high at 171.61, so we reject the null of normality in the error term since the statistic is higher than the critical value of 3.84.

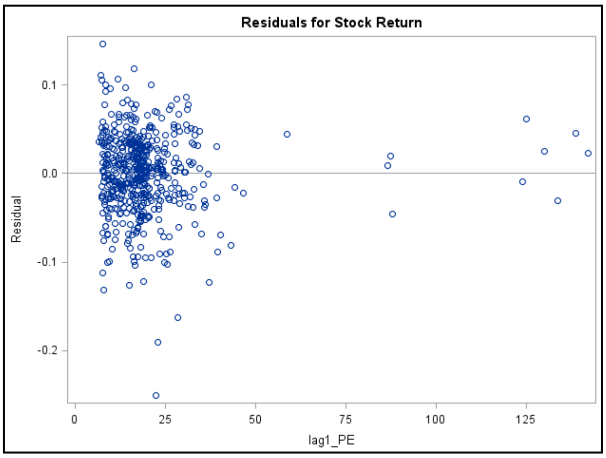


Figure 24: Residuals for stock return against the first lag of PE

As previously mentioned, the large outliers will possibly have an effect on the residuals, and here it is confirmed by the scatter plot above. Despite the outliers, the scatter plot still shows that the observations are randomly scattered around and there is no sign of heteroscedasticity.

The ACF and PACF of stock return illustrate nicely behaved observations again, not surprising since the observations are almost identical to the ones used when testing DY. For PE the lags are decreasing quickly, but several lags in both the ACF and PACF are outside the significance band. Accordingly, this is an indication that stock returns are stationary, while PE is nonstationary.

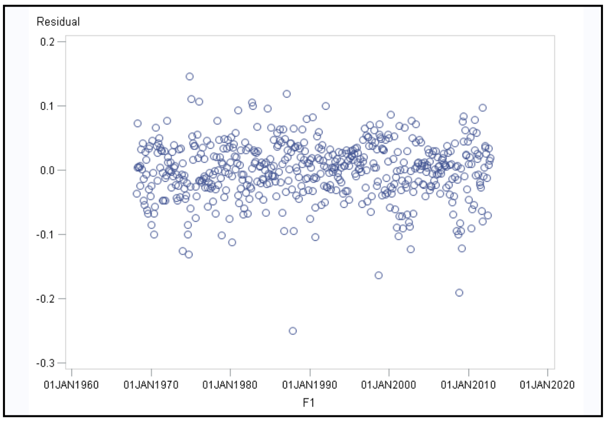


Figure 25: Residuals over time

The residuals plotted against time show no indication of autocorrelation, and the BG statistic is 1.64 for ρ=1 and 3.00 for ρ=3, hence BG confirms that there is no autocorrelation. The t-statistic is 0.34, therefore we cannot reject the null hypothesis of no predictability in PE.

The next sample period we tested was the individual sample without the financial crisis. When we plot the scatter plot between the first lag of PE and stock returns the large outliers from the previous sample are no longer present, and we can see that the value of the x-axis changes.

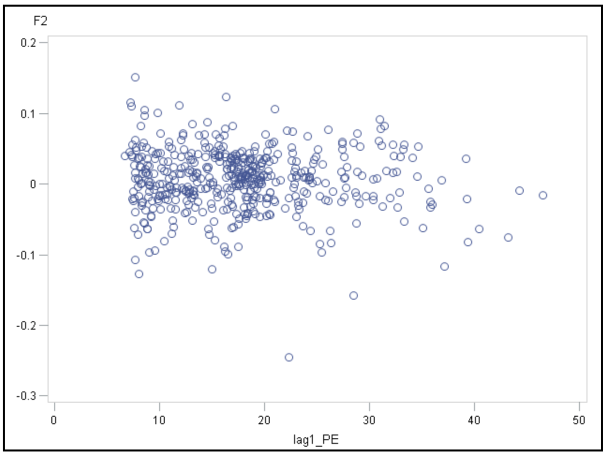


Figure 26: Relationship between stock returns and the first lag of PE

The coefficient of PE is -0.0004, and the R^2 is 0.51%, which means that PE explains 0.51% of stock returns.

We observe a slight improvement in the JB at 164.74, but it is still notably higher than 3.84, so we reject the null hypothesis of normal distribution in the error term. The observations we made from the scatter plot of the first lag of PE on stock returns are the same observations we make when looking at the scatter plot of the first lag of PE on the residuals. The large outliers are no longer present in the plot, and there is no sign of pattern in the variance, so we conclude with homoscedasticity.

The ACF and PACF of the stock return is once again nicely behaved, while for PE they are declining slower than in the case with the individual sample and at a more constant pace. All the lags are outside the significance band. The stock returns are stationary and the PE seem to be nonstationary. We plot the residuals against time to get a sense of whether the error terms are independent over time. The result is quite similar to the plot from the individual sample, and autocorrelation is not present. The BG statistic is 0.11 for ρ=1 and 0.69 for ρ=3, so we conclude with no autocorrelation. The t-statistic is -1.56, and we reject the null of no predictability at the 10% significance level where the one-sided alternative hypothesis is $H_1 = \beta_1 < 0$.

In the sample excluding the financial crisis, PE ratio shows signs of predictability, while the individual sample indicates that the PE ratio does not have any predictive ability in stock returns.

4.1.3 CONSUMPTION WEALTH RATIO

Individual Sample	
Cay	
B	0.0457
Standard Error	0.0425
t-value	1.07
p-value	0.2837
R-square	0.0058
F	1.16
Jarque-Bera	47.52
Durbin Watson	1.83
Breusch-Godfrey	
- p=1	1.31
- p=4	2.59

Individual sample excl. Financial Crisis	
Cay	
B	0.0667
Standard Error	0.0477
t-value	1.40°
p-value	0.1635
R-square	0.0112
F	1.96
Jarque-Bera	47.91
Durbin Watson	1.94
Breusch-Godfrey	
- p=1	0.15
- p=4	0.98

Table 6: OLS estimates for CAY

We are now working with quarterly observations, so the stock returns are different from the monthly returns. Cay is also calculated in log, so this is a log-log model while the others are log-lin models. The individual sample is used to test for predictive ability in cay. We plot the first lag of cay on stock return in a scatter plot to get an idea of what kind of relationship exists between cay and stock returns.

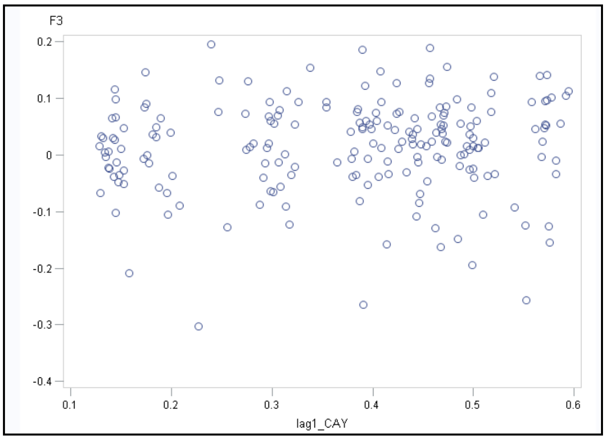


Figure 27: Relationship between stock returns and the first lag of CAY

Our impression is that the relationship is influenced by random observations and does not take any specific form. Our prediction regarding the coefficient of cay is that it is small and positive. When we run the linear regression, the result is that the coefficient value is 0.0457, and the R^2 is 0.58%.

The JB statistic is 47.52, which is significantly smaller than the other variables we have checked for this far, but it continues to be much larger than 3.84 and we have to reject the null hypothesis of normality in the error term.

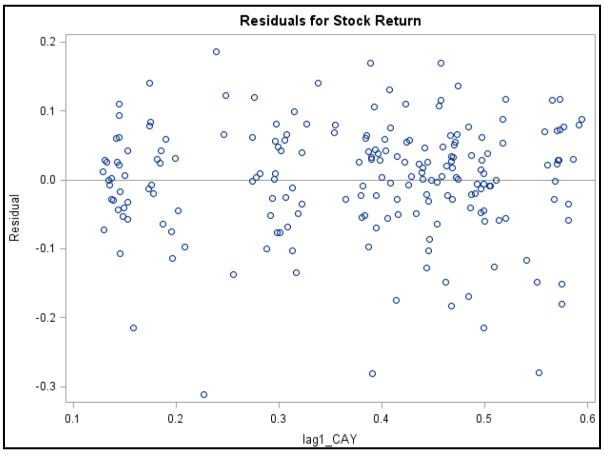


Figure 28: Residuals against the first lag of CAY

The scatter plot of the first lag of cay on the residuals is very familiar to the scatter plot of the first lag of cay on the stock returns. The observations are randomly scattered, and there are no indication of heteroscedasticity since there is no pattern in variance. The ACF and PACF of stock returns are also nicely behaved for the quarterly observations, while the ACF and PACF of cay are slowly decreasing with a large number of lags significant different from zero. There is a possibility that cay is nonstationary.

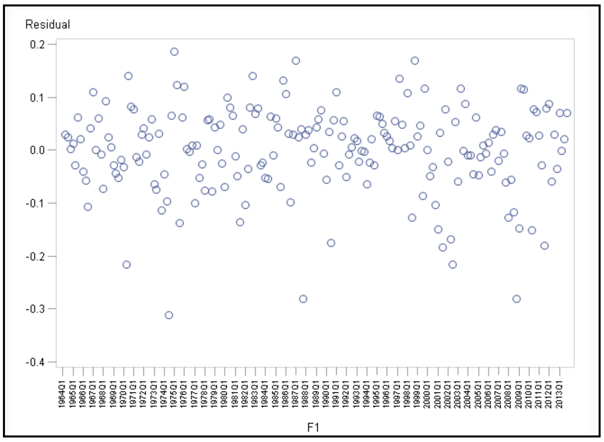


Figure 29: Residuals over time

The residuals are plotted against time to check for autocorrelation. There does not exist any dependence between the residuals and time, hence there are no sign of autocorrelation. We run the BG test to get a better indication of the presence of autocorrelation, and the statistic is 1.31 for ρ=1 and 2.59 for ρ=4, so we keep the null hypothesis of no autocorrelation. We look at ρ=1 and ρ=4 since cay is observed on a quarterly basis. The t-statistic is 1.07, which is below the critical values, and consequently we do not reject the null of no explanatory ability.

When we exclude the financial crisis from the individual sample, we get a scatter plot where the observations that had a cay value over 0.5 seem to have disappeared. This may result in a better coefficient of cay and t-value than in the previous sample period. The coefficient is 0.0667, and the R^2 is 1.12%.

The normality assumptions is controlled for by looking at the JB statistic at 47.91, which is a bit higher than in the individual sample, and significantly higher than 3.84, so we reject the null hypothesis. The scatter plot of the first lag of cay on residuals is identical to the one from the previous sample, excluding some observations. There is still no sign of heteroscedasticity.

For stock returns, the ACF and PACF behave nicely and look normal. The ACF and PACF of cay are very similar to the ones for cay in the individual sample, which are slowly decreasing lags and hints of nonstationary time series. The residuals plotted against time show indication of no autocorrelation, and so does a BG statistic of 0.15 for ρ=1 and 0.98 for ρ=4. We keep the null hypothesis of no autocorrelation. The t-statistic is 1.40, which is higher than 1.289 when we test for the one-sided alternative hypothesis $H_1 = \beta_1 > 0$, so we reject the null hypothesis of no predictive ability.

4.1.4 INTEREST RATE

Individual Sample		Individual sample excl. Financial Crisis	
Interest Rate		**Interest Rate**	
B	-0.0023	B	-0.0037
Standard Error	0.0018	Standard Error	0.0018
t-value	-1.25	t-value	-2.04**
p-value	0.2120	p-value	0.0421
R-square	0.0032	R-square	0.0099
F	1.56	F	4.16
Jarque-Bera	182.65	Jarque-Bera	168.04
Durbin Watson	1.91	Durbin Watson	2.02
Beusch-Godfrey		Breusch-Godfrey	
- p=1	0.94	- p=1	0.03
- p=3	3.03	- p=3	0.99

Table 7: OLS estimates for interest rate

First, we look for predictability in the individual sample. We plot the first lag of interest rate on stock returns in a scatter plot, and the observations are concentrated around specific values, interest rates are rarely extremely high or extremely low. So the scatter plot below makes a cloud of observations around 0 in stock returns and 0 in interest rate.

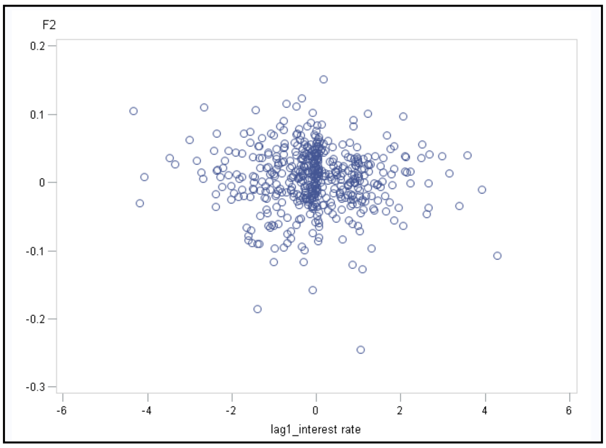


Figure 30: Relationship between the stock returns and first lag of interest rate

Our prediction for the coefficient of interest rate is that it is slightly low and positive. It turns out that the coefficient of the first lag of interest rate is -0.0023, and the R^2 is 0.32%. The JB statistic is high at 182.65, which is notably higher than 3.84, and we reject the null of normality in the error term.

We make a scatter plot of the first lag of interest rate on residuals and the observations resemble a cloud in the plot. The observations seem to concentrate around 0 on both the vertical and horizontal axis. There is no sign of a pattern in variance, and hence no sign of heteroscedasticity.

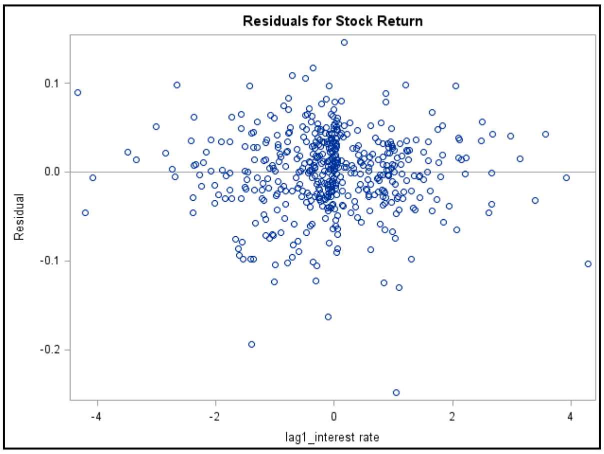


Figure 31: Residuals against the first lag of interest rate

Next, we look at the ACF and PACF of both stock returns and interest rate. As in previous results the ACF and PACF of stock return are nicely behaved. The ACF of interest rate is decreasing in a wave pattern, and we can conclude with signs of nonstationary time series in interest rate.

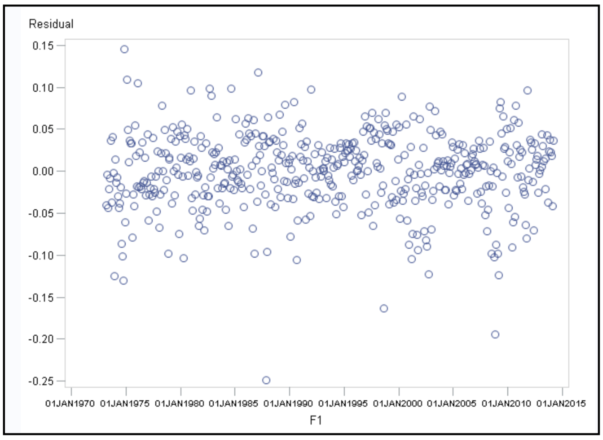


Figure 32: Residuals over time

The scatter plot of residuals on time show no indication of dependence between residuals and time, hence the conclusion is no autocorrelation. The BG statistic of 0.94 for ρ=1 and 3.03 for

ρ=3 confirms that autocorrelation is not believed to be present. The t-statistic is -1.25, and once again we fail to reject the null hypothesis of no predictability in the individual sample.

Further, we test for predictability of the interest rate in the financial crisis sample. The scatter plot between the first lag of interest rate on stock returns looks very much the same as the one from the individual sample. The only difference is fewer observations around zero in interest rate.

We run the linear regression and the result is a coefficient of -0.0037 for interest rate and a R^2 of 0.99%. The JB statistic of 168.04 is an improvement from the individual sample, but still drastically higher than 3.84, and the null hypothesis of normality in the error term is rejected. The scatter plot between the first lag of interest rate and the residuals are almost identical to plot for the individual sample, except for a few observations around zero in interest rate. Heteroscedasticity is not present.

The resemblance to the individual sample continues in the ACF and PACF of both stock returns and interest rate. The ACF of interest rate is declining in a wave pattern, while it looks normal for stock returns. Nonstationary time series may be the case for interest rate, whereas the time series for stock returns seems stationary. Both the scatter plot of residuals on time and the BG test show no indication of autocorrelation.

The t-statistic is -2.04, which is larger than -1.96 in a two-sided test at the 5% significance level and we reject the null hypothesis of no predictive ability. Hence, when the financial crisis is excluded from our sample period, the interest rate has predictive ability on stock returns.

4.1.5 MOMENTUM

Individual Sample	
Momentum	
B	-0.0083
Standard Error	0.0133
t-value	-0.63
p-value	0.5309
R-square	0.0008
F	0.39
Jarque-Bera	179.15
Durbin Watson	1.89
Breusch-Godfrey	
- p=1	1.23
- p=3	2.92

Individual sample excl. Financial Crisis	
Momentum	
β	-0.0052
Standard Error	0.0152
t-value	-0.34
p-value	0.7335
R-square	0.0003
F	0.12
Jarque-Bera	177.68
Durbin Watson	1.99
Breusch-Godfrey	
- p=1	0.01
- p=3	0.70

Table 8: OLS estimates for momentum

The individual sample is applied first to test whether the momentum effect in the market can predict stock returns in the future. The scatter plot between the first lag of momentum effect on stock returns makes us believe there is no clear relationship between the two variables. The observations gather in a cloud around zero on both axis, and there are a few large outliers which pushes the cloud of observations to the right in the plot.

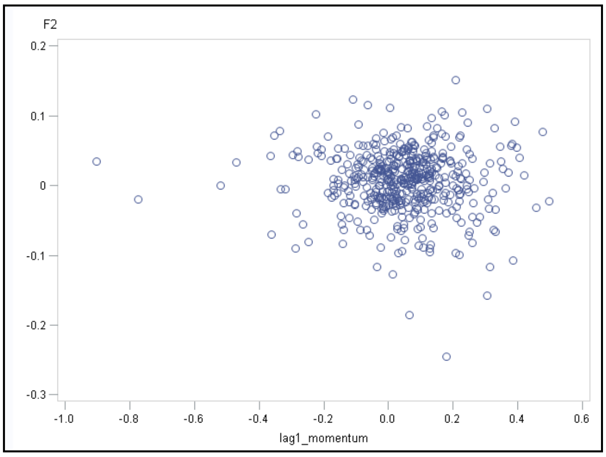


Figure 33: Relationship between stock returns and the first lag of momentum

Our prediction for the coefficient of momentum effect is that it will be small and positive. The linear regression, on the other hand, predict the coefficient to be -0.0083 and the R^2 to be

0.08%. The JB statistic is 179.15, which is very high and significantly higher than 3.84, so we reject the null of normal distribution.

The scatter plot between the first lag of momentum and the residuals is also affected by the outliers, but the majority of the observations lie around zero on both axis and form a cloud of observations. There is no patter in variance, and no sign of heteroscedasticity.

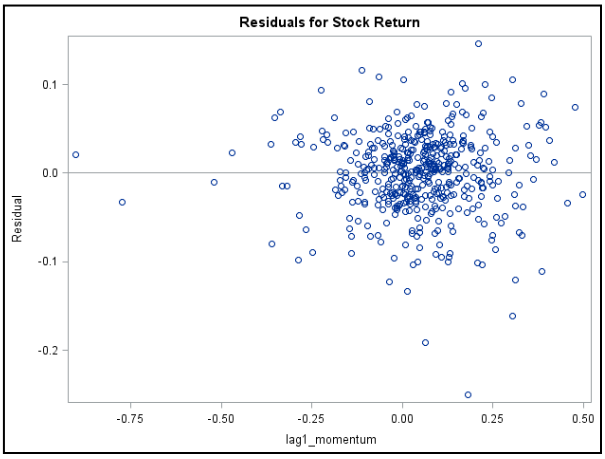


Figure 34: Residuals against the first lag of momentum

The ACF and PACF of stock returns are properly behaved, and for the momentum effect the ACF has four lags outside the significance band, which indicate weak signs of nonstationary time series. We make a scatter plot of residuals over time, and the observations are randomly scattered around with seemingly a mean close to zero. The BG statistic confirms that we can exclude the case of autocorrelation with a value of 1.23 for ρ=1 and 2.92 for ρ=3.

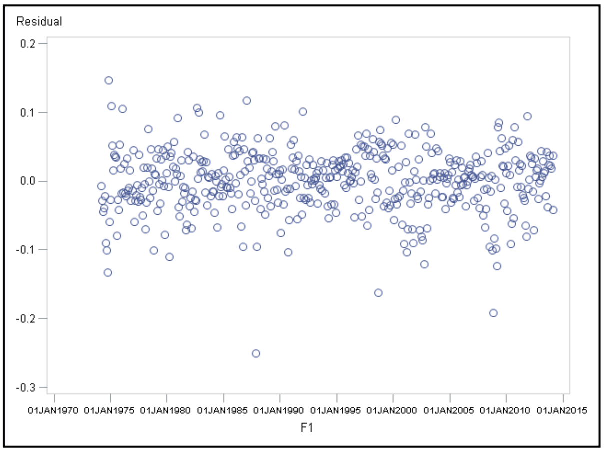


Figure 35: Residuals over time

The t-statistic is -0.63 which is lower than 1.282, consequently we cannot reject the null hypothesis of no explanatory ability in stock returns. Next, we replace the individual sample with the financial sample and look for predictability of momentum effect on stock return in a shorter sample period.

We make a scatter plot of the first lag of momentum effect on stock returns to investigate the relationship between the two variables. By using a shorter sample period we have managed to eliminate the outliers from the sample. The observations are more spread around due to smaller values along the horizontal axis.

The coefficient of momentum is -0.0052, and the R^2 is 0.03%. The JB statistic of 177.68 is an improvement from the individual sample, but still significantly higher than the critical value at 3.84, hence we reject the null of normal distribution. The scatter plot between the first lag of momentum and the residuals show the same tendencies as the scatter plot between momentum and stock returns. The outliers have been removed, and the values on the axis have changed. Heteroscedasticity is not present.

This time the ACF of momentum has a notch better lags, but there are still four outside the significance band. The ACF and PACF for stock returns are still looking good and the time series seem stationary. The residuals show no dependence over time in the financial crisis sample either, and the BG statistic is 0.01 for ρ=1 and 0.70 for ρ=3. Hence, we keep the null of no autocorrelation. The t-statistic is -0.34, and we fail to reject the null hypothesis in this sample as well.

4.1.6 SHARPE RATIO

Individual Sample	
Sharpe Ratio	
B	0.0002
Standard Error	0.0001
t-value	1.28
p-value	0.1995
R-square	0.0034
F	1.65
Jarque-Bera	180.87
Durbin Watson	1.92
Breusch-Godfrey	
- p=1	0.83
- p=3	2.95

Individual sample excl. Financial Crisis	
Sharpe Ratio	
β	0.0003
Standard Error	0.0001
t-value	1.82*
p-value	0.0699
R-square	0.0079
F	3.30
Jarque-Bera	169.62
Durbin Watson	2.02
Breusch-Godfrey	
- p=1	0.05
- p=3	1.04

Table 9: OLS estimates for Sharpe ratio

The individual sample is the starting point when we test for predictability on stock returns in the Sharpe ratio. The scatter plot between the first lag of Sharpe ratio and stock returns is another case of gathering of observations in a cloud shape. There is no clear relationship between stock returns and the Sharpe ratio.

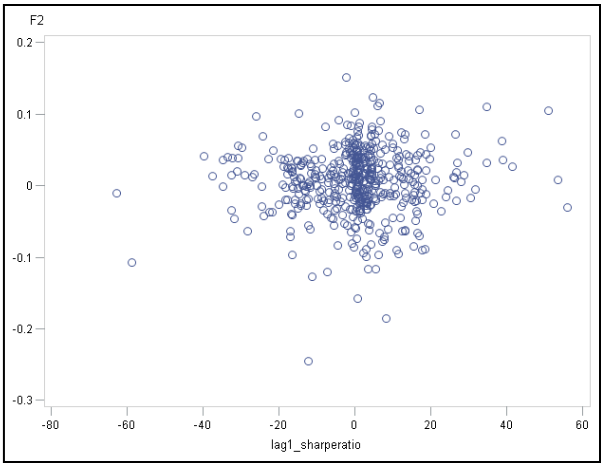


Figure 36: Relationship between stock returns and first lag of Sharpe ratio

Our prediction of the coefficient of Sharpe ratio is that it is small and positive. The predictions were accurate as the coefficient of Sharpe ratio is 0.0002, and the R^2 is 0.34%. With a JB statistic of 180.87, we reject the null hypothesis of normality in the error term since it is much larger than 3.84.

The first lag of Sharpe ratio is also plotted in a scatter plot against the residuals to get an idea of whether heteroscedasticity is present. The observations are randomly scattered around and there is no indication of pattern in variance, so we conclude with no heteroscedasticity.

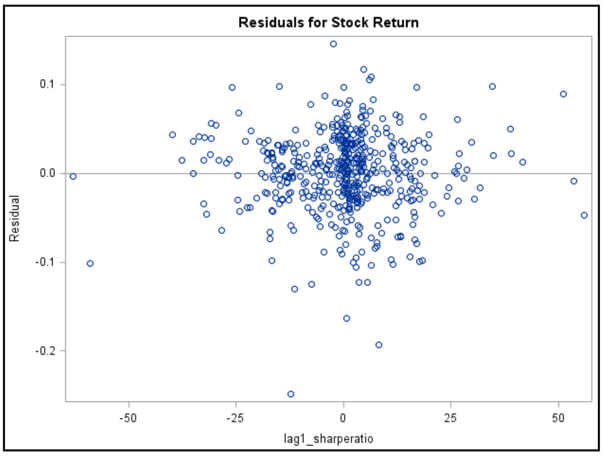


Figure 37: Residuals against the first lag of Sharpe ratio

The ACF and PACF of stock returns are stationary, while the ACF of the Sharpe ratio is declining in a wave pattern and show signs of nonstationary time series. We make a scatter plot of the residuals on time, and the observations do not indicate dependence between residuals and time. The BG test confirms our prediction of no autocorrelation with a statistic of 0.83 for ρ=1 and 2.95 for ρ=3.

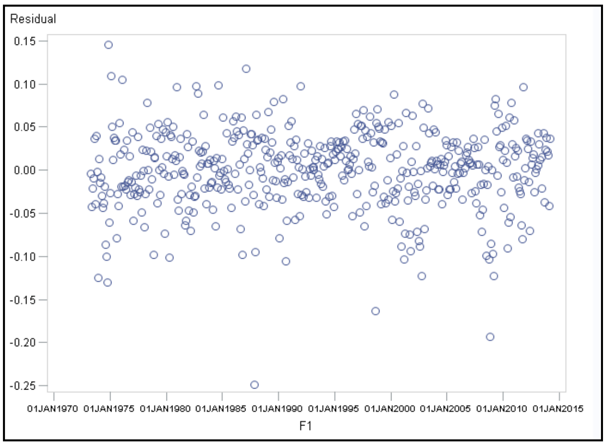


Figure 38: Residuals over time

The t-statistic is 1.28, which means that the Sharpe ratio has no ability to predict stock returns in the individual sample period. Further, we test for predictive ability in the financial crisis sample period. The scatter plot between Sharpe ratio and stock returns is similar to the one from the individual sample except for a few observations around zero in Sharpe ratio.

The coefficient of Sharpe ratio is 0.0003, the R^2 is 0.79%, and the JB statistic is 169.62. The statistic for normality has improved compared to the previous sample period tested, but is still drastically larger than 3.84, and we reject the null of normal distribution in the error term.

The new scatter plot between the first lag of Sharpe ratio and the residuals is very much the same as in the individual sample, apart from a few observations around zero in Sharpe ratio. Hence, heteroscedasticity is not present here either. The ACF of Sharpe ratio is identical to the one from the individual sample, which means that there are signs of a nonstationary time series. For stock returns, both the ACF and PACF are well behaved.

The residuals are plotted against time and the observations are randomly scattered around indicating no autocorrelation between the residuals. This is confirmed by a BG of 0.05 for ρ=1 and 1.04 for ρ=3, so we do not reject the null hypothesis of no autocorrelation. The t-statistic is 1.82, which means we can reject the null of no predictability at the 10% significance level.

We conclude with no predictive ability in Sharpe ratio on stock returns in the individual sample and predictability in the financial crisis sample.

4.1.7 SUMMARY

The financial crisis sample proved to be the only sample to show any predictability in some of the variables. The financial crisis sample excludes observations after December 31th 2007 from the individual sample, and hence there are indications that the financial crisis of 2008 may affect our results. Variables such as PE have some very large outliers in the period after 2007, which may cause different results and may be explained by the unusual circumstances that occurred during the financial crisis.

Both of the sample periods have not satisfied the assumption of normality in the error term for any of the variables. There have been no signs of heteroscedasticity or autocorrelation. Autocorrelation have been checked for by looking at the plot between residuals and time, in addition to running the BG test. If we had been able to rely on the results by the DW test, it would have made the same conclusions as the BG test does in both the samples on all of the variables.

Stationary time series have also been a problem, but according to Gujarati and Porter (2009) there is only reason to be concerned if both of the variables, the explanatory and the dependent, have nonstationary time series. We want to avoid the spurious regression problem that may arise from regressing a nonstationary time series on one or more nonstationary time series (Gujarati & Porter, 2009, p.760). In our case, the time series of stock returns are always stationary, hence the spurious regression problem is not an issue. In inspiration from Lewellen (2004) we ran the following regression:

(4.2) $$x_t = \varphi + \rho x_{t-1} + \mu_t$$

Empirical tests rely heavily in on the assumption that ρ is not greater than one. According to statistics ρ should preferably be equal to one, but prior studies emphasis that as long as it is below one it should be valid. We run the regression for both individual and financial crisis samples on all variables, and ρ never exceeds one. We conclude with stationarity in the explanatory variables.

4.2 U.S. SUBSAMPLES

In the first tests we look at quite a large time sample, the longest period was almost 50 years. To test the robustness of the first completed tests we choose to look at subsamples, as Lewellen (2004) choose to do when he tested whether DY, B/M, and E/P can forecast stock return. When using this method with subsamples we can check whether the result from the long period is consistent with the results for a smaller time period, and also if some of the periods may have biased the results from the long sample.

Except from the change in time period the same explanatory variables and regression are used. We run the same tests as we did in the long sample to check the fulfillment of the OLS assumptions and if we have significant t-statistics. We have looked at the full sample period, the period where we have data for all six variables, and divided into three subsamples. Thus, we are left with three samples of 12 year each, 1974-1986, 1987-1999, and 2000-09:2012.

4.2.1 DIVIDEND YIELD

One of the econometrics issues we can possibly meet during our testing is autocorrelation. As described in the theory, one of the ways to remove the autocorrelation issue is to use the GLS method to transform our model. For DY, autocorrelation is present in the last subsample, so we do the transformation from OLS to GLS. Below the results for all the three subsamples is presented, the original results for the two first periods and the results from the GLS model for the last subsample.

Subsample 1974 - 1986	
Dividend Yield (DY)	
B	0.0077
Standard Error	0.0049
t-value	1.54°
p-value	0.1245
R-square	0.0154
F	2.39
Jarque-Bera	2.18
Durbin Watson	2.04
Breusch-Godfrey	
- p=1	0.07
- p=3	1.16

Subsample 1987 - 1999	
Dividend Yield (DY)	
β	-0.0018
Standard Error	0.0046
t-value	-0.40
p-value	0.6878
R-square	0.0011
F	0.16
Jarque-Bera	441.14
Durbin Watson	1.88
Breusch-Godfrey	
- p=1	0.19
- p=3	3.36

Subsample 2000 - 09:2012	
Dividend Yield (DY) *GLS*	
β	0.0125
Standard Error	0.0117
t-value	1.07
p-value	0.2847
R-square	0.0077
F	1.15
Jarque-Bera	14.82
Durbin Watson	1.93
Breusch-Godfrey	
- p=1	0.05
- p=3	3.57

Table 10: OLS and GLS estimates for DY

From the table we can see that the results vary over the different subsamples. The first period is the only case where we have a significant coefficient. The t-statistic is not high enough to exceed the critical value of the 10% two-sided test, but is higher than the critical value for the one-sided 10% at 1.289, so we conclude with significance at a 10% level for the one-sided test. In the other two subsamples we do not have t-statics to reject the critical value in neither of the significance levels, thus there is not any indication that DY predicts stock return. For the first subsample we see no evidence of autocorrelation and heteroscedasticity from the different statistics and plots we have made. By looking at the JB statistic we can say whether the error term is normally distributed. If the statistic is lower than the critical value of 3.84, we do not reject the normality assumption and can conclude that the error term is normally distributed. This subsample has a JB statistics of 2.18, which is lower than the critical value, and thus we have normally distributed error term in the first period. Compared to the first test with the long sample period for DY, the period 1974-1986, gives the best result looking at the t-value, JB statistics, R^2, and p-value.

Looking at the other two subsamples we observed that these samples have error terms that are not normally distributed based on the JB statistics, 441.14 and 14.82 are both higher than the critical value. The middle period stand out as the least favorable period and has a negative beta coefficient. There is no sign of heteroscedasticity and the heteroscedasticity assumption for the OLS regression is fulfilled.

Regarding autocorrelation we look at the figure where we have plotted the residual against time. From the plots we observe random observations and this indicate that there is no problem related to autocorrelation. Most of the observations are located between 0.1 and -0.1. In addition to the graphical investigation we run the numerical BG test. Based on the statistics reported in table 10 we keep the hypothesis of no autocorrelation in the middle period. As mentioned above this is not the case for the last period, and based on the original results we had to reject the null hypothesis for ρ=1 and ρ=3, both on a 5% and 10% level. After the transformation from OLS to GLS the autocorrelation problem is removed and based on the statistics from the BG test we can now keep the null for both ρ=1 and ρ=3. Below the BG test results from both the OLS regression and GLS regression are presented.

Godfrey's Serial Correlation Test		
Alternative	LM	Pr > LM
AR(1)	3.8784	0.0489
AR(2)	4.6681	0.0969
AR(3)	8.4928	0.0369
AR(4)	9.1830	0.0567

Godfrey's Serial Correlation Test		
Alternative	LM	Pr > LM
AR(1)	0.0529	0.8181
AR(2)	1.1460	0.5638
AR(3)	3.5662	0.3123
AR(4)	4.6411	0.3261

Table 11: BG-statistics before and after GLS transformation

A common feature for all the three samples is that the DY time series seems to be nonstationary. Nonstationary time series are characterized by lags outside of the significant area in the autocorrelation table. The three samples all have high values at lag 1 and then there is a slow decline in the lags, which is a typical sign of nonstationarity. The case that shows the weakest sign of nonstationarity is the first subsample with 8 lags outside the significant area and fastest decline in the lags, then period three with 9 lags, and the middle period is the worst case where we have 14.

4.2.2 PRICE-EARNINGS

Subsample 1974 - 1986	
Price-Earnings (PE)	
β	-0.0012
Standard Error	0.0015
t-value	-0.79
p-value	0.4333
R-square	0.0040
F	0.62
Jarque-Bera	2.18
Durbin Watson	2.06
Breusch-Godfrey	
- p=1	0.14
- p=3	0.94

Subsample 1987 – 1999	
Price-Earnings (PE)	
β	-0.0005
Standard Error	0.0006
t-value	-0.77
p-value	0.4416
R-square	0.0039
F	0.60
Jarque-Bera	420.72
Durbin Watson	1.87
Breusch-Godfrey	
- p=1	0.29
- p=3	2.68

Subsample 2000 - 09:2012	
Price-Earnings (PE)	
β	0.0002
Standard Error	0.0002
t-value	1.13
p-value	0.2598
R-square	0.0084
F	1.28
Jarque-Bera	17.23
Durbin Watson	1.72
Breusch-Godfrey	
- p=1	2.74
- p=3	7.42

Table 12: OLS estimates for PE

The scatter plot of stock returns against the lagged PE tells us that this time series consist of some outliers, and this is the situation for all the three subsamples. In the time period 2000-2012 we do observe some especially large outliers as shown in the figure below, which probably will lead to biased results. These big outliers are from the period between April 2009 and January 2010 and are an effect of the financial crisis in 2008. Apart from the obvious outliers, the observations looks randomly scattered around.

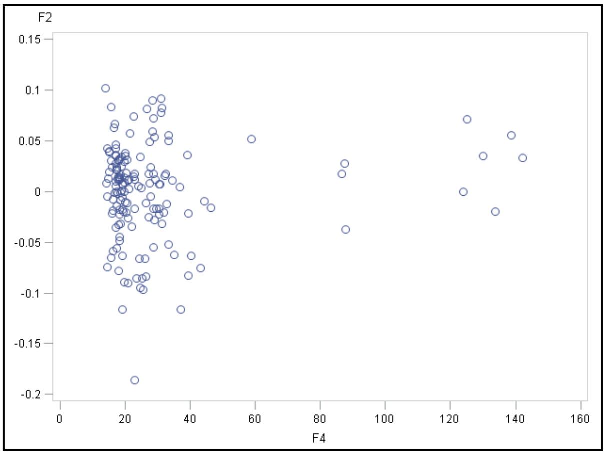


Figure 39: Relationship between stock return and the first lag of PE

After we have done the graphical investigation of the two variables we run the regression and look at the t-statistic. When analyzing PE as a predictive factor for stock return, we do not get a significant t-value in any of the three subsamples. Thus, we keep the null hypothesis, which means that PE cannot explain some of the variation in stock return. The coefficient is negative in the first two samples, but in the last period it is positive and the statistics is better than the two other periods.

To control for the heteroscedasticity assumption we look at the scatter plot of the residuals against the lagged PE variable to check if there is a constant variance for all of the observations. Looking at the plot, the observations seem to be randomly plotted around and there is no tendency to a pattern, hence we have a constant variance and no heteroscedasticity. As with the scatter plot above, we observe some outliers, but except from these the rest of the observations are random, and the outliers do not change the conclusion about heteroscedasticity.

The error term for PE is only normally distributed in the first subsample, in this sample we observe a JB statistic of 2.18 which is under the critical value and we can keep the null hypothesis of normally distributed error terms. In the two other periods the statistic is too high for us to keep the hypothesis.

From the graphical inspection of the ACF and PACF correlogram we observe patterns that indicate nonstationarity. In the first and second period the lags decline slowly at a constant rate, for the first period there is a switch in direction and the lags start to increase again at lag 20 and onwards. For the third period there is a relatively fast decline and the lags become negative after lag 10.

The autocorrelation plot of the residuals against time gives the indication that autocorrelation is not present. There is no pattern in the plot and the observations look random, so there is no sign that indicates that the error term seems to be correlated and dependent over time. The last period has random observations, but the plot has a butterfly shape as opposed to the two first periods. In addition we look at the BG test, to get some more concrete numbers to investigate autocorrelation. From the BG test we observed that we have the same problem regarding autocorrelation as we saw in the regression with DY. The first two subsamples display strong indication of no autocorrelation, but in the last period the BG test is only able to keep the null hypothesis at a 5% significance level.

4.2.3 CONSUMPTION WEALTH RATIO

Subsample 1974 - 1986	
Cay	
B	0.3365
Standard Error	0.2479
t-value	1.36°
p-value	0.1808
R-square	0.0362
F	1.84
Jarque-Bera	4.33
Durbin Watson	1.75
Breusch-Godfrey	
- p=1	0.71
- p=4	5.07

Subsample 1987 - 1999	
Cay	
B	0.5166
Standard Error	0.3676
t-value	1.41°
p-value	0.1661
R-square	0.038
F	1.98
Jarque-Bera	27.98
Durbin Watson	2.22
Breusch-Godfrey	
- p=1	1.71
- p=4	3.18

Subsample 2000 - 09:2012	
Cay	
β	0.4416
Standard Error	0.2939
t-value	1.5°
p-value	0.1393
R-square	0.0441
F	2.26
Jarque-Bera	7.79
Durbin Watson	1.76
Breusch-Godfrey	
- p=1	0.72
- p=4	1.58

Table 13: OLS estimates for Cay

The plots of stock return against cay do not indicate any significant relationship between these two variables, since the quarterly observations are randomly plotted around. From the plot, shown in figures 40, 41, and 42 below, it seems to be a slightly positive relationship, especially for the two first periods.

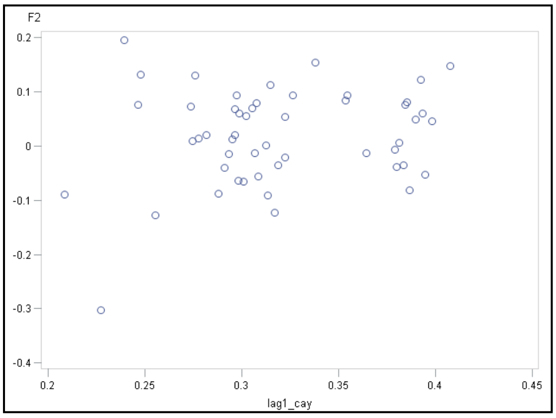


Figure 40: Relationship between X and Y 1974-1986

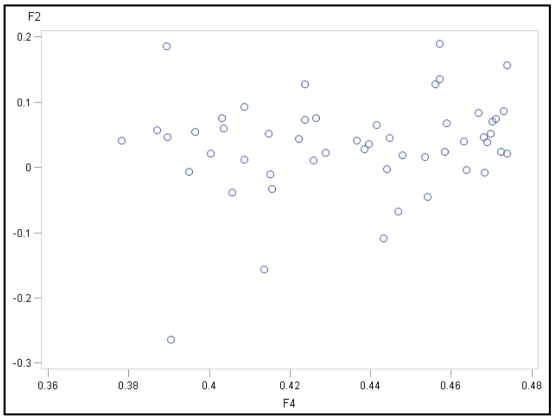


Figure 41: Relationship between X and Y 1987-1999

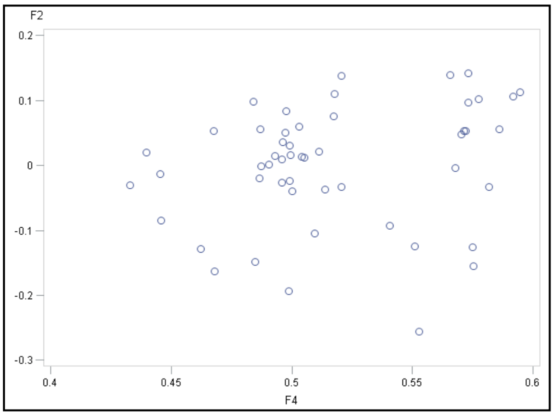


Figure 42: Relationship between X and Y 2000-Sept 2012

Looking at the t-statistics from the linear regression we see that we have positive coefficients in all the subsamples and high t-statistics, but not high enough to exceed the critical value of 2 for the two-sided 5% significance level. It is first when we run the one-sided test at a 10% significance level that we get significance results, on this level we reject the null hypotheses in all of the periods.

Cay is the variable that has the best results, compared to the other variables, for all the three periods and it is also the variable with minimum variability within the different subsamples. Among the variables we are testing, the R^2 reported for the cay variables is the highest R^2 we get from the regressions we have run, and the only case where we have a significant variable in all of the three subsamples.

Further we also check for heteroscedasticity and autocorrelation and there is no evidence of a systematic pattern in the different scatter plots. The BG test confirms this conclusion for autocorrelation based on the graphical test. For ρ=1 and ρ=4 we have values that correspond to keeping the null hypothesis. Thus, the assumption for the OLS regression regarding heteroscedasticity and autocorrelation is fulfilled. Looking at the JB statistics regarding normality we need a statistic that is lower than 3.84 to keep the null about normal distribution in the error term. For cay the following values, 4.33, 27.98, and 7.79 are reported for each of the corresponding periods. As we can see each of the values are higher than 3.84 and thus we reject the hypothesis of normal distribution in the error term.

The ACF correlogram tells us that the time series of cay are nonstationary. In all samples we observe a high first lag, and then the lags decrease slowly with a constant factor, and this is equal for all the three periods. For the first two periods there are 5 lags outside the significant area, and the last period is slightly better with 4 lags outside.

4.2.4 INTEREST RATE

	Subsample 1974 - 1986	Subsample 1987 - 1999	Subsample 2000 - 09:2012
	Interest Rate	Interest Rate	Interest Rate
β (B)	-0.0056	-0.0021	0.0088
Standard Error	0.0023	0.0047	0.0043
t-value	-2.42**	-0.45	2.04**
p-value	0.0168	0.6546	0.0430
R-square	0.0368	0.0013	0.0269
F	5.85	0.20	4.17
Jarque-Bera	2.29	409.26	9.19
Durbin Watson	2.16	1.87	1.76
Breusch-Godfrey			
- p=1	1.09	0.25	2.05
- p=3	1.96	2.84	6.31

Table 14: OLS estimates for interest rate

We plot stock return against the lagged cay and observe that the observations are randomly plotted around, without a clear tendency in the relationship between the two factors. The results we get when running the regressions show that we have a significant t-value in two of the periods, the first and the last. We have both a positive and negative significant coefficient, and they are significant at a 5% level, which also means that they would be significant at a 10% level.

The middle period performs poorly compared to the two others. The t-statistic is not close to being significant, and there is an extremely high JB statistic. With this high JB statistic we have to reject the hypothesis about normally distribution in the residual, the same is the case with the last period, where the statistic is 9.19 and the hypothesis has to be rejected. It is only the first subsample where we have a low enough JB statistic to keep the null hypothesis about normal distribution in the residual.

By looking at the residual plot we observe randomly scattered observations, and thus there is no sign of heteroscedasticity. From the DW test we do not observe any sign of autocorrelation, but the more appropriate BG test show sign for ρ=3 on a 10% level. The critical value for ρ=3 is 6.25, so we are at the limit where we have to reject the null hypothesis. Looking at the figures we again observe a butterfly shape of the observations in the scatter plot where we have residuals against time in the last period, while the first two have random observations without any indication of a pattern. With only a weak sign of autocorrelation there should not be any problem.

The interest rate time series are nonstationary in all of the subsamples. The autocorrelation correlogram has different patterns, but common to all is that they have a high lag 1, and the lags outside of the shaded significant area are both positive and negative lags. The first period has a wave-like pattern with 5 lags outside, while the two other periods have a slow decrease in the lags with respectively 8 lags outside for both periods, and both of these patterns are signs of nonstationary time series.

4.2.5 MOMENTUM

Subsample 1974 - 1986	
Momentum	
β	0.0343
Standard Error	0.0276
t-value	1.24
p-value	0.2153
R-square	0.0100
F	1.55
Jarque-Bera	2.20
Durbin Watson	2.10
Breusch-Godfrey	
- p=1	0.46
- p=3	1.09

Subsample 1987 - 1999	
Momentum	
B	-0.0212
Standard Error	0.0238
t-value	-0.89
p-value	0.3736
R-square	0.0051
F	0.80
Jarque-Bera	388.21
Durbin Watson	1.86
Breusch-Godfrey	
- p=1	0.36
- p=3	2.70

Subsample 2000 - 09:2012	
Momentum	
β	-0.0244
Standard Error	0.0204
t-value	-1.2
p-value	0.2337
R-square	0.0094
F	0.003
Jarque-Bera	16.83
Durbin Watson	1.73
Breusch-Godfrey	
- p=1	2.66
- p=3	7.39

Table 15: OLS estimates for momentum

The results from the regressions where momentum is the explanatory variable shows a positive coefficient for the first period and negative coefficients for the last two periods. It turns out that momentum has no explanatory power of stock return, none of the t-statistics are significant and we have to keep the null hypothesis. It is important to remember that we only have tested one of many strategies of momentum effect, so this conclusion may not be the same if we had tested for some of the other strategies.

Regarding the assumption about heteroscedasticity, the scatter plot of lagged interest rate on residuals shows that observations are random, and we can conclude that heteroscedasticity is not present. It seems that the momentum time series are nonstationary, for the different subsamples we get high and decreasing lags when looking at the ACF diagram. The lags have a waveform in all of the subsamples, but only a few lags outside the significance area, 4 lags for the first sample and 3 for the last two.

Based on the BG test we have no evidence for the presence of autocorrelation in the error term in the first two periods, we keep the null hypothesis for both levels of ρ and significance. When running the DW test we end up in the zone of indecision for the last subsample, so we might have a weak sign of autocorrelation. Checking this with the BG test, we have to reject the null at the 10% level for ρ=3, and we have evidence of autocorrelation. Looking at the residuals against time scatter plot the observations look randomly scattered around in a butterfly shape for this last period, with less volatile observations in 2006/2007. Again this is a pattern in the observations we have observed earlier when autocorrelation has been present.

For momentum as the explanatory variable we also have an extremely high JB statistic for the middle period which, in combination with the associated p-value for this statistic <.0001, are sufficient ground to reject the hypothesis of normal distribution in the error term. The last subsample has a considerably lower JB statistic, but it is still over the critical value of 3.84, so we have to reject the hypothesis also in this sample. Once again we have a situation where only the first period is normally distributed in the residuals.

4.2.6 SHARPE RATIO

Subsample 1974 - 1986	
Sharpe Ratio	
β	0.0004
Standard Error	0.0002
t-value	2.19**
p-value	0.0298
R-square	0.0305
F	4.81
Jarque-Bera	2.03
Durbin Watson	2.15
Breusch-Godfrey	
- p=1	1.00
- p=3	1.82

Subsample 1987 - 1999	
Sharpe Ratio	
B	0.0001
Standard Error	0.0004
t-value	0.35
p-value	0.7245
R-square	0.0008
F	0.12
Jarque-Bera	415.57
Durbin Watson	1.87
Breusch-Godfrey	
- p=1	0.23
- p=3	2.85

Subsample 2000 - 09:2012	
Sharpe Ratio	
β	-0.0007
Standard Error	0.0004
t-value	-1.72*
p-value	0.0882
R-square	0.0191
F	2.95
Jarque-Bera	12.03
Durbin Watson	1.73
Breusch-Godfrey	
- p=1	2.65
- p=3	6.85

Table 16: OLS estimates for Sharpe ratio

The scatter plot for stock return against Sharpe ratio is a plot of randomly plotted observations, with a few outliers. In the figures below you see the scatter plot with stock return on the vertical axis and lagged Sharpe ratio on the horizontal axis. In the first period we have the highest spread in the value of our observations of Sharpe ratio, the values go from -62.92 and up to 55.94.

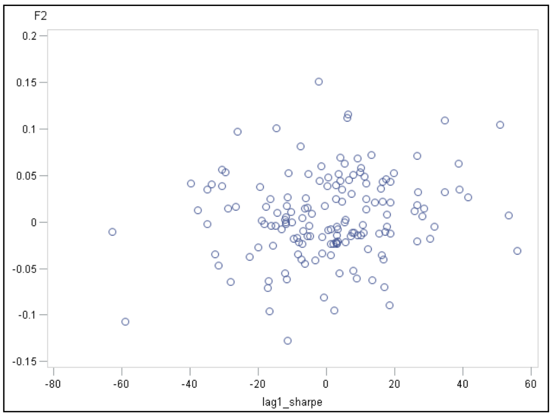


Figure 43: Relationsship between X and Y 1974-1986

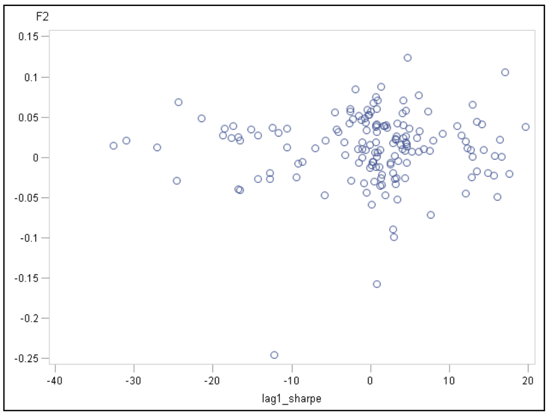


Figure 44: Relationship between X and Y 1987-1999

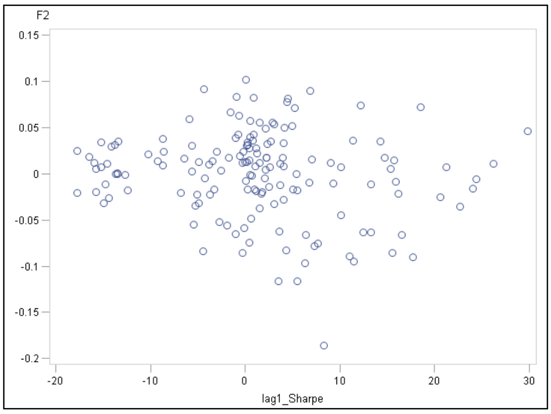


Figure 45: Relationship between X and Y 2000-Sept 2012

Checking our monthly observations on Sharpe ratio we have values between -63 and 56, -33 and 20, -18 and 30 in each of the three subsamples. The biggest difference is in the first where the spread is as high as 118 between the lowest and highest value. The Sharpe ratio coefficients are extremely low, slightly positive in the first two periods and slightly negative in the last period. One explanation behind this low level can be the high differences in the value of the Sharpe ratio which leads to a low average of the sample.

In the first subsample we reject the null hypothesis because of the significant t-statistic of 2.19. This statistic is significant both for a two-sided and one-sided test at a 5% significance level. We can also reject the null hypothesis in the last sample, here the t-statistic of -1.72 is significant at the 10% level for the two-sided and one-sided test. The result from the t-test in the middle period is not significant and we have to keep the hypothesis, and hence Sharpe ratio did not have any explanatory power on stock returns.

By plotting the residuals against the lagged Sharpe ratio the observations look random and there is no sign of heteroscedasticity. We also find these random observations when we look at the residuals against the predicted value. Sharpe ratio is also a variable where we have poor results about autocorrelation for the last sample period, but not in the first two. The residuals against time plot looks random without a specific pattern in the observations in the period 1974-1987 and 1988-1999, but for the last period we see the butterfly-like shape pattern. The poor result for autocorrelation is again the case with the 10% significance level and ρ=3, where the BG test statistic of 6.85 is higher than the critical value, indicating that we have to reject the null hypothesis. When we ran the DW test we got ambiguous results about autocorrelation, the DW statistics of 1.73 is in the zone of indecision. The DW test check for a first order autoregressive process (ρ=1), so this can explain why we do not get equal results from these two tests, in BG we only have to reject when ρ=3 and not ρ=1.

By observing the lags in the autocorrelation correlogram, the time series for Sharpe ratio looks nonstationary through all the three subsamples. The samples have lags both on the positive and negative side of the axis and the correlogram is characterized by positive lags gradually decreasing down to negative lags. The nonstationary time series get worse from period to period. The first period has 5 lags outside the significant area, and decrease with an uneven pace. The middle period has 7 lags outside and the last has 9, both have a slow constant decrease.

4.2.7 SUMMARY

From the discussion above we can summarize that the results are different between the three subsamples, except from cay that is fairly constant. The results from these tests are also different and not consistent with the first test we run with the long subsamples. From the results we have discussed above we have observed some common features for the different periods, we take a closer look at these below.

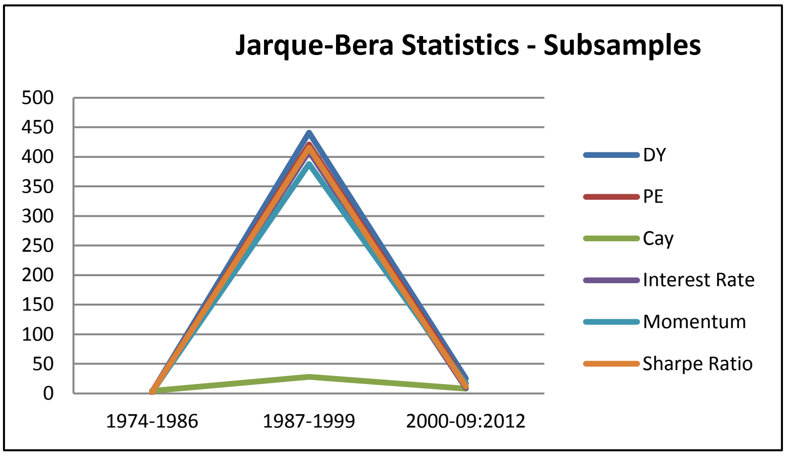
In total of all of our 18 scenarios, six variables and three periods, we have a significant variable in eight of the cases, this is a percentage of 44%. From the tests we ran for the long sample we have significant in four out of twelve cases, and some of the significant variables are not the significant variables we get when testing the subsamples. In the table below you get an overview of the variables and the periods. For PE and momentum we have not observed any significant t-statistics, and thus we have to keep the null hypothesis which states that the explanatory variable has none predictability on stock return. The period from 1974 to 1986 proves to be the subsample with the best results based on significant t-statistics with four cases of significance.

Summary Significant t-statistics			
	1974 - 1986	1987 - 1999	2000 - 09:2012
DY	°		
PE			
Cay	°	°	°
Interest Rate	**		**
Momentum			
Sharpe Ratio	**		*

Table 17: Summary of significant t-statistics U.S. data

A commonality across the six variables is the poor performance in the second period. Except from cay, the middle period shows the lowest R^2 and the highest p-value for all variables. In addition, cay is the only variable that is significant in this period, so in the period 1987-1999 cay is the only variable that could predict stock return based on our results. For cay we only have quarterly data so it is not directly comparable with the other variables.

A special salient observation from this period is the JB statistic which is extremely high, and on a different scale compared to the results from the two other subsamples. Looking at figure 46 we see a high peak, which stands out from the normal level in the JB statistics in period 1 and 3. The reason behind the expression extremely high is because we compare it to the critical value 3.84, from the χ^2-distribution. So the null hypothesis about normal distribution in the residual is rejected with certainty, which also is consistent with the low p-value reported with the JB statistics in SAS. If we look back at the results from the long sample testing, reported in figure 46 below, we also observe high statistics, but the statistics in the second subsample are even higher. The high values can help explain why we also observe quite high values in the long individual sample. Despite the low values in the beginning and end of the long sample, the extreme level in the middle period affects the result considerably.



Jarque-Bera Statistics Individual Sample	
DY	175.3796
PE	171.6145
Cay	47.5236
Interest Rate	182.6468
Momentum	179.1511
Sharpe Ratio	180.8653

Figure 46: Jarque-Bera statistics development

Regarding stationarity we check for this in the stock return time series and in the time series for the explanatory variables. To detect stationarity we can look at a graph or the ACF. Below you can see both the graphical line plot and the correlogram for stock return in the three subsamples. From the graphical plot below there is no sign of trend, neither an upward or downward trend. This is a strong indication that the time series are stationary and that the mean of the time series is around zero. Further we look at the correlogram of the ACF, this is the figure in the top right of the figures below. For the time series to be stationary the lags should move around zero and be inside of the blue significant area. This is the case for all of the subsamples and this confirms that we have stationary time series for stock return.

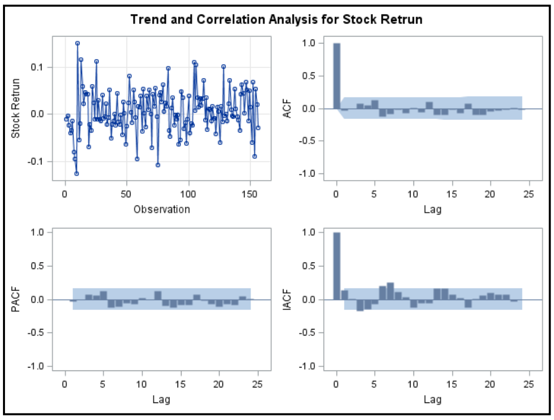


Figure 47: ACF and PACF for stock return 1974-1986

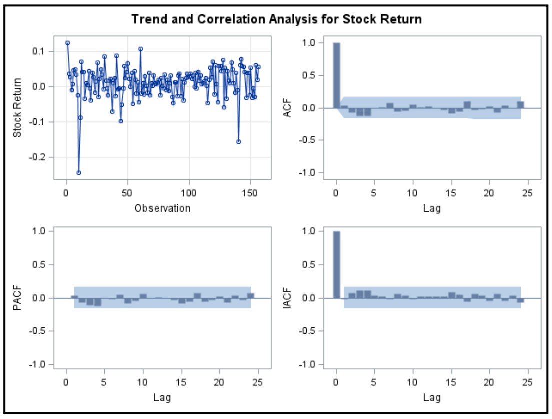


Figure 48: ACF and PACF for stock return 1987-1999

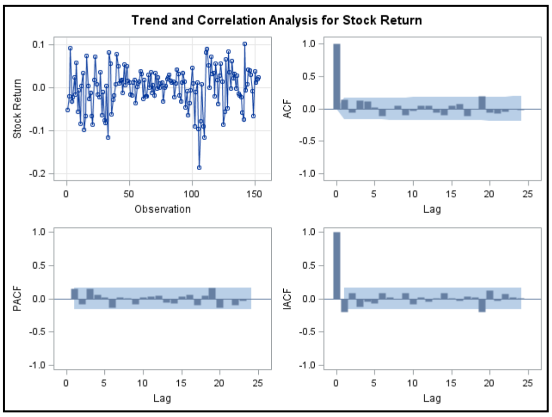


Figure 49: ACF and PACF for stock return 2000-Sept 2012

For our explanatory we have seen a different situation regarding stationarity. In all three sample periods for the entire set of explanatory variables there are signs of nonstationary time series in different degrees. We take a further look into stationarity inspired by Lewellen (2004), and run the lagged of our explanatory variables on our explanatory variables. The coefficients on the explanatory variables are in all cases <1, and with that we conclude that the presence of nonstationarity in the time series should not be a problem. The method is presented in detail in the discussion section.

Overall, the time series we have seem to fulfill the OLS assumption about heteroscedasticity and autocorrelation in almost all cases. We have some problem with the presence of autocorrelation in the last subsample for DY, so the OLS regression is transformed to a GLS regression and the autocorrelation problem is removed.

4.3 NORWAY

To check whether our explanatory variable has some explanatory power in the Norwegian stock market, we have collected data on the OSEAX. We run the same linear regression and tests as earlier, and check for the predictability for the following variables; DY, PE, interest rate and Sharpe ratio. The time periods applied are the two last periods from our subsample testing, 1987-1999 and 2000 – September 2012.

4.3.1 DIVIDEND YIELD

When running our original regression model for DY the results show that we have one significant t-value at the 10% significance level, this was for the first period tested 1987-1999. After we have checked the model for the classical OLS assumption we found no evidence of heteroscedasticity but we found that the assumption for autocorrelation was not fulfilled, this was the case for both periods. With the lack of fulfillment of this assumption we cannot rely on the reported t-statistics and the results are not valid. To eliminate the autocorrelation problem we created the new model with the GLS method. Below the results from the GLS regression will be reported.

Subsample 1987 - 1999	
Dividend Yield (DY) *GLS*	
β	0.0212
Standard Error	0.0114
t-value	1.85*
p-value	0.0666
R-square	0.0220
F	3.41
Jarque-Bera	127.07
Durbin Watson	1.95
Breusch-Godfrey	
- p=1	0.04
- p=3	2.17

Subsample 2000 - 09:2012	
Dividend Yield (DY) *GLS*	
B	-0.0027
Standard Error	0.0061
t-value	-0.45
p-value	0.6563
R-square	0.0013
F	0.20
Jarque-Bera	52.48
Durbin Watson	1.95
Breusch-Godfrey	
- p=1	0.08
- p=3	0.79

Table 18: GLS estimates for DY

We start our investigation of the relationship between DY and stock return by plotting stock return against the lagged variable in a scatter plot, the results are in figure 50 and 51 below. The observations look randomly scattered for both of the periods with some outliers. In the last period we have a few observations with high values compared to the first period. The horizontal axis has the value 0-3 in the first period, and 0-7 in the last period, which means that we have outliers with values that are higher than the double of the values we have in the rest of the observations in our dataset.

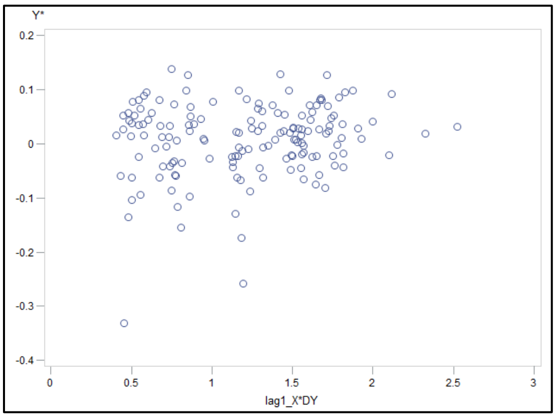


Figure 50: Relationship between X and Y 1987-1999

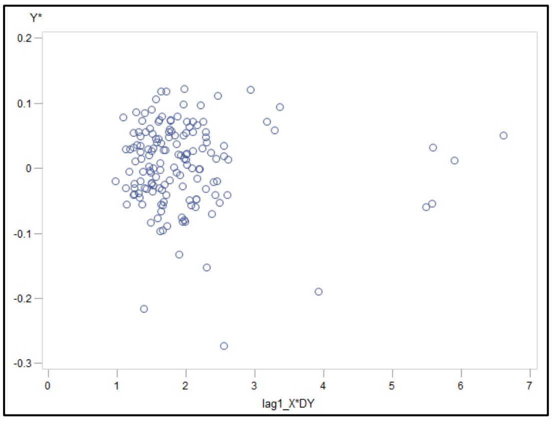


Figure 51: Relationship between X and Y 2000-Sept 2012

The t-test shows that for DY we have one period with a significant coefficient, which is similar to the result before the GLS transformation. With a t-statistic of 1.85 we reject the null hypothesis on a 10% significance level for the two-sided test and then also for both levels of significance in the one-sided test. In the last period the statistic is low and with the corresponding high p-value we keep the hypothesis of no explanatory power. The large outliers we observe in the scatter plot may explain the different results for the two periods.

There are no sign of heteroscedasticity when looking at the residual plot, as expected since the OLS regression showed no sign of heteroscedasticity. The plots have the same random pattern as the scatter plots above, and again we observe the large outliers in the last period. For both periods we reject the hypothesis about normality in the error term. In the first period we get quite a high JB statistic that is considerably reduced from 127.07 to 52.48 in the second period, but it is still too high to keep the null.

Regarding autocorrelation we have done the GLS transformation to get rid of the problem which is confirmed by the results from the BG test. In the tables below the old results and the new results are presented. As the table shows, by applying the new model the AR(1) for ρ=1 and ρ=3 is greatly reduced. AR(1) and AR(3) are smaller than the respective critical value by a wide margin and we keep the null hypothesis about no autocorrelation in the error term.

Godfrey's Serial Correlation Test		
Alternative	LM	Pr > LM
AR(1)	6.4025	0.0114
AR(2)	7.4037	0.0247
AR(3)	8.1540	0.0429
AR(4)	9.0765	0.0592

Godfrey's Serial Correlation Test		
Alternative	LM	Pr > LM
AR(1)	0.0449	0.8322
AR(2)	1.4904	0.4746
AR(3)	2.1675	0.5384
AR(4)	2.7369	0.6028

Table 19: BG-statistics before and after GLS transformation

Godfrey's Serial Correlation Test		
Alternative	LM	Pr > LM
AR(1)	7.2453	0.0071
AR(2)	7.6392	0.0219
AR(3)	7.9555	0.0469
AR(4)	7.9813	0.0923

Godfrey's Serial Correlation Test		
Alternative	LM	Pr > LM
AR(1)	0.0777	0.7804
AR(2)	0.3863	0.8244
AR(3)	0.7871	0.8525
AR(4)	0.7876	0.9401

Table 20: BG-statistics before and after GLS transformation

The investigation of the ACF and PACF correlogram indicates that the time series for DY for both periods are nonstationary. In the first period there are 9 lags outside the blue shaded area, and the lags are slowly decreasing, which is a typical pattern for a time series that is not stationary. For the last period the ACF has 5 lags outside and decrease rapidly, which is a sign of nonstationarity.

4.3.2 PRICE-EARNINGS

For PE we observed the same problem with autocorrelation for the last period, 2000-September 2012, when we ran our original OLS regression. To eliminate this problem we apply the GLS method to the last subsample.

Subsample 1987 - 1999	
Price-Earnings (PE)	
β	0.0002
Standard Error	0.0002
t-value	0.80
p-value	0.4269
R-square	0.0041
F	0.63
Jarque-Bera	165.11
Durbin Watson	1.59
Breusch-Godfrey	
- p=1	5.97
- p=3	7.64

Subsample 2000 - 09:2012	
Price-Earnings (PE) *GLS*	
B	-0.0007
Standard Error	0.0004
t-value	-1.5°
p-value	0.1352
R-square	0.0149
F	2.26
Jarque-Bera	58.40
Durbin Watson	2.04
Breusch-Godfrey	
- p=1	0.06
- p=3	0.33

Table 21: OLS and GLS estimates for PE

For S&P 500 we observed a few very large outliers in the sample from 2000 to September 2012. When looking at the Norwegian data we get a similar scatter plot, see figure 52, but for the period 1987-1999. Looking closer into the dataset these outliers are from the period between March 1989 and February 1990. For the last period the observations are randomly scattered with most of the observations concentrated between the value 10-40 for PE and a few outliers where the highest value is 82.93.

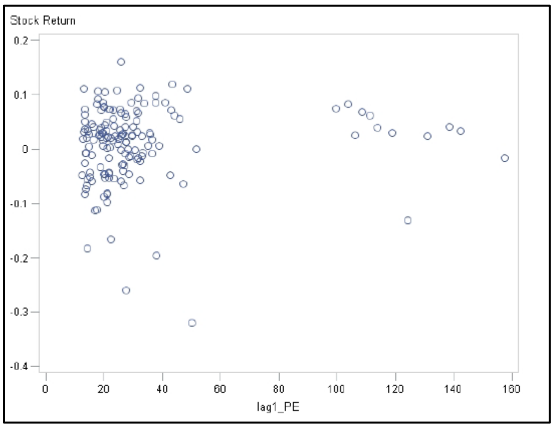


Figure 52: Relationship between stock return and the first lag of PE, 1987-1999

Displayed by the t-statistics we also have one period with significant statistic. For PE this is the case for the last period where we have a negative statistic of -1.5, which is significant for the one-sided test at a 10% level. For both PE and DY we have significant coefficients in the period where we do not have the largest outliers.

None of the periods show sign of heteroscedasticity. Again we get high values when running the JB test for normality, thus we have to reject the null hypothesis for normality. The highest value is observed in the first period.

In the process where we transformed our original OLS model with the GLS method, we assured that autocorrelation was not present in the new model. Again we could confirm this by the BG test, and the old and new results are shown in table 22 below. For the first period we have kept our original and by looking at the plot where the residuals are plotted against time we observe a quite random plot, and no clear pattern stands out. The BG result is poor and autocorrelation is present in almost all the different cases, but we have a case where the autocorrelation assumption is fulfilled. We keep the hypothesis of no autocorrelation in the first period for ρ=3 at a 5% level.

Godfrey's Serial Correlation Test		
Alternative	LM	Pr > LM
AR(1)	7.5873	0.0059
AR(2)	7.9045	0.0192
AR(3)	8.0597	0.0448
AR(4)	8.0760	0.0888

Godfrey's Serial Correlation Test		
Alternative	LM	Pr > LM
AR(1)	0.0554	0.8139
AR(2)	0.1807	0.9136
AR(3)	0.3337	0.9536
AR(4)	0.3338	0.9875

Table 22: BG-statistics before and after GLS transformation

Testing the PE time series for stationarity the result is an indication of nonstationarity. For the first period there are 6 lags outside and they decrease at a constant factor until lag 10 where they go under the axis and down to the negative scale, and move around with an uneven factor. Lag number 3, 10, and 13 are also outside the significance band in the PACF. Compared to the other variables tested here, the last period for PE have the ACF that is closest to be stationary, the results are presented in figure 53. Therefore only 3 lags outside in the ACF, and only lag number 3 outside in the PACF, but there are still weak signs of nonstationarity.

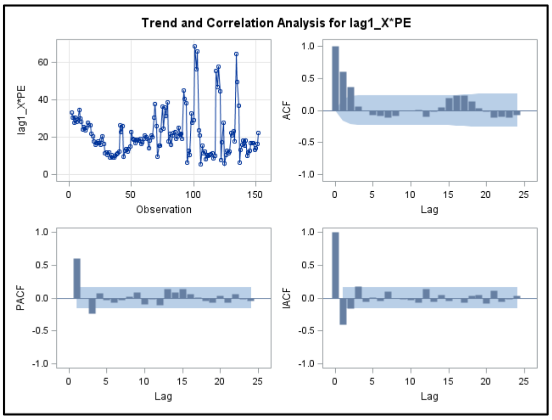


Figure 53: ACF and PACF for the first lag of PE (GLS), 2000-Sept 2012

4.3.3 INTEREST RATE

Subsample 1987 - 1999		Subsample 2000 - 09:2012	
Interest Rate		**Interest Rate**	
β	-0.0086	B	-0.0141
Standard Error	0.0049	Standard Error	0.0056
t-value	-1.75*	t-value	-2.53**
p-value	0.0824	p-value	0.0126
R-square	0.0196	R-square	0.0405
F	3.06	F	6.38
Jarque-Bera	142.85	Jarque-Bera	73.92
Durbin Watson	1.65	Durbin Watson	1.61
Breusch-Godfrey		Breusch-Godfrey	
- p=1	4.44	- p=1	5.77
- p=3	6.57	- p=3	6.29

Table 23: OLS estimates for interest rate

The scatter plot of the two variables, stock return and interest rate, shows randomly scattered observations. In the first period most of the observations are gathered in a cloud-like pattern, in the last period the observations are a bit more scattered throughout the plot. For both periods it looks like there is a slightly predominance of negative observations. This prediction is confirmed by the results we get when we run the linear regression. In both periods the coefficient is negative, which also leads to negative t-statistics. The t-statistics is high enough to reject the null hypothesis in both periods, at 10% in the first and at 5% in the last period. By rejecting the null hypothesis we conclude that the interest rate can predict stock return in the time period 1987-1999 and 2000-2012.

Looking at the residuals plot there are no signs of a pattern in the variance, and again there is no sign of heteroscedasticity, and we have homoscedasticity as we desired. For normality the situation is undesirable because of the high statistic from the JB test. Again the first period has the highest value of 142.85, almost twice as high as the value in the last period. Both are way too high to be under the critical value of 3.84, which leads to the rejection of the hypothesis of normal distribution in the error term.

Testing for autocorrelation we detect the presence of autocorrelation at the first-order autoregressive scheme, ρ =1, at both levels of significance and in both periods. The only situation where we can keep the null of no autocorrelation is for the third-order autoregressive scheme at a 5% level of significant, this is equal for both periods.

As with the other variables we have investigated, the time series for interest rate in the two subsamples are nonstationary. The ACF follows the same pattern; the lags are decreasing rather rapidly toward zero and continue decreasing so the lags get negative values on the vertical axis. Although the lags decrease rapidly there is sign of nonstationarity because of a significant number of lags outside the significant areas, respectively 6 and 7. In the PACF we observe 2 lags outside for both periods.

4.3.4 SHARPE RATIO

Subsample 1987 - 1999	
Sharpe Ratio	
B	0.0011
Standard Error	0.0006
t-value	1.80*
p-value	0.0732
R-square	0.0208
F	3.25
Jarque-Bera	137.39
Durbin Watson	1.66
Breusch-Godfrey	
- p=1	4.36
- p=3	6.21

Subsample 2000 - 09:2012	
Sharpe Ratio	
β	0.0018
Standard Error	0.0008
t-value	2.35**
p-value	0.0200
R-square	0.0353
F	5.53
Jarque-Bera	72.84
Durbin Watson	1.62
Breusch-Godfrey	
- p=1	5.42
- p=3	5.93

Table 24: OLS estimates for Sharpe ratio

The relationship between stock return and the lagged Sharpe ratio looks quite random. In the first period the mean is close to zero, with a slightly positive direction. For the last subsample the observations are even more random and also with a slightly positive tendency. The results show that the Sharpe ratio coefficient is positive in both periods, and also significant. We can reject the null hypothesis at a 10% level for a two-sided test in the first period where the t-statistic is 1.8. For the last period we find a higher t-value and thus stronger sign of significance, which is confirmed with a lower p-value, and we can reject at a 5% level.

The observation in the residual plot looks nice and the observations are randomly scattered, which is a strong sign of no heteroscedasticity. In the results table above we observe a high JB statistics for both period, thus despite the reduction in the last period it is still to high to keep the hypothesis. The rejection indicates that the error term for Sharpe ratio is not normally distributed.

For Sharpe ratio we have the best result regarding autocorrelation among the other variables tested for the Norwegian data. We still have to reject the null hypothesis for ρ =1 at both levels of significance, but when checking for ρ=3 we keep at both levels in both periods. Summarized, there is no problem with autocorrelation when testing for the third order autoregressive scheme.

Based on the autocorrelation correlogram we can conclude that the Sharpe ratio time series are nonstationary. In the first period the ACF has 6 lags outside, and decrease rapidly down to the negative side after lag 9, and in the PACF there are lags number 4 and 9 outside. For the last period the lags decrease a bit slower and become negative after 15 lags, in the ACF there are 8 lags outside, and we find 3 of the lags outside in the PACF.

4.3.5 SUMMARY

Significant t-statistics Norway		
	1987 - 1999	2000 - 09:2012
DY	*	
PE		°
Interest Rate	*	**
Sharpe Ratio	*	**

Table 25: Summary of significant t-statistics, Norwegian data

Summarized we have six out of eight cases with significant t-statistics for the Norwegian market. As mentioned under the affected variable we have transformed our model with the GLS method for 3 of our cases. This transformation has not changed the results in any material way, apart from the important removal of the presence of autocorrelation.

For S&P 500 we saw extremely high values in the period 1987-1999 for the JB normality test. In the results for the Norwegian data we also observe high values for this period, but not as extreme. The S&P 500 had an average of 421.6743 and the same average for the Norwegian samples is 141.99. Regarding normality we have no situation where we can conclude that we have normal distribution in the error term based on the results above. In S&P 500 we only found normality for the period 1974-1986, which is the one period we do not test here.

When investigating the U.S. data for the S&P 500 we detected that the stock return time series were stationary in all periods, but the time series for the all the explanatory variables were nonstationary. We see the same tendency when we are looking at the Norwegian time series. As mentioned above for each of the explanatory variables the time series have a pattern that indicates nonstationarity. Checking this with the method done by Lewellen (2004) we observe $\rho<1$ but close to 1 for all of the variables in the different time periods. Thus, we conclude with valid t-test results and no problem related to the stationarity assumption.

For stock return, the results from both periods are shown in figure 54 and 55 below. Looking at the graphical line plot top left in each of the figures, the stock return time series do not seem to follow a trend. In both of the periods the observations fluctuate around the mean, and when there are some outliers they tend to return back to the mean, also known as mean reversion (Gujarati & Porter, 2009, p.741). Looking at the ACF and PACF we have 1 significant lag for both functions in both periods.

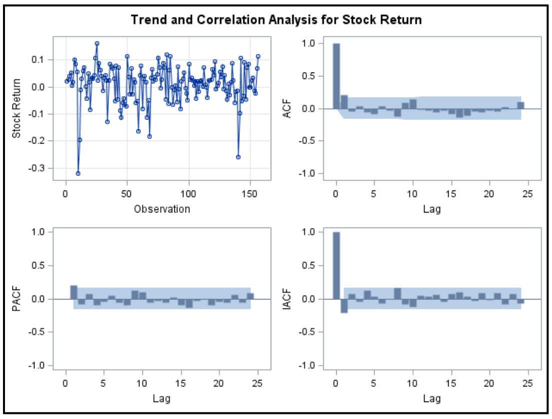


Figure 54: ACF and PACF for stock return, 1987-1999

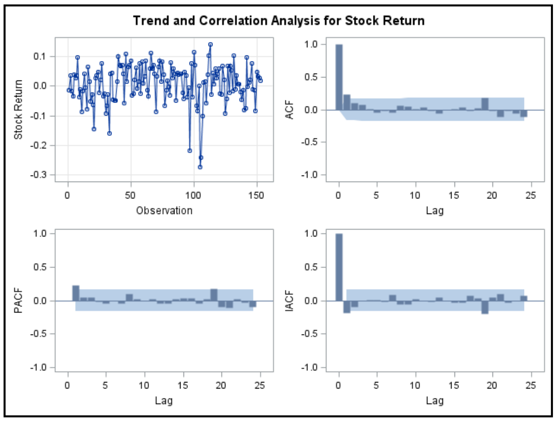


Figure 55: ACF and PACF for stock return, 2000-Sept 2012

4.4 SWEDEN

As with the Norwegian data we run our linear regression to check if we find some explanatory power in the Swedish stock market by looking at the OMXS index in connection with our variables. We test for the same explanatory variables as with the Norwegian data and the same two subsamples.

4.4.1 DIVIDEND YIELD

Subsample 1987 - 1999	
Dividend Yield (DY)	
β	0.0029
Standard Error	0.0052
t-value	0.56
p-value	0.5747
R-square	0.0021
F	0.32
Jarque-Bera	46.84
Durbin Watson	1.56
Breusch-Godfrey	
- p=1	6.15
- p=3	6.94

Subsample 2000 - 09:2012	
Dividend Yield (DY)	
B	-0.0033
Standard Error	0.0067
t-value	-0.49
p-value	0.6219
R-square	0.0016
F	0.24
Jarque-Bera	8.37
Durbin Watson	1.75
Breusch-Godfrey	
- p=1	2.38
- p=3	12.12

Table 26: OLS estimates for DY

From the scatter plot of stock return and DY we observe a quite random relationship between the variables. Most of the observations for DY are concentrated between the value of 1 and 3. In the first period the observations are slightly more scattered throughout the plot compared to the last period, and we observe a few outliers in both cases. From the t-test we do not have significant values, and there are no indications that DY can predict stock returns. This is also consistent with the low R^2 obtained in the test, respectively 0.21% and 0.16%.

The residual plot created to control for heteroscedasticity has the same pattern as the plot where we look at stock return against DY. The plot shows no sign of a pattern in the variance and the disturbances seem to be homoscedastic. We start by looking at residuals against time to check if there is some indication of autocorrelation in the disturbances. Figure 56 and 57 shows the plot for both periods, and there is not a clear pattern that stands out, thus no clear signals of autocorrelation. We take the investigation further by checking the statistics from the BG test. We are able to keep the null hypothesis of no autocorrelation in both periods, but for different orders of autocorrelation. In the first period we have to reject the hypothesis for ρ=1, but we can keep at ρ=3 at a 5% significance level. For the last period we can keep ρ=1 at both levels of significance.

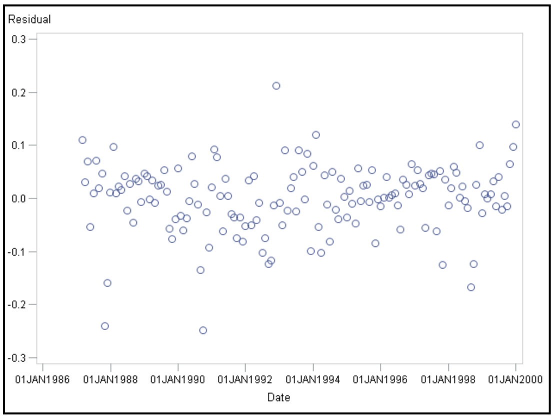


Figure 56: Residuals over time, 1987-1999

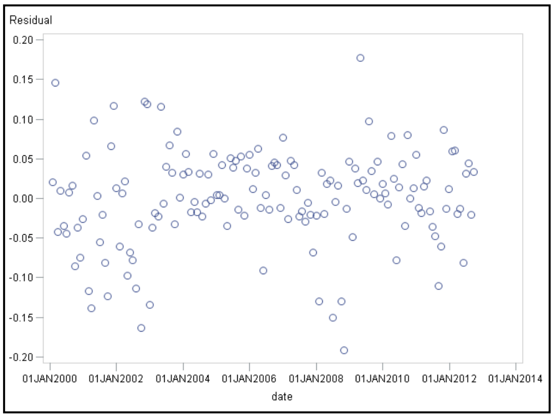


Figure 57: Residuals over time, 2000-Sept 2012

From our test on the U.S. and Norwegian data the error term has never been normally distributed for DY in these two periods we are testing here. This is also the case when checking the normality assumption for the Swedish data on DY. The critical value following the χ^2-distribution is 3.84 and neither of the obtained JB statistics are below this value, and that leads to the rejection of the normality hypothesis.

The time series for the lagged DY show the classical pattern for a nonstationary time series. In figure 58 and 59 below the result from our graphical investigation of stationarity is presented. Both periods show similar pattern for the ACF and PACF. The ACF is rapidly decreasing down towards zero and down on the negative axis, there are respectively 8 and 7 lags outside the blue shaded area. In the graph in the top left corner there is a line plot where the earlier discussed outliers are apparent.

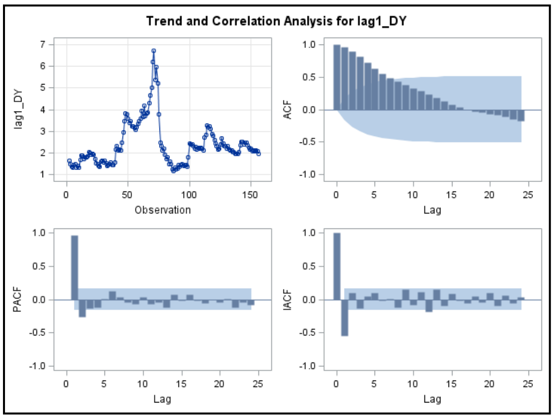


Figure 58: ACF and PACF for the first lag of DY, 1987-1999

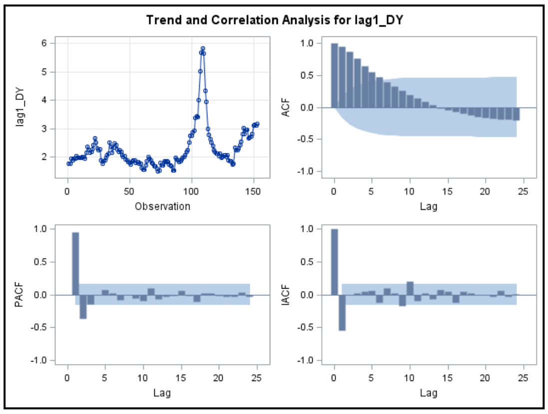


Figure 59: ACF and PACF for the first lag of DY, 2000-Sept 2012

4.4.2 PRICE-EARNINGS

Subsample 1987 - 1999	
Price-Earnings (PE)	
B	-0.0001
Standard Error	0.0002
t-value	-0.61
p-value	0.5417
R-square	0.0024
F	0.37
Jarque-Bera	50.79
Durbin Watson	1.57
Breusch-Godfrey	
- p=1	5.99
- p=3	6.91

Subsample 2000 - 09:2012	
Price-Earnings (PE)	
β	-0.0002
Standard Error	0.0003
t-value	-0.56
p-value	0.5793
R-square	0.0020
F	0.31
Jarque-Bera	10.73
Durbin Watson	1.73
Breusch-Godfrey	
- p=1	2.75
- p=3	12.51

Table 27: OLS estimates for PE

Earlier we have observed that PE is a variable where the dataset have contained some extreme outliers in one of the subsamples. For S&P 500 we saw these in the last sample period, but for the Norwegian data they were observed in the time period 1987-1999. In the Swedish dataset we see outliers with rather high values in both of the sample, shown clearly in figure 60 and 61. These abnormal high values do impact the results from our regression. The results showed in the tables above report slightly negative values on the coefficient in the two periods. Our reported t-values are also negative and low with corresponding high p-values. None of the t-values are higher than the critical values and we have to keep the null hypothesis about no predictability on stock return.

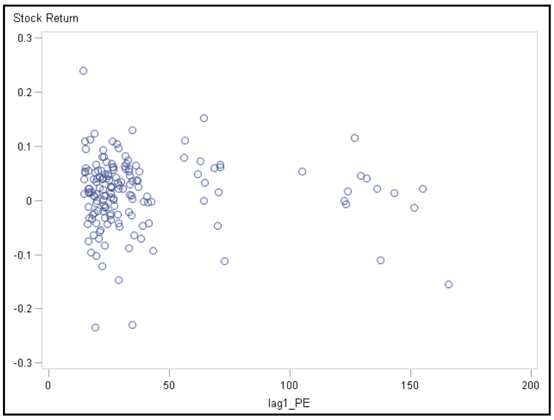


Figure 60: Relationship between X and Y, 1987-1999

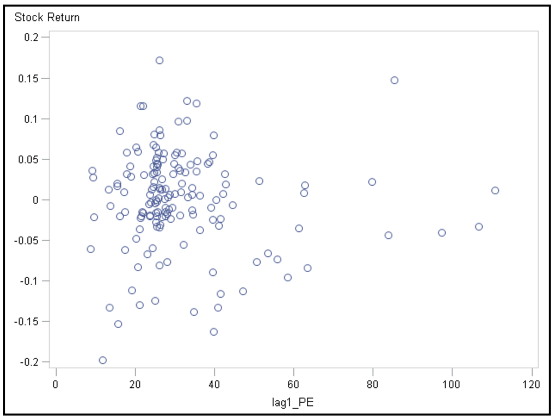


Figure 61: Relationship between X and Y, 2000-Sept 2012

There is no sign of a pattern in the variance, and we can conclude with no indication of heteroscedasticity for PE. The reported JB statistics are above the critical value, and once again we have to reject the hypothesis about normal distribution in the error term. Plotting the residuals against time gives us a similar plot as we got for DY, and no strong indication of autocorrelation. When running the BG test we are able to keep the null in one of the cases in both of the period, and both at the 5% significance level. In the first period we keep for ρ=3 and in the last period we keep for ρ=1.

Looking at the ACF we observe multiple significant lags in both periods which indicate that the time series are nonstationary. In the first period the lags slowly decrease from lag 1 down to lag 25 which is almost equal to zero, and in total there are 9 lags outside. For the last period there are 6 lags outside and they decrease rapidly and after lag 13 they move around zero.

4.4.3 INTEREST RATE

Subsample 1987 - 1999	
Interest Rate	
B	-0.0016
Standard Error	0.0033
t-value	-0.49
p-value	0.6214
R-square	0.0016
F	0.24
Jarque-Bera	52.87
Durbin Watson	1.58
Breusch-Godfrey	
- p=1	5.82
- p=3	7.14

Subsample 2000 - 09:2012	
Interest Rate	
B	-0.0196
Standard Error	0.0063
t-value	-3.09**
p-value	0.0024
R-square	0.0596
F	9.56
Jarque-Bera	7.32
Durbin Watson	1.83
Breusch-Godfrey	
- p=1	1.02
- p=3	10.61

Table 28: OLS estimates for interest rate

The relationship between stock return and the lagged interest rate seems quite random, and it is hard to draw any conclusion about whether the predicted coefficient will be negative or positive. In the first period most of the observations are between the values of -2 and 2 for the interest rate and between -1 and 1 in the second period. In the first period there is one extreme outlier, seen in figure 62, which may bias the results. In both periods the coefficient from our linear regression is negative. This is the first case where we have a significant result in the Swedish data. The significant result is observed in the last period, and we reject the null hypothesis at the 5% level for a two-sided test. Following this strong significant t-value the observed p-value is low, and from this we can conclude that interest rate has predictability power on the OMXS index.

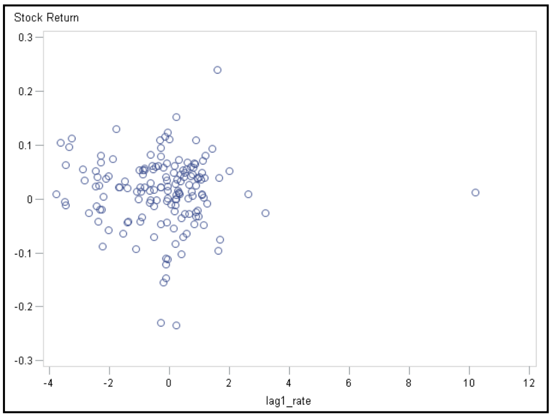


Figure 62: Relationship between stock return and interest rate, 1987-1999

The disturbances have no indication of heteroscedastitcy. By plotting the residuals against the lagged interest rate the observations are randomly plotted around and there is no sign of a pattern in variance. Interest rate is the variable that has the highest JB statistic for the Swedish data. This high value is from the first period and indicates that the error term is not normally distributed and the null hypothesis is rejected. For the second period the statistic is reduced but is still over the critical value of 3.84 and the conclusion about normality is the same for both periods.

Again we see a similar pattern for the autocorrelation plot where we have the residuals plotted against time. The results from the BG test follow the same pattern as we have seen for the other variables. We are able to keep the hypothesis about no autocorrelation for ρ=3 at a 5% significance level in the first period. In the last period we are now able to keep the null for both significance levels for ρ=1.

Looking at the graphical representation of the ACF and PACF the patterns of the lags indicate the presence of nonstationary time series for interest rate. The first period has 5 significant lags in the ACF and the lags decrease rapidly in a wave pattern which moves form positive to negative and back to positive values. In the last period the lags decrease with a constant rate from positive to negative values after lag number 13, and there are 7 significant lags in the ACF.

4.4.4 SHARPE RATIO

Subsample 1987 - 1999	
Sharpe Ratio	
B	0.0002
Standard Error	0.0003
t-value	0.44
p-value	0.6575
R-square	0.0013
F	0.20
Jarque-Bera	52.54
Durbin Watson	1.57
Breusch-Godfrey	
- p=1	5.80
- p=3	7.04

Subsample 2000 - 09:2012	
Sharpe Ratio	
β	0.0023
Standard Error	0.0008
t-value	3.08**
p-value	0.0025
R-square	0.0591
F	9.48
Jarque-Bera	6.85
Durbin Watson	1.86
Breusch-Godfrey	
- p=1	0.77
- p=3	9.84

Table 29: OLS estimates for Sharpe ratio

Figure 63 and 64 on the next page show the scatter plots of the variables in the regression for both periods. As the figures show the observations are randomly scattered, and it may look like it is a slightly positive relationship between stock return and the lagged Sharpe ratio. A noteworthy observation is the difference in the values on the x-axis. The observations in the first sample on Sharpe ratio are located between – 57.39 and 40.09, for the last period the observations are between -12.07 and 20.91. This difference is reflected in the linear regression results. In the first sample we have a wide range in the values, the coefficient is small, and the t-value is not significant. For the last period where the observations are more concentrated on a smaller scale, the coefficient is bigger compared to the one in the first sample period, and we have a significant t-value. The t-value is significant at a 5% level for a two-sided test.

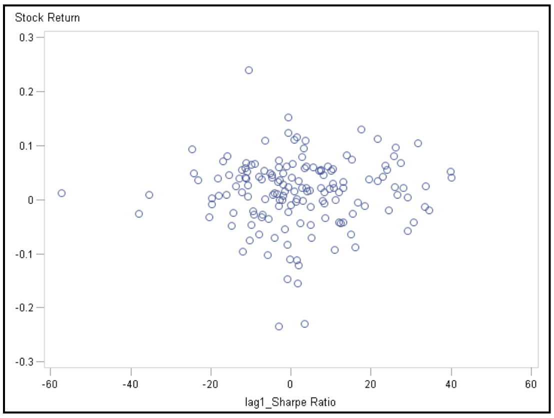


Figure 63: Relationship between X and Y, 1987-1999

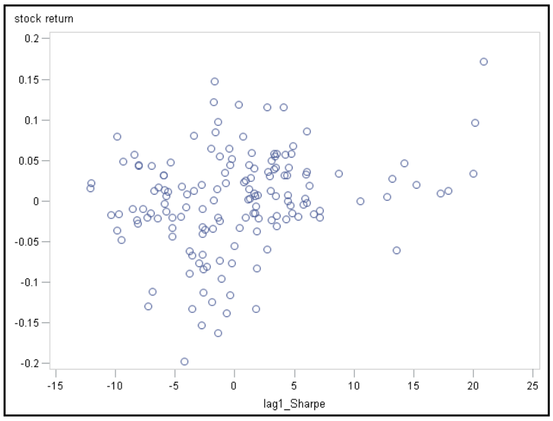


Figure 64: Relationship between X and Y, 2000-Sept 2012

In resemblance to the plots above, the plot with the residuals against the first lag of the explanatory variable shows randomly scattered observations. There is no sign of a pattern in the variance, and hence no indication of heteroscedasitcity. Sharpe ratio is the variable in the Swedish data that has the lowest JB statistic. Despite this, the JB statistic of 6.85 in the last period is still too high to keep the null hypothesis, so we have to conclude with no normality in the error term for Sharpe ratio.

Regarding autocorrelation, the outcome from the BG test is exactly the same as we get when we are testing for the interest rate. For the first period we keep at ρ=3 for the 5% significance level, and in the last period we keep ρ=1 for both 10% and 5%. Not surprisingly, the time series for Sharpe ratio are also nonstationary. The ACF has 6 and 7 lags outside in the respective two periods, and the lags decrease rapidly towards zero and further down on the negative side of the y-axis.

4.4.5 SUMMARY

Significant t-statistics Sweden		
	1987 - 1999	2000 - 09:2012
DY		
PE		
Interest Rate		**
Sharpe Ratio		**

Table 30: Summary of significant t-statistics, Swedish data

From the table above we have an overview of the significant results for the Swedish data. As we can see there is only two out of eight cases that have significant coefficients, and these are both in the last period. Compared to the Norwegian data, the Swedish results are poorer with only two significant results, while for the Norwegian we had six significant cases. Again, interest rate and Sharpe ratio stands out as the two explanatory variables with the strongest results.

The pattern we have observed from our other tests is that the JB statistic has been extremely high in the period 1987-1999. For the Swedish data the statistic is clearly highest in this period, but not at the same level we have seen for the U.S. and Norwegian data.

The OMXS time series seem to be stationary, the line plot for the observations and the ACF and PACF are presented in figure 65 and 66. The line plot to the top left shows the observations of stock returns from 1987 to 1999 and from 2000 to September 2012. As the graphs indicate there is no clear trend in the movements of the stock returns on OMXS throughout the subsample periods. In the first period we observe two lags that are at the limit to being significant in the ACF, apart from this the lags move nicely around zero. For the last period there is 1 lag outside in the ACF and PACF, but otherwise no troubling pattern. For the explanatory variables we have tested the additional regression done by Lewellen (2004), which concludes with no problem regarding the validity of our t-test result due to the presence of nonstationarity in the time series.

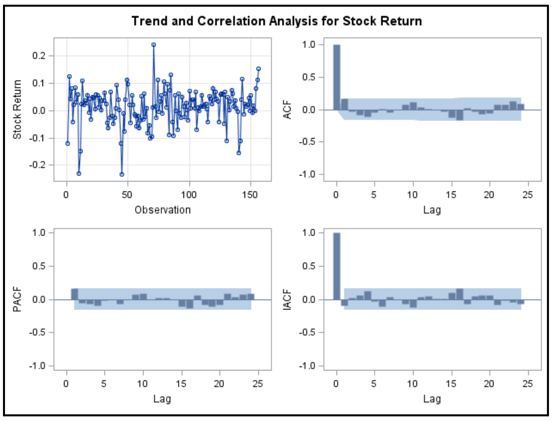


Figure 65: ACF and PACF for stock return, 1987-1999

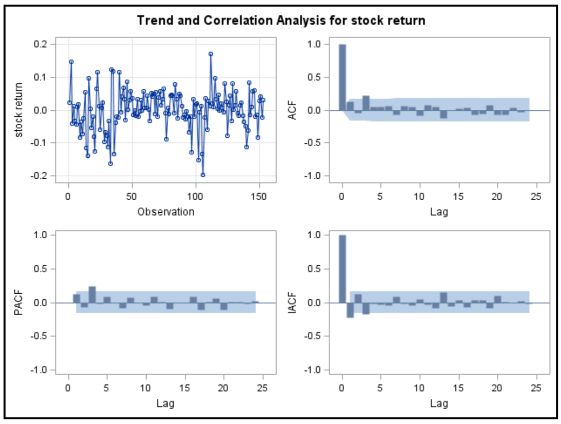


Figure 66: ACF and PACF for stock return, 2000-Sept 2012

For autocorrelation there are similarities in the BG test results in the two time periods for the four variables. In the first period we see that we reject the hypothesis about no autocorrelation for ρ=1 for all varaibles, but we are able to keep for ρ=3 at a 5% significance level. In the last period we have the opposite situation, here we are able to keep for ρ=1 but we have to reject for all variables at ρ=3. For ρ=3 we observe values that are bigger than the critical value, and also noteworthy is the great jump in the BG value from ρ=1 to ρ=3 in this last period.

4.5 OUT-OF-SAMPLE TESTING

The OOS test was performed on our last subsample which spans from January 2000 to September 2012, and only on the U.S data. The sample period was divided into an IS (R) and OOS (P) period, where R consists of 2/3 of the period and P consist of 1/3. This leaves us with the following sample split in number of observations;

	T	R (IS)	P (OOS)
Monthly data	153	102	51
Quarterly data	51	34	17

Table 31: Sample split OOS testing

Both the MSE-F and ENC-NEW statistics follows an asymptotic distribution and the critical values are obtained from the corresponding tables. In table 32 below an extract from these two distributions is presented. The extract displays the critical values which correspond to the parameter that belong to our sample, the whole table can be found in appendix F. *k2* takes the value 1 since we look at stock return related to the explanatory variable one period back. Pi is defined as π =P/R and is equal to 2 for both the monthly and quarterly data (McCracken, 2007, p.11). Similarly to our IS testing the statistics is checked for both a 5% and 10% significance level.

MSE-F				**ENC-NEW**		
	π	2			**π**	2
k2	**significance**			***k2***	**Significance**	
1	0.90	0.616		1	0.90	1.28
	0.95	1.518			0.95	2.085

Table 32: Critical values MSE-F and ENC-NEW

We run our OLS regression for each observation from time period R up to R+P as described in the theory to get the estimates to calculate the forecast values. This process leaves us with 52 forecast for the monthly data and 18 for the quarterly data. The forecast from the unrestricted and restricted model is used further to calculate the MSE-F and ENC-NEW statistics which is reported in table 33 below, in addition see Appendix G for all calculations. In addition to these calculations we have tested the OOS sample with our original OLS regression where we have replaced our original values for stock return with the forecast from the unrestricted model in the time period R to P. This means that the stock return data up to observation 101 in the period is the same as in the IS test, and from observation 102 to 51 the original data is replaced with the new forecasted values. The OLS assumptions are fulfilled, and the t-statistic is valid. The results from this OOS regression is marked with (OOS) in the table below.

Out-of-sample testing, subsample 2000 - 09:2012, U.S.						
	DY	PE	Cay	Interest Rate	Momentum	Sharpe Ratio
β (IS) t-value (IS)	0.01253 1.07	0.00017978 1.13	0.44157 1.5°	0.00879 2.04**	-0.02438 -1.20	-0.00067737 -1.72*
β (OOS) t-value (OOS)	0.01261 1.86*	0.00012573 1.13	0.39316 1.93*	0.00504 1.69*	-0.02151 -1.53°	-0.00041794 -1.53°
MSE-F ENC-NEW	0.0523 4.0683**	9.9539** 6.7515**	325.3684** 317.9846**	2.1258** 1.2333	1.2486* 0.7824	1.3317* 0.7319

Table 33: OOS Results, subsample 2000 – 09:2012, U.S

The IS results in the table is equal to the last subsample results which are presented in section 4.2, for the respective variable. The OOS regression results show significant results for two more variables, DY and momentum, compared to the original IS results. However, the level of significance varies, and on a general basis the OOS regression results is significant at a lower level compared to the IS results.

The MSE-F statistic is significant in five out of six tested scenarios. This indicates that the unrestricted model forecast is better than the restricted model forecast. For cay, interest rate, and the Sharpe ratio the MSE-F results matches our IS results, while it performed better than IS for PE and momentum. For the ENC-NEW statistic the results indicate three out of six significant variables, and for those three OOS tests performed better than the reported IS results. While IS tests have stronger results for interest rate and Sharpe ratio. The significant results for the ENC-NEW indicate that DY, PE, and cay are variables that contain useful information with respect to stock return predictability.

In the already existing literature there is different results regarding IS and OOS testing, this situation is also observed in this thesis based on results presented above. In addition there are different results within the OOS testing, which make it hard to draw any sure conclusion about the power of the test.

4.6 MULTIPLE REGRESSION

Until now we have tested if each of our explanatory variables, one by one, are able to predict stock returns. In addition to check for the individual predictability we perform a multiple regression analysis, where we include all of the variables to check if they have any joint explanatory power. See Appendix H for complete overview of the parameter estimates for the different versions of multiple regressions tested below. Since cay is a variable where we apply quarterly data we do not include it here in the multiple regression. By excluding cay, all the variables in the regression have the same type of data, specifically monthly data. First the long samples of U.S. data are applied before testing the subsamples and then the Scandinavian data. The natural first regression would be the linear regression where we have all our variables on stock return.

(4.3) $$Stock\ return_t = \beta_0 + \beta_1 * DY_{t-1} + \beta_2 * PE_{t-1} + \beta_3 * Interest\ Rate_{t-1} + \beta_4 * Momentum_{t-1} + \beta_5 * Sharpe\ ratio_{t-1} + u_t$$

The test results in table 34 below show that we do not have any strong significant results when running them all together, and all of the beta values are quite low. This is as expected since none of the variables were significant when we tested them individually for the long sample. Despite the poor results, DY is surprisingly significant at a 10% level for a one-sided test. The regression gives a $R^2$of 0.81%, which is relatively low considering that the regression consist of 5 explanatory variables.

U.S Individual sample (All)					
Parameter Estimate					
R-square: 0.81%					
Variable	**B**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.00332	0.00824	-0.4	0.6873	0
DY	0.00240	0.00181	1.32°	0.1859	1.34901
PE	0.00010	0.00016	0.65	0.5172	1.48631
Interest	0.00337	0.00907	0.37	0.7105	23.11206
Momentum	-0.00599	0.01446	-0.41	0.6789	1.14679
Sharpe ratio	0.00045	0.00074	0.61	0.5428	23.13204

Table 34: OLS estimates for all explanatory variables

When the regression consists of more than one regressor we have to be aware of the possible problem regarding multicollinearity. One of the typical signs could be a high R^2 which is not the situation for this case. To check further we look at the VIF, as a rule of thumb this should be under 10 (Gujarati & Porter, 2009, p. 340). We observe a VIF > 10 for interest rate and the

Sharpe ratio. A further graphical investigation of interest rate plotted against Sharpe ratio show a clearly correlated relationship, presented in figure 67. Since interest rate is included in the computation of Sharpe ratio the existence of a correlated relationship between the two variables is expected.

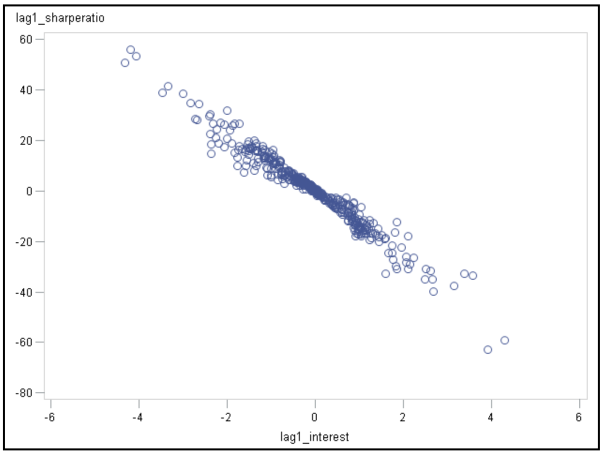


Figure 67: Correlation between interest rate and Sharpe ratio

Although the variables all together do not predict stock returns there may be another combination with fewer variables that gives us some significant results. A natural next step would be to eliminate the variables that perform weakest to see if results improve. Interest rate and Sharpe ratio have the lowest t-values, and by eliminating these from the regression we also remove multicollinearity.

(4.4) $Stock\ return_t = \beta_0 + \beta_1 * DY_{t-1} + \beta_2 * PE_{t-1} + \beta_3 * Momentum_{t-1} + u_t$

The new regression provides an R^2 of 0.48% which is still quite low and lower compared to the first regression, which is natural since two variables are removed. The coefficient for the variable is slightly higher, but not noticeably. DY is still the only variable that shows significance.

4.6.1 U.S. SUBSAMPLE 1974-1986

When running the regression including all the variables, the results show that three out of five variables are significant. These are DY, interest rate, and Sharpe ratio which are the same variables that were significant when the variables were tested individually for this subsample. DY and Sharpe ratio are significant based on a one-sided test at a 10% significance level, and the interest rate is significant for a two-sided test at a 10% level. The reported R^2 is 8.35% which is considerably higher compared to the long sample, but still not a high explanatory power. We continue the testing by removing PE and momentum since they are not significant in the first regression. The three remaining variables are significant, the t-statistic has increased for DY and decreased for interest rate and Sharpe ratio. Regarding multicollinarity we have the same problem in this subsample as we experienced in the long sample. To remove the problem of multicollinarity, we exclude Sharpe ratio and are left with DY and interest rate. The new results show that both variables are significant at the 5% significant level for a two-sided test, and there is no sign for multicollinarity.

4.6.2 U.S. SUBSAMPLE 1987-1999

For this subsample we did not observe any significant variables in the individual testing except for cay, which we do not include in the multiple regression. Based on our previous results, it was surprising that DY and PE show strong significant results at the 5% level when running the multiple regression. For the next regression we remove the variables that have a t-value < 1, which is only the Sharpe ratio in this case, by doing so we also remove the multicollinarity problem. The results from the regression without Sharpe ratio are poorer for all of the variables except from the interest rate which gets slightly better, but is still not significant.

4.6.3 U.S. SUBSAMPLE 2000-2012

Running the multiple regression for the last subsample results in two significant variables, DY and interest rate. The significance level is 10% for a one-sided test, DY has a positive coefficient, and the interest rate has negative values. The new VIF values are smaller compared to the values from the first subsample which were 37 for both, in this sample it is 13 and 12. The values are still over 10 so we can conclude that multicollinearity is present. Momentum is the variable with the lowest t-value and highest p-value so we exclude this from the next regression. In addition we choose to remove Sharpe ratio to eliminate the multicollinarity problem with the interest rate. With the new results the DY is no longer significant, with a t-value of 1.27 it is slightly under the critical value of 1.289. PE becomes significant just above the critical value for a one-sided test at a 10% level and the interest show stronger significant signs.

For the Scandinavian market we only look at four of the six variables, thus the first regression will be the following:

(4.5) $$Stock\ return_t = \beta_0 + \beta_1 * DY_{t-1} + \beta_2 * PE_{t-1} + \beta_3 * Interest\ Rate_{t-1} + \beta_4 * Sharpe\ ratio_{t-1} + u_t$$

4.6.4 NORWAY 1987-1999

From the first regression where all the four variables are included we observe a R^2 of 5.43%, and the fundamental valuation factors have the greatest results. DY is significant at a 5% level and PE is slightly below the critical value for the 10% one-sided test. The two other variables show relatively low t-values with related high p-values. This finding is different from the results we got when we tested each of the variables one by one since interest rate and Sharpe ratio were significant at the same level as DY. PE performs better in combination with other variables than alone. Regarding the multicollinarity problem we saw for the U.S. data, it seems that this is also the case for the Norwegian data. The reported VIF values from SAS are not as high as those for the U.S. data, and Sharpe ratio is the only variable with a VIF over 10. In the next regression, interest rate and Sharpe ratio are dropped as a natural consequence based on the poor results. Without these variables the t-values get stronger for both DY and PE, and PE is now significant at a 10% level in the two-sided test.

4.6.5 NORWAY 2000-2012

The next subsample show stronger results for all of the variables, except from PE, compared to the last period. All of the variables have t-values over 1/-1, and the observed R^2 is 9.74%. DY and interest rate show the strongest results and are significant at the 5% two-sided level. Sharpe ratio is also significant but at a lower level of significance. When PE is removed from the regression both DY and Sharpe ratio get slightly lower t-values, while interest rate increases.

4.6.6 SWEDEN 1987-1999

When we regressed the variables alone on stock returns there were not any significant results for the Swedish data in this time period. This also happen to be the case when we run the multiple regression, all of the variables have relatively low coefficients, t-values, and R^2 is only 0.59%. The VIF values are 7 for both interest rate and Sharpe ratio and taking the rule of thumb into consideration multicollinarity is not a problem in this regression. If we try to exclude Sharpe ratio, which has the poorest result, the new results for the other variables did not improve. It seems that in general there is no predicting power for the Swedish data in this period, neither alone or in combination with each other.

4.6.7 SWEDEN 2000-2012

Looking at the subsample with the most recent data, the results are quite different. The results from the regression show that DY is significant at a 5% level and the interest rate at a one-sided 10% level. R^2 is 9.49% which is a difference of 8.9% compared to the explanatory power in the first subsample. In this sample the VIF values are higher, respectively 17.8 and 15.7 which is over the "limit" of 10. By excluding Sharpe ratio from the regression we eliminate the problem. In addition we remove PE since it is not significance. DY gets slightly reduced but not noticeably, and the interest rate gets strengthened results and is now significant at the same level as DY. R^2 slightly decline down to 9.01% which is a small decrease considering that we remove two variables.

4.6.8 SUMMARY

Multiple regression								
	USA				Norway		Sweden	
	Long sample	74-86	87-99	00-12	87-99	00-12	87-99	00-12
DY	°	°	**	°	**	**		**
PE			**					
Interest Rate		*		°		**		°
Momentum					-	-	-	-
Sharpe ratio		°				°		

Table 35: Summary of significant t-values from the multiple regression

The table above summarizes the results from our multiple regression testing. The results are from the regression where all variables are included. An interesting observation is the behavior of DY when it is regressed in combination with the other variables. When we look at DY alone there are only two situations where there are weak signs of significance among all the periods and markets we are looking at. In the multiple regression, DY is significant at different levels in seven of the tested scenarios, in addition DY is the variable that performs the best considering the level of significance.

In addition to control for multicollinarity between the explanatory variables we have checked for the other econometrics issues as well. The fulfillment of the OLS assumption follows the patterns as we have seen in our earlier research. The heteroscedasticity assumption is fulfilled in all scenarios, and we are able to reject the hypothesis about no autocorrelation for either ρ=1 and ρ=3. Regarding stationarity the results will be the same as observed earlier since stationarity only take one time series into account at each time, and not several time series as a group. The assumption about normal distribution is rejected most of the time because of high JB statistics, which also was the case when looking at the variables one by one.

Overall the multiple regression do not deliver any strong significant results. The maximum number of significant variables at once is three and we never experience that all of the variables are significant at the same time.

5. DISCUSSION

5.1 OUR RESULTS COMPARED TO PREVIOUS STUDIES

We look to previous studies on stock return predictability to compare our U.S. results. Our regressions, sample periods, methods, and sources are not completely identical to any of the previous studies, but we will be capable of comparing some of our results with parts of other authors' studies.

LEWELLEN (2004)

Lewellen (2004) estimates regressions for NYSE equal- and value-weighted returns and for nominal and excess returns (measured net of the one-month T-bill rate). For us the results for the nominal returns are the most relevant, but we will also take the results for excess returns into consideration.

Lewellen check whether the financial ratios DY, B/M, and E/P ratio have predictive power with respect to stock returns. For us, the results on DY and E/P are important. We test for PE and not E/P, but since both of these ratios are based on the same parameters, there is reason to believe that they will have the same ability to predict stock returns.

In the article by Lewellen (2004), the following regressions are used:

(5.1) $$r_t = \alpha + \beta x_{t-1} + \varepsilon_t$$

(5.2) $$x_t = \varphi + \rho x_{t-1} + \mu_t$$

The explanatory variable is lagged by one period, in this case one month. The first equation is the predictability regression, where the explanatory variable is tested for predictive ability with respect to stock returns. We apply the same regression in our tests. The second equation is run to check for stationarity in the explanatory variable. The empirical tests rely heavily on the assumption that ρ is not greater than one. Statistically, the tests remain valid if ρ equals one, but Lewellen assumes that ρ is strictly less than one to be consistent with prior studies. It also makes little sense to predict returns with a nonstationary variable. Economically, x_t should be stationary unless there is an explosive bubble in stock prices. For Lewellen's purposes, the empirical tests should be relatively insensitive to this type of nonstationarity as long as ρ remains below one (Lewellen, 2004). Inspired by Lewellen, we ran the same regression to check for nonstationarity in our explanatory variables since stock returns have stationary time series and the explanatory variables do not, according to our results. For all of the explanatory variables over all periods and sample periods ρ is close to one or lower than one, so we conclude with stationary time series in the explanatory variables.

In similarity with us, DY is tested for predictability over a large sample period, 1946 to 2000, and subsamples. The predictive regressions use the natural log of DY, rather than the raw series, because it is expected to have better time-series properties. Raw DY, measured as a ratio, is likely to be positively skewed and its volatility depends mechanically on its level, log solves both of these problems. Since DY is a ratio, log DY should better approximate a normal distribution (Lewellen, 2004). In contrast to Lewellen we use the raw DY as a predictor in our regressions. Before deciding on our explanatory variables we read through several previous studies, and the majority include DY in its natural form. Therefore we choose to apply the raw DY. B/M and E/P are tested for a sample period from 1963 to 2000, including subsamples, and are also calculated in logs. Again, we use the raw PE ratio, so the results may differ. To check for the difference between applying DY and PE in log and in raw numbers we run the regressions using log, and the results are very much the same, see Appendix I. The coefficient of the explanatory variable always improves when running with log, but the t-statistics do not change substantially. DY is significant in the last sample period for U.S. data, it was not before, but again DY show weaker predictability in the first subsample for Norwegian data. For PE, the results are the same. Hence, there is not a significant difference in the results depending on whether the explanatory variable is computed in log or raw numbers.

When Lewellen regressed NYSE returns on log DY from 1946 to 2000, the OLS slope estimate is 0.92 with a standard error of 0.48. For the first half of the sample, 1946-1972, the bias-adjusted estimate is 0.84 with a p-value less than 0.001. For the second half of the sample, 1973-2000, the bias-adjusted estimate is 0.64 with a p-value of 0.000. The point estimate, standard error, and p-value are all adjusted for bias using the marginal distribution of $\hat{\beta}$, obtained from Monte Carlo simulations. Our results do not display as large slope estimates as Lewellen's results. The largest slope estimate on DY is in the subsample period from January 2000 to September 2012 when it is 0.1253. The differences in the results may be caused by the fact that Lewellen applies log DY while we use the raw DY, or it may be because he uses NYSE returns and we use S&P 500 returns. The NYSE index consists of more than 1900 stocks, which is a significantly higher number than the 500 stocks in S&P 500, and may cause smaller variations in the overall return of the index. In addition, the NYSE index is a global diversified index since it contains of all stocks noted on NYSE, while S&P 500 only consists of U.S. stocks. Another explanation to the differences in results may be that Lewellen's sample does not take the newer financial crisis into account, but on the other hand his sample includes figures from the years following the Second World War.

He finds that B/M and E/P forecast both equal-and value-weighted NYSE returns over the period 1963-1994, but they only predict the equal-weighted index once data for 1995-2000 are included. The evidence for both periods is much stronger than in previous studies. To ensure that the tests are predictive, Lewellen does not update accounting numbers until four months after the fiscal year. Our results on the predictability of PE show that PE only have predictive ability in the individual sample excluding the financial crisis. This sample period stretches from 1968 to 2007, making our period long enough to compare to Lewellen's results. The dissimilarity between these two periods is the current financial crisis, which may have a negative impact on the predictive ability of PE, making our results different from his.

When testing for predictability, Lewellen runs conditional and unconditional (Stambaugh) tests. The test developed in this article is useful only when the predictive variable's sample autocorrelation is close to one (otherwise, high values of ρ are unlikely anyway, so the constraint ρ <1 provides little information). DY, B/M, and E/P at monthly frequency have an autocorrelation near one and most of their movement is caused by price changes in the denominator. Lewellen solely look at the one-sided test, with $\beta > 0$. The predictive variable is assumed to follow a stationary AR1 process where $\rho < 1$. We use the t-statistic checked for OLS assumptions and we run both one-sided and two-sided tests at both the 5% and 10% significance level. We are searching for variables that may have any types of predictability on stock returns, whether it is a negative relationship or a positive.

For simplicity, Lewellen assumes the variables to be normally distributed. We run the JB test in our regressions to check for normality, and the subsample from 1974 to 1986 is the only period where we keep the null of normality in the error term. It would have been interesting to see how high JB statistic Lewellen would get on his sample. Since Lewellen makes the assumption of normal distribution, we choose to not put too much emphasis on the JB statistic and believe that no harm is done to our t-statistic despite that normality is not present. The OLS assumptions can be very sensitive to large outliers, and we believe that this could be one of the reasons to why we reject the null of normality so often. We tested the JB statistic in a sample where we removed some of the large outliers, and the statistic improved immediately. As we can see from the first subsample, the normality is good because the outliers are probably present in the later subsamples.

When testing the DY on equal-weighted NYSE returns the results are very similar for the nominal and excess returns in the period from 1946 to 2000, which is positive for us since we only test for the nominal returns. For the subsamples, 1946-1972 and 1973-2000, the tests strongly reject the null in most cases even though the periods are quite short. In our subsamples DY only reject the null of no predictability in one scenario, the one-sided test where $\beta > 0$ at the 10% significance level, which is not a strong rejection. Hence, Lewellen finds stronger evidence of predictability in DY than we do. Lewellen recognizes the period from 1995 to 2000 as different from the rest in values of DY, but when he only uses 1946-1994 the t-statistic is almost unchanged. Fama and French find that DY predicts monthly NYSE returns from 1941-1986 with t-statistics between 2.20 and 3.21 depending on the definition of returns, so Lewellen conclude that he has stronger results than previous research. Lewellen show that the small-sample distribution studied by Stambaugh (1986; 1999) and Nelson and Kim (1993), which has become standard in the literature, can substantially understate, in some circumstances, DY's predictive ability (Lewellen, 2004).

E/P appears to forecast nominal returns, but there is little evidence that it forecasts excess returns, which again is positive for us since we only test for nominal returns. For both nominal and excess equal-weighted NYSE returns, the addition of 1995-2000 again strengthens the case for predictability. The financial crisis sample is the only scenario in our regressions where P/E shows predictive power. We reject the null of no predictability when we run a one sided test with $\beta > 0$ at the 10% significance level, which does not indicate a strong rejection. Hence, Lewellen again finds stronger signs of predictability in the explanatory variable than we do.

AVRAMOV (2002)

Avramov (2002) uses a Bayesian model averaging to analyze the sample evidence on return predictability in the presence of model uncertainty. Given that the true set of predictive variables is virtually unknown, this paper proposes a Bayesian model averaging approach to analyze stock return predictability. In the context of predictive regressions, the Bayesian methodology is attractive. It explicitly incorporates model uncertainty, and is therefore robust to model misspecification, at least within the universe of linear forecasting models (Avramov, 2002).

The Bayesian model averaging procedure computes posterior probabilities for the collection of 2^M models, where M is the number of explanatory variables. It then uses the probabilities as weights on the individual models to obtain one composite weighted forecasting model, which summarizes the dynamics of future stock returns. The weighted model is employed to investigate the sample evidence on predictability and to analyze investment implications of model uncertainty (Avramov, 2002). We compute a completely different model, where one and one regression is run at a time. Still it will be interesting to compare our results with Avramov's to determine whether the choice of model have a significant effect on the results.

In applications studying stock return predictability there are many possible explanatory variables but only a limited number of observations. Avramov thus believe that the traditional single predictive regression paradigm offers little help in identifying useful predictors, which are the type of regressions that we run. Hence, the Bayesian methodology is applied to find the predictable variables among a great selection of financial variables (Avramov, 2002). Some of the variables that are included in the analysis are DY, B/M, earnings yield, momentum effect, default spread risk, monthly rate, default risk premium, term premium, monthly inflation rate, cay, and value premium.

Avramov uses monthly and quarterly observations on stock returns and information variables from April 1953 through December 1998. The variables that are comparable with ours are DY on the value-weighted NYSE index, earnings yield on the Standard and Poor's Composite, the winners-minus-losers one-year momentum in stock returns, monthly rate of a three-month Treasury bill, and cay. Avramov reports that when accounting for estimation and model risks, the perceived distribution of future returns departs from normality, and may be affected by higher-order moments such as skewness and fat tails. The normality assumption is rarely obtained in our regressions as well, and we believe the reason may be the same as in Avramov's case with large outliers causing skewness.

The analysis performed by Avramov finds that in the monthly observations term premium (defined as the rate of return differential between long-term and short-term treasuries) and the market premium are useful predictors of future stock returns. On the other hand, the DY and book-to-market, among others, have relatively small posterior probabilities of being correlated with future returns. This may coincide with our results for DY in the individual sample from January 1965 to September 2012 where we find a t-statistic of 1.26 for DY and hence no predictive power. The financial crisis sample is also long enough to be comparable to Avramov's sample period, and here we get a t-statistic of 1.14, which again means no predictability in DY.

The only case where momentum effect has tendencies of showing predictability is when Avramov run a regression with stock returns as the dependent variable and all the explanatory variables are included in the regression. Other than that, momentum effect rarely shows any signs of being a strong predictor. Our results when testing momentum effect for predictability are similar to the results of Avramov. Momentum effect fails to reject the null of no predictability in all of our sample periods, hence we conclude with no predictability in momentum effect. Avramov test for the one-year momentum and we test for the 6-month momentum, therefore maybe a shorter horizon momentum will have signs of predictability, like a 3-month momentum. We cannot reject momentum to be unpredictable overall, since there may be other momentum strategies that shows stronger signs of predictability.

For the quarterly observations the variable cay dominates DY, the market premium, default-risk spread, and term spread, predictive variables also studied by Lettau and Ludvigson (2001). Lettau and Ludvigson (2001) introduced cay as a powerful predictor of quarterly returns at short and intermediate horizons (Lettau & Ludvigson, 2001). Additionally, the variable outperforms B/M, momentum, value premium, and inflation. Cay has proven to be a strong predictor in our tests as well. In our financial crisis sample and in all of our three subsamples, cay show predictive ability when running the one-sided test for $\beta > 0$ at the 10% significance level.

Term premium and cay are close to significant or significant in forecasting quarterly returns based on t-ratios that ignore model uncertainty, but not when such uncertainty is incorporated. Avramov finds that taking model uncertainty into account appears to substantially diminish the predictive power of some explanatory variables. Model uncertainty is proven to be an important factor to take into consideration. For example, when based on variance decompositions, the analysis shows that model uncertainty is more important than estimation risk for short-horizon investors (Avramov, 2002).

PESARAN AND TIMMERMANN (1995)

An alternative approach to evaluating the economic significance of stock market predictability would be to see if the evidence could have been exploited successfully in investment strategies, like Pesaran and Timmermann do in their article from 1995. This can be done by either evaluating the track records of portfolio managers in "real time", and see if these portfolios systematically generate excess returns, or by simulating investors' decisions in real time using publicly available information on a set of factors thought a priori to have been relevant to forecasting stock returns (Pesaran & Timmermann, 1995).

The purpose of the article by Pesaran and Timmermann is to assess the economic significance of the predictability of U.S. stock returns, explicitly accounting for the forecasting uncertainty faced by investors who only have access to historical information. They assume that agents establish a base set of potential forecasting variables and, at each point in time, search for a reasonable model specification, capable of predicting stock returns, across this set. Selection criteria such as $\bar{R}^2$, Akaike information criterion, or Bayesian information criterion (BIC) are applied to prove which variable show predictability at that time and will be included in the model. The primary aim is to select a forecasting equation that could be viewed at the time as being a reasonable approximation to the data generating process. It also captures the possibility than an investor may switch from one model to another in light of new empirical evidence obtained as the sample size expands (Pesaran & Timmermann, 1995).

All variables were measured at monthly frequencies over a period from 1954 to 1992. Stock prices came from the S&P 500 index. Excess return is applied as the dependent variable. Explanatory variables they test that may be relevant for us are DY, earnings-price ratio, and 1-month T-bill rate. The dividend and earning yields are based on 12-month moving averages, and a one period lag of these variables was included in the base set. For the interest variables both a two month lagged and a one month lagged value is included. Even though we do not calculate the 12- month moving average on our variables, we lag our variables with one month, so the results from these regressions are appropriate for us.

Pesaran and Timmerman's results show that the DY variable is selected in most periods from 1970 onward. The E/P ratio lagged one month is not selected as often in the forecasting models, in fact, this regressor is never chosen by the BIC. The results cannot be directly compared to our case, since we search for signs of any form of predictability in the explanatory variables while they are interested in which variable show the strongest predictive ability at a specific point in time. It is interesting that Pesaran and Timmermann would select DY in most of their periods, since this is opposite from our results. DY only shows signs of predictability on stock returns in our financial crisis sample, in all other scenarios there is no indication of predictive ability in DY. The results on E/P ratio on the other hand are coincided with our results since we find predictive ability in PE in only one of our sample periods.

The only variable to be included in the forecasting models throughout the entire sample period from 1960 to 1992 is the one-month lagged value of the one-month T-bill rate. We do not check for predictability in such a short rate as the one-month T-bill, but we check for the three-month T-bill and our results are somewhat similar to the results of Pesaran and Timmermann (1995). The three-month T-bill is the variable that shows the strongest predictability in most periods, we reject the null of no predictability in two of our subsamples and in the financial crisis sample period when we run the two-sided test for $\beta \neq 0$ at the 5% significance level. Together with Sharpe ratio, the interest rate is the only variable that rejects the null when running a two-sided test at the 5% significance level.

Pesaran and Timmermann conclude that there does not seem to be a robust forecasting model in the sense that the determinants of the predictability of stock returns in the U.S. seem to have undergone important changes throughout the period under consideration (Pesaran & Timmermann, 1995).

RAPACH AND WOHAR (2006)

Rapach and Wohar (2006) undertake an extensive analysis of IS and OOS tests of stock return predictability to better understand the nature of the empirical evidence on stock return predictability. They test for return predictability by applying a bootstrap procedure that explicitly accounts for data mining when calculating critical values, and they find that certain financial variables display significant IS and OOS predictive ability with respect to stock returns (Rapach & Wohar, 2006). The key to controlling for data mining is the use of appropriate critical values for both IS and OOS predictability tests. We check for both OOS and IS in the last subsample for U.S. data, hence the results from this testing will be relevant to compare with Rapach and Wohar's results.

Annual data from 1927 to 1999 and two stock return series are applied; log real returns on the S&P 500 and CRSP equal-weighted portfolios. Variables that are tested for predictive ability, that we also look at, are D/P ratio, PE ratio, and the short-term interest rate. They follow the literature and assess IS predictability via the t-statistic corresponding to the slope coefficient in a predictive regression model. The alternative hypothesis is one-sided with $\beta > 0$. Rapach and Wohar test for OOS predictability by comparing OOS forecasts generated by a model of constant returns to forecasts generated by a model that utilizes a given financial variable using two recently developed t-statistics. When we compare our results it is important to keep in mind that Rapach and Wohar test for annual data in these regressions, and not monthly, and their sample period stretches over a longer time period than ours.

First, they test for IS predictability. For the S&P 500 stock returns PE have significant IS predictive ability at the 10-year horizon. We test the predictability of PE over a 1-month horizon, so the results will not be fully comparable to a horizon as long as 10 years. For this data set, there is no discrepancy between the IS and OOS test results once they use powerful OOS tests, so if we rely on these results it is sufficient enough to look solely on IS predictability. In the CRSP equal-weighted index the D/P ratio display significant IS predictive ability at the 5-year horizon, and the short-term interest rate evinces significant IS predictive ability at the 10-year horizon. Again, the horizon they look at is too long to be comparable with our 1-month horizon. In this data set there is also little discrepancy between IS and OOS (Rapach & Wohar, 2006).

When testing for OOS predictability in S&P 500 the results are that in the long horizon of 10 years, PE is a significant predictor. For CRSP, the D/P ratio has significant predictive power at the 5-year horizon, while the short-term rate has significant predictive ability at the 10-year horizon. In our results, PE show predictability when we look at the MSE-F and ENC-NEW statistic, but the t-statistic from the OOS regression fail to reject the null hypothesis which is in line with our IS testing. DY is estimated to contain helpful information by the ENC-NEW statistic and by the OOS regression, but the MSE-F statistic conclude with no predictability. For the short-term rate our results indicate predictability in both IS and OOS testing, though the IS testing had the strongest indication of predictive power. Again, Rapach and Wohar find significant predictive ability in the short-term rate over a long horizon of 10 years when testing for OOS predictability, while we find stronger predictability in IS testing over a one month horizon. The difference in long and short horizon may partly explain the variance in results.

Rapach and Wohar also computed IS and OOS tests using postwar quarterly data covering from the second quarter in 1953 to the fourth quarter in 2000 for S&P 500 and CRSP equal-weighted log real returns. Here they include a new variable, cay, which we also test for. In S&P 500, cay is significant at the 1-quarter horizon, 8-quarter horizon, and 16-quarter horizon when applying IS tests, and the ENC-NEW conclude with OOS predictability in all three periods while MSE-F collaborates on the 8-quarter and 16-quarter horizon. When applying CRSP, cay has significant t-statistics at the 1-quarter horizon, 8-quarter horizon and 16-quarter horizon, and again they find that the ENC-NEW concludes with similar OOS predictive ability in cay in all three horizons. We test cay over the 1-quarter horizon, and conclude with predictability in four of five cases when testing for IS predictability. Hence our results are congruent with Rapach and Wohar. When applying OOS tests the predictability of cay seems even stronger, and both ENC-NEW and MSE-F concludes with predictive power in cay.

When they control for data mining in the S&P 500, they still find significant evidence of predictability at the 10-year horizon, but not in any of the variables we use. None of the IS or OOS maximal statistics in the CRSP equal-weighted returns are significant at the 5-year and 10-year horizons, so that robust evidence of predictability is limited to a short horizon (Rapach & Wohar, 2006). This may indicate that the case of data mining is stronger for the longer horizons, and we can assume that it is not threatening in our short horizon of one month.

5.2 SHARPE RATIO

Sharpe ratio is included in the analysis out of curiosity from the authors as to whether it could be a predictor that has not been tested for earlier. Hence, there are not many previous studies to compare our results with. Our results draw some counterintuitive conclusions. In the long samples, the individual and the financial crisis samples, Sharpe ratio only shows predictability in the sample that excludes the financial crisis. But when we tested for predictability in the subsamples, and the Sharpe ratio is significant in the sample from 1974 to 1986, and from 2000 to September 2012.

These findings could mean that the period between 2007 and 2014 has such poor predictability that it ultimately neutralized the predictability that was in the period from 2000 to 2012 in the individual sample. Something to notice is that Sharpe ratio shows predictability in the subsample between 2000 and September 2012 when we run the two-sided test at the 10% significance level. The t-value of Sharpe ratio is negative, meaning that we test for $\beta < 0$, so Sharpe ratio has a negative predictability in this subsamples. Hence, when the Sharpe ratio increases, the stock returns are expected to decrease, which is not intuitive according to literature.

5.3 FINANCIAL CRISIS

We exclude the financial crisis from our individual sample to test for the effects of the recent financial crisis. By looking at the graphs over the different variables' development over the years we observe some changes around 2007 to 2009. DY has an increase in value around 2008 after a rather stabile progress from 2003 to 2006. PE ratio experiences a dramatic increase in 2009 at an all-time high around 40, when it usually circles around 20 in value. Both the interest rate and the momentum effect are decreasing in 2008 and 2009. Cay continues to rise as the trend has been in the variable, and the Sharpe ratio has a peak around 2007 before it decreases again.

Our purpose of excluding the financial crisis from the sample was to see if the events after 2007 can change the predictive power of the variable. In four of our six variables, PE ratio, interest rate, cay, and Sharpe ratio, the predictability is present when applying the truncated sample. These findings indicate that the observations from the period during the financial crisis and after are contributing to changes in predictability in the variables that we test for.

5.4 U.S. RESULTS COMPARED TO SCANDINAVIAN RESULTS

To check whether the results obtained for U.S. data also are present in other markets we expanded our analysis with the Norwegian and Swedish market. We have looked at four of the six variables we used in our main analysis; DY, PE ratio, interest rate, and Sharpe ratio. The two time periods we are looking at are equal to the last two subsamples which were tested for in the U.S. data, 1987-1999 and 2000-September 2012. Below our significant results for the two periods for all of the three markets are summarized.

	1987 - 1999			2000 - 09:2012		
	U.S.	Norway	Sweden	U.S	Norway	Sweden
DY		*				
PE					*	
Cay	°	-	-	°	-	-
Interest Rate		*		**	**	**
Momentum		-	-		-	-
Sharpe Ratio		*		*	**	**

Table 36: Summary of significant t-values for subsamples 1987-1999 and 2000-Sept 2012

From our first test with the U.S. data we got poor results for the period from 1987-1999 with only one significant variable, which was cay, the only variable where we looked at quarterly data. The most surprising result from our analysis of the Scandinavian market was the results from Norway in this first period, three of the four tested variables were significant and shows predictability. The observation of the result within the Scandinavian market shows that the results are not similar and robust for all the variables. The Swedish results are poor with no significant results in the first period and two in the last. In contrast, the Norwegian results show significance in four more cases.

PE is determined by the performance of a company, and DY is determined by the dividend policy in a company and might be quite different between companies. In total the DY and PE for each of the indices are influenced by the performance of all the shares included in the index. The indices we are looking at differ with regard to the number of stocks, size of the company, liquidity, and industries. Taking these differences into account it is no wonder that we do not see robust significant results across the three markets. For these two periods the two factors only show significant result for Norway, DY in 1987 - 1999 and PE for 2000 – September 2012.

Further investigation of the DY show that there is a great jump in DY for all the markets, this is the case from the second half in 2008 to the first half in 2009. The Norwegian and Swedish index experience a larger jump compared to the U.S., the S&P 500 and OSEAX index peaks in February 2009 and OMXS peak in December 2008. To get an idea of the reason behind this happening we investigated the share price to see if there have been some unusual movements during that time period. In figure 68 below the share price for the three indices are presented for the time period between January 1987 and September 2012, the period we test our Scandinavian data for. As the graph shows, there is a decrease in the share price around the period where the DY increased, two movements that seem to belong together. The fall in stock price could be connected to the financial crises in 2008, and as expected it is S&P 500 that declines the most. Since we observe these outliers in DY, it is no surprise that we do not have any significant results in the last subsample, since we know that the regression we run is sensitive to outliers in the dataset.

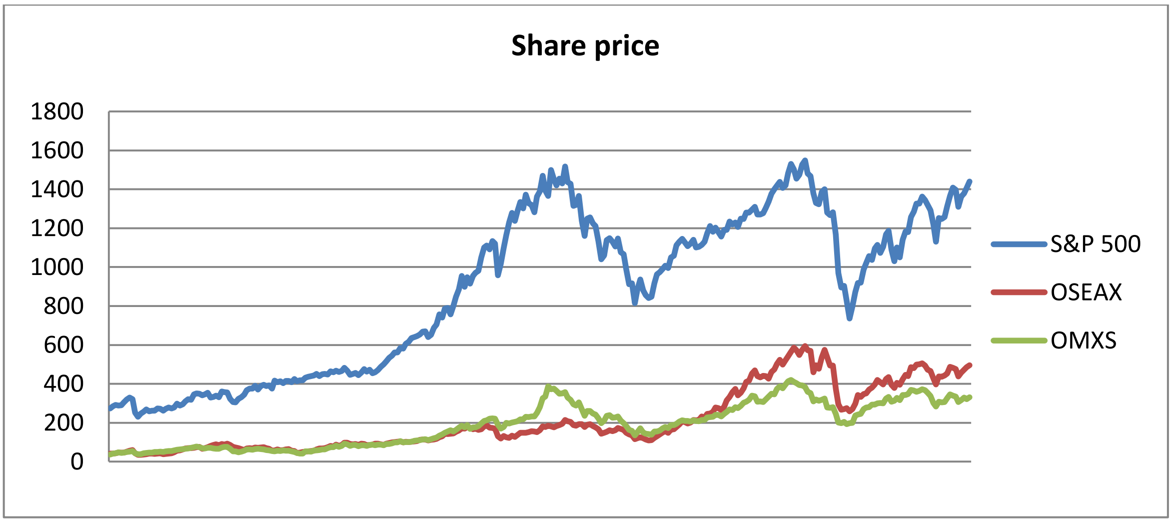


Figure 68: Development in share price from Jan 1987 to Sept 2012

Outliers in the dataset may be a good explanation to the poor result on PE as a significant ratio. Looking back to the test where we exclude the financial crisis for the individual long sample testing, we got a significant result for the U.S. We observe some large outliers from the financial crisis, which are present in the last subsample that probably destroys the results. As an effect of the financial crises the firms probably perform poorer which results in lower earnings, and for companies with an increase in share price the PE escalated to abnormal high values. For the Scandinavian market we also observe these outliers but at different points in time than the financial crisis. For Norway these are present in the beginning of the period, March 1989 – February 1990, and the test show no significant result for this period. These outliers may be due to the banking crisis that occurred in Norway in the late 1980s and lasted until the beginning of 1990s. However, in the last subsample PE is significant for Norway at a 10% significance level. In Sweden the outliers are present in the middle of the whole sample, which means that both subsamples are affected and we observe no significant results.

The results from our test of U.S. data show that interest rate and the Sharpe ratio are the variables with the strongest significant results. This also proves to be the case for the Norwegian and Swedish data in the last period, thus the results are robust across the three markets.

The effect on stock prices from movement in interest rate would normally be opposite. In other words, if the interest rate increases we should expect a decrease in stock price and thus also a negative return, and reverse if the interest rate decreases. Based on this theory we should expect a significant relationship between the interest rate and stock return, and from our tests we see that this is the case in the last period for all markets. For the Norwegian market the interest rate variable is significant also in the other period, 1987-1999. Our findings are consistent with Schrimpf (2010) who finds that interest rate-related variables usually are among the most prominent predictive variables (Schrimpf, 2010). Looking at the coefficient in our linear regression with the lagged interest rate as the explanatory variable, the coefficient is negative in 5 out of the 6 cases. This means that a one unit increase in interest rate leads to a decline in stock return, which matches the theory mentioned above. The only case with a positive coefficient is in the U.S. market in the last period.

Looking into the movement of our variables, the time series for interest rate and Sharpe ratio are more volatile than the time series for DY and PE. The volatile time series are characterized by frequent fluctuations in the observations and a graph that indicate a mean close to zero. Since the observations have frequent fluctuation, a clear increase or decrease would not make any significant impact on the time series as a whole. A clear deviation from the mean would almost always be compensated with a deviation later in the opposite direction, thus the mean would be unaffected. Fluctuations in horizons as short as one month can cause great effect to the results, but since the values are adjusted by the 12-month moving average there is reason to believe that the influence from the deviation is minimized. For DY and PE a deviation from the normal level has a greater impact on the time series, because these deviations are more unusual and happens infrequent. This difference in pattern of the time series may help to explain why we have a robust significant result for interest rate and Sharpe ratio, and not for DY and PE.

6. CONCLUSION

Since the breakthrough in 1988 by Fama and French, stock return predictability has become a well-researched subject engaging several new studies on variables that can possess predictive ability on stock returns. Campbell and Schiller (1989) were early involved in the research, and in similarity with Fama and French they tested for predictive ability in different financial variables. The discoveries of this research inspired other scholars to test for predictability in additional categories of variables. Macroeconomic factors such as interest rate, cay, inflation, and exchange risk have been regressed on stock returns in search for explanatory power, in addition technical factors including momentum effect has been applied as an independent variable on stock returns.

In light of the previous research we desired to investigate some of the variables in more recent sample periods, and include the latest financial crisis in our sample. DY and PE were included as our fundamental valuation variables, interest rate and cay entail macroeconomic features, and momentum effect was a technical contribution to our variables. Based on the properties of Sharpe ratio we wanted to include this parameter as a possible new explanatory variable of stock return.

Our findings from running the main IS regression of each control variable on stock return show different signs of predictable power. The fundamental valuation factors, DY and PE, show weak signs of predictive power on stock returns since they are only significant at the lowest level of significance in one sample period each. Cay is the only variable that is run in log and on a quarterly basis, and it is a variable that turns out to be significant in four out of five periods. Therefore we conclude that cay is a very stable variable with predictive ability on stock returns. The other macroeconomic factor, interest rate, turns out to be our strongest predictor. We reject the null hypothesis at the 5% significance level in three of five scenarios when regressing interest rate. Based on our results momentum effect proves to be our weakest predictor since it does not manage to reject the null hypothesis in any of our scenarios. From our OOS tests we investigate whether predictability is present when running a different type of test than IS. Our results indicate OOS predictability in all the variables which is concluded by the MSE-F statistic, the ENC-NEW statistic, or both.

Sharpe ratio is our contribution to the thesis as this is a variable that we believe may have predictability, and has not received much attention in previous studies. Sharpe ratio can be applied as a parameter for investors when deciding among portfolios, hence we anticipate Sharpe ratio to be able to predict stock returns. Given that interest rate is a part of the calculation of Sharpe ratio, the two parameters are expected to show similar predictive ability. This expectation is consistent with our results since Sharpe ratio proves to be one of our strongest predictors, and is significant in the same sample periods as interest rate. Therefore, our contribution is considered to be beneficial.

We extended our analysis by including Norwegian and Swedish data on DY, PE, interest rate, and Sharpe ratio in the last two subsample periods. The reason for including Scandinavian data is to check for robustness in the results across international markets. After comparing the three markets Norway stands out as the country with the highest number of significant parameters. The resemblance between the markets is the strong prediction of Sharpe ratio and interest rate. DY and PE in both U.S. and Swedish data show no indication of ability to predict stock returns, while Norwegian data has weak signs of predictability in the two variables. For that reason we conclude that the explanatory variables have a rather equal predictable power in the three markets.

The financial crisis' effect on stock return predictability is tested for by excluding the data from 2008 up until today from the individual sample. The results improve notably when we exclude the financial crisis, and four of our variables are now showing significance at different levels. In contrast none of the variables were able to reject the null hypothesis in the original sample. The interest rate values could for example shift from high to very low during the financial crisis due to default risk spread. Therefore we conclude that the data after 2008 contribute to destroy the variables predictive ability in the individual sample.

The research on stock return predictability involves testing of several variables, over various time periods, applying a variety of methods, and they are regressed on different indices. When comparing our results to previous studies we take into account that differences will be present. Hence, we search for commonalities in our results with previous research and find that sometimes there are similarities and sometimes not.

The research on stock return predictability will most likely continue to expand in the future. We consider the financial crisis to be of high relevance and think it would be interesting to see further research performed on newer sample periods. Since Sharpe ratio turned out to be a strong predictor in our results, it would be fascinating to see if other authors find predictability in this parameter. We believe that there is still uncovered potential on this subject, and more predictable variables are probably still to be discovered.

LIST OF REFERENCES

Ang, A., & Bekaert, G. (2007). Stock return predictability: Is it there? *Review of Financial Studies, 20*(3), 651-707.

Avramov, D. (2002). Stock return predictability and model uncertainty. *Journal of Financial Economics, 64*(3), 423-458.

Avramov, D., & Chordia, T. (2006). Predicting stock returns. *Journal of Financial Economics, 82*(2), 387-415.

Berk, J., & DeMarzo, P. (2014). *Corporate finance* (3. ed. ed.). Harlow: Pearson Education Limited.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton. N. J:

Campbell, J. Y., & Shiller, R. J. (1989). *Stock prices, earnings and expected dividends*. Cambridge, Mass: National Bureau of Economic Research.

Campbell, J. Y., & Yogo, M. (2006). Efficient tests of stock return predictability. *Journal of Financial Economics, 81*(1), 27-60.

Chui, A. C. W., Titman, S., & Wei, K. C. J. (2010). Individualism and momentum around the world. *Journal of Finance, 65*(1), 361-392.

Clark, T. E., & McCracken, M. W. (2009). Nested forecast model compariosons: A new approach to testing equal accuracy.(Research Working Paper 09-11, The Federal Reserve Bank of Kansas City Economic Research Department)

Clark, T. E., & McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics, 105*(1), 85-110.

Eitrheim, Ø, Klovland, J. T., & Qvigstad, J. F. (2007). *Historical monetary statisitcs for norway - part II.* (No. 38). Oslo: Norges Bank.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 1989, 25; Vol.25*(1; 1), 23; 23-49; 49.

Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal, 21*(5), 55-59.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*(2), 383-417.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*(1), 3-25.

Federal Reserve Board. (2014). Financial accounts of the united states - historical data. Retrieved April 11, 2014, Retrieved from http://www.federalreserve.gov/releases/z1/Current/data.htm

Gao, P., & Huang, K. X. D. (2004). Aggregate consumption-wealth ratio and the cross-section of stock returns: Some international evidence. *Research Division, Federal Reserve Bank of Kansas City,*

Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5. ed. ed.). London:

Ho, J. C., & Hung, C. (2012). Predicting stock market returns and volatility with investor sentiment: Evidence from eight developed countries. *Journal of Accounting & Finance (2158-3625), 12*(4), 49-65.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance, 48*(1), 65-91.

Johnston, K., Hatem, J., & Scott, E. (2013). A note on the evaluation of long-run investment decisions using the sharpe ratio. *Journal of Economics & Finance, 37*(1), 150-157.

Kendall, M. G., & Hill, A. B. (1953). The analysis of economic time- series-part I: Prices. *Journal of the Royal Statistical Society.Series A (General), 116*(1), 11-34.

Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance, 2001, 56; Vol.56*(3; 3), 815; 815-849; 849.

Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics, 74*(2), 209-235.

Lo, A. W. (2002). The statistics of sharpe ratios. *Financial Analysts Journal, 2002, 58; Vol.58*(4; 4), 36; 36-52; 52.

McCracken, M. W. (2007). Asymptotics for out of sample tests of granger causality. *Journal of Econometrics, 140*(2), 719-752.

Norges Bank.Statskasseveklser, månedsgjennomsnitt. Retrieved May 27, 2014, Retrieved from http://www.norges-bank.no/no/prisstabilitet/rentestatistikk/statskasseveksler-rente-manedsgjennomsnitt-av-daglige-noteringer/

Oslo Børs.Aksjeindekser. Retrieved May 26, 2014, Retrieved from http://www.oslobors.no/Oslo-Boers/Produkter-og-tjenester/Markedsdata/Indekser/Aksjeindekser/(tab)/2

Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *Journal of Finance, 50*(4), 1201-1228.

Rapach, D. E., & Wohar, M. E. (2006). In-sample vs. out-of-sample tests of stock return predictability in the context of data mining. *Journal of Empirical Finance, 13*(2), 231-247.

Rapach, D. E., Wohar, M. E., & Rangvid, J. (2005). Macro variables and international stock return predictability. *International Journal of Forecasting, 21*(1), 137-166.

Schrimpf, A. (2010). International stock return predictability under model uncertainty. *Journal of International Money and Finance, 29*(7), 1256-1282.

Stambaugh, R. F. (1986). Biases in regressions with lagged stochastic regressors.(Working Paper No. 156, Graduate School of Business, University of Chiago)

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics, 54*(3), 375-421.

Tsai, C., & Hsiao, Y. (2010). Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. *Decision Support Systems, 50*(1), 258-269.

Zvi, B. (2011). In Kane Alex, Marcus Alan J (Eds.), *Investments and portfolio management*. New York.

APPENDIX

APPENDIX A

DATASTREAM

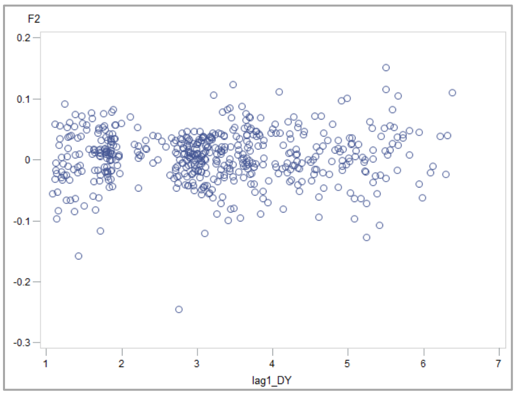
Variable	Datastream code	Frequency	Start date
U.S. stock index prices	S&PCOMP(PI)	Monthly	31.01.1964
U.S. stock individual prices	"Name of company"(P)	Monthly	31.01.1973
U.S. interest rate	USGBILL3	Monthly	31.01.1972
U.S. dividend yield	S&PCOMP(DY)	Monthly	29.01.1965
U.S. price-earnings	S&PCOMP(PE)	Monthly	31.01.1968
Norwegian dividend yield	"Name of company"(DY)	Monthly	31.01.1983
Norwegian price-earnings	"Name of company"(PE)	Monthly	31.01.1973
Swedish stock index prices	SWESALI(PI)	Monthly	31.12.1979
Swedish stock individual prices	"Name of company"(P)	Monthly	29.01.1982
Swedish interest rate	SDGBILL3	Monthly	15.01.1982
Swedish dividend yield	"Name of company"(DY)	Monthly	29.01.1982
Swedish price-earnings	"Name of company"(PE)	Monthly	29.01.1982

"Name of company" means that information on each stock included in the index is collected

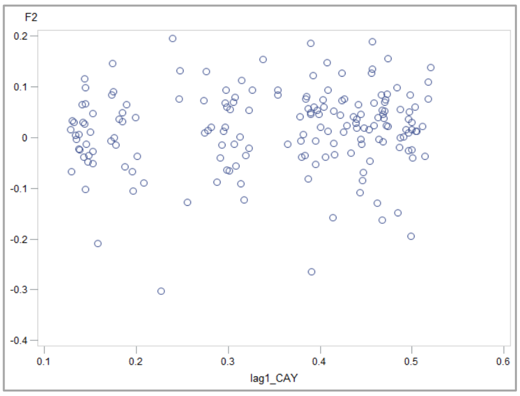
APPENDIX B

RELATIONSHIP BETWEEN X AND Y

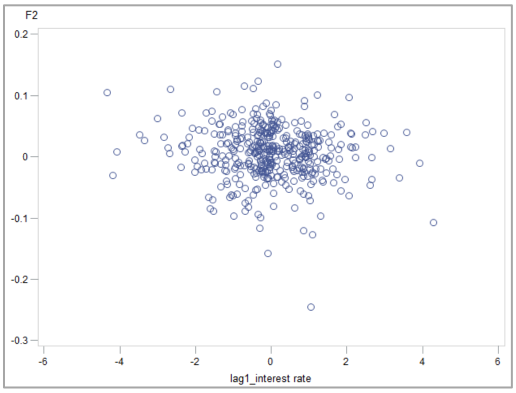
U.S. FINANCIAL CRISIS



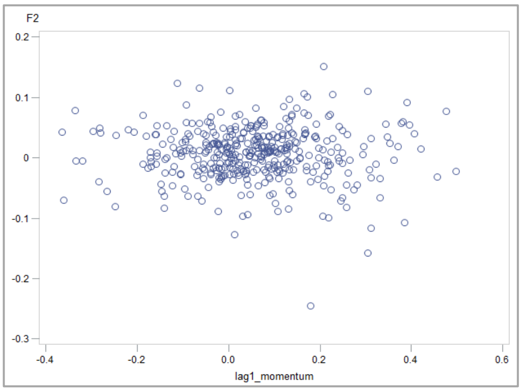
Dividend yield



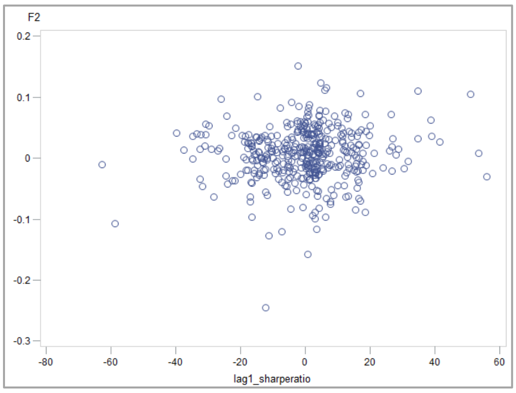
Cay



Interest rate

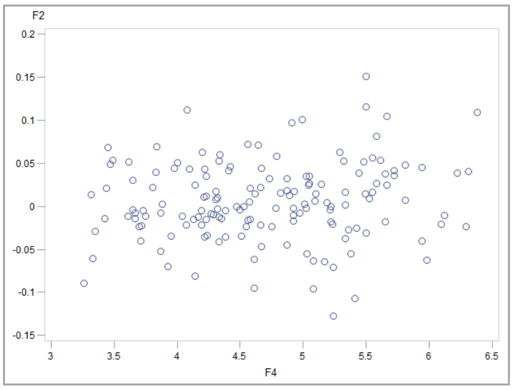


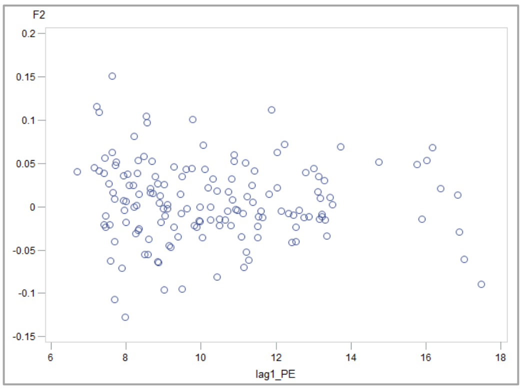
Momentum



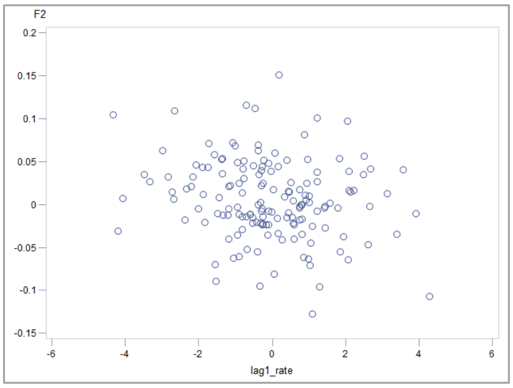
Sharpe ratio

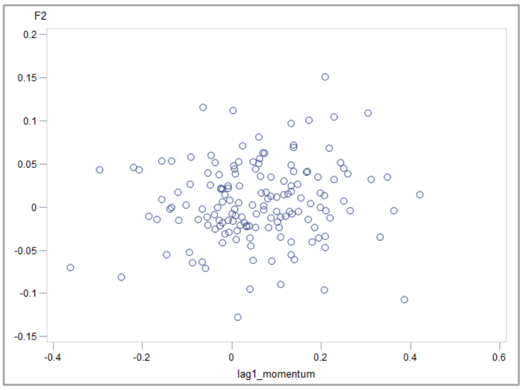
U.S. SUBSAMPLE 1974-1986





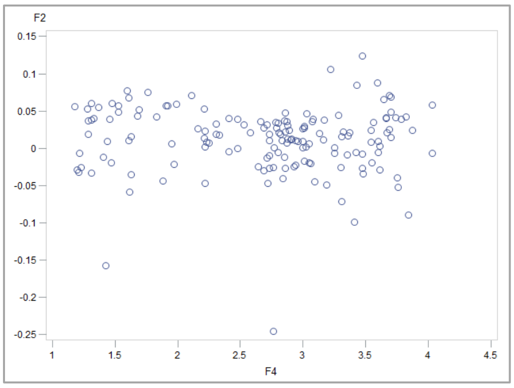
Dividend Yield *Price-earnings*

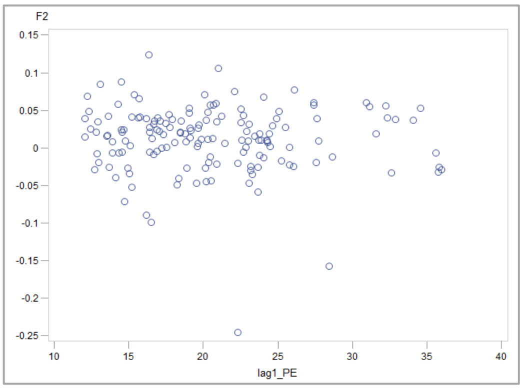




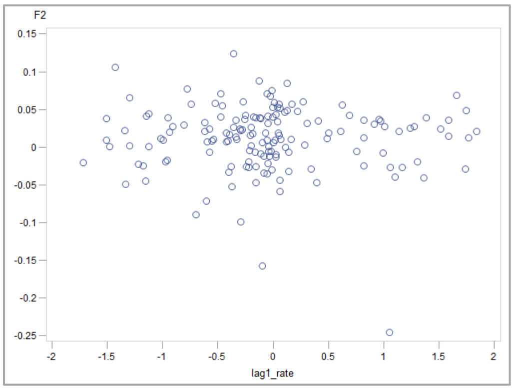
Interest rate *Momentum*

U.S. SUBSAMPLE 1987-1999

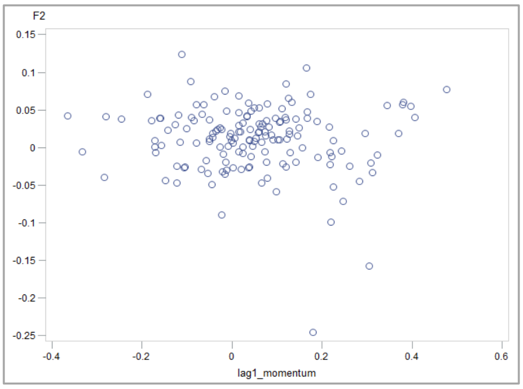




Dividend Yield *Price-earnings*

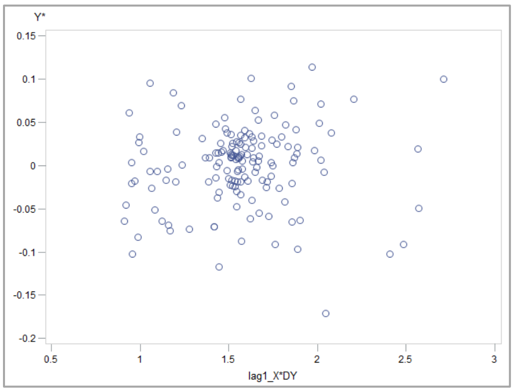


Interest rate

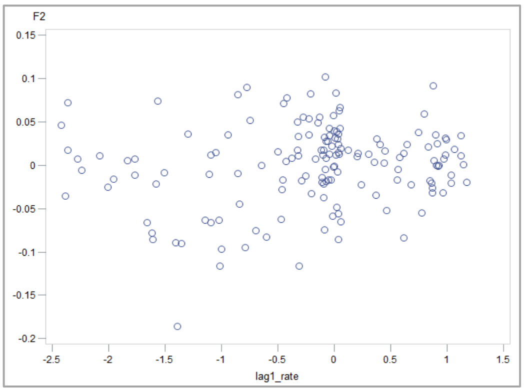


Momentum

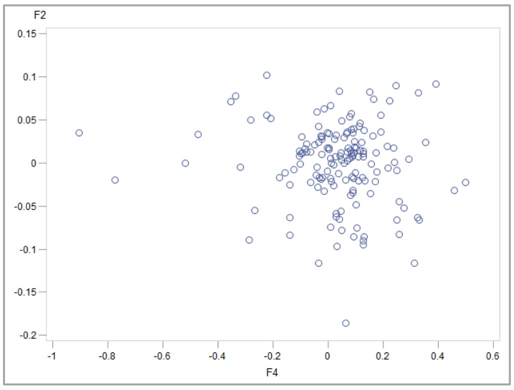
U.S. SUBSAMPLE 2000-09:2012



Dividend Yield (GLS)

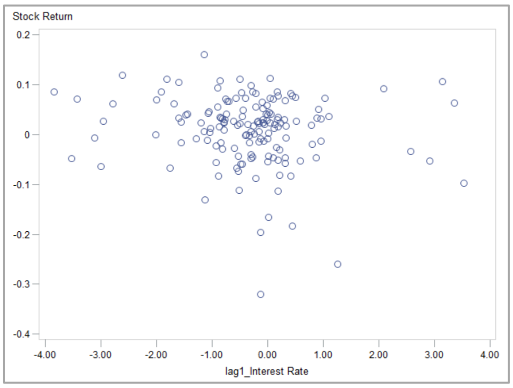


Interest rate

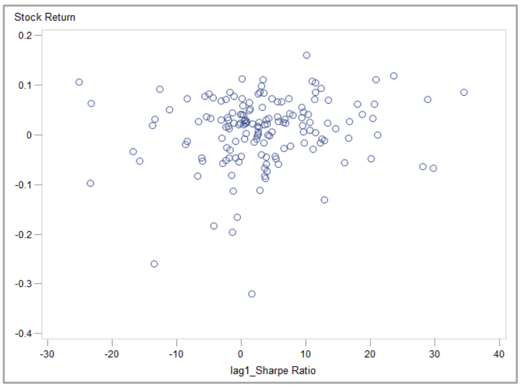


Momentum

NORWAY SUBSAMPLE 1987-1999

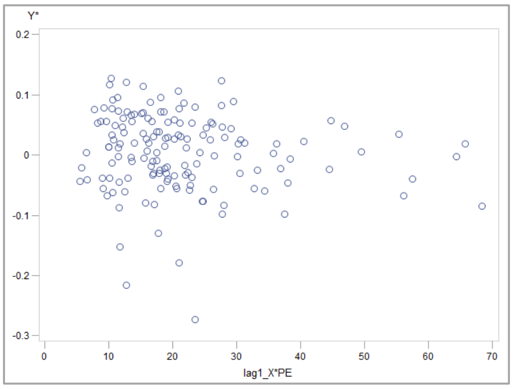


Interest rate

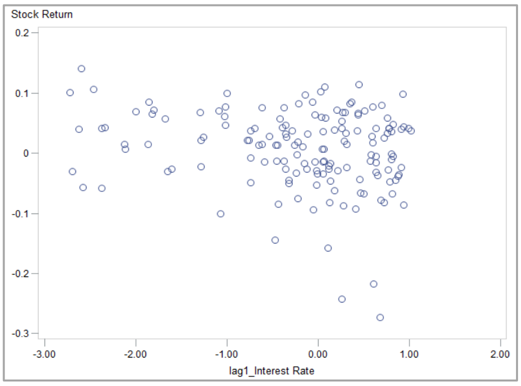


Sharpe Ratio

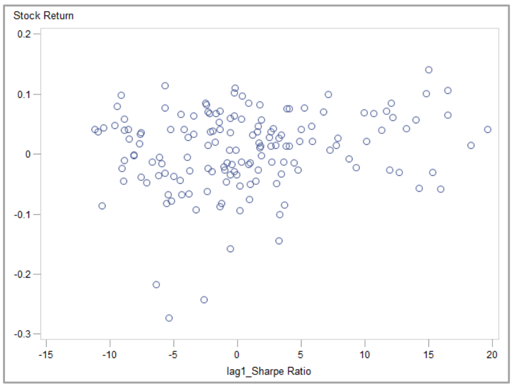
NORWAY SUBSAMPLE 2000-09:2012



Price-earnings (GLS)

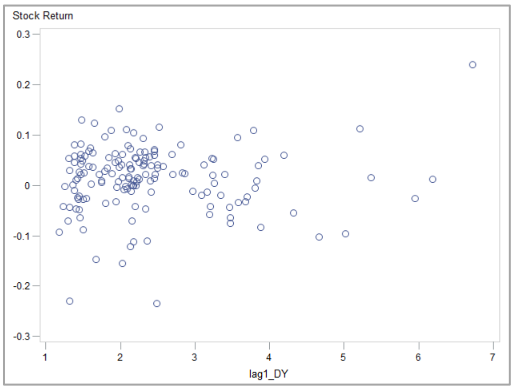


Interest rate



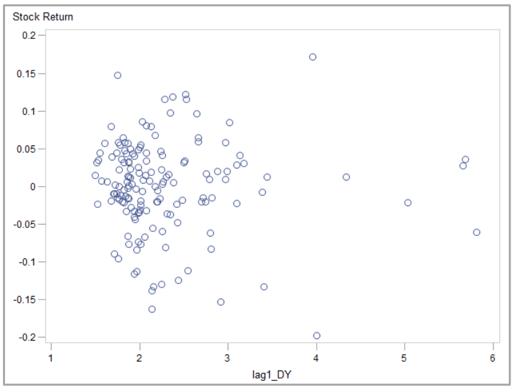
Sharpe ratio

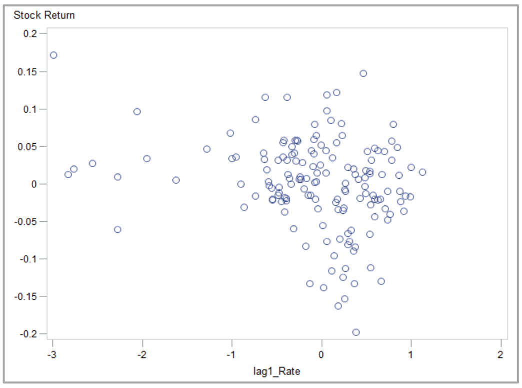
SWEDEN SUBSAMPLE 1987-1999



Dividend Yield

SWEDEN SUBSAMPLE 2000-09:2012



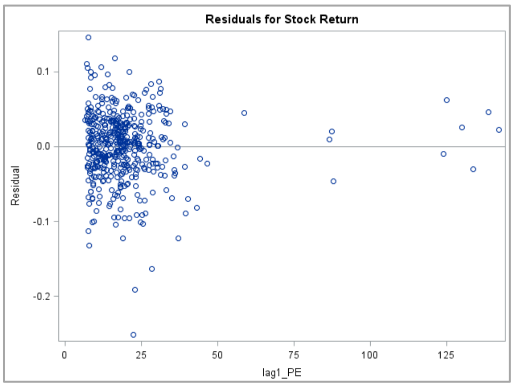


Dividend Yield *Interest Rate*

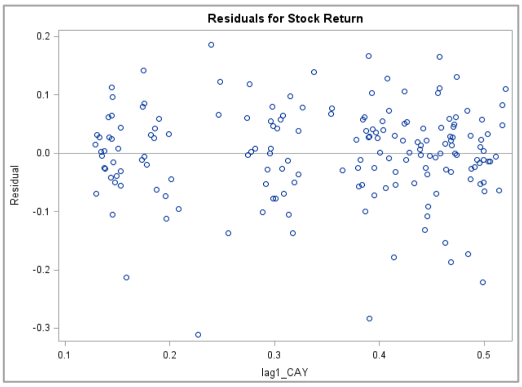
APPENDIX C

HETEROSCEDASTICITY

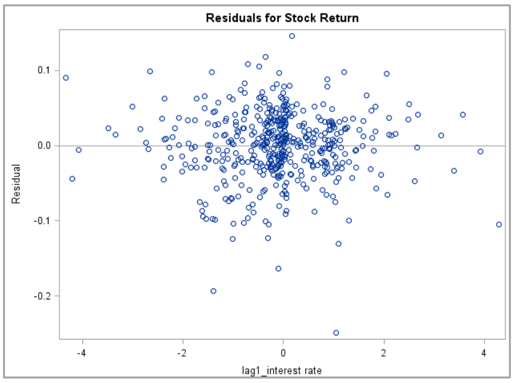
U.S. FINANCIAL CRISIS



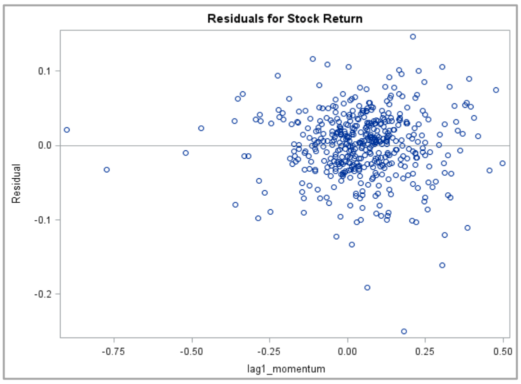
Price-earnings



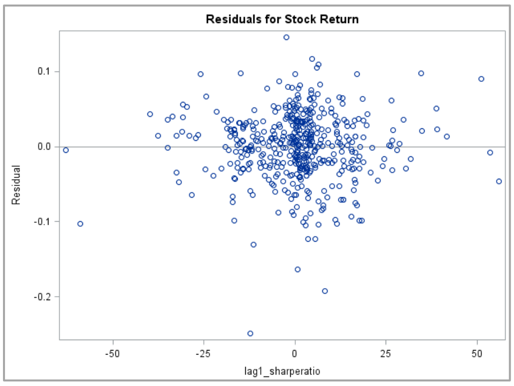
Cay



Interest rate

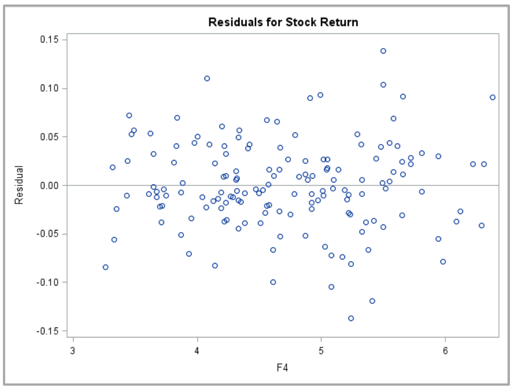


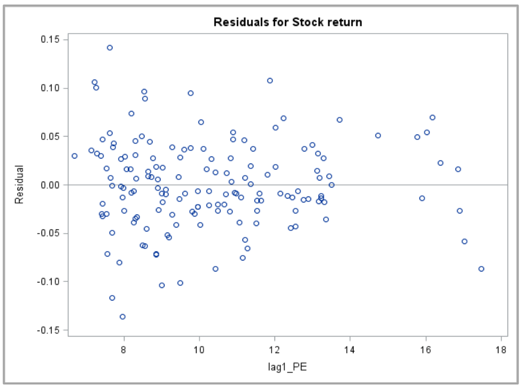
Momentum



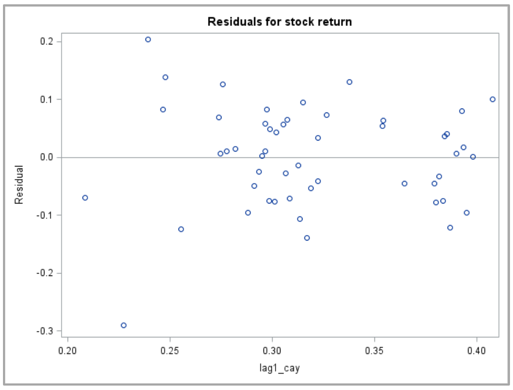
Sharpe ratio

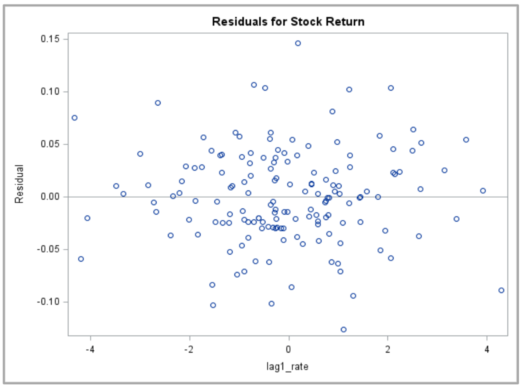
U.S SUBSAMPLE 1974-1986



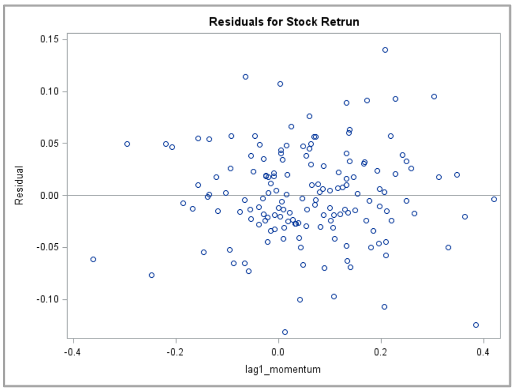


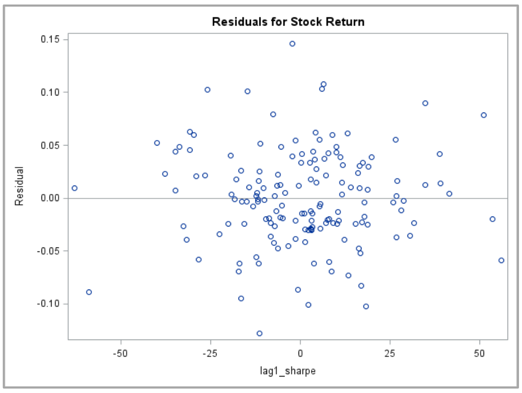
Dividend Yield *Price-earnings*





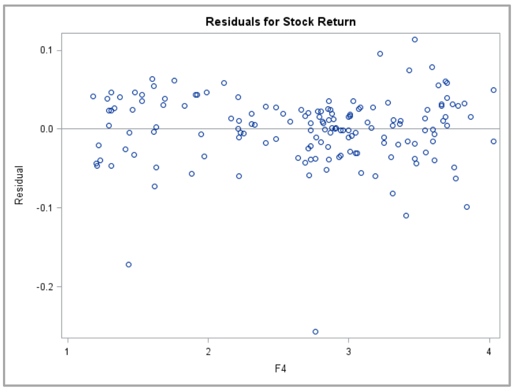
Cay *Interest rate*



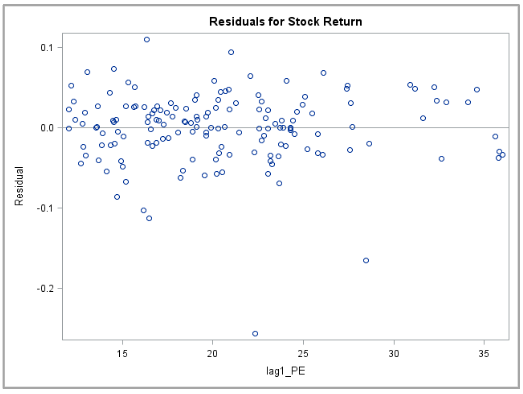


Momentum *Sharpe ratio*

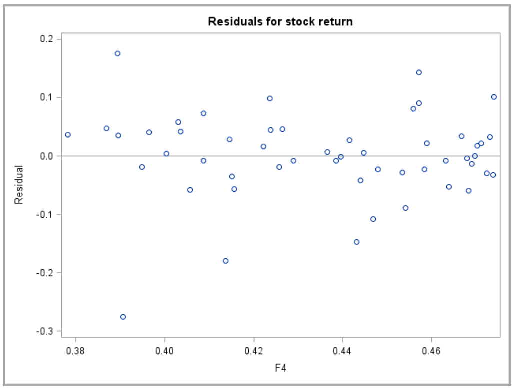
U.S. SUBSAMPLE 1987-1999



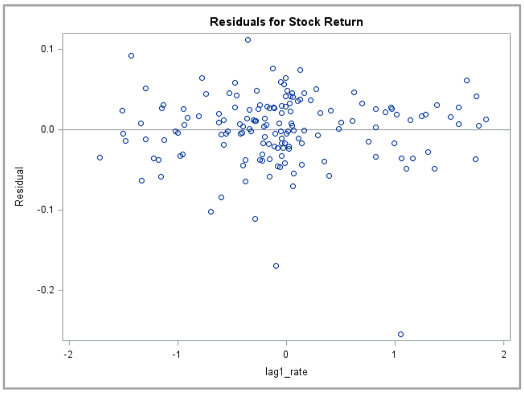
Dividend Yield



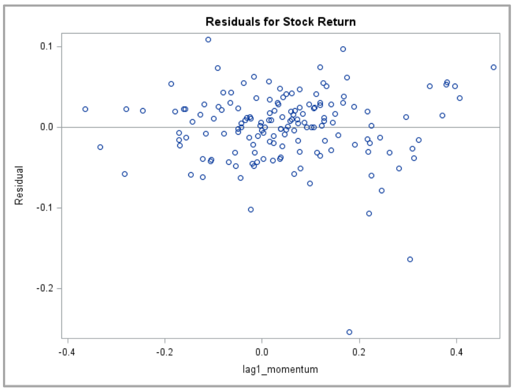
Price-earnings



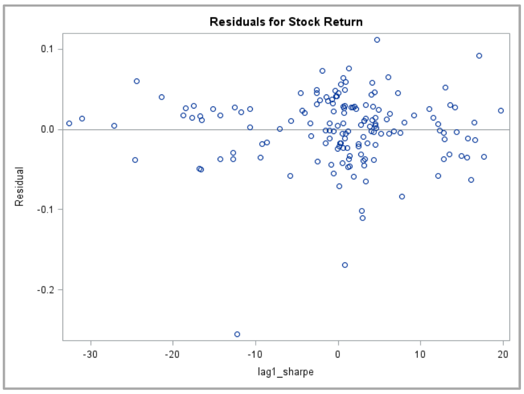
Cay



Interest rate

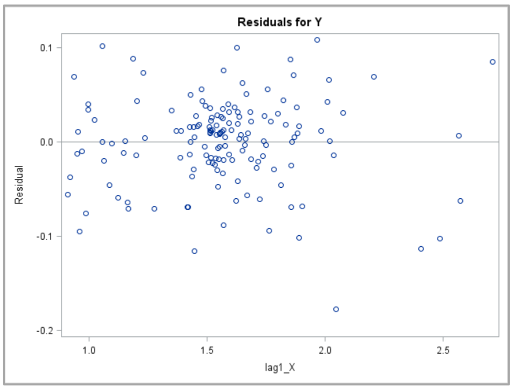


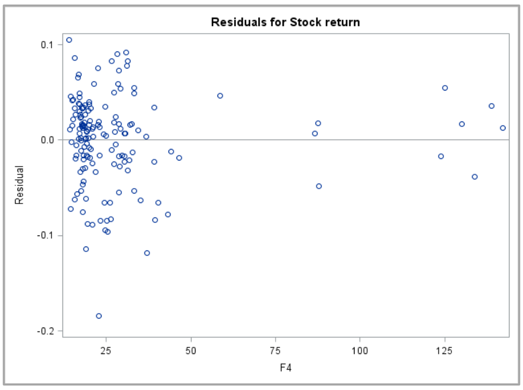
Momentum



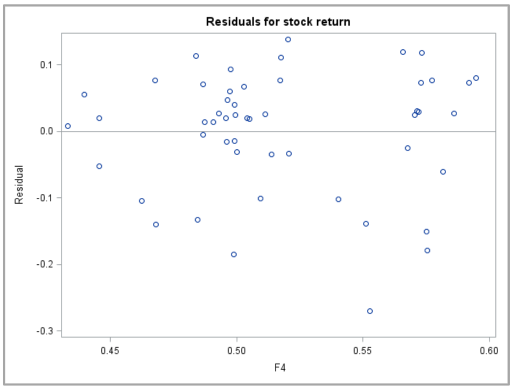
Sharpe ratio

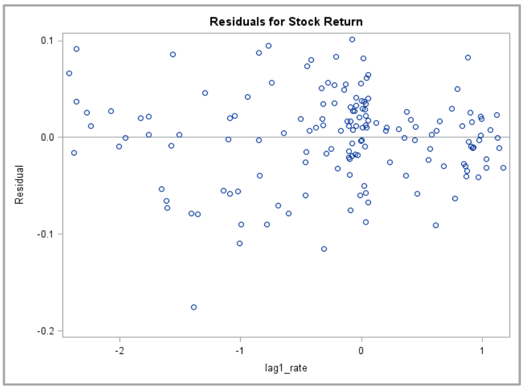
U.S SUBSAMPLE 2000 – 09:2012



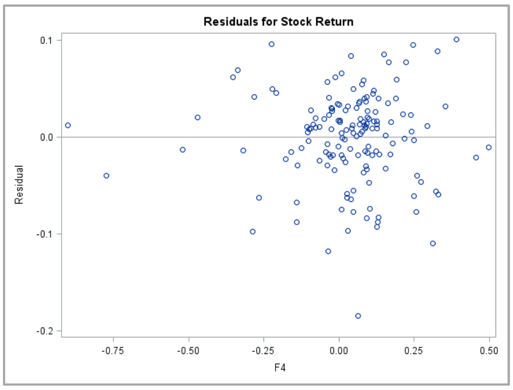


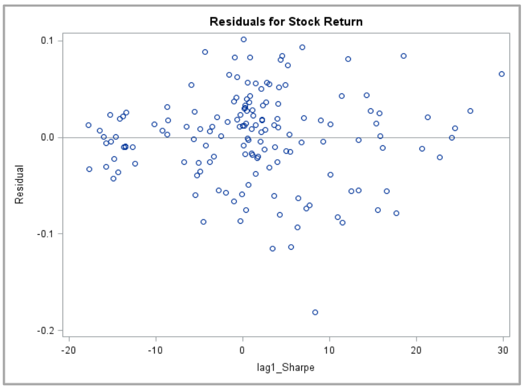
Dividend Yield (GLS) *Price-earnings*





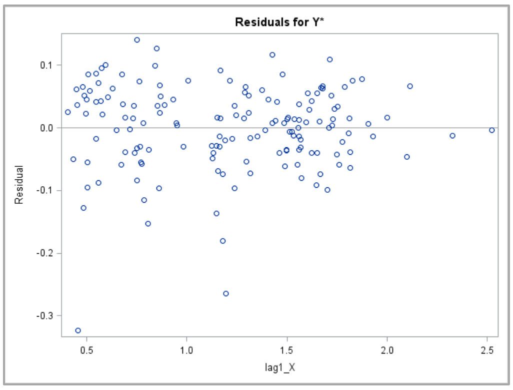
Cay *Interest rate*



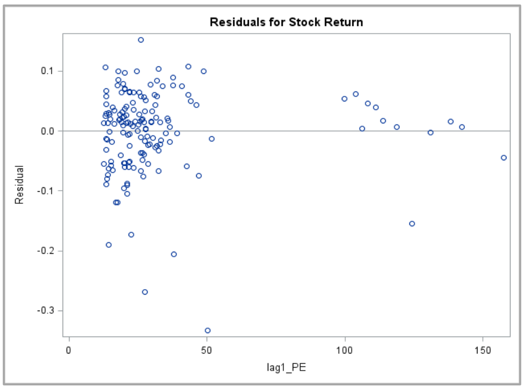


Momentum *Sharpe ratio*

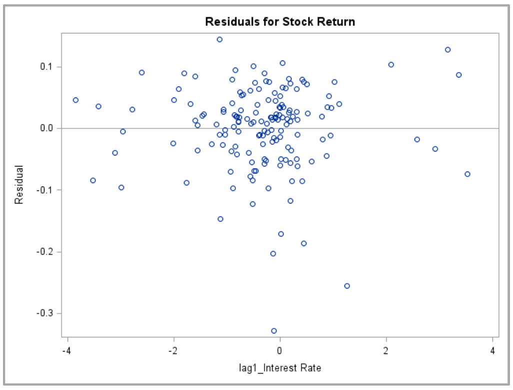
NORWAY SUBSAMPLE 1987 – 1999



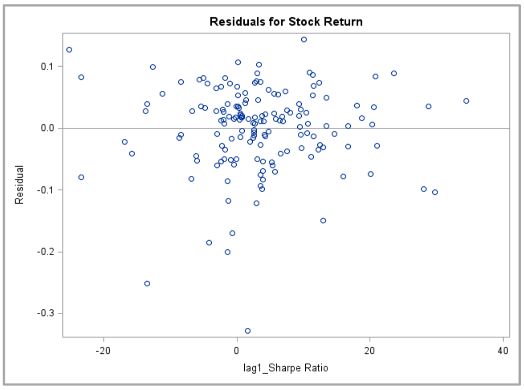
Dividend Yield (GLS)



Price-earnings

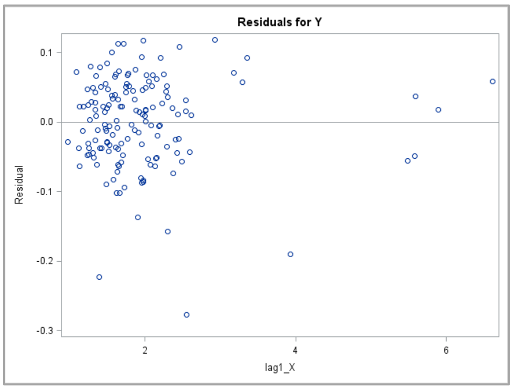


Interest rate

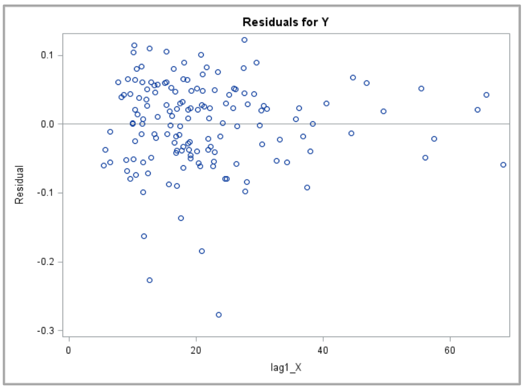


Sharpe ratio

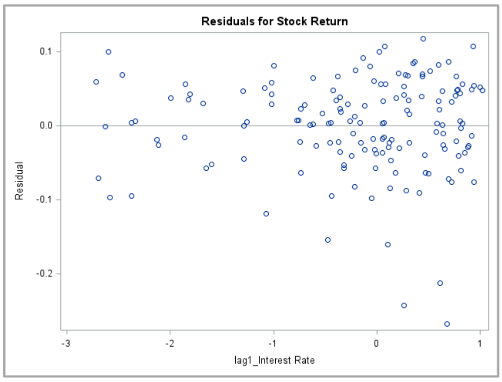
NORWAY SUBSAMPLE 2000 – 09:2012

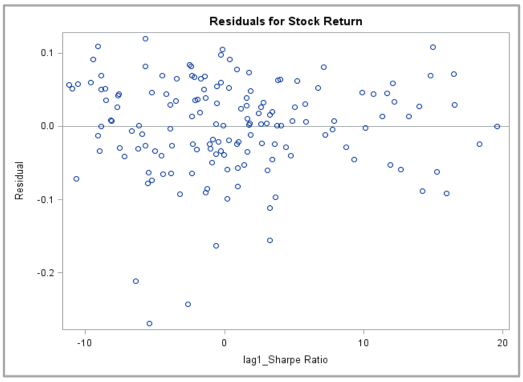


Dividend Yield (GLS)



Price-earnings (GLS)

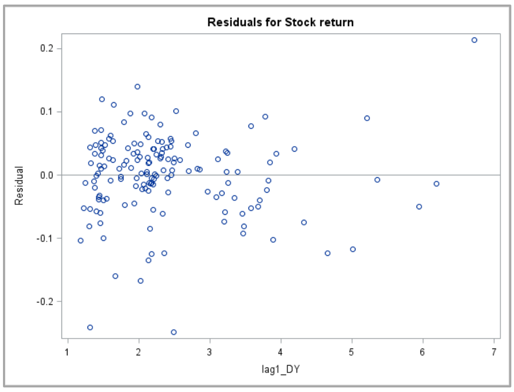




Interest rate

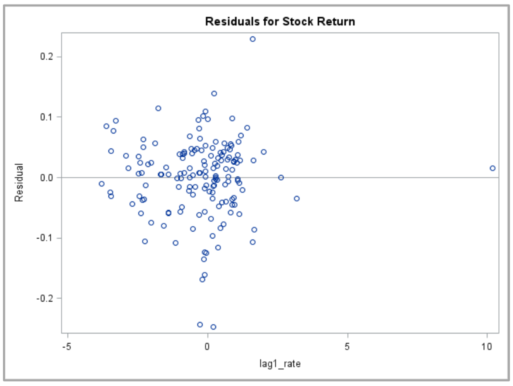
Sharpe ratio

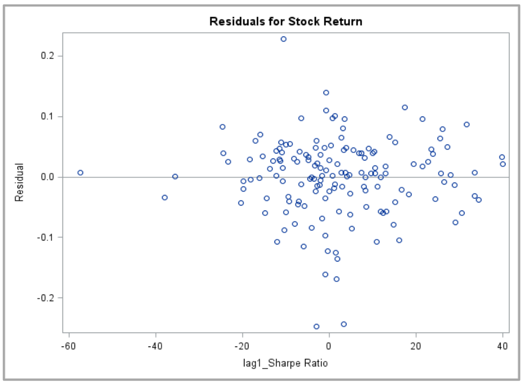
SWEDEN SUBSAMPLE 1987-1999



Dividend Yield

Price-earnings

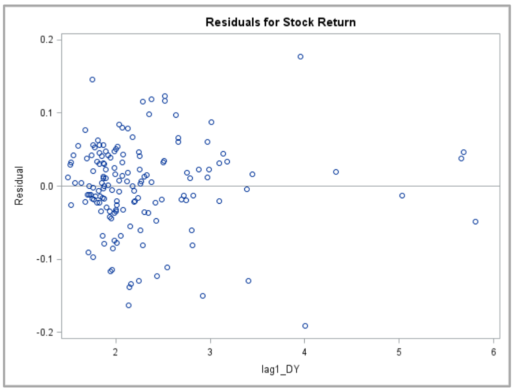




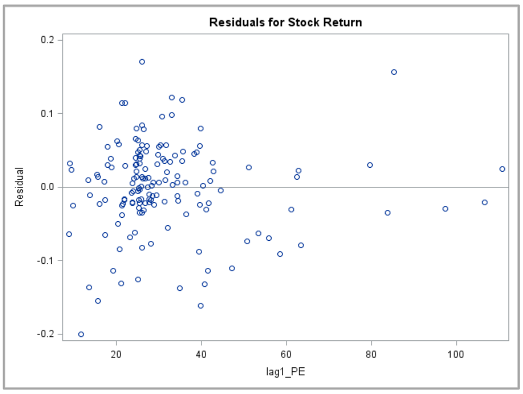
Interest rate

Sharpe ratio

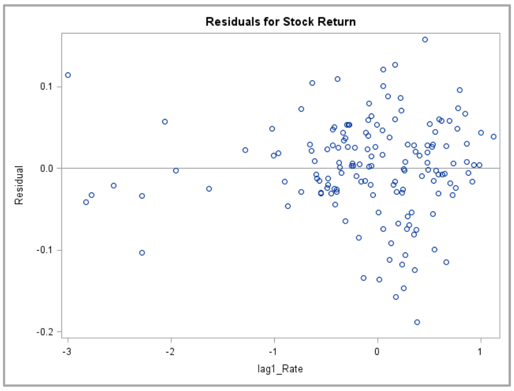
SWEDEN SUBSAMPLE 2000 – 09:2012



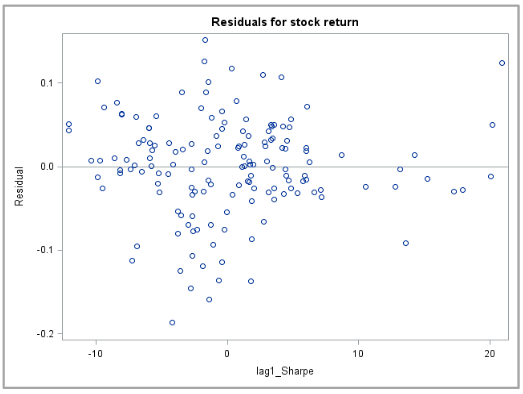
Dividend Yield



Price-earnings



Interest rate

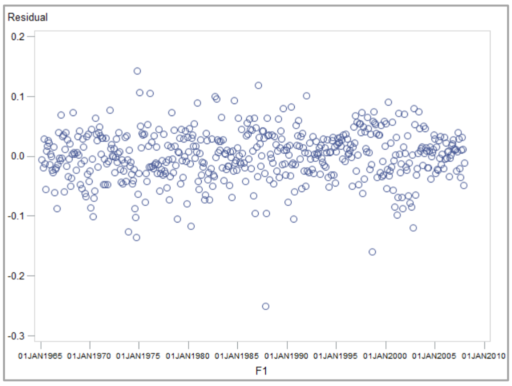


Sharpe ratio

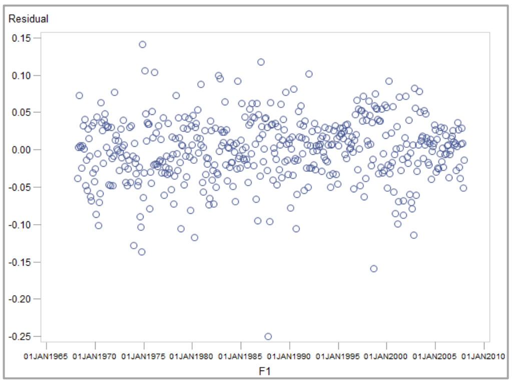
APPENDIX D

AUTOCORRELATION

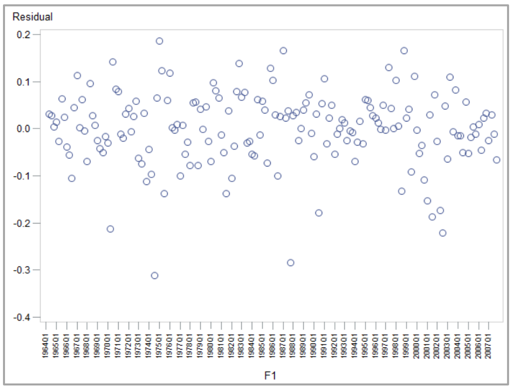
U.S. FINANCIAL CRISIS



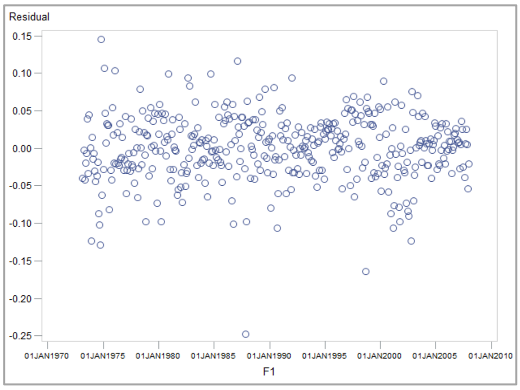
Dividend yield



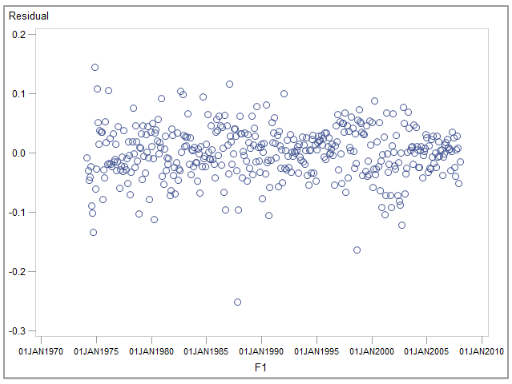
Price-earnings



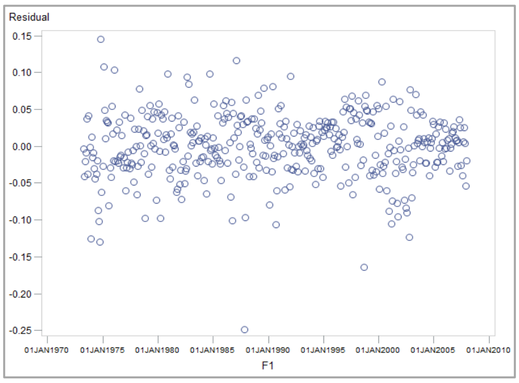
Cay



Interest rate

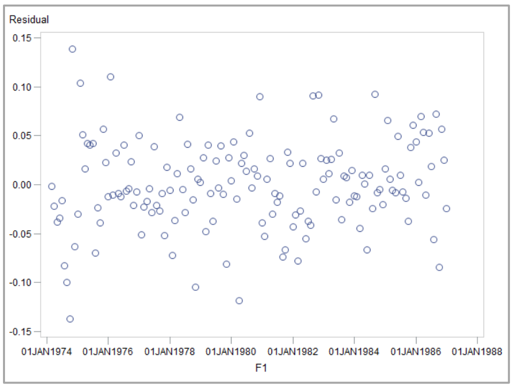


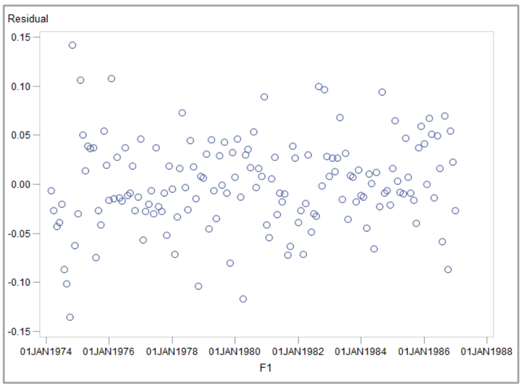
Momentum



Sharpe ratio

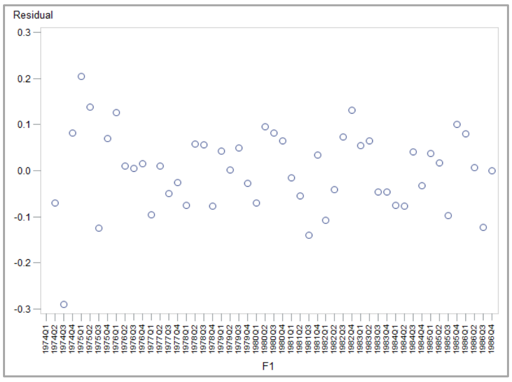
U.S SUBSAMPLE 1974-1986

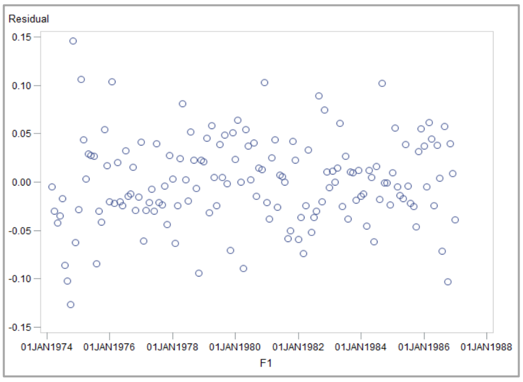




Dividend yield

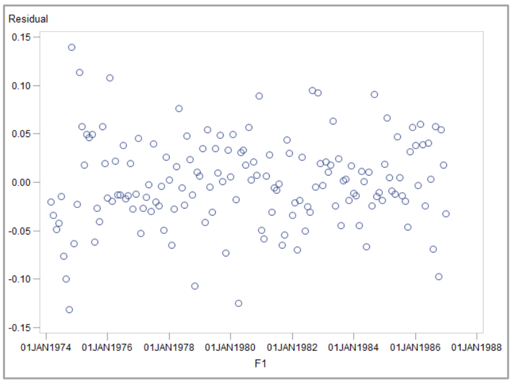
Price-earnings

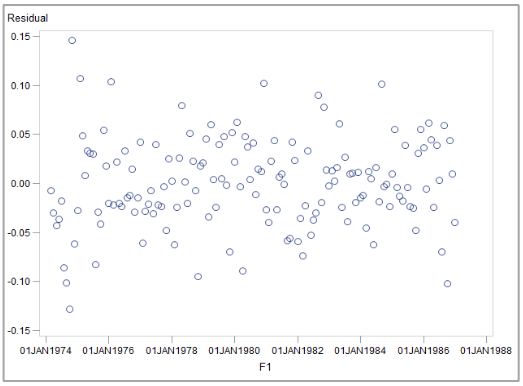




Cay

Interest rate

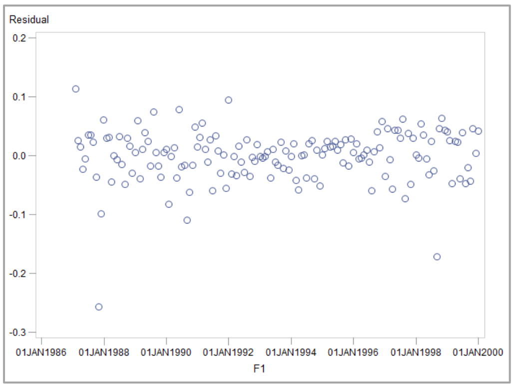




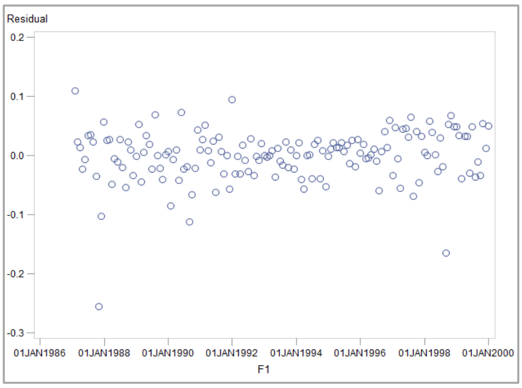
Momentum

Sharpe ratio

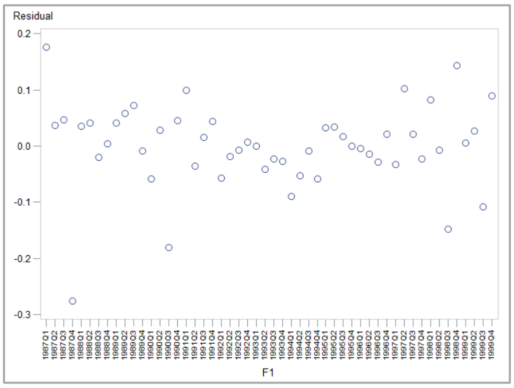
U.S. SUBSAMPLE 1987-1999



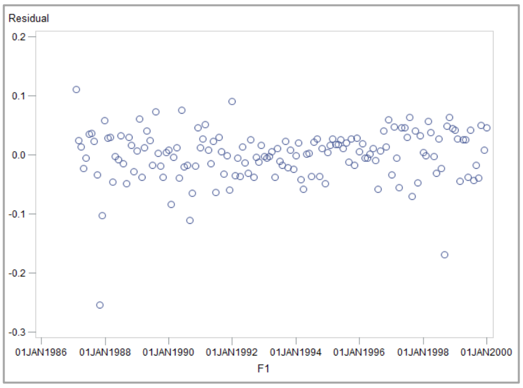
Dividend yield



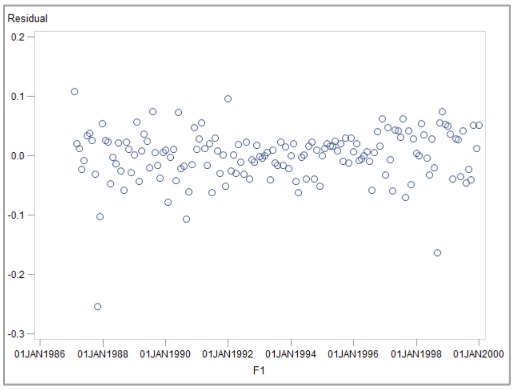
Price-earnings



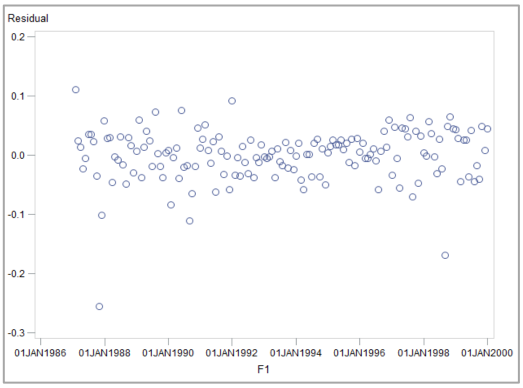
Cay



Interest rate

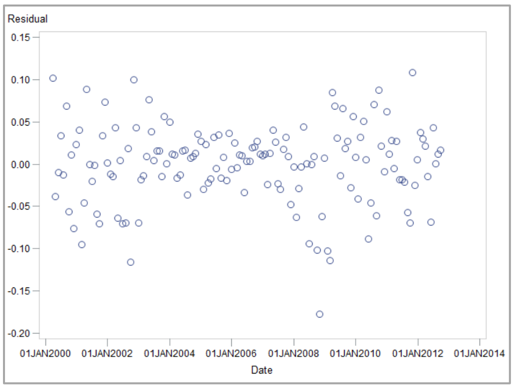


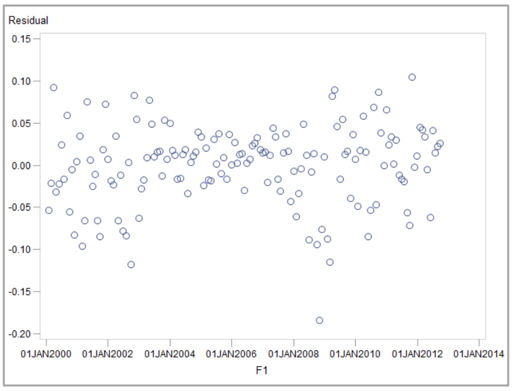
Momentum



Sharpe ratio

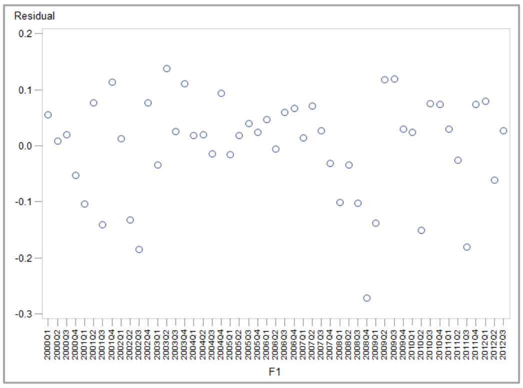
U.S. SUBSAMPLE 2000-09:2012

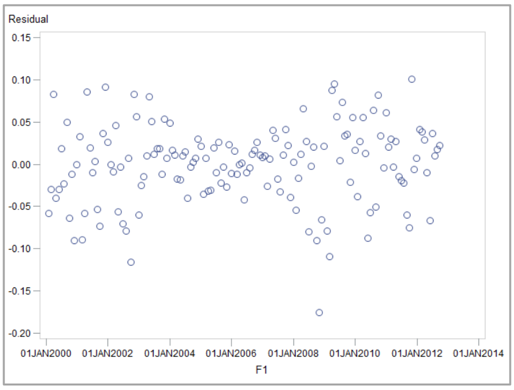




Dividend Yield (GLS)

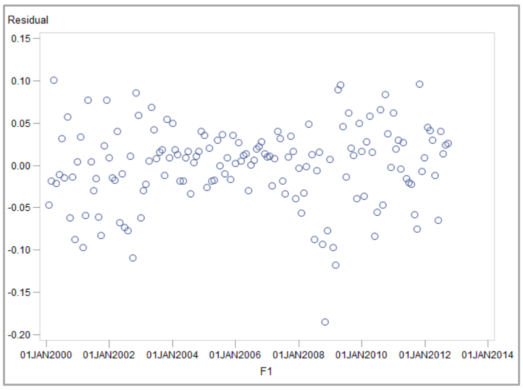
Price-earnings

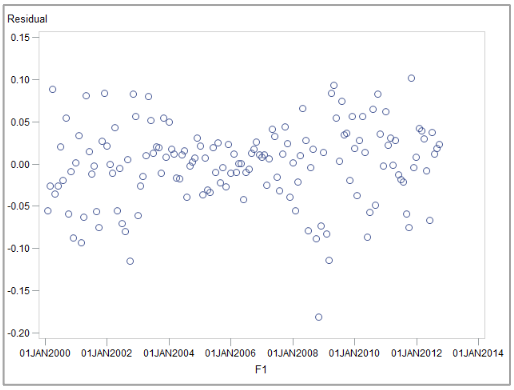




Cay

Interest rate

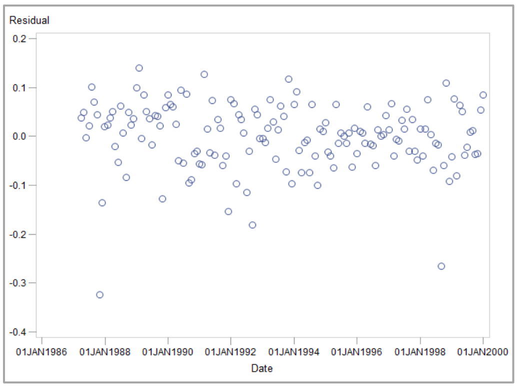




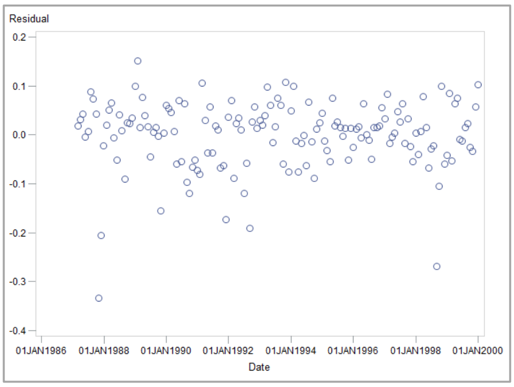
Momentum

Sharpe ratio

NORWAY SUBSAMPLE 1987-1999



Dividend yield (GLS)

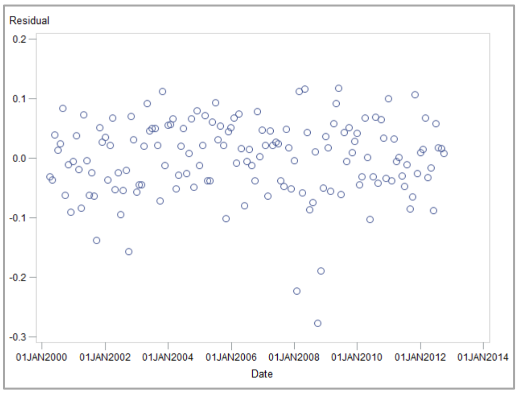


Price-earnings

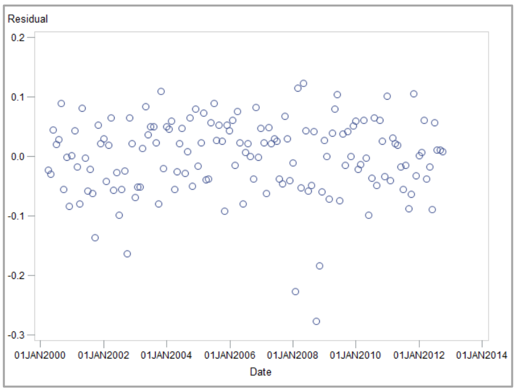
Interest rate

Sharpe ratio

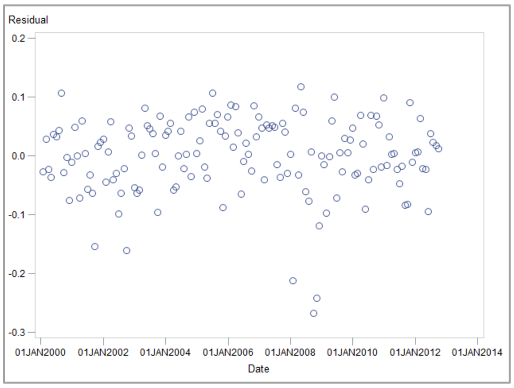
NORWAY SUBSAMPLE 2000-09:2012



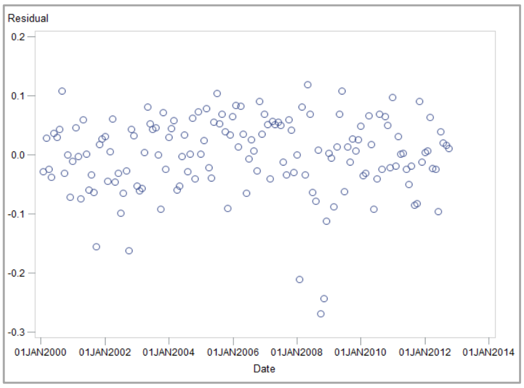
Dividend yield (GLS)



Price-earnings (GLS)

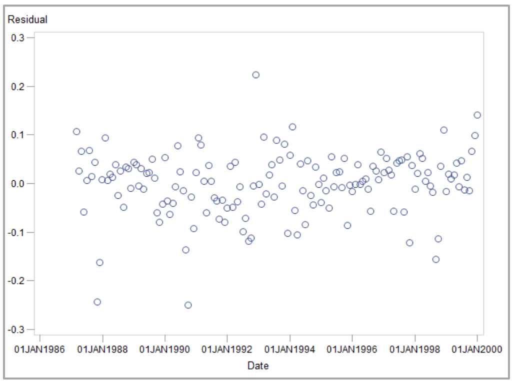


Interest rate

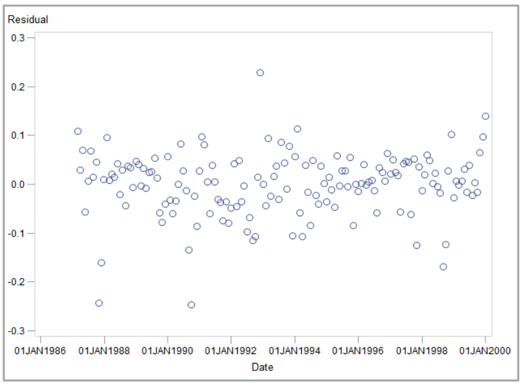


Sharpe ratio

SWEDEN SUBSMAPLE 1987-1999



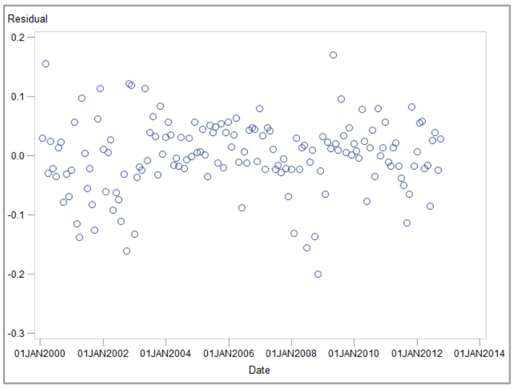
Price-earnings



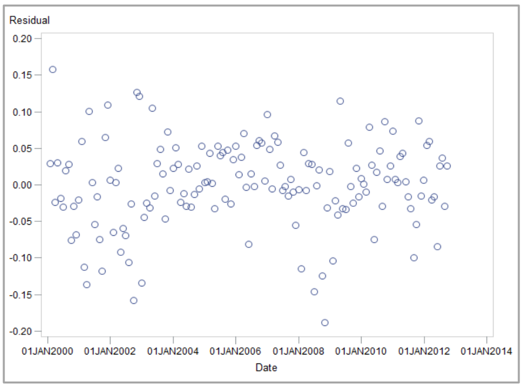
Interest rate

Sharpe ratio

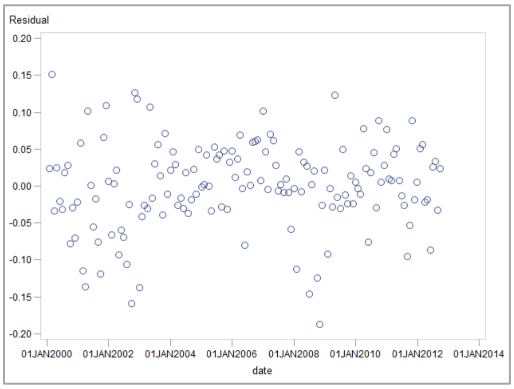
SWEDEN SUBSAMPLE 2000-09:2012



Price-earnings



Interest rate

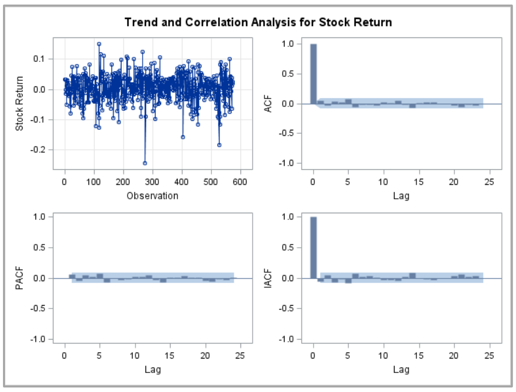


Sharpe ratio

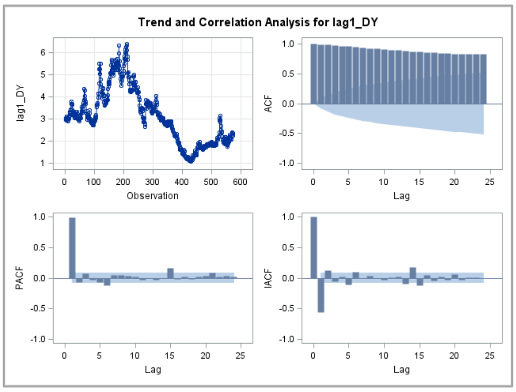
APPENDIX E

STATIONARITY

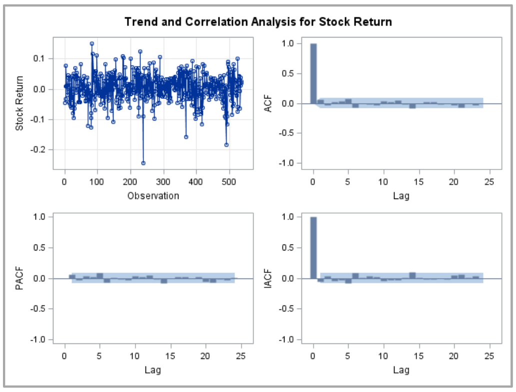
U.S. INDIVIDUAL SAMPLE



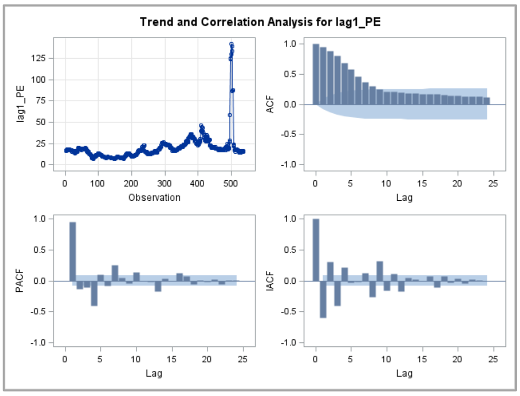
Stock return (dividend yield)



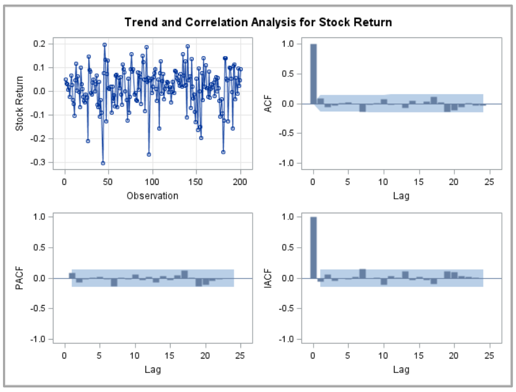
Dividend yield



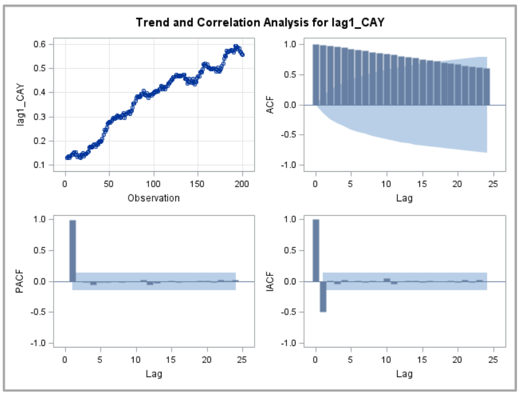
Stock return (price-earnings)



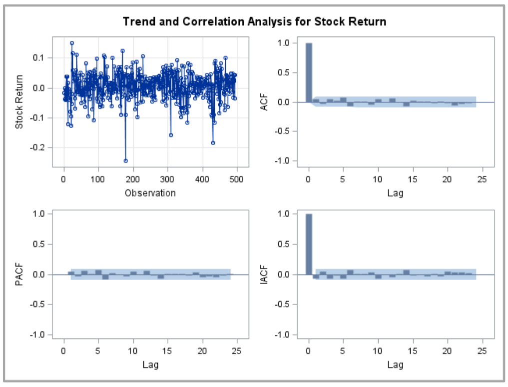
Price-earnings



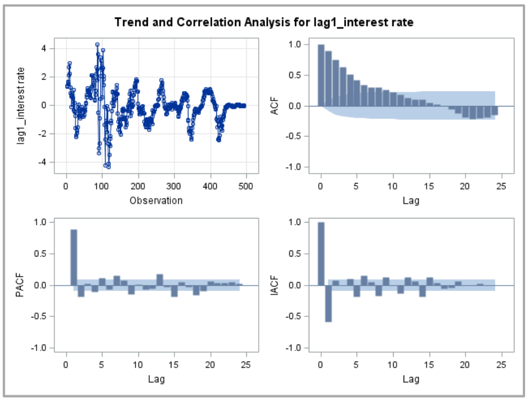
Stock return (cay)



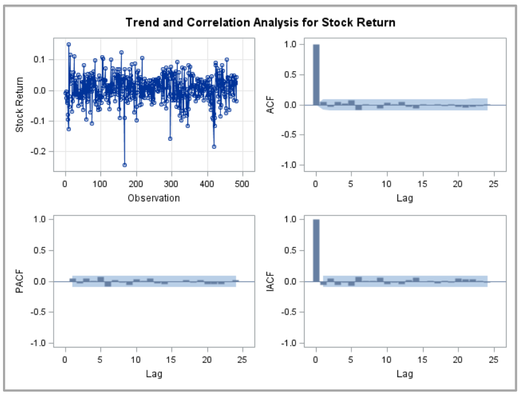
Cay



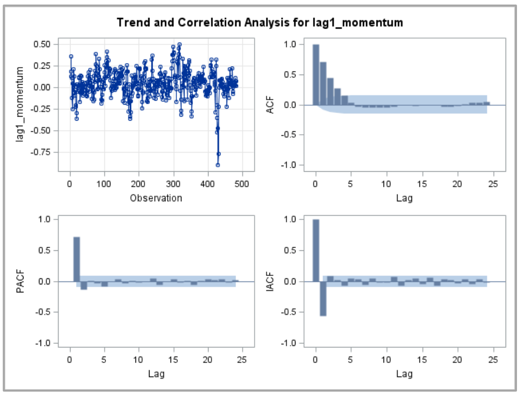
Stock return (interest rate)



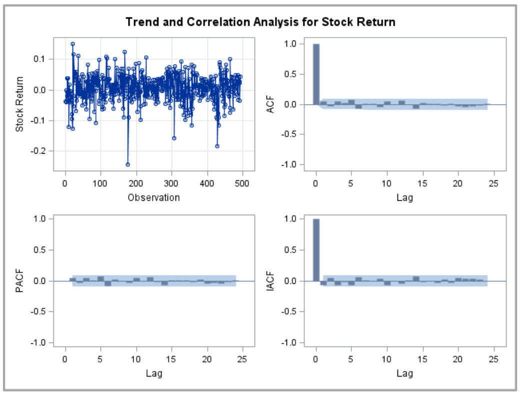
Interest rate



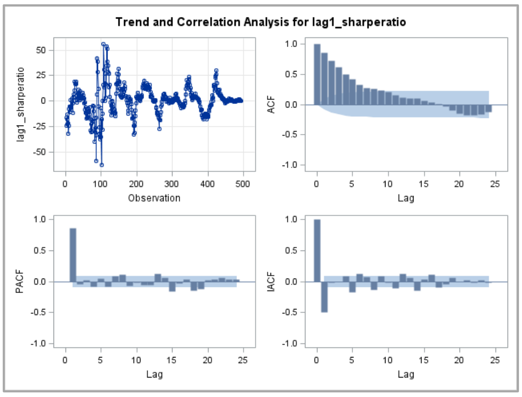
Stock return (momentum)



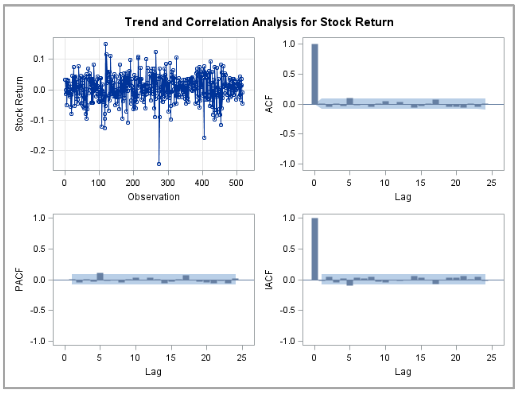
Momentum



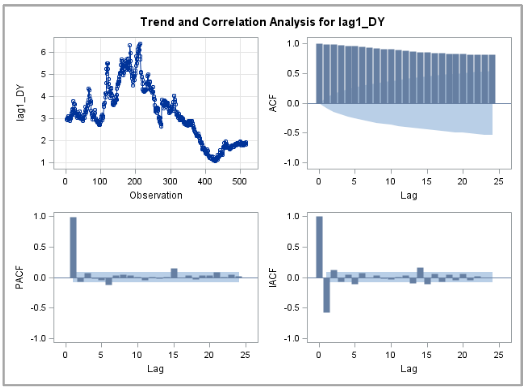
Stock return (Sharpe ratio)



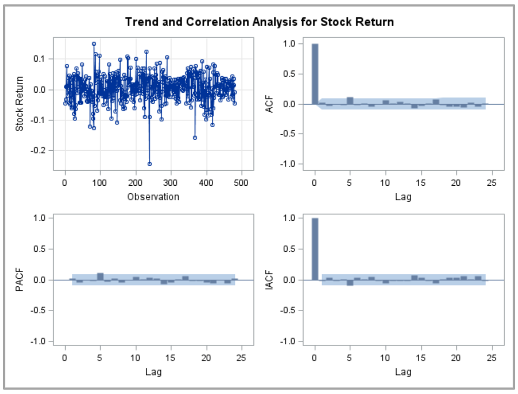
Sharpe ratio



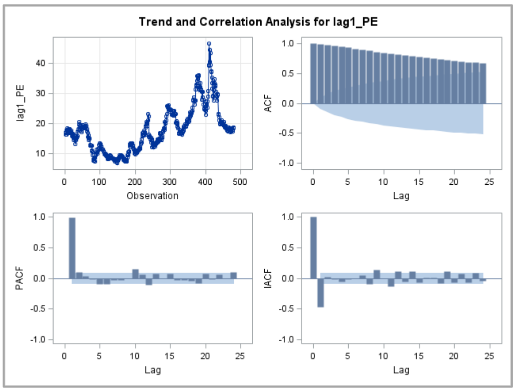
Stock return (dividend yield)



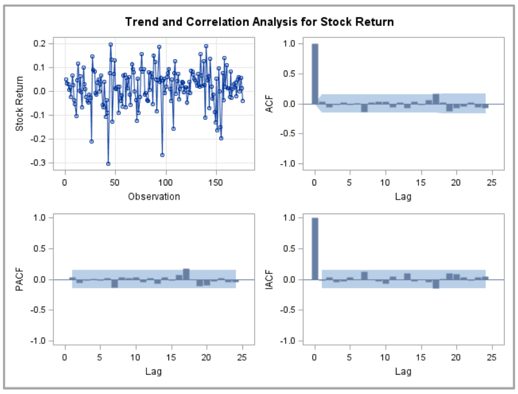
Dividend yield



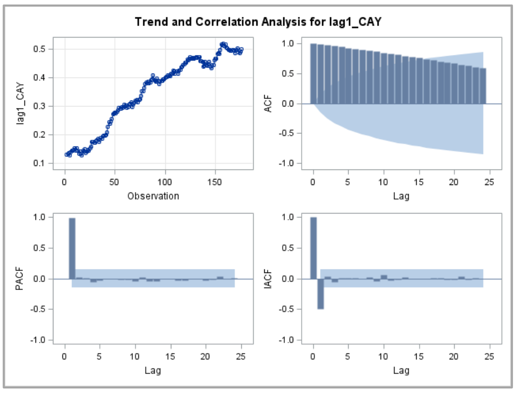
Stock return (price-earnings)



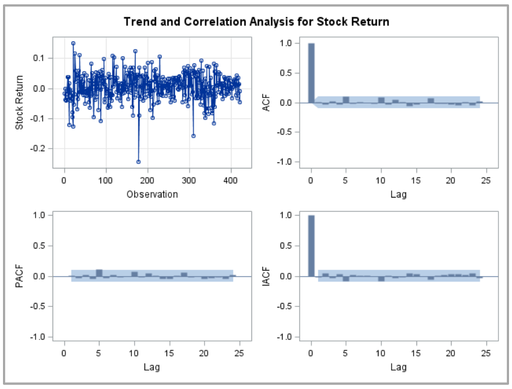
Price-earnings

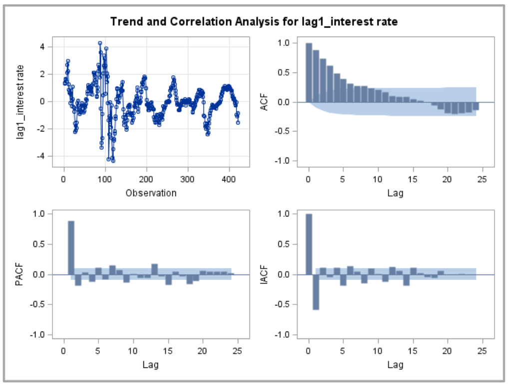


Stock return (cay)



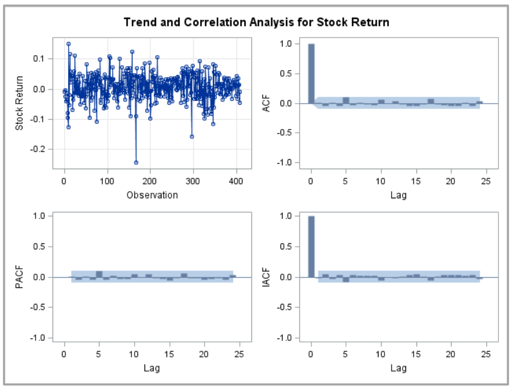
Cay

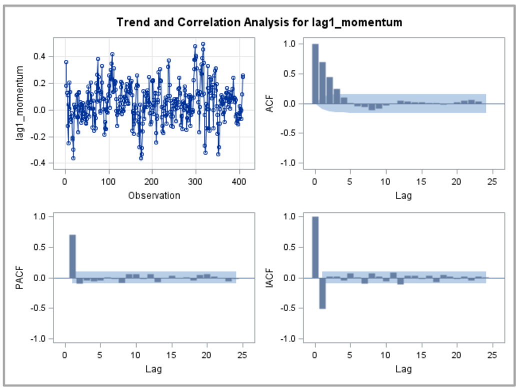




Stock return (interest rate)

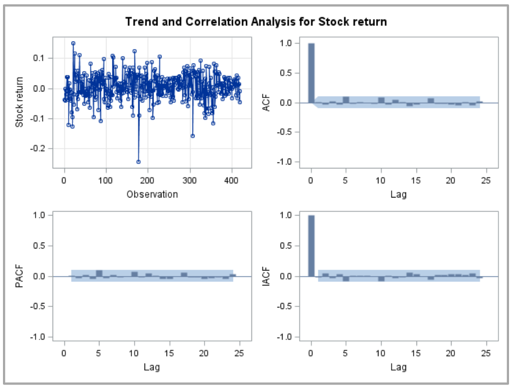
Interest rate

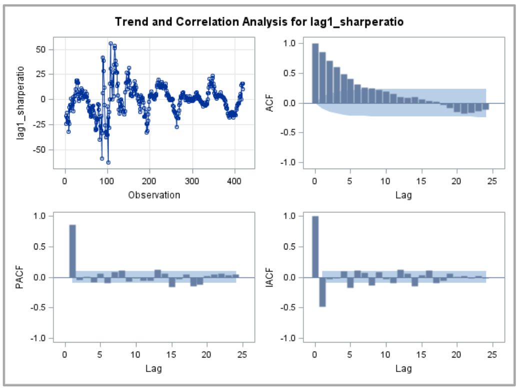




Stock return (momentum)

Momentum

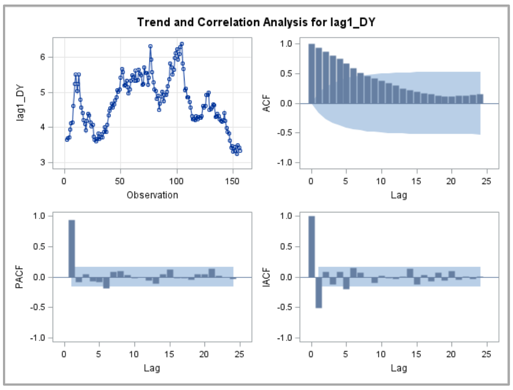




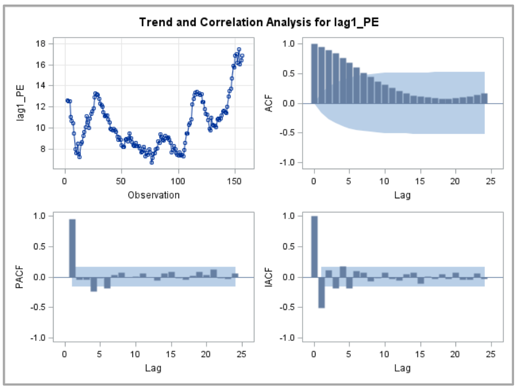
Stock return (Sharpe ratio)

Sharpe ratio

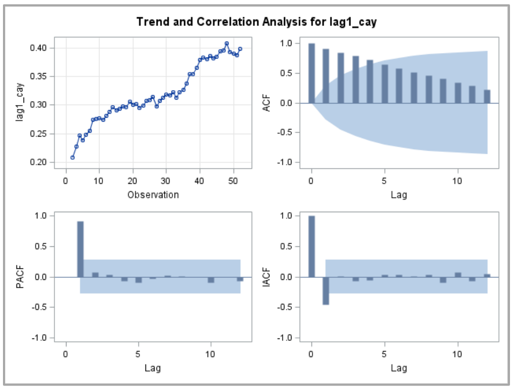
U.S. SUBSAMPLE 1974-1986



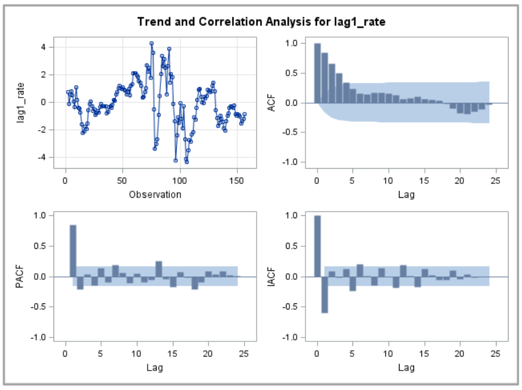
Dividend Yield



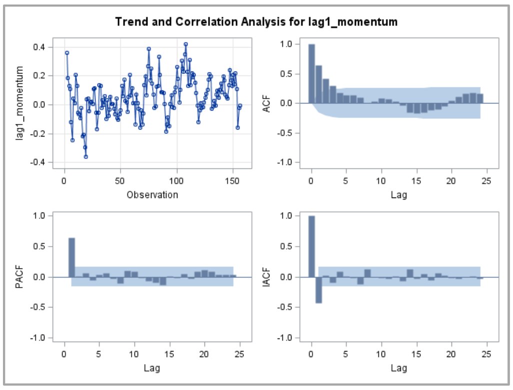
Price-earnings



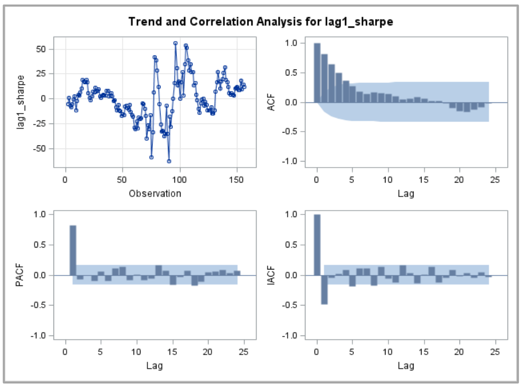
Cay



Interest rate

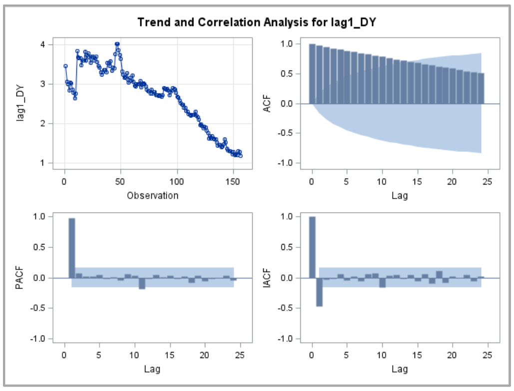


Momentum

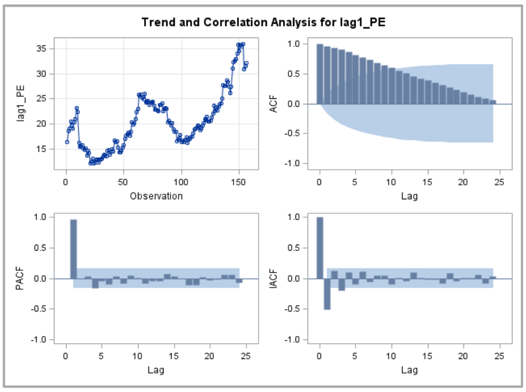


Sharpe ratio

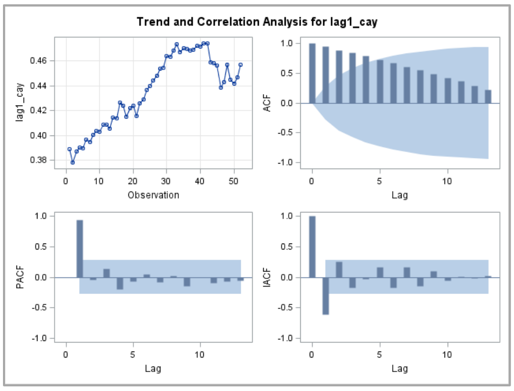
U.S. SUBSAMPLE 1987-1999



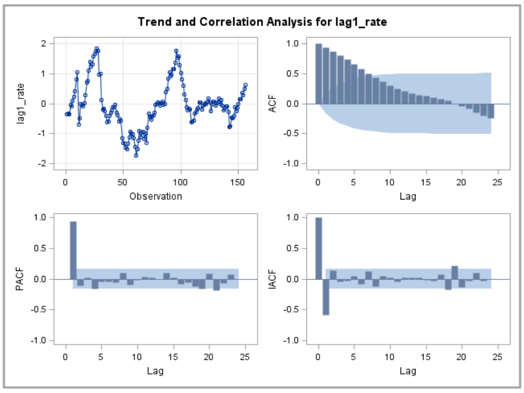
Dividend Yield



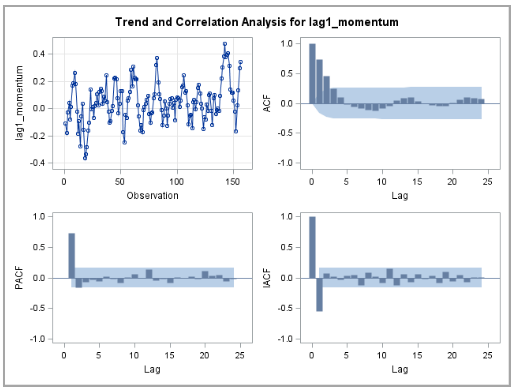
Price-earnings



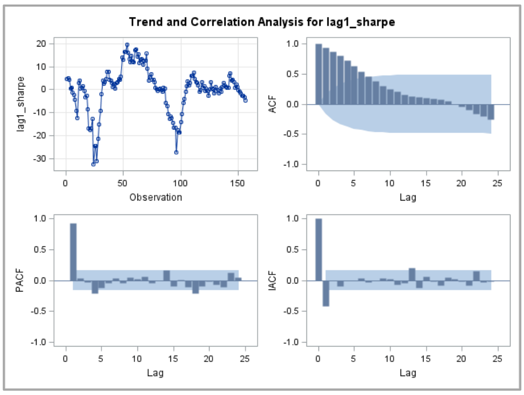
Cay



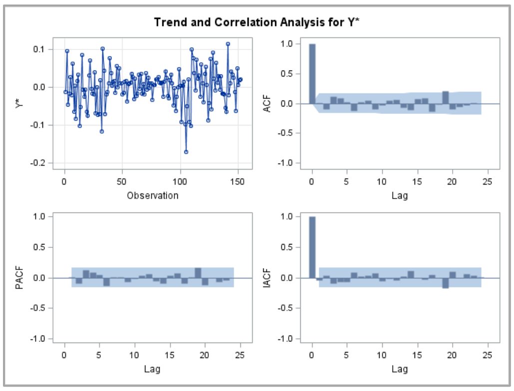
Interest rate

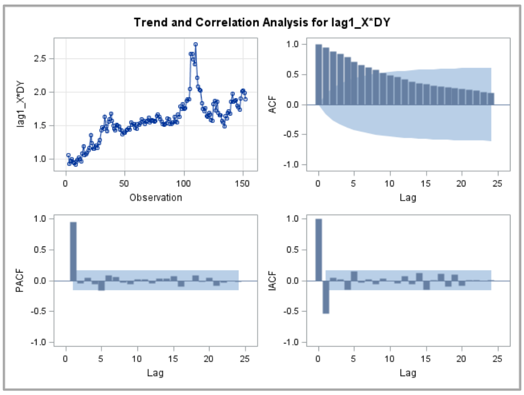


Momentum



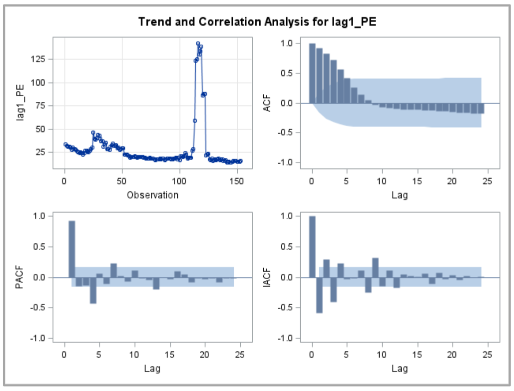
Sharpe ratio

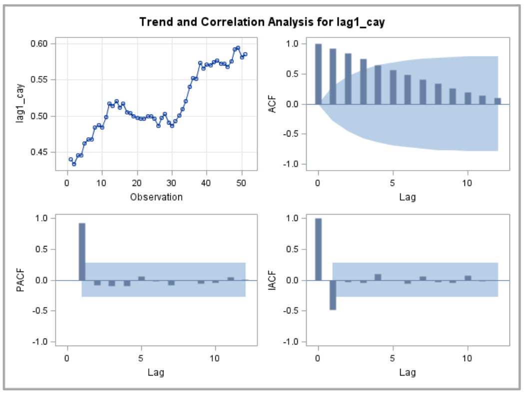




Stock return (GLS)

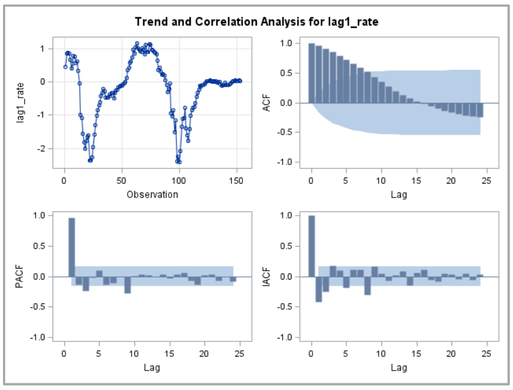
Dividend Yield (GLS)

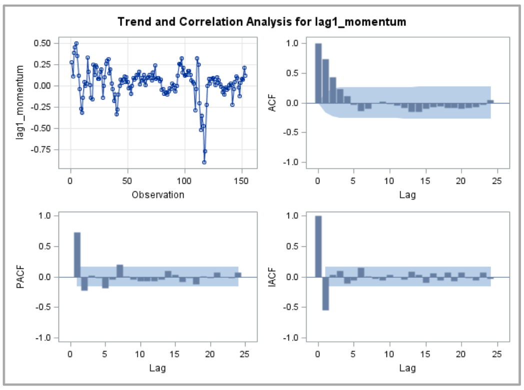




Price-earnings

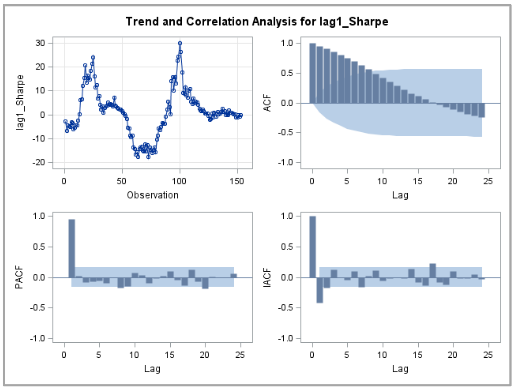
Cay





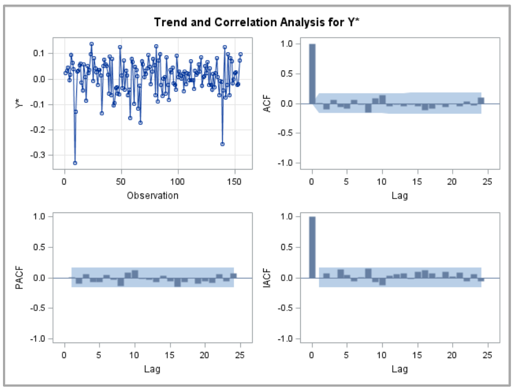
Interest rate

Momentum

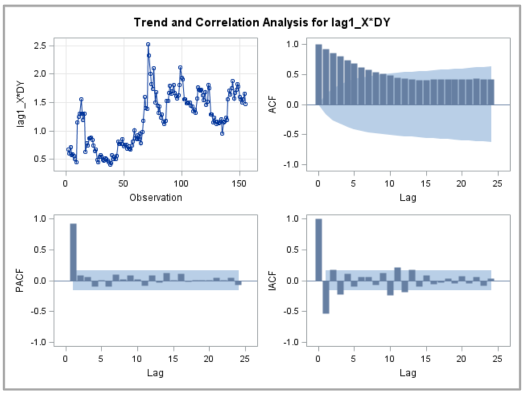


Sharpe ratio

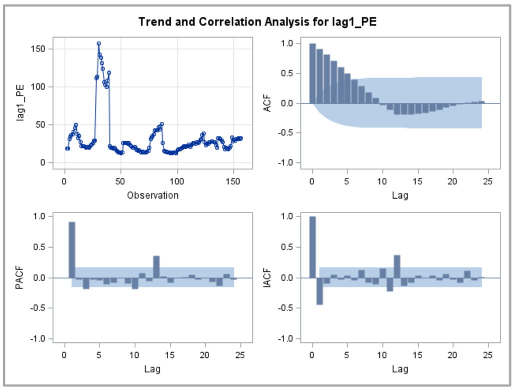
NORWEGIAN SUBSAMPLE 1987-1999



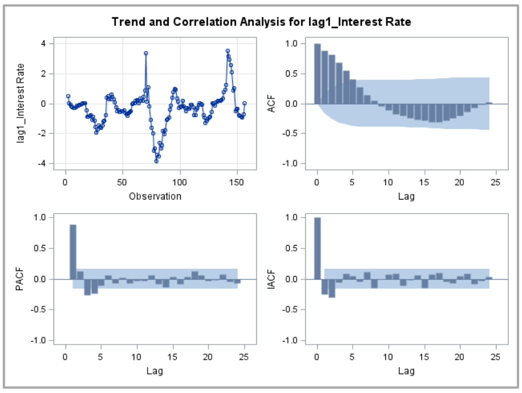
Stock return (GLS, Dividend yield)



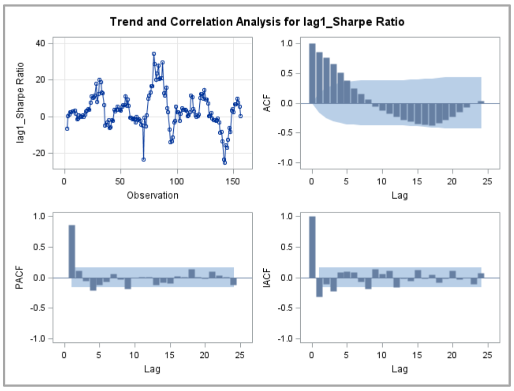
Dividend yield (GLS)



Price-earnings

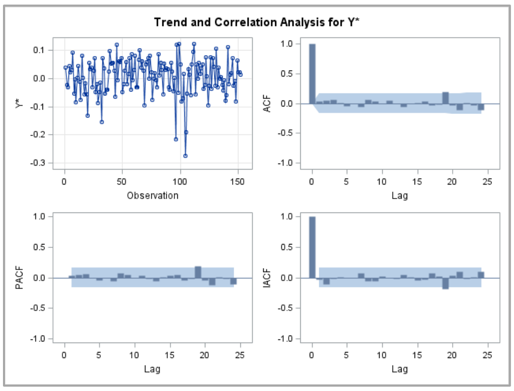


Interest rate

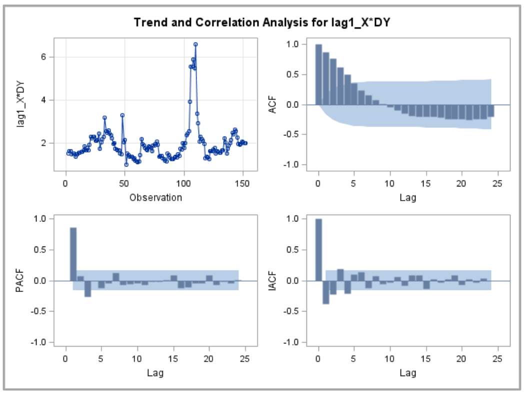


Sharpe ratio

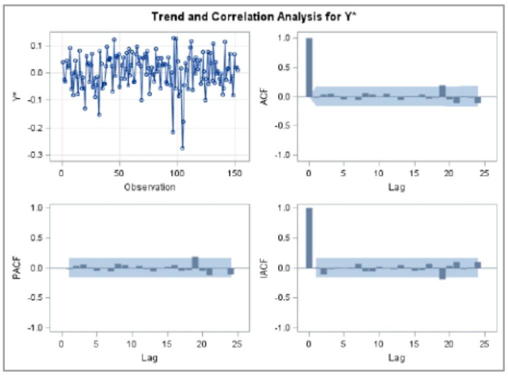
NORWEGIAN SUBSAMPLE 2000-09:2012



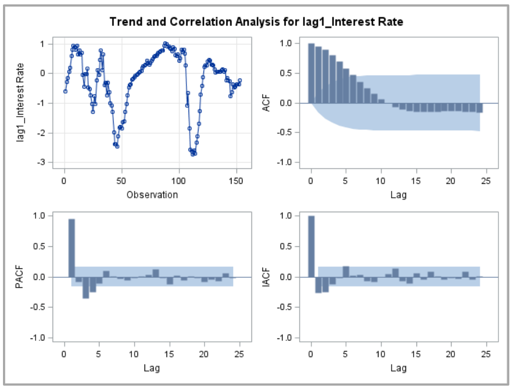
Stock return (GLS, Dividend yield)



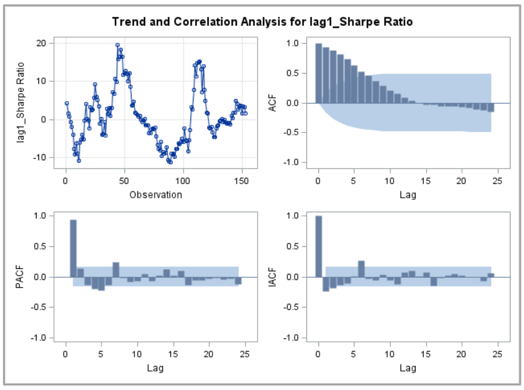
Dividend yield (GLS)



Stock return (GLS, Price-earnings)

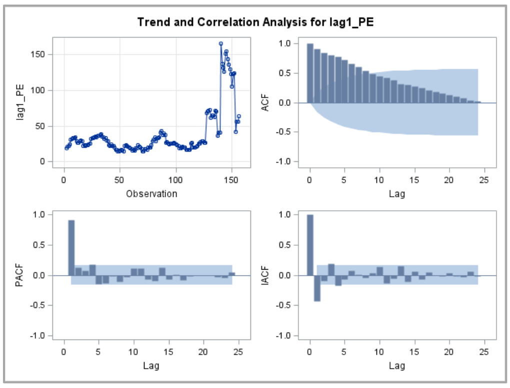


Interest rate

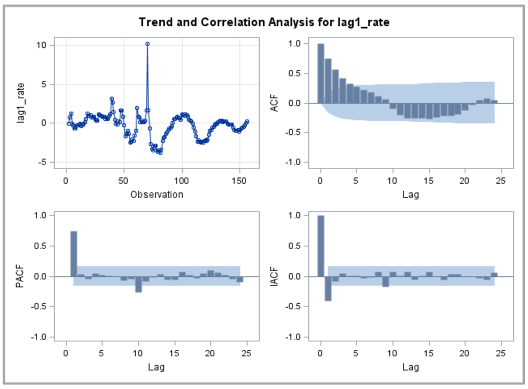


Sharpe ratio

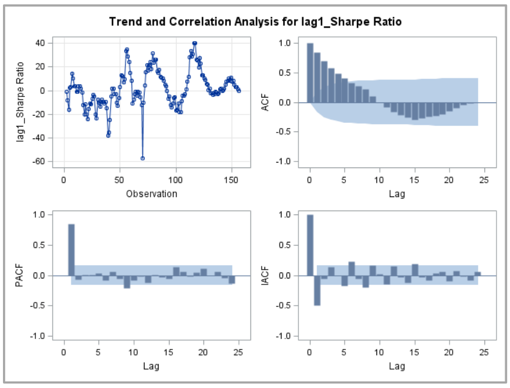
SWEDISH SUBSAMPLE 1987-1999



Price-earnings

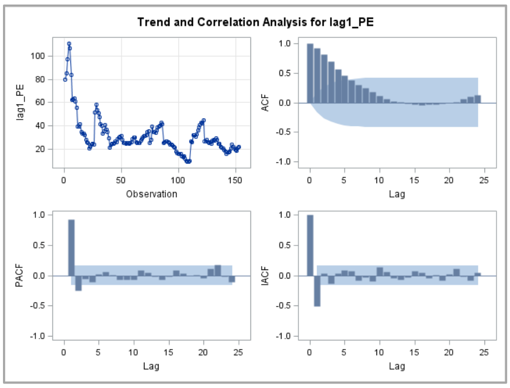


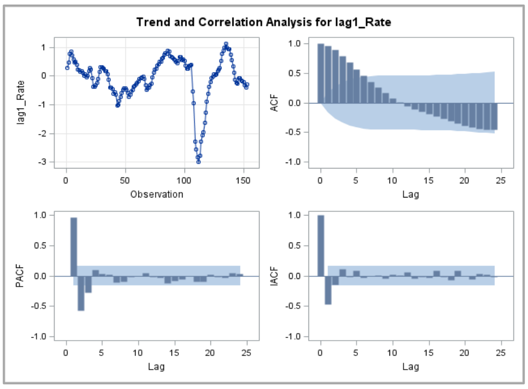
Interest rate



Sharpe ratio

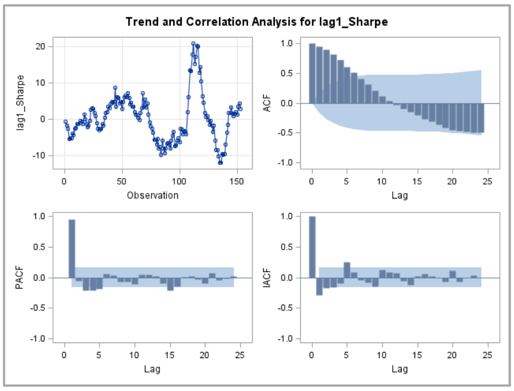
SWEDISH SUBSAMPLE 2000-09:2012





Price-earnings

Interest rate



Sharpe ratio

APPENDIX F

ASYMPTOTIC CRITICAL VALUES FOR MSE-F

Table 4
Percentiles of OOS-*F*: recursive

k_2	%-ile	π 0.0	0.1	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
1	0.99	5.902	1.608	2.129	2.768	3.179	3.459	3.584	3.771	3.589	3.838	3.882	3.951
	0.95	3.270	0.850	1.038	1.298	1.554	1.567	1.548	1.583	1.623	1.599	1.553	1.518
	0.90	2.210	0.530	0.659	0.814	0.796	0.798	0.751	0.759	0.698	0.685	0.687	0.616
2	0.99	7.910	1.996	2.691	3.426	3.907	4.129	4.200	4.362	4.304	4.309	4.278	4.250
	0.95	4.826	1.184	1.453	1.733	1.891	1.820	1.802	1.819	1.752	1.734	1.692	1.706
	0.90	3.324	0.794	0.912	1.029	1.077	1.008	0.880	0.785	0.697	0.666	0.587	0.506
3	0.99	9.230	2.418	3.092	4.080	4.136	4.322	4.341	4.337	4.192	4.089	4.365	4.184
	0.95	5.946	1.434	1.710	2.062	2.073	1.978	1.909	1.930	1.795	1.715	1.710	1.612
	0.90	4.216	0.970	1.064	1.117	1.121	0.960	0.857	0.691	0.599	0.386	0.276	0.127
4	0.99	10.472	2.714	3.440	4.541	4.609	4.378	4.202	4.586	4.477	4.337	4.247	4.096
	0.95	6.712	1.566	1.964	2.246	2.194	1.900	1.809	1.578	1.376	1.256	1.122	1.029
	0.90	5.048	1.060	1.225	1.313	1.184	0.829	0.545	0.354	0.197	−0.058	−0.234	−0.456
5	0.99	11.398	2.902	3.673	4.466	4.434	4.249	4.351	4.349	4.187	3.945	3.783	3.783
	0.95	7.404	1.688	2.082	2.235	2.242	1.773	1.449	1.316	1.045	0.718	0.502	0.459
	0.90	5.568	1.130	1.277	1.228	0.958	0.614	0.241	−0.099	−0.361	−0.656	−0.820	−1.072
6	0.99	12.434	3.212	3.846	4.545	4.676	4.637	4.703	4.286	4.144	3.981	3.525	3.321
	0.95	8.164	1.828	2.124	2.217	2.121	1.660	1.360	1.181	0.761	0.413	0.299	−0.109
	0.90	6.216	1.220	1.313	1.164	0.890	0.419	−0.044	−0.405	−0.776	−1.072	−1.395	−1.664
7	0.99	13.212	3.450	4.098	4.508	4.419	4.271	4.312	4.150	3.677	3.155	3.090	2.880
	0.95	8.930	2.000	2.239	2.424	2.057	1.604	1.282	0.928	0.378	−0.008	−0.199	−0.591
	0.90	6.584	1.272	1.333	1.118	0.799	0.242	−0.363	−0.728	−1.194	−1.657	−2.033	−2.507
8	0.99	13.886	3.408	4.130	4.645	4.625	4.202	4.147	3.912	3.185	2.933	2.952	2.484
	0.95	9.346	2.136	2.312	2.373	1.895	1.390	0.943	0.587	0.131	−0.372	−0.680	−1.140
	0.90	7.014	1.338	1.369	1.058	0.552	0.014	−0.632	−1.076	−1.633	−2.174	−2.731	−3.160
9	0.99	15.010	3.540	4.388	4.703	4.873	4.122	4.066	3.753	3.027	2.925	2.802	2.186
	0.95	9.946	2.168	2.440	2.219	1.714	1.286	0.631	0.198	−0.356	−0.851	−1.241	−1.696
	0.90	7.480	1.354	1.432	0.920	0.393	−0.327	−1.007	−1.595	−2.229	−2.666	−3.250	−3.794
10	0.99	15.586	3.646	4.433	4.813	4.718	3.944	3.645	3.194	2.578	2.282	2.152	1.436
	0.95	10.414	2.202	2.489	2.157	1.536	1.055	0.205	−0.431	−1.071	−1.459	−1.988	−2.378
	0.90	7.862	1.458	1.401	0.884	0.155	−0.600	−1.341	−2.008	−2.782	−3.348	−3.839	−4.437

Notes: The critical values are those for the OOS-*F* when the recursive scheme is used for various values of k_2 and π. Note that the definition of the OOS-*F* changes when $\pi = 0$. In particular the usual statistic is rescaled by $(R/P)^{1/2}$. For this reason the critical values that change smoothly across columns for $\pi = 0.1$ to $\pi = 2.0$ do not do so between the columns for $\pi = 0.0$ to $\pi = 0.1$. Note also that those critical values for $\pi = 0.0$ are identical to those in Tables 5 and 6. This arises since each of the recursive, rolling and fixed schemes are asymptotically equivalent in probability when $\pi = 0.0$.

Source: "Asymptotics for Out of sample test of Granger Causality", M.W. McCracken (2007)

Table 1
Percentiles of the ENC-T, ENC-REG, and ENC-NEW statistics, $\pi > 0$: recursive scheme[a]

k_2	%-ile	π 0.1	0.2	0.4	1.0	2.0	3.0	5.0
ENC-T and ENC-REG								
1	0.95	1.422	1.360	1.338	1.331	1.322	1.329	1.336
	0.90	1.056	1.002	1.005	0.955	0.939	0.937	0.922
2	0.95	1.505	1.467	1.445	1.413	1.443	1.409	1.380
	0.90	1.166	1.101	1.086	1.066	1.035	1.034	1.028
3	0.95	1.574	1.525	1.529	1.476	1.473	1.469	1.436
	0.90	1.227	1.138	1.105	1.113	1.114	1.083	1.074
4	0.95	1.594	1.596	1.552	1.463	1.481	1.474	1.445
	0.90	1.219	1.175	1.192	1.132	1.111	1.091	1.090
ENC-NEW								
1	0.95	0.520	0.744	1.079	1.584	2.085	2.374	2.685
	0.90	0.335	0.473	0.685	0.984	1.280	1.442	1.609
2	0.95	0.766	1.028	1.481	2.234	2.889	3.293	3.627
	0.90	0.524	0.716	1.019	1.471	1.914	2.074	2.428
3	0.95	0.940	1.273	1.865	2.709	3.564	3.989	4.384
	0.90	0.686	0.890	1.285	1.905	2.366	2.664	3.132
4	0.95	1.060	1.526	2.181	3.007	3.894	4.542	4.957
	0.90	0.776	1.062	1.528	2.169	2.727	3.032	3.513

[a]Notes:
1. The test statistics ENC-T, ENC-REG, and ENC-NEW are defined in Section 3.
2. The upper panel of Table 1 reports estimates of the 90th and 95th percentiles of the asymptotic distribution of both the ENC-T and ENC-REG statistics when the recursive scheme is used and $\pi > 0$. The lower panel reports the corresponding percentiles of the asymptotic distribution of the ENC-NEW statistic.
3. The estimates were constructed based upon 5000 simulated draws from the relevant distribution for given values of k_2 and π. See Section 3.1 of the text for further detail on how the simulations were conducted.

Source: “Tests of Equal Forecast Accuracy and Encompassing for Nested Models”, Clark and McCracken (2001)

APPENDIX G

OUT-OF-SAMPLE CALCULATION, CAY AS AN EXAMPLE

	Alpha	Beta	Stock return	x, CAY t-1	Beta*x	y1, unrestricted forecast	u1, forecast error	y0, restricted forecast	u0, forecast error	u0^2	u1^2	d(hat)	c(hat)
R	-0.4898	0.9903	-0.0328	0.5206	0.5156	0.0258	-0.0586	-0.4898	0.4570	0.2089	0.0034	0.2054	0.2356
R+1	-0.3290	0.6554	-0.0930	0.5404	0.3542	0.0251	-0.1181	-0.3290	0.2361	0.0557	0.0140	0.0418	0.0836
R+2	0.0233	-0.0746	-0.2556	0.5527	-0.0412	-0.0179	-0.2377	0.0233	-0.2790	0.0778	0.0565	0.0213	0.0115
R+3	0.1330	-0.3018	-0.1241	0.5510	-0.1663	-0.0333	-0.0908	0.1330	-0.2571	0.0661	0.0082	0.0578	0.0428
R+4	-0.0767	0.1295	0.1417	0.5732	0.0742	-0.0025	0.1442	-0.0767	0.2184	0.0477	0.0208	0.0269	0.0162
R+5	-0.2035	0.3907	0.1396	0.5657	0.2210	0.0176	0.1221	-0.2035	0.3431	0.1177	0.0149	0.1028	0.0758
R+6	-0.2312	0.4478	0.0534	0.5713	0.2558	0.0246	0.0288	-0.2312	0.2847	0.0810	0.0008	0.0802	0.0728
R+7	-0.2479	0.4208	0.0476	0.5703	0.2400	-0.0079	0.0554	-0.2479	0.2955	0.0873	0.0031	0.0842	0.0709
R+8	-0.1437	0.2681	-0.1263	0.5750	0.1542	0.0105	-0.1368	-0.1437	0.0174	0.0003	0.0187	-0.0184	0.0027
R+9	-0.1994	0.3824	0.1018	0.5773	0.2208	0.0213	0.0805	-0.1994	0.3012	0.0907	0.0065	0.0843	0.0665
R+10	-0.2398	0.4655	0.0972	0.5729	0.2667	0.0269	0.0703	-0.2398	0.3370	0.1136	0.0049	0.1086	0.0899
R+11	-0.2523	0.4911	0.0528	0.5719	0.2809	0.0286	0.0242	-0.2523	0.3051	0.0931	0.0006	0.0925	0.0857
R+12	-0.2399	0.4655	-0.0039	0.5676	0.2643	0.0244	-0.0283	-0.2399	0.2360	0.0557	0.0008	0.0549	0.0624
R+13	-0.1616	0.3050	-0.1546	0.5754	0.1755	0.0139	-0.1685	-0.1616	0.0070	0.0000	0.0284	-0.0283	0.0012
R+14	-0.2043	0.3916	0.1057	0.5918	0.2317	0.0275	0.0783	-0.2043	0.3100	0.0961	0.0061	0.0900	0.0718
R+15	-0.2433	0.4708	0.1133	0.5947	0.2800	0.0366	0.0767	-0.2433	0.3566	0.1272	0.0059	0.1213	0.0998
R+16	-0.2206	0.4242	-0.0334	0.5816	0.2467	0.0262	-0.0596	-0.2206	0.1871	0.0350	0.0036	0.0315	0.0462
R+17	-0.2305	0.4445	0.0560	0.5860	0.2605	0.0300	0.0261	-0.2305	0.2865	0.0821	0.0007	0.0814	0.0746
									Sum	1.4360	0.1979	1.2381	1.2100

T-R+1	52
MSE1	0.0038
MSE0	0.0276
d(bar)	0.0238
MSE0-MSE1	0.0238
MSE-F	325.3685
c(bar)	0.0233
ENC-NEW	317.9846

APPENDIX H

MULTIPLE REGRESSION

U.S Individual sample (DY, PE and Momentum)					
Parameter Estimate					
R-square: 0.48%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.00397	0.00818	-0.48	0.6281	0
DY	0.00244	0.00181	1.35°	0.1772	1.34349
PE	0.00012401	0.00016041	0.77	0.4399	1.47198
Momentum	-0.00426	0.01438	-0.3	0.7671	1.13587

U.S Subsample 1974-1986 (All)					
Parameter Estimate					
R-square: 8.35%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.10799	0.08724	-1.24	0.2177	0
DY	0.0175	0.01141	1.53°	0.1271	5.53915
PE	0.00283	0.00351	0.81	0.4214	5.5581
Interest	-0.0258	0.01391	-1.86*	0.0655	37.36883
Momentum	0.02946	0.02913	1.01	0.3136	1.17498
Sharpe ratio	-0.00162	0.00112	-1.45°	0.1502	37.57983

U.S Subsample 1974-1986 (DY, Interest Rate and Momentum)					
Parameter Estimate					
R-square: 6.93%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.04138	0.02341	-1.77	0.0792	0
DY	0.00991	0.00492	2.02**	0.0457	1.02663
Interest Rate	-0.02144	0.01352	-1.59°	0.1149	35.2498
Sharpe Ratio	-0.00123	0.00108	-1.14	0.2577	35.1898

U.S Subsample 1974-1986 (DY and Interest Rate)					
Parameter Estimate					
R-square: 6.14%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.04059	0.02343	-1.73	0.0851	0
DY	0.00981	0.00492	1.99**	0.048	1.0263
Interest Rate	-0.0063	0.00231	-2.73**	0.0071	1.0263

U.S Subsample 1987-1999 (All)					
Parameter Estimate					
R-square: 4.29%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	0.10902	0.04243	2.57	0.0112	0
DY	-0.01807	0.0081	-2.23**	0.0272	3.22499
PE	-0.00235	0.00111	-2.11**	0.0369	3.24945
Interest Rate	-0.02619	0.02464	-1.06	0.2895	28.19283
Momentum	-0.03009	0.02579	-1.17	0.2451	1.18901
Sharpe ratio	-0.00157	0.00191	-0.82	0.4135	27.78822

U.S Subsample 1974-1986 (DY, PE, Interest rate and Momentum)					
Parameter Estimate					
R-square: 3.86%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	0.10576	0.04219	2.51	0.0133	0
DY	-0.01691	0.00796	-2.12**	0.0354	3.12663
PE	-0.00232	0.00111	-2.09**	0.0387	3.24675
Interest Rate	-0.00641	0.00499	-1.28	0.2013	1.15938
Momentum	-0.02621	0.02532	-1.04	0.3023	1.1489

U.S Subsample 2000-09:2012 (All)					
Parameter Estimate					
R-square: 5.49%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.02787	0.02033	-1.37	0.1724	0
DY	0.01436	0.01032	1.39°	0.1661	1.12934
PE	0.00016077	0.00018837	0.85	0.3948	1.43445
Interest Rate	0.02412	0.01558	1.55°	0.1237	13.10761
Momentum	-0.00489	0.02443	-0.2	0.8416	1.46505
Sharpe ratio	0.00135	0.00141	0.95	0.3418	12.97147

U.S Subsample 2000-09:2012 (DY, PE and Interest rate)					
Parameter Estimate					
R-square: 4.86%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.02684	0.01867	-1.44	0.1526	0
DY	0.01242	0.00976	1.27	0.2053	1.01825
PE	0.00020317	0.00015748	1.29°	0.199	1.00953
Interest Rate	0.01	0.00434	2.31**	0.0225	1.02427

Norway Subsample 1987-1999 (All)					
Parameter Estimate					
R-square: 5.43%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.03869	0.02072	-1.87	0.0639	0
DY	0.02388	0.01045	2.29**	0.0237	1.25675
PE	0.00029569	0.00023461	1.26	0.2095	1.3838
Interest Rate	-0.00512	0.01536	-0.33	0.7392	9.86131
Sharpe Ratio	0.00029363	0.00184	0.16	0.8735	10.10972

Norway Subsample 1987-1999 (DY and PE)					
Parameter Estimate					
R-square: 4.08%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	-0.04084	0.02061	-1.98	0.0493	0
DY	0.02498	0.01037	2.41**	0.0172	1.23659
PE	0.0003953	0.00022189	1.78*	0.0768	1.23659

Norway Subsample 2000-09:2012 (All)					
Parameter Estimate					
R-square: 9.74%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	0.05503	0.01937	2.84	0.0051	0
DY	-0.018	0.00612	-2.94**	0.0038	1.53819
PE	-0.00044631	0.00039907	-1.12	0.2652	1.2052
Interest Rate	-0.04996	0.02291	-2.18**	0.0308	17.51708
Sharpe Ratio	-0.00417	0.00296	-1.41°	0.1601	15.60234

Norway Subsample 2000-09:2012 (DY, Interest rate and Shapre ratio)					
Parameter Estimate					
R-square: 8.98%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	0.04031	0.01422	2.84	0.0052	0
DY	-0.0172	0.00608	-2.83**	0.0053	1.51757
Interest Rate	-0.0515	0.02289	-2.25**	0.0259	17.45359
Sharpe Ratio	-0.00407	0.00296	-1.38°	0.1708	15.58691

Sweden Subsample 1987-1999 (All)					
Parameter Estimate					
R-square: 0.59%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	0.01022	0.01669	0.61	0.5413	0
DY	0.0028	0.00557	0.5	0.6168	1.14412
PE	-0.00008737	0.00016579	-0.53	0.599	1.04049
Interest Rate	-0.00322	0.00887	-0.36	0.7173	7.07247
Sharpe Ratio	-0.00015949	0.0009193	-0.17	0.8625	7.13773

Sweden Subsample 1987-1999 (DY, PE and Interest rate)					
Parameter Estimate					
R-square: 0.57%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	0.01106	0.01591	0.7	0.4879	0
DY	0.00249	0.00528	0.47	0.6374	1.03354
PE	-0.00008941	0.00016484	-0.54	0.5883	1.03523
Interest Rate	-0.00179	0.00333	-0.54	0.5913	1.00355

Sweden Subsample 2000-09:1012 (DY, PE and Interest rate)					
Parameter Estimate					
R-square: 9.49%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	0.0491	0.0239	2.05	0.0417	0
DY	-0.01996	0.00842	-2.37**	0.0191	1.69869
PE	-0.00020119	0.00031884	-0.63	0.529	1.23727
Interest Rate	-0.03989	0.02645	-1.51°	0.1337	17.82653
Sharpe Ratio	-0.00152	0.00295	-0.51	0.6076	15.71344

Sweden Subsample 2000-09:1012 (DY and Interest rate)					
Parameter Estimate					
R-square: 9.01%					
Variable	**β**	**Standard Error**	**t-value**	**p-value**	**VIF**
Intercept	0.03521	0.01698	2.07	0.0399	0
DY	-0.01633	0.00728	-2.24**	0.0263	1.2805
Interest Rate	-0.02696	0.00706	-3.82**	0.0002	1.2805

APPENDIX I

LOG REGRESSION

U.S.	Individual		Financial crisis		1974-1986		1987-1999		2000-09:2012	
DATA	Beta	T-value	Beta	T-value	Beta	T-value	Beta	T-value	Beta	T-value
DY	0.0053	1.2	0.00484	1.11	0.03402	1,48°	-0.0052	-0.49	0.02346	1,33°
PE	-0.00257	-0.63	-0.00618	-1,31°	-0.01517	-0.91	-0.01118	-0.85	0.00384	0.49

NORWEGIAN	1987-1999		2000-09:2012	
DATA	Beta	T-value	Beta	T-value
DY	0.01963	1,57°	-0.01699	-1.02
PE	0.01028	1.02	-0.0185	-1,56°

SWEDISH	1987-1999		2000-09:2012	
DATA	Beta	T-value	Beta	T-value
DY	0.0077	0.54	-0.01185	-0.61
PE	-0.00363	-0.41	-0.00064	-0.06