SCHOOL OF ECONOMICS AND MANAGEMENT
Lund University

Bachelor Thesis
Financial Economics
15 ECTS credits
February 2006

A multi-factor model for the Swedish stock market

- Is the market efficient or not?

Supervisor:
Erik Norrman

Authors:
Pär Karlsson
Patrik Nilsson

Abstract

In 1996, Robert A. Haugen and Nardin L. Baker published a study where they had made a multi-factor model based on micro-variables on the American stock market between 1979 and 1993. In testing the model on the examined period they made an average annual return of 30,9 percent. They also tested the model on several other markets around the world with similar results. This study is a somewhat simpler remake of that study, but for the Swedish stock market from 1999 to 2005 instead. By testing this model, we wanted to see if the Swedish stock market of today is efficient or not. We collected financial data via the SIX Trust database and used Eviews to make regressions with the aim of discerning which variables are significant over time. By doing so, we hoped to make the model work and thereby find ways of constructing portfolios with possible excess return due to an inefficient market. Our results show, however, that none of the variables was significant over time and, hence, the model does not work with acceptable results. Consequently, our conclusion is that the Swedish stock market *is*, in fact, efficient.

Keywords

Efficient market hypothesis, regression, multi-factor model, statistical theory, prediction, fundamental variables

Table of content

Introduction

In this chapter, we will describe the background, the study we are basing this essay on, some relevant earlier research and the goals and aims of this essay.

1.1 Background

Predicting returns to stocks along with prediction of risks are two areas that have been subject to intensive research ever since Harry Markowitz founded the principals of modern portfolio theory in his paper titled "Portfolio selection" in 1952[1]. Markowitz work was followed by the development of the capital asset pricing model (CAPM) by Sharpe (1964), Lintner (1965) and Mossin (1966)[2]. This model came to reign as the primary model in finance for close to 15 years. The first persons to really delve into the possibilities of a multi-factor model were Barr Rosenberg and Walter McKibben. They first proposed a model for prediction of systematic and specific risk in stocks in an article from 1973[3]. The following year Rosenberg presented another article regarding multi-factor models, this time with the prediction of returns in mind[4]. The multi-factor model was later incorporated into the asset pricing theory (APT) which was developed and introduced by Steve Ross in 1976[5]. This essay will focus on a multi-factor model for predicting returns and will be based on a study made by Haugen and Baker in 1996[6].

---

1 Markowitz, Harry, 1952, *Portfolio Selection*, The Journal of Finance, Vol. 7, No. 1, pp. 77-91
2 Haugen, Robert A., 2001, *Modern investment theory,* Prentice Hall, p. 2
3 Rosenberg, Barr, McKibben, Walt, 1973, *The prediction of Systematic and Specific Risk in Common Stocks,* The Journal of Finance and Quantitative Analysis, Vol. 8, No. 2, pp. 317-333
4 Rosenberg, Barr, 1974, *Extra-market components of covariance in security returns,* The Journal of Financial and Quantitative Analysis, Vol. 9, No. 2, pp. 263-274
5 Haugen, Robert A., 2001, *Modern investment theory,* Prentice Hall, p. 2
6 Haugen, Robert A., Baker, Nardin L, 1996, *Commonality in the determinants of expected stock returns*, Journal of financial Economics, Vol. 41, No. 3, pp. 401-440

1.2 Problem definition

The main obstacle in creating a multi-factor model is finding the relevant variables and the corresponding sensitivities of stock returns to these numbers. There are many advanced statistical models for creating multi-factor models available today. Haugen, however, showed in his paper that the use of a relatively simple model built on multiple regressions could yield impressive results when tested on large markets. Our point of interest is whether this model can give good predictions over time on a small market with a lesser statistical base, like the Swedish, and by this see if the market of today is efficient or not.

1.3 Purpose

Our goal is to test if the hypothesis of efficient markets is in play in Sweden. We will do so by replicating the basics of an earlier study made on several of the world's largest stock markets which results showed that the market was inefficient. By replicating the basics of this study, we aim to create a simple multi-factor model for the Swedish market and then benchmark it against the Swedish AFGX index over a series of years.

1.4 Delimitation

Considering our limited amount of time this essay will not be as detailed as the one we are basing it on. When creating the model, Haugen used an index containing 3000 stocks. We will have to limit ourselves to the stocks listed on the A-list and O-list (including Attract40) on the Stockholm Stock Exchange.

1.5 Previous research

1.5.1 Haugen, Robert A. & Baker, Nardin L.

This study, on which we have based our essay, was made by Haugen and Baker in 1996.[7] Their intention was to determine whether any good estimations of expected return could be done using a multi-factor model built on a multiple regression of a wide selection of variables. The variables used can be categorized into five groups: *risk factors, liquidity factors, price factors, growth factors* and *technical factors*. The characteristics of these groups will be described in a later chapter. The results are surprising when compared to the relative simplicity of the method they used. By adjusting their portfolio each month over 14 years, they realized a 30.9% average annual return for the US stock market (costs of trade excluded). Similar results were found when testing the model on the Japanese, French, British and German stock markets. A certain commonality amongst the more relevant factors was also observed across the markets. The correlation between the returns to the different factors across the markets was, however, quite low. This implies that although the tested markets share many of the factors with higher predictive power, each market values them differently.

1.5.2 Connor, Gregory

This paper by Gregory Connor[8] explores the differences between three approaches to predicting returns using multi-factor models. The approaches described are the macroeconomic, the statistical and the fundamental model. Connor concludes that the fundamental and the statistical models clearly outperform the macroeconomic model in terms of prediction of returns. Another interesting finding is that the addition of

[7] Haugen, Robert A., Baker, Nardin L, 1996, *Commonality in the determinants of expected stock returns*, Journal of financial Economics 41, pp. 401-439

[8] Connor, Gregory, 1995, *The three types of factor models: A comparison of their explanatory power*, Financial Analysts Journal, Vol. 51, No. 3, pp. 42-46

macroeconomic variables to a model already using fundamental variables does not add to the predictive power of the model. This implies that the fundamental variables also capture the risk characteristics of the macroeconomic variables thus rendering a combination of these two models more or less pointless.

1.5.3 Fama, Eugene F. & French, Kenneth R.

The capital asset-pricing model of Sharpe[9], Lintner[10] and Mossin[11] (CAPM) has shaped the way practitioners as well as academics think about average stock returns for decades. Their study implies that expected returns on securities are positively related to the market beta and that market betas suffice to describe the cross-section of expected return. Nevertheless, in the study made by Fama and French the authors conclude that the average stock returns are *not* positively related to market betas.[12] Furthermore, they show that beta does not seem to explain the cross-section of average stock return. Instead they mean that the univariate relations between average return and size, leverage, E/P (earnings to price) and book-to-market equity (BE) are very strong. In addition, the combination of size and book-to-market equity seems to absorb the roles of leverage and E/P in average stock returns (between 1963 and 1990 at least), which implies that if assets are priced rationally, the results suggests that stock risks are multidimensional. One dimension of risk is approximated by size (market equity, ME). The other dimension of risk is approximated by the ratio of the common equity to market equity (i.e. BE/ME; the ratio of book value of common equity to market equity). The conclusion of the study is that beta does not have the explanatory power suggested by CAPM.

[9] Sharpe, William F., 1964, *Capital asset prices: A theory of market equilibrium under conditions of risk*, The Journal of Finance, Vol. 19, No. 3, pp. 425-442

[10] Lintner, John, 1965, *The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets*, The Review of economics and statistics, Vol. 47, No. 1, pp. 13-37

[11] Mossin, Jan, 1966, *Equilibrium in a Capital Asset Market*, Econometrica: Journal of the Econometric Society, Vol. 34, No. 4

[12] Fama, Eugene F., French, Kenneth R., 1992, *The cross-section of expected stock returns*, The Journal of Finance, Vol. 47, No. 2, pp. 427-465

1.5.4 Jegadeesh, Narasimhan & Titman, Sheridan

This paper by Jegadeesh & Titman[13] explores the theory that investors that buy stocks that have performed well in the past and sell stocks that have performed poorly generates a significant positive rate of return for the next three to twelve months. However, half of the returns generated the first twelve months dissipate over the forthcoming two years. They also present evidence that the profits are not due to the systematic risk of the different trading strategies or to lead-lag effects as a result from delayed stock price movement due to common factors. Nevertheless, the results presented are consistent with delayed price reactions to firm-specific information. Two interpretations of the result can be made. One is that the transactions made by investors who buy past winners and sell past losers move prices away from their long run values temporarily and thereby cause prices to overreact. Another is that the market under-reacts to information about the short-term prospects of firms but overreact to information about their long-term prospects. The authors do however suggest that their evidence is overly simplistic and that a more sophisticated model of investor behavior is needed to fully explain the patterns shown.

1.5.5 Kothari, S. P., Sloan, Richard G. & Shanken, Jay

The study by Kothari, Sloan and Shanken[14] is a reexamination of the study made by Eugene F. Fama and Kenneth R. French[15] discussed earlier in this paper, using a somewhat different data source. In this study the authors uses Standard & Poor's (S&P) industry level data from 1947 to 1987 while Fama & French used the COMPUSTAT database. The conclusion drawn is, among others, that the Fama & French results are influenced by a combination of survivorship bias in the COMPUSTAT database and extraordinary periodic-specific performance of past "winner" stock as well as past

[13] Jegadeesh, Narasimhan, Titman, Sheridan, 1993, *Returns to buying winners and selling losers: Implications for stock market efficiency*, The Journal of Finance, Vol. 48, No. 1, pp. 65-91
[14] Kothari, S. P., Sloan, Richard G., Shanken, Jay, 1995, *Another look at the cross-section of expected stock returns*, The Journal of Finance, Vol. 50, No. 1, pp. 185-224
[15] Fama, Eugene F., French, Kenneth R., 1992, *The cross-section of expected stock returns*, The Journal of Finance, Vol. 47, No. 2, pp. 427-465

"loser" stock. Using S&P industry level data instead, we find that BE/ME is at best weakly related to average stock return. The effect shown is about 40 percent lower then the one obtained using COMPUSTAT. The authors also show that average returns *do* indeed reflect substantial compensation for beta risk, provided that beta is measured at the annual interval. However, the study still shows that not all variation of expected return can be explained by beta and that deviations from the linear CAPM risk-return trade-off are related to such variables as firm size, earnings yield, leverage, and the ratio of a firm's book value of equity to its market value.

1.6 Disposition

After this introductory chapter containing background, problem definition, purpose, previous research and delimitation, the essay will have the following disposition:

Theory: This chapter will describe some basic concepts, formulas and theories of portfolio selection and risk management. It will also discuss two widely used risk management methods, the Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT). Furthermore, it contains some basic statistical methods and concepts.

Method: The method chapter contains the statistical methods used as well as the operationalisation of the different variables and the gathering of data.

Empirical results and analysis: In this chapter, we will present the result of the empirical study together with the analysis based on the theoretical frame of reference presented earlier in the essay.

Conclusions: The chapter contains a presentation of the conclusions made from the empirical analysis as well as some of the authors' reflections over the results and some proposals for further research.

2. Theory

This chapter will describe some basic concepts, formulas and theories of portfolio selection and risk management. It will also discuss two widely used risk management methods, the Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT). Furthermore, it contains some basic statistical methods and concepts.

2.1 Portfolio Theory

The method of diversifying your assets by combining them into portfolios is based on portfolio theory and its variables return and risk. The objective of the investor is to maximize the return and at the same time minimize the risk. To minimize the risk, the investor should diversify by buying different kinds of assets at different times and to maximize the return, he should invest only in those assets with the highest expected return given his selected risk profile. It is not possible to dispose of all risk by diversifying because there is always some covariance between all assets returns. However, there is a way of finding the optimal portfolio based on its risk and return. Harry Markowitz wrote about this in his article "Portfolio Selection" already in 1952.[16]

2.1.1 Expected rate of return

The expected rate of return of an asset is calculated by multiplying its return by the returns probability to occur. The expected rate of return of a portfolio is calculated by multiplying the expected return of the different assets by its weight in the portfolio, called portfolio-weights.[17] The formula for the portfolios expected rate of return is:[18]

[16] Markowitz, Harry, 1952, *Portfolio Selection*, The Journal of Finance, Vol. 7, No. 1, pp. 77-91
[17] Markowitz, Harry, 1952, *Portfolio Selection*, The Journal of Finance, Vol. 7, No. 1, pp. 77-91
[18] Haugen, Robert A., 2001, *Modern investment theory*, Prentice Hall, p. 60

$$E(r_p) = \sum_{J=1}^{M} x_J E(r_J)$$

$E(r_P)$: *Expected rate of return on the portfolio*

M : *Number of assets in the portfolio*

x_J : *Portfolio weight for asset J*

$E(r_J)$: *Expected return for asset J*

2.1.2 Standard deviation and variance

The variance is a quadratic measure showing how much the average return deviates from the mean return. It's quadratic because this way we don't suffer the risk of negative and positive results counteracting each other. The formula for variance is:[19]

$$\sigma^2 = \frac{1}{N}\sum_{t=1}^{N}\left(r_t - \bar{r}\right)^2$$

σ^2 : *Variance*

N : *Number of months over which you take the sample*

r_t : *Rate of return at time t*

$\bar{r}$: *Mean return for the duration of the estimation-time N*

The standard deviation is a measure of how much the rate of return, by average, deviates from the mean rate of return. To calculate the standard deviation, take the square root of the variance.

$$\sigma = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(r_t - \bar{r}\right)^2}$$

[19] Haugen, Robert A., 2001, *Modern investment theory*, Prentice Hall, p. 34

2.1.3 Covariance and correlation

Covariance is a measure of how the assets covariate during given conditions. It measures the linear relationship between the assets returns. When producing a well-diversified and efficient portfolio it's not only important to invest in many assets, but also very important to make sure that the assets don't covariate too much. The lower the covariance, the lower the risk. The covariance is calculated as:[20]

$$\sigma_{i\,j} = \frac{1}{N-1}\sum_{t=1}^{N}\left[\left(r_{i,t}-\overline{r_i}\right)\left(r_{j,t}-\overline{r_j}\right)\right]$$

$\sigma_{i,j}$: *Covariance between asset i and asset j*

$r_{i,t}$: *Rate of return on asset i at time t*

$\overline{r_i}$: *Average rate of return on asset i*

$r_{j,t}$: *Rate of return on asset j at time t*

$\overline{r_j}$: *Average rate of return on asset j*

Whereas the covariance measure stretches from minus infinity to plus infinity, the correlation coefficient only stretches from -1 to +1, even though they measure the same thing. If two assets have a correlation of +1, they are perfectly correlated with each other. This means that they react the same way on the market. If two assets have a correlation -1, they are perfectly un-correlated which means they react contrarily each other. When minimizing the portfolio risk one wants to have as un-correlated assets as possible in the portfolio. The formula for correlation is:[21]

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$

[20] Haugen, Robert A., 2001, *Modern investment theory*, Prentice Hall, p. 36
[21] Haugen, Robert A., 2001, *Modern investment theory*, Prentice Hall, p. 40

ρ_{ij} : *Correlation coefficient between asset i and asset j*

σ_{ij} : *Covariance between asset i and asset j*

σ_i : *Standard deviation for asset i*

σ_j : *Standard deviation for asset j*

2.1.4 Minimum variance and efficient sets

Given several available assets, it's possible to combine these into numerous different portfolios. Moreover, given the desired level of risk, investors always want to invest in portfolios with the highest possible return and these combinations form a bullet shaped curve called the minimum variance set.[22] All portfolios on the minimum variance set represent the combination of assets with the lowest standard deviation of all available assets. The edge of the bullet, which divides the minimum variance set in two, has the lowest standard deviation possible and is called the Minimum Variance Portfolio (MVP)[23]. The top half of the bullet is called the efficient set, because they meet the following criterion: "*Given a particular level of standard deviation, the portfolios in the efficient set have the highest attainable expected rate of return.*"[24]

2.1.5 CAPM

The Capital Asset Prices Model (CAPM) was independently but simultaneously developed by William Sharpe[25], Jan Mossin[26] and John Lintner[27] in the sixties and is perhaps the most widely used model for calculating asset and portfolio risk. It's based on

[22] Haugen, Robert A., 2001, *Modern investment theory*, Prentice Hall, p. 81
[23] Haugen, Robert A., 2001, *Modern investment theory*, Prentice Hall, p. 82
[24] Haugen, Robert A., 2001, *Modern investment theory*, Prentice Hall, p. 82
[25] Sharpe, William F., 1964, *Capital asset prices: A theory of market equilibrium under conditions of risk*, The Journal of Finance, Vol. 19, No. 3, pp. 425-442
[26] Mossin, Jan, 1966, *Equlibrium in a Capital Asset Market*, Econometrica: Journal of the Econometric Society, Vol. 34, No. 4, pp. 768-783
[27] Lintner, John, 1965, *The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets*, The Review of economics and statistics, Vol. 47, No. 1, pp. 13-37

the homogenous expectation hypothesis and investors, who are all characterized as risk-averse, are supposed to maximize their utility function based only on the mean and variance of their future wealth. This prompts them, by definition, to hold the same perfectly diversified portfolio at all times and to consider only the associated risk of this portfolio. Thus the beta determines each stock's return.

The Capital Market Line (CML), also called the risk-reward trade-off line, shows the tangency portfolio (i.e. the optimal portfolio) of risky and risk-less assets.[28] It's a straight line connecting the risk-free asset and the market portfolio. CAPM says that in equilibrium, CML represents the best combination of assets available to all investors. The CML formula is:[29]

$$E(r_i) = r_f + \frac{E(r_M) - r_f}{\sigma_M}\sigma_i$$

$E(r_i)$: *Expected rate of return for asset i*

r_f: *Risk-free rate*

$E(r_M)$: *Expected rate of return for the market portfolio*

σ_M: *Standard deviation for the market portfolio*

σ_i: *Standard deviation for asset i*

The CML shows that the larger the standard deviation of the return the larger the equilibrium expected return, and hence, the larger the risk. However, in CAPM the standard deviation generally does not measure the risk of assets. We therefore use the term beta instead. Beta measures the marginal contribution of the assets return to the standard deviation of the market portfolio. If the asset has a beta of 1 it varies just like the market portfolio. If the beta is <1 it varies less and if the beta is >1 it varies more then the market portfolio. The formula for beta is:[30]

---

[28] Bodie, Zvie, Merton, Robert C., 2000, *Finance*, Prentice Hall, p. 345
[29] Bodie, Zvie, Merton, Robert C., 2000, *Finance*, Prentice Hall, p. 345
[30] Bodie, Zvie, Merton, Robert C., 2000, *Finance*, Prentice Hall, p. 348

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2}$$

β_i : *Beta for asset i*

σ_{iM} : *Covariance between asset i and the market portfolio*

σ_M^2 : *Variance for the market portfolio*

In equilibrium, the Security Market Line (SML) shows the risk premium on the market portfolio, hence the assets beta times the risk premium on the market portfolio. The formula for SML is:[31]

$$E(r_i) = r_f + \beta_i \lfloor E(r_m) - r_f \rfloor$$

$E(r_i)$: *Expected return on asset i*
r_f : *Risk free interest rate*
β_i : *Beta for asset I; i.e. the asset risk*
$E(r_m)$: *Expected return on the market portfolio*

However, the CAPM is a very static measure because all investments are restricted to one period. This implicitly assumes stability in the firm's financial structure and even a small change in this can modify the share capital risk and thus the expected return.

2.1.6 APT

Another model for asset pricing is the Arbitrage Pricing Theory (APT), first introduced by Ross[32], which takes multiple micro- and macroeconomic variables into account. Such models are called multifactor models. The APT multifactor model is not an equilibrium model and does not try to explain the stocks expected return. Instead, it rests on the

[31] Bodie, Zvie, Merton, Robert C., 2000, *Finance*, Prentice Hall, p. 348
[32] Haugen, Robert A., 2001, *Modern Investment Theory*, Prentice Hall, p. 255

hypothesis that stock prices are influenced by several non-correlated factors. The expected return is a linear combination of each factor's beta and each beta corresponds to a systematic risk measure. The multifactor formula for expected return is:[33]

$$r_{J,t} = A_J + \beta_{1,J} I_{1,t} + \beta_{2,J} I_{2,t} + \ldots + \beta_{n,J} I_{n,t} + \varepsilon_{J,t}$$

$r_{J,t}$: *Rate of return on asset J in a given period of time, t*

A_J : *Company specific rate of return conditioned on the fact that all factors take on a value of zero*

$\beta_{n,J}$: *Beta for factor n on stock J*

$I_{n,t}$: *Value of factor n in the given period of time, t*

$\varepsilon_{J,t}$: *Residual of stock J in the given period of time, t*

For the APT we assume that there are no restrictions in short selling and that there are an unlimited number of securities. Furthermore, we assume that the general relationship between expected return and factor risk is approximately linear. Therefore, given this relationship, we can construct any riskless portfolio containing any of the stocks available and still have the same expected return. The equation for the risk-expected return is given by [34]

$$E(r_J) \approx E(rZ) + \sum_{I=1}^{n} \lambda_I \beta_{I,J}$$

$E(r_J)$: *Expected rate of return for asset J*

$E(r_Z)$: *Expected rate of return for the risk free asset*

λ_I : *Factor price of factor I*

$\beta_{I,J}$: *Beta for factor I on asset J*

[33] Haugen, Robert A., 2001, *Modern Investment Theory*, Prentice Hall, p. 255
[34] Haugen, Robert A., 2001, *Modern Investment Theory*, Prentice Hall, p. 260

The equation is an approximation since the relationship is not completely linear all the time. This means that unlimited arbitrage opportunities can be possible, which is the central message of the APT. However, the APT gives us no direction as to the choice of factors.[35]

2.2 Statistical theory

2.2.1 Regression analysis

Regression analysis is used to determine a relationship between the dependent variable and the explanatory variables. It's important to understand that the results are approximations only, but regression analysis can be a helpful tool to prognosticate the effect of the different variables.

In a simple regression, we study the relationship between the dependent variable, y, and one explanatory variable, x. The formula used is:[36]

$$y = \alpha + \beta x + \varepsilon$$

α : The intercept, i.e. where x = 0

β : The beta, i.e. how much y changes when x changes by 1

ε : The residual, i.e. the change in y the can't be explained by the equation

If a variable is presumed to be dependent on more than one explanatory variable, a multiple model is used. Furthermore, this model lets us separate and distinguishes the sway of the different explanatory variables. The multiple regression formula is:[37]

[35] Haugen, Robert A., 2001, *Modern Investment Theory*, Prentice Hall, p. 266
[36] Haugen, Robert A., 2001, *Modern Investment Theory*, Prentice Hall, p. 135
[37] Haugen, Robert A., 2001, *Modern Investment Theory*, Prentice Hall, p. 145

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$$

α : The intercept, i.e. where $x = 0$

β_1 : The Beta for variable 1, i.e. how much y changes when x_1 changes by 1

β_2 : The Beta for variable 2, i.e. how much y changes when x_2 changes by 1

β_n : The Beta for variable n, i.e. how much y changes when x_n changes by 1

ε : The residual, i.e. the change in y the can't be explained by the equation

2.2.2 Multi-colinearity

If a variable is highly correlated with another variable, which already is included in the model, multi-colinearity prevails. This means that there are two variables with approximately the same effect on the dependent variable. To avoid this, the correlation coefficient is calculated. If the correlation coefficient is, or close to, +1 or -1 you should only use one of them or exclude both.

2.2.3 Autocorrelation

When a regression model, like the multi-factor model above, suffers from correlation across time between the different factors it is said to be autocorrelated. This means that if there is a false in one period it is likely that there will be a false in the next period as well. In this study, if autocorrelation occurs in for example month twelve of the trailing mean it will have impact on the following twelve months. The degree of autocorrelation of a time series is calculated as the correlation of the function against a time shifted version of itself. The coefficient is always between +1 to -1 and the formula is as follows:

$$R(k) = \frac{E\left[\left(x_t - \bar{x}\right)\left(x_{t+k} - \bar{x}\right)\right]}{\sigma^2}$$

$R(k)$: Coefficient of autocorrelation

k: The time series lag

x_t: Value of factor X at time t

$\overline{x}$: Mean value of factor x

σ^2: Variance

2.2.4 Heteroskedasticity

When using statistical methods, such as OLS, several assumptions are made. One of these is that the error terms in the model have constant variances. If this is not true, we have heteroskedasticity. When applying OLS to such a model the estimated variances become biased estimator of the true variances. The values will be either over- or underestimated. To test for heteroskedasticity in a model White's test, developed in 1980 by Halbert White, is used.[38]

2.2.5 Significance

When making a regression, a value for the regression coefficients significance emerges. This is done by doing a hypothesis test where we test whether the null-hypothesis, i.e. that the regression coefficient is equal to nil, is true. Based on that test, we either accept or reject the null-hypothesis. If the null-hypothesis is rejected we accept the alternative hypothesis, which means that the value of the regression coefficient in the regression analysis is significant (i.e. the explanatory variable has significant effect on the dependent variable).[39]

Two different measures can be used to congregate the level of significance of the regression coefficient, either a t-value or a p-value. With the t-value, a table of normal

[38] White, Halbert, 1980, *A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity*, Econometrica: Journal of the Econometric Society, Vol. 48, No. 7, pp. 817-838
[39] Körner, Svante, 2001, *Statistisk Dataanalys*, Studentlitteratur, pp. 254-256

distribution is used to congregate the significance. A rule of thumb is that if fewer then 30 observations are present, a special t-distribution table is used instead.[40] Statistical programs, such as Eviews or Minitab, normally generate the p-value. If we for example would like to test the null-hypothesis on a level of significance of 5 percent the p-value should be less than 0,05 or the null-hypothesis must be accepted.[41]

2.2.6 Coefficient of determination

The coefficient of determination, denominated R^2, is a measure of the explanatory power, i.e. how much of the result that can be explained by the regression. The explanatory power varies between one and nil and is expressed in percent. In a multiple regression, R^2 is the quota between the part of the variation explained by the regression and the total amount of variation. If the coefficient of determination is 100% all of the variation in the dependent variable can be explained by the explanatory variable. When using a single regression model, the coefficient of determination is the square sum of the correlation of the two variables.

Adjusted R^2 is a modification of R^2 that adjusts for the number of terms in the model. R^2 always increases when a new term is added to the model, whilst adjusted R^2 only increases if the new term improves the model more than would be expected by chance.

[40] Körner, Svante, 2001, *Statistisk Dataanalys*, Studentlitteratur, pp. 211-213
[41] Körner, Svante, 2001, *Statistisk Dataanalys*, Studentlitteratur, pp. 269-270

3. Method

The method chapter contains the statistical methods used as well as the operationalisation of the different variables employed and the gathering of data.

3.1 Methods of estimation

When building a multi-factor model for prediction of returns we have to estimate the sensitivity of the returns to each factor. This is done using an ordinary least squares (OLS), cross-sectional, multiple regression analysis which is run during a series of years using monthly data for our range of variables. This results in monthly estimates of the payoffs to each factor. The regression model looks as follows:

$$r_{j,t} = \sum_i \widehat{P}_{i,t} * F_{j,i,t-1} + u_{j,t},$$

$r_{j,t}$: *Rate of return to stock j in month t*

$\widehat{P}_{i,t}$: *Regression coefficient or exposure to factor i in month t*

$F_{j,i,t-1}$: *Value of factor i for stock j at the end of month t-1*

$u_{j,t,}$: *Unexplained component of return*

When later testing the model we will use a slightly modified formula to predict the returns:

$$E(r_{j,t}) = \sum_i E(\widehat{P}_{i,t}) * F_{j,i,t-1},$$

$E(r_{j,t})$: *Expected rate of return to stock j in month t*

$E(\hat{P}_{i,t})$: *Exposure to factor i in month t (the arithmetic mean of the estimated payoff over the trailing 12 months)*

$F_{j,i,t-1,}$: *Value of factor i for stock j based on information available at the end of month t-1*

The trailing 12-month mean enables us to take advantage of possible variations in sensitivities over time.

3.2 Firm characteristics

The factors used in the regression model can be divided into five different groups of variables: factors related to *risk, liquidity, pricing, growth potential and price history.*

3.2.1 Risk factors

The focus on risk in widely used models like the CAPM and the APT leads us to believe that risk factors should have a certain impact on the returns in a multi-factor model. We will, however, due to limited time exclude risk factors relying on macro variables. We feel that our study will not suffer from this exclusion since Haugen's[42] study proved macro variables to have low predictive power with regard to returns. Connor[43] drew the same conclusion in his study regarding different factor models. What's more, he reaches the conclusion that the risk characteristics of macro variables are captured quite well by fundamental firm characteristics. Thus, we feel quite safe excluding these variables. We expect the overall payoff to risk factors to be positive. The risk factors included are as follows:

[42] Haugen, Robert A., Baker, Nardin L, 1996, *Commonality in the determinants of expected stock returns*, Journal of financial Economics 41, pp. 401-439
[43] Connor, Gregory, 1995, *The three types of factor models: A comparison of their explanatory power*, Financial Analysts Journal, Vol. 51, No. 3, pp. 42-46

- *Market beta*; trailing 12 months regression on excess return
- *Volatility of return*; trailing 36 months average
- *Earnings risk*; standard error of trailing 36 months
- *Solidity*; Shareholders' equity to total assets
- *Solidity trend*; trailing 36 months average
- *Times interest earned*; net operating income to total interest charges
- *Times interest earned trend*; trailing 36 months average
- *Yield variability*; 36 months trailing volatility in earnings, dividend and cash flow yield

3.2.2 Liquidity factors

Stock liquidity has impact on investor decisions in the sense that a highly liquid stock makes for quicker entries and exits than a stock with low liquidity. Liquidity has a certain bearing on the risk of a stock considering that it, for example, represents the possibilities of quickly ridding a portfolio of bad positions. Liquidity should therefore have a negative impact on the return with low liquidity stocks giving greater premiums.
The factors included are:

- *Market capitalization*; current market price times the most recently available number of outstanding share
- *Market price per share*
- *Trading volume to market capitalization*; trailing 12 months average monthly trading volume to market capitalization
- *Trading volume trend*; trailing 36 months average

3.2.3 Factors of pricing

These factors relate the current market price to various accounting figures and in doing so determines whether a stock is relatively cheap or pricy. Higher cash-flows relative to market price should, for example, result in greater expected returns. The factors we have included are:

- *Earnings to price*; most recently available four quarters, earnings to current market price
- *Earnings to price trend*; trailing 36 months average
- *Book to price*; most recently available book value to current market price
- *Book to price trend*; trailing 36 months average
- *Dividend to price*; most recently available dividend to current market price
- *Dividend to price trend*; trailing 36 months average
- *Cash flow to price*; most recently available ratio of earnings per share to current market price
- *Cash flow to price trend*; trailing 36 months average
- *Sales to price*; most recently available four quarters, total sales per share to current market price
- *Sales to price trend*; trailing 36 months average

3.2.4 Factors of growth potential

In line with the assumptions of Haugen, we expect firms with high current earnings to have good possibilities of future growth. Therefore, we include several measures of profitability in the model:

- *Profit margin*; net operating income to total sales
- *Profit margin trend*; trailing 36 months average
- *Return to assets*; net operating income to total assets

- *Return to assets trend*; trailing 36 months average
- *Return to equity*; net income to total book value of total equity capital
- *Return to equity trend*; trailing 36 months average
- *Earnings growth*; trailing 36 months in earnings per share

3.2.5 Factors of historical returns

Previous research have shown observable trends based on prior performance in stocks.[44] Therefore, several trailing averages measuring excess returns relative a suitable index will be included:

- *Excess return*; relative to the AFGX index in previous one month
- *Excess return*; relative to the AFGX index in previous two months
- *Excess return*; relative to the AFGX index in previous three months
- *Excess return*; relative to the AFGX index in previous six months
- *Excess return*; relative to the AFGX index in previous twelve months

3.3 Data gathering

The period we have examined stretches from January 1999 to December 2005. We gathered the data through SIX Trust, a Swedish supplier of financial information. Only companies with full data coverage during the entire period of 1994-2005 were included. The gap between the examined period 1999-2005 and the data period 1994-2005 is due to the lag, the trend-variables and the trailing mean employed in the estimations of return. The companies included are all members of the A-list or the O-list (including Attract40). The accounting key ratios we used are based on yearly financial reports. Ideally, we would have liked to use quarterly data but unfortunately, SIX Trust doesn't

[44] Jegadeesh, Narasimhan, Titman, Sheridan, 1993, *Returns to buying winners and selling losers: Implications for stock market efficiency*, The Journal of Finance, Vol. 48, No. 1, pp. 65-91

provide key ratios on quarterly basis. We didn't have access to any other systems with acceptable data coverage during this period. This forced us to make do with yearly accounting data. We acknowledge the fact that this might pose problems later on considering that the market bases its valuation on the latest available data. The model, if successful, might therefore show declining performance during the year as more recent data is published.

3.4 Regressions

The regressions were made using Eviews, a statistical software primarily focused on forecasting and estimation. We made 96 monthly regressions containing data from 64 companies during the period of 1998-2005. The reason that we used only 64 companies was that we wanted a long period of examination to capture the market under different conditions. We applied a 2-month reporting lag to the variables based on accounting. This is due to the fact that most annual financial reports are released during February and therefore, as a group, cannot be used for estimations until March. All other variables were used as reported at the end of each month. The t-values received for the estimated coefficients each month were averaged for the entire 96 months to make it possible to distinguish which variables were statistically significant during the entire period of regression. The variables were then sorted after the absolute value of the t-statistics to determine which factors had the largest bearing on the expected rate of return.

3.5 Sources of bias

Because of the fact that the Swedish stock market is quite small (Haugen and Baker used over 3000 stocks) our model might be subject to a certain amount of bias. The possibility of bias is also due to the fact that we had to exclude large parts of the examined market because of lack of historical key data for some stocks. Furthermore, the lack of historical quarterly reports in the information systems used have forced us to use only yearly key data. Our hopes, though, are that the sample will provide an acceptable representation of the market and that the information loss thereby won't give to great of a bias.

In their paper, Haugen and Baker present several possible sources of bias proposed in a variety of earlier reports. The biases mentioned are *survival bias, look-ahead bias, bid-ask bounce* and *data-snooping bias*.

3.5.1 Survival bias

According to the authors, survival bias is a bias created by excluding firms that go inactive during the estimation period. The solution Haugen and Baker proposed is filling the spots of the firms that have gone inactive with an average based on the active firms. Considering our relatively small sample of firms, we will only use firms that were active during the entire period of estimation.

3.5.2 Look-ahead bias

This bias occurs when predictions use data that wasn't actually available at the date of forecast. For instance using accounting figures that was reported for the first quarter of the year to predict first quarter returns would yield such a bias. The actual figures for the first quarter usually aren't available until several weeks after the actual period. We try to minimize this bias by applying a two-month reporting lag to the accounting figures.

3.5.3 Bid-ask bounce

Stocks trade at bid or ask prices. Assume that the market value of a stock doesn't change during the month t, but the last trade of that month is at the bid price. The market value is still constant during month $t+1$, and there is an even chance that it will close at the ask price at the end of month $t-1$ or $t+1$. Thus, assuming no change in the bid-ask spread, the measured return will be either zero or negative for period t, and either zero or positive for period $t+1$. Therefore, returns measured over closing prices can appear to be negatively auto-correlated even when they are not. Thus, bid-ask bounce can lead to the false conclusion that the last periods return has predictive power even when successive stock returns are completely uncorrelated.

3.5.4 Data-snooping bias

This bias occurs when researchers first examine the results of other studies of a database, then build predicting-models employing good and promising factors based in these results, and then test the power of their model on the same database. Since most researchers in, for example, a country or a specific area of research uses the same databases for research and then publish and discuss their results in the same forums, it makes it a difficult problem to address. Nevertheless, it can be partially solved by employing data from market that have not yet been extensively studied or by predicting using periods of time that are new to analysis.

4. Empirical results and analysis

In this chapter, we will present the result of the empirical study together with the analysis of the results found.

4.1 Results

4.1.1 Regression results

Most of the variables used in the estimation showed significant t-values in the monthly regressions from time to time. However, large fluctuations in size as well as between positive and negative values were observed and none of the variables showed significance at the 95% level of confidence when we averaged the absolute values across the 96 months. Running the regressions with a function adjusting for heteroskedasticity didn't yield better results and the differences were on average very small. To illustrate the fluctuations of the variables we calculated an average of the t-values of the regressions. These can be found in table 1 in the appendix along with the average of the absolutes and the average coefficients with the average of their absolutes. The R^2, the adjusted R^2 and the standard error of regression are also included. The fact that we based the accounting figures on the annual financial reports instead of on quarterly reports might have had an impact on the results. Comparing the results from the months directly following the annual report with the results from the months later during the year, however, showed no greater difference in t-values. The reason that we found no differences might have been that the market only takes account of the latest available data and therefore would disregard the annual report in favor of the 4th quarter report.

To totally exclude possibilities of bias stemming from the use of old accounting data we ran a regression over the same period, this time excluding the accounting variables. The result of this regression was, however, similar to the one using all variables with no variables reaching significance over time. The results of this regression are included in table 2 in the appendix.

One possible other source of bias stems from our use of a reference index that assumes re-investment of dividends when calculated. Our use of this index was due to the fact that there were no appropriate indexes excluding reinvestment of dividends available. Our option was to calculate a value-weighted index ourselves. This option was, however, not feasible considering the small amount of stocks used and the fact that Ericsson with its huge dominance in value was included. Basing a value-weighted index on this selection of stocks would have created an index largely dependant on the returns of the Ericsson stock. One way of getting around this problem would have been to exclude Ericsson entirely but we felt that this in turn would have created a bias considering the large impact Ericsson has on the Swedish stock market.

To examine whether the possibility of bias from the use of this index could have affected our results we ran another regression. This time we used the market beta and excess return factors only. These are the only factors associated with the index in the regression. The excess return variables have been proved both by Haugen & Baker[45] and Jegadeesh & Titman[46] to be useful in prediction of returns during the 60's, 70's and 80's and the market beta has long been known to have positive effects on expected return although its uses in this aspect has been questioned during later years. To get a wider selection of companies and thereby the possibility of creating a feasible index we used the period of 2003-2005. The index was weighted according to market value and comprised of 236 stocks from the A-list, the O-list and the Attract40. The regression using 236 stocks over 36 months returned results similar to the earlier regressions. The

[45] Haugen, Robert A., Baker, Nardin L, 1996, *Commonality in the determinants of expected stock returns*, Journal of financial Economics, Vol. 41, No. 3, pp. 401-440

[46] Jegadeesh, Narasimhan, Titman, Sheridan, 1993, *Returns to buying winners and selling losers: Implications for stock market efficiency*, The Journal of Finance, Vol. 48, No. 1, pp. 65-91

only variable showing somewhat consistent results with high t-values was the market beta. On an average, the returns to this variable were positive although the returns were negative in 8 out of 36 months. The average t-value of this variable during the period was 2,0 while the average of the absolute t-value amounted to 4,1. This result differs from the result of the earlier regressions using the AFGX index. The predictive power of the excess returns, however, still seems weak. Therefore, we conclude that the use of the AFGX probably did not create any significant bias in the earlier regression. The results of these regressions are included in table 3 in the appendix.

4.1.2 Benchmark

Due to the inconsistency of the t-values and coefficients, we decided not to create a benchmark.

4.2 Analysis

Several possible explanations as to why we didn't find results similar to Haugen's and Baker's are available. Most of them have to do with the efficiency of the markets then and now. The period examined by Haugen and Baker stretches from 1979 to 1993 while our period stretches from 1999 to 2005. There is one major thing that separate these two periods: The Internet. The internet had its commercial break-through in Sweden during mid to late 90's and during this period it became a common household-tool for obtaining information. In other words, it made it possible for a larger mass of potential investors to gain fast and easy access to vast amounts of information, there among financial data. Larger amounts of data spread among a wider selection of investors should make for a more efficient market as long as the data, on average, is handled correctly. Another product of the extended use of the internet is the online brokerage firms. These make it easier for the average investor to, on his own, buy and sell stocks and other securities without having to deal with banks and other intermediaries. Internet brokers have gained larger and larger market shares during the last years along with the populations increased interest in stocks.[47] Both the number of trades and the value of the trades made on the Stockholm Stock Exchange have increased with approximately 500% since mid 90's.[48] Markets with higher activity should faster eliminate perceived errors in valuation, which in turn would lead to a more efficient market.

[47] Stockholm Stock Exchange Monthly Report, December issue, 1999-2005, http://domino.omgroup.com/www/xsse-statistik.nsf/(html)/ReportNew
[48] OMX Exchanges Annual Statistics 2004, http://domino.omgroup.com/.../arsstatistik_omx_exchanges_2004.pdf

5. Conclusions

The chapter contains a presentation of the conclusions made from the empirical analysis as well as some of the authors' reflections over the results and some proposals for further research

5.1 Conclusions

The conclusions drawn are that the Swedish market of today is largely efficient. The only variable that shows any kind of consistent statistical significance when regressed against the return is the market beta. As expected the R^2 and the adjusted R^2 sinks with the number of variables while the standard error of regression increases. As shown in our account of previous research, different results appear at different times depending on methods used, period of examination, statistical base etc. Further research on the subject might therefore yield different results. Suggestions of possible improvements are presented in the following section.

5.2 Further research

For further research, the use of quarterly data is an essential improvement. Use of a wider statistical base with a larger number of companies might also contribute to a "better" result. This can be done using either a shorter period of examination or by using a larger market like, for example, the British, German or French market.

6. References

6.1 Published sources

Bodie, Zvie, Merton, Robert C., 2000, *Finance*, Prentice Hall

Connor, Gregory, 1995, "The three types of factor models: A comparison of their explanatory power", *Financial Analysts Journal*, Vol. 51, No. 3

Fama, Eugene F., French, Kenneth R., 1992, "The cross-section of expected stock returns", *The Journal of Finance*, Vol. 47, No. 2

Haugen, Robert A, 2001, *Modern investment theory,* Prentice Hall

Haugen, Robert A, Baker, Nardin L., 1996, "Commonality in the determinants of expected stock returns*", Journal of financial Economics* , Vol. 41, No. 3

Jegadeesh, Narasimhan, Titman, Sheridan, 1993, "Returns to buying winners and selling losers: Implications for stock market efficiency", *The Journal of Finance*, Vol. 48, No. 1

Kothari, S. P., Sloan, Richard G., Shanken, Jay, 1995, "Another look at the cross-section of expected stock returns", *The Journal of Finance*, Vol. 50, No. 1

Körner, Svante, 2001, *Statistisk Dataanalys*, Studentlitteratur

Lintner, John, 1965, "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets", *The Review of Economics and Statistics*, Vol. 47, No. 1

Markowitz, Harry, 1952, "Portfolio Selection", *The Journal of Finance*, Vol. 7, No. 1

Mossin, Jan, 1966, "Equlibrium in a Capital Asset Market", *Econometrica: Journal of the Econometric Society*, Vol. 34, No. 4

Rosenberg, Barr, McKibben, Walt, 1973, "The prediction of Systematic and Specific Risk in Common Stocks", The *Journal of Finance and Quantitative Analysis*, Vol. 8, No. 2

Rosenberg, Barr, 1974, "Extra-market components of covariance in security returns", *The Journal of Financial and Quantitative Analysis*, Vol. 9, No. 2

Sharpe, William F., 1964, "Capital asset prices: A theory of market equilibrium under conditions of risk", *The Journal of Finance*, Vol. 19, No. 3

White, Halbert, 1980, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica: Journal of the Econometric Society*, Vol. 48, No. 7

6.2 Electronic sources

OMX Exchanges Annual Statistics 2004, http://domino.omgroup.com/.../arsstatistik_omx_exchanges_2004.pdf, 2006-01-05

SIX Trust, http://www.six.se/

Stockholm Stock Exchange Monthly Report, December issue, 1999-2005, http://domino.omgroup.com/www/xsse-statistik.nsf/(html)/ReportNew, 2006-01-10

7. Appendix

Table 1 – Regression results including all variables

Variable	Absolute t-value	T-value	Absolute coefficient	Coefficient
Market Beta	0,8969	-0,0855	0,0499	-0,0031
Volatility of return	0,9630	0,0099	2,9010	-0,1718
Earnings risk	0,8357	0,0547	0,0055	0,0006
Solidity	0,9119	0,3466	0,0021	0,0006
Solidity Trend	0,9386	-0,2276	0,0023	-0,0004
Times interest earned	0,9344	0,1414	0,0007	0,0001
Times interest earned Trend	0,9422	0,1368	0,0008	0,0000
Earnings yield	0,8986	0,0010	2,1215	0,3162
Dividend yield	0,9446	-0,2607	34,1467	3,8179
Cash flow yield	1,0474	0,0546	1,5937	0,2017
Market capitalization	0,8242	0,0352	0,0000	0,0000
Market price	0,7928	0,0793	0,0000	0,0000
Trading vol. to market cap.	0,8971	-0,1528	0,3151	-0,0540
Trading volume trend	0,9009	-0,0449	0,0000	0,0000
Earnings to price	1,0886	-0,4526	0,4282	-0,2025
Earnings to price Trend	1,0793	0,4235	0,6759	0,2774
Book to price	1,3340	-0,9702	0,1335	-0,1006
Book to price Trend	1,1558	0,7787	0,1046	0,0783
Dividend to price	0,9408	-0,1182	0,9606	-0,1851
Dividend to price Trend	1,0157	0,5267	1,3599	0,5040
Cash flow to price	1,0302	-0,0824	0,3021	0,0043
Cash flow to price Trend	0,9503	0,2937	0,3399	0,0858
Sales to price	1,1225	-0,6644	0,0248	-0,0140
Sales to price Trend	0,9918	0,5950	0,0231	0,0132
Profit margin	0,9805	-0,1860	0,1330	-0,0499
Profit margin Trend	0,9060	0,2210	0,1194	0,0367
Return to total assets	1,0142	0,1773	2,6256	0,3263
Return to total assets Trend	1,0116	0,1019	2,2221	0,3211
Return to equity	0,9689	0,2608	0,2768	0,0875
Return to equity Trend	0,9868	-0,0962	0,3557	-0,0654
Earnings per share Trend	0,9545	-0,2824	0,0045	-0,0013
1 month excess return	1,0192	-0,3411	0,2467	-0,0826
2 month excess return	0,8802	-0,1953	0,2194	-0,0423
3 month excess return	1,0130	-0,1230	0,1923	-0,0270
6 month excess return	1,0348	0,6695	0,3652	0,2099
12 month excess return	1,1522	-0,7605	0,0784	-0,0556

R-squared	0,7348
Adjusted R-squared	0,4034
S.E. of regression	0,0736

Table 2 – Regression results without accounting variables

Variable	Absolute t-value	T-value	Absolute coefficient	Coefficient
Market Beta	1,1806	0,1875	0,0336	0,0085
Volatility of return	1,2600	0,0306	1,7815	0,1684
Dividend yield	0,8325	-0,1885	11,3935	-2,9476
Market capitalization	0,7281	-0,0582	0,0000	0,0000
Market price	0,8224	-0,3742	0,0002	-0,0001
Trading vol. to market cap.	0,9887	-0,0249	0,3402	0,0023
Trading volume trend	0,8478	0,1068	0,0000	0,0000
Dividend to price	1,0048	-0,0083	0,6835	-0,0043
Dividend to price Trend	0,9109	0,2228	0,8470	0,1756
1 month excess return	1,2286	-0,4854	0,2571	-0,0769
2 month excess return	1,0054	-0,1748	0,1985	-0,0425
3 month excess return	0,8658	0,1402	0,1547	0,0236
6 month excess return	0,9578	0,1272	0,1038	0,0156
12 month excess return	1,1044	0,1638	0,0603	0,0066

R-squared	0,3528
Adjusted R-squared	0,1845
S.E. of regression	0,0932

Table 3 – Regressions results including index associated variables only

Variable	Absolute t-value	T-value	Absolute coefficient	Coefficient
Market Beta	4,0805	2,0046	0,0282	0,0206
1 month excess return	1,8274	-0,1638	0,1997	-0,0594
2 month excess return	1,3789	0,3815	0,1395	0,0497
3 month excess return	1,2291	-0,6029	0,0959	-0,0537
6 month excess return	1,3740	0,7793	0,0596	0,0402
12 month excess return	1,3293	-0,3773	0,0290	-0,0058

R-squared	0,0269
Adjusted R-squared	0,0057
S.E. of regression	0,1243